統計学者・数理工学者のための統計物理入門

格子スピン模型とマルコフ連鎖モンテカルロ法を中心にして

$Version\ 3.02$

1997 年 2 月 20 日

統計数理研究所

伊庭幸人

もくじ

1 平衡統計物理の基礎概念 **5**

1.1 ギブス分布 5

1.2 例 6

1.2.1 ２状態系 (イジングスピン) 6

1.2.2 イジング模型 6

1.2.3 調和振動子とその変形 7

1.3 熱力学的極限と巨視的変数 8

1.4 ゆらぎと感受率 9

2 マルコフ連鎖と緩和 **10**

2.1 マルコフ連鎖と定常分布 10

2.2 例 13

2.2.1 ２状態遷移模型 13

2.2.2 動的イジング模型 13

2.3 マルコフ連鎖モンテカルロ法 14

2.3.1 計算手段としてのマルコフ連鎖 14

2.3.2 狭義のメトロポリス法 16

2.3.3 Gibbs Sampler(熱浴法) 17

2.3.4 積分それ自体の計算 18

2.3.5 simulated annealing 19

3 相転移 **21**

3.1 ２次元イジング模型の相転移 21

3.2 相転移とダイナミクス 22

3.3 平均場近似 23

3.4 １次転移と２次転移 25

4 格子スピン模型 (格子上のマルコフ場モデル) の例 **26**

4.1 ポッツ模型 26

4.2 反強磁性イジング模型 28

4.3 ANNNI 模型 28

4.4 離散ガウス模型・ＳＯＳ模型 29

4.5 平面回転子模型 (古典 XY 模型) 29

5 ランダム系の統計物理 **31**

5.1 ランダム系の物理 31

5.2 スピングラス模型 31

5.3 ランダム磁場模型 32

5.4 ランダム系の平均場理論 33

5.4.1 空間ゆらぎを完全に無視した場合 33

5.4.2 ランダムネスの空間ゆらぎを残した場合 34

5.4.3 相互作用の範囲を無限大にした場合 34

6 付録: マルコフ連鎖モンテカルロ法の技術的側面 **37**
6.1 補足的注意 37
6.2 新しいアルゴリズム 41
6.2.1 非局所的ダイナミクス 41
6.2.2 ダイナミクスの学習 42
6.2.3 拡張アンサンブル 42

7 参考文献 **47**

お願い

Ver.2.x より小改訂しましたが、まだ相当数の誤りがあると思います。単純ミスや考え方の不備、アルゴリズムの間違いなど、どんどん指摘していただければ幸いです。また、質問や感想も歓迎します。内容の訂正は筆者のホームページに掲載する場合があります。筆者のメイルアドレスおよびホームページは下記の通りです。

E mail: iba@ism.ac.jp
URL: http://www.ism.ac.jp/~iba/

本稿は **ISM Research Memorandum No.635** として登録の予定です。

はじめに

近年，統計物理の方法や概念が統計学や情報処理の分野で注目を集めている．この総合報告は，それらの分野の研究者のために統計物理の基本的概念と用語を解説したものである．

大部分の統計物理の教科書には，当面不用な題材が多く含まれており，逆に，ランダム系の統計物理など最近の重要な進歩は書かれていない．そのため，他分野の研究者が知識を得るためには何冊ものレベルの違う本や総合報告を読む必要が生じ，面倒である．また，普通の教科書は数理科学一般に共通の知識を前提としないで書かれているため，そういう知識をすでにもっている一線の研究者 (working statistician, working engneer, etc.) にとっては冗長に過ぎると思われる．この総合報告では，平衡系の統計物理を，高次元の非ガウス分布についての概念や技術の集成としてとらえなおすことを試みた．この分野をひとつの “部品”(モジュール) として他分野に転用のきく形にまとめるのが狙いである．また，数理科学一般の常識と思われることは，省くか，簡単に触れるかだけにとどめた．逆に，統計物理としては特殊な知識でも，他分野の研究者からみて興味があると思われる話題は積極的に取り上げてある．用語については，主として物理の用語を用いた．これは，これを読んだ読者が自分で統計物理の論文を読んで理解できるようにするためである．やや脇道にそれると思われる部分は細字にし，通読が容易になるように配慮した．

統計物理のモデルとして取り上げたのは，“(格子) スピン模型” とよばれるモデルが主である．計算機時代の統計物理入門としては，これらがもっとも理解しやすいと考えたためである．また，これらに似たモデルが統計的情報処理において重要な役割を演じるためでもある．統計物理のモデルにはこれ以外にも重要なものがあり，統計的情報処理への応用に限っても，液体や気体のモデルが空間点過程の表現として有効なことが知られている．この意味では，平衡系の統計物理入門としての本稿は相当偏ったものである．

統計的情報処理の分野では，マルコフ連鎖モンテカルロ法が流行になっているため，それのみに絞った解説が望ましいと考える読者もいるかも知れない．それはそれで有意義と思われるが，狭義の方法論のみを採り入れるのは，やや行き過ぎた技術主義とも考えられる．“部品” として転用するとしても，周辺的知識や各分野特有の直観を全く切り離すのは必ずしも良いことではないであろう．また，物理の用語や概念を全く理解しないでは，物理の論文を読むことはできない．学際的交流という点からいっても，周辺的知識を含めた理解が望ましいと考えられる．

他分野の研究者を対象にした総合報告であるから，統計物理の専門家が読んで興味ある点はあまりないと思われるが，マルコフ連鎖モンテカルロ法の “新しいアルゴリズム” を扱った部分 (付録の後半) は例外である．逆に専門家でない読者にとってはこの部分はやや難解かもしれない．参考文献の指示は末尾にまとめた．ただし，“新しいアルゴリズム” の項については，本文中に直接指示した．全体に，参考文献としてオリジナルや重要な総説を網羅するような努力はしていない．統計学や情報処理への応用について，直接説明することは本稿の目的ではない．それらについては別の機会に報告するつもりであるが，とりあえず興味のある読者のために参考文献のところで応用に関する文献をあげておいた．

1　平衡統計物理の基礎概念

ここでは，平衡系の統計物理の基礎概念といくつかのモデルについて論じる．前者の内容のかなりの部分は，普通の統計学のなかにも含まれていると思われるが，その部分は用語の説明と理解されたい．以下では，系の内部状態を定めるものに対しても，外から与えるものに対しても “変数 (variable)” という言葉を用いた．混乱を避けるために，後者をとくに “外部変数 (external variable)” と呼ぶことにする．

1.1　ギブス分布

平衡統計物理の基礎は，温度 T の熱浴と接触している系が状態 α にある確率 P_α が $\exp(-E_\alpha/T)$ に比例するということにある．ここで E_α は状態 α のエネルギーである．これは，

$$Z = \sum_\alpha \exp(-E_\alpha/T) \tag{1}$$

を用いて，

$$P_\alpha = \frac{\exp(-E_\alpha/T)}{Z} \tag{2}$$

と書くこともできる．これをギブス分布 (Gibbs distribution) と呼ぶ．

規格化定数 Z を分配関数 (partition function)，その対数に温度をかけて負号をつけたもの

$$F = -T\log Z \tag{3}$$

を自由エネルギーという．

温度 T が大きいときはエネルギーの値によらずどの状態もほぼ同じ確率で出現するが，T が小さくなると分布はエネルギーの小さいところに集中する．最低エネルギーの状態 (基底状態) がひとつしかない場合，ギブス分布は $T \to 0$ でその上の δ 分布になる．

量 A のギブス分布による期待値を $\langle A\rangle$ と書くと，

$$\langle A\rangle = \frac{\sum_\alpha A_\alpha \exp(-E_\alpha/T)}{\sum_\alpha \exp(-E_\alpha/T)} \tag{4}$$

となる．

外部から操作できる変数 a があって，エネルギーが

$$E_\alpha = E_\alpha^0 - aA_\alpha \tag{5}$$

と書けるとき，変数 A と a は共役 (conjugate) であるといわれる (たとえば，このあとの例では磁化を A とすると，磁場が a にあたる)．このとき，(1，3，5) の各式より，(4) 式の右辺は自由エネルギー F の微分であらわせて，

$$\langle A\rangle = -\frac{dF}{da} \tag{6}$$

となる．

同様にして，エネルギーの期待値については，逆温度 $\beta = 1/T$ を用いて

$$\langle E\rangle = -\frac{d}{d\beta}\log Z = \frac{d}{d\beta}\left(\frac{F}{T}\right) \tag{7}$$

が成り立つ．

ギブス分布 (2) はたとえば，$\sum_\alpha P_\alpha = 1$ のもとで ，

$$L_g(\{P_\alpha\}) = \sum_\alpha P_\alpha E_\alpha + T \sum_\alpha P_\alpha \log P_\alpha \tag{8}$$

を最小にする分布 $\{P_\alpha\}$ として導くことができる．このとき， L_g の最小値は，自由エネルギー F に一致する．

$$F = L_g(\{\frac{e^{-E_\alpha/T}}{Z}\}) \tag{9}$$

統計学と統計物理との関連についてはいろいろな視点から論じられているが，そのなかには，(8) 式に着目するもの (最大エントロピー法) もある．われわれはこれと違って，統計物理を高次元の非ガウス分布を扱う技術および概念の集合体として捉えるので，むしろ (2) 式を出発点と考えることにする．

1.2 例

以下の説明で使うため，典型的な物理系の例をあげる．

1.2.1 2 状態系 (イジングスピン)

最も簡単な例として， 2 つの状態しかもたない系を考える．一方の状態にいるときのエネルギーを $+1$，もう一方の状態にいるときのエネルギーを -1 としよう． 2 つの状態のどちらにいるかを 2 値変数 $S = \pm 1$ であらわすことにすると，エネルギー E は

$$E(S) = -hS \tag{10}$$

とかける．

すると系が $S = +1$ に存在する確率 $P(1)$ と $S = -1$ に存在する確率 $P(-1)$ は $\exp(h/T)$ 対 $\exp(-h/T)$ であり，

$$Z = \exp(h/T) + \exp(-h/T) \tag{11}$$

$$P(1) = \frac{\exp(h/T)}{\exp(h/T) + \exp(-h/T)} \tag{12}$$

$$P(-1) = \frac{\exp(-h/T)}{\exp(h/T) + \exp(-h/T)} \tag{13}$$

となる．

このような 2 状態系のことをイジングスピン (Ising spin) と呼ぶことがある． スピンという名は磁性理論に由来するが，本論文の範囲では単に局所的変数という意味と考えて良い．イジングは人名である．なお，磁性理論では h が磁場 (magnetic field) にあたり， $p(1)-p(-1)$ が磁化 (magnetization) に当たる．

1.2.2 イジング模型

こんどは，多数の 2 状態系 $\{S_i\}$ が大きさ $L \times L$ の正方格子上に配列している系を考えよう． i でラベルされた格子点の上に， $S_i = +1$ ならば黒い碁石が， $S_i = -1$ の時白い碁石が，それぞれのっていると考えればよい (図 1)． 2 状態系の間に相互作用が無ければ，この系は単に多数のイジングスピンをよせ集めたのと変わらない．そこで，スピン間に相

互作用のあるモデルのうち最も簡単なものとして，各スピンが隣接する 4 個とのみ相互作用し，相互作用の形が格子上の位置によらないモデルを考えよう．

エネルギーを式で書くと，

$$E(\{S_i\}) = -J\sum_{nn} S_i S_j - h\sum_i S_i \tag{14}$$

となる．$\sum_{nn}$ は正方格子上で隣接する (i,j) 対に関する和を表わす．このモデルをイジング模型，正確には 2 次元正方格子上の最隣接相互作用のイジング模型と呼ぶ．

ここで境界条件が問題になるが，単純に打ち切る (自由境界条件)，周囲のスピンを固定する (固定境界条件)，上下左右をつなげてトーラス状にする (周期境界条件) などいくつかの決め方が可能で，目的によって使い分ける．以下，本稿ではとくに問題になる場合のみ境界条件について触れる．

このエネルギーから導かれるギブス分布は，

$$P(\{S_i\}) = \frac{1}{Z}\exp(-\frac{J}{T}\sum_{nn} S_i S_j - \frac{h}{T}\sum_i S_i) \tag{15}$$

$$Z = \sum_{config.} \exp\,(-\frac{J}{T}\sum_{nn} S_i S_j - \frac{h}{T}\sum_i S_i) \tag{16}$$

となる．ここで，$\sum_{config.}$ は可能な 2^{L^2} 個の $\{S_i\}$ に関する和を意味するものとする．

分配関数 Z の計算はこの場合極めて難しくなる．このような系を調べるひとつの方法は，乱数を使ったシミュレーションである．その方法については 2節 で扱う．また，さまざまな近似解法も知られている．そのうち，もっとも簡単な近似法である平均場近似については 3節 で触れる．

なお，(15,16) 式のモデルで $h = 0$ の場合については，巧妙な方法を用いて $\log Z$ を解析的に評価することができる．このような “厳密解” は非常に特別な場合にしか得られないが，相転移の研究などでは重要な役割を演じた．

格子上の確率場で，各格子点に対して場所によらない一定の近傍が定まっていて，ある変数がとる値の確率分布がその近傍の変数の値を固定すれば決まるものを，一般にマルコフ場と呼ぶ．正方格子上の最隣接相互作用をするイジング模型はその最も簡単な例である．マルコフ場の条件は統計物理の言葉でいえば相互作用が局所的で，エネルギーが局所的なものの和でかけることを意味している．統計物理でいう格子スピン模型は格子上のマルコフ場の良い例になっている．

1.2.3 調和振動子とその変形

連続変数のモデルの内でもっとも簡単なものは，調和振動子の集まりである．

1 個の調和振動子に関しては，

$$E = \frac{J}{2}x^2 \tag{17}$$

となる．但し運動エネルギーの部分は無視してある．

沢山の質点を，ばねで鎖あるいは網状に結び合わせたものについては，やはり，運動エネルギーの部分を無視すると，

$$E = \frac{J}{2}\sum_{nn} (x_i - x_j)^2 \tag{18}$$

となる．

なお，ベイズ統計に良く出てくる滑らかさを表わす事前分布に対応するエネルギーは，もうひとつ高階の差分を考えたもので，鎖の場合でいえば，

$$E = \frac{J}{2}\sum_i \left(x_{i-1} - 2x_i + x_{i+1}\right)^2 \tag{19}$$

の形のものである．

これらのエネルギーはすべて x_i の 2 次形式なので，生成されるギブス分布は高次元のガウス分布になり，いろいろな量をたやすく計算することができる．

連続モデルと離散モデルの関係について考えるために (18) 式のモデルを変形して，

$$E = \frac{J}{2}\sum_{nn} \left(x_i - x_j\right)^2 + \sum_i f(x_i) \tag{20}$$

$$f(x_i) = -2Jx_i^2 + \lambda(x_i^2 - 1)^2 \tag{21}$$

というモデルを考えてみる．

このモデルは 4 次の非線型項を含んでいるので ϕ^4 モデルともいわれるが，イジング模型はここで非線型性を強くした極限 $(\lambda \to \infty)$ に相当している．イジング模型はエネルギーに S_i の 2 次の項しか含んでいないが，連続モデルの極限としてみれば，非ガウス性の非常に強い場合にあたっていることが分かる．

1.3　熱力学的極限と巨視的変数

統計物理では，系の大きさが大きい極限での性質を問題にすることが多い．これを熱力学的極限と呼ぶ．

もっとも簡単な例として，相互作用のないイジングスピンの集まり $\{S_i\}$ を考える．この場合，系の大きい極限とはスピンの数　$N \to \infty$ を意味している．物理量として，1 スピンあたりの磁化 $m = (1/N)\sum_i S_i$ を考えると，(12)，(13) 式から，

$$\langle m \rangle = \tanh(h/T) \tag{22}$$

となる．ここで，$\langle\ \ \rangle$ はギブス分布による期待値を表わす．また，2 項分布の分散の公式から，ゆらぎは

$$\sqrt{\langle (m - \langle m \rangle)^2 \rangle} = 2\sqrt{\frac{p(1)p(-1)}{N}} = \frac{1}{\sqrt{N}}\frac{1}{\sqrt{\cosh^2(h)}} \tag{23}$$

と表わせる．

N が小さいときでもこれらの式は意味をもつが，その場合には，(22) 式の左辺はあくまで期待値として意味を持つにすぎない．これに対して，N が非常に大きい場合，m のゆらぎは $\sqrt{1/N}$ のオーダーになるので，m はほぼ確定値をとることになる．このような場合に m は巨視的な変数 (マクロな変数) であるという．これに対して，ひとつひとつのスピンは微視的な変数 (ミクロな変数) である．

独立な系の集まり以外でも，同様にして熱力学的極限，巨視的変数の概念を定義できる．たとえば，2 次元イジング模型の場合，$N = L^2$ として $N \to \infty$ を考えればよい．この場

合，1スピンあたりの磁化のゆらぎが漸近的に $1/\sqrt{N}$ に比例することは自明でないが，格子上で十分離れた場所のスピンが独立と見なせるという仮定をおけば成り立つと考えられる．

このような仮定は一般には成り立つとは限らないが，仮定が破れる場合でも，巨視的変数が別の値をとるような熱力学的極限が複数個あると解釈することで，巨視的変数の概念を導入できる場合もある．簡単な例については (3.1) 節で触れる．ランダム系のある種のモデル，たとえば SK 模型 ((5.4.1) 節) ではさらに複雑なことが起こる．

熱力学的極限や巨視的変数の定義に問題のあるような系では，有限サイズの場合にもいろいろと風変わりなことが起きること予想される．

1.4 ゆらぎと感受率

変数 A が，それと共役な外部変数 a を変化させたときどう変わるかを考えてみる．これは (4)，(5) 式よりすぐ計算できて，

$$\frac{d\langle A\rangle}{da} = \frac{1}{T}(\langle A^2\rangle - \langle A\rangle^2) \geq 0 \tag{24}$$

となる．

(24) 式は系に外場をかけた時の感受率 (左辺) をゆらぎ (右辺) で表わす式と解釈できる．これはいわゆる線型応答理論のもっとも簡単な例になっている．シミュレーションの場合，左辺を計算するためには，a をいろいろ変化させて数値微分を行う必要があるが，右辺はその必要がなく便利である．この式は熱力学的極限をとらなくても，任意の (有限サイズの) 系で成り立つ．

なお，(24) 式の右辺がつねに非負であることを用いると，(6) 式より，自由エネルギーの凸性

$$\frac{d^2F}{da^2} \leq 0 \tag{25}$$

がいえる．

相互作用のないイジングスピンの場合，(24) 式が成り立っていることは，(22),(23) 式から，$(\tanh(x))' = 1/\cosh^2(x)$ より，すぐにわかる．

2　マルコフ連鎖と緩和

以上述べた範囲 (平衡統計物理) では，どのようにしてギブス分布が実現され，維持されるかというダイナミカルな側面には全く触れなかった．この節ではそれを考えよう．

マルコフ連鎖に関する一般論は統計物理に固有の内容とはいえないが，基本事項についてはあらためて説明した．系を大きくしたときの緩和時間の漸近形など，物理でとくによく調べられている内容についても触れた．

マルコフ連鎖であらわされるダイナミクスを利用して，与えられた確率分布からのサンプルを得る方法はマルコフ連鎖モンテカルロ法 (動的モンテカルロ法) と呼ばれる．これは統計物理で 40 年にわたって重要な役割を果たしてきたが，近年になって統計的情報処理への応用が注目されている．

2.1　マルコフ連鎖と定常分布

物理の問題としては，系のダイナミクスは決定論的な力学 (ニュートンの運動方程式あるいはシュレディンガー方程式) から導かれるべきものであるが、実際の問題を扱うのに有効な立場のひとつは，はじめから確率を含んだダイナミクス (粗視化されたダイナミクス) から出発することである．この場合，状態 $\{\alpha\}$ とその間の遷移確率 $\{W(\alpha \to \gamma)\}$ を与えることによって決まるマルコフ連鎖を用いたモデルが重要な手段になる．

定常分布がギブス分布 $\{P_\alpha^\infty\} = \{\exp(-E_\alpha/T)/Z\}$ になるようなマルコフ連鎖を作るには，$\{W(\alpha \to \gamma)\}$ を次の条件 (詳細釣合条件) をみたすように定めればよい．

$$P_\alpha^\infty W(\alpha \to \gamma) = P_\gamma^\infty W(\gamma \to \alpha) \tag{26}$$

有限系の場合，(26) のほかに，$W(\alpha \to \gamma) \neq 0$ なる (α, γ) を結んだグラフが連結であるなどの条件 (あとの注参照) がみたされれば，定常分布が唯一に定まり，$\{P_\alpha^\infty\}$ に一致することが示せる．

(26) 式は，

$$\frac{P_\alpha^\infty}{P_\gamma^\infty} = \frac{W(\gamma \to \alpha)}{W(\alpha \to \gamma)} \tag{27}$$

とも書けるが，これだけからは $\{W(\alpha \to \gamma)\}$ は一意には定まらない．

そこで，可能な W の中から物理的に自然なものを選ぶわけであるが，状態を定義する変数のうち同時に変化するのが 1 つもしくは少数のみと仮定するのが普通である．W でいえば，$W(\alpha \to \gamma)$ のうちほとんどを零とし，状態 α と γ が “近い” ごく少数のもののみを非零とするわけである．具体的な例はあとに示す．このようなダイナミクスを “局所的なダイナミクス” と呼ぶことにする (これは状態を定義する変数の取り方に依存するので、本当は曖昧な定義である。後で扱う格子系では、各格子点の上にひとつずつ変数があるとして、1 つまたは少数の格子点を除いて変数の値が同じ状態の間だけが直接に遷移できるような場合を “局所的なダイナミクス” と考える).

なお，(27) 式の左辺が比の形になっているために，遷移確率の表式は計算の難しい分配関数 Z を含む必要がないことを注意しておく．計算技術としてマルコフ連鎖が有効なのはこのおかげである．

マルコフ連鎖で重要なのは，緩和 (relaxation) の概念である．以下，W を単位時間の遷移確率として連続時間モデルで考える．(従って，厳密にはマルコフ連鎖でなくマルコフ過

程というべきである．) 本質は離散時間の場合でも同じである．適当な初期分布 $P_\alpha(0)$ から出発して，$t > 0$ 時間後に状態 α にいる確率を $P_\alpha(t)$ とすると，これは

$$\frac{dP_\alpha(t)}{dt} = \sum_\gamma P_\gamma(t) W(\gamma \to \alpha) - P_\alpha(t) \sum_{\tilde{\gamma}} W(\alpha \to \tilde{\gamma}) \tag{28}$$

をみたす．リウビル演算子 L を

$$L_{\alpha\gamma} = W(\gamma \to \alpha) - \delta_{\alpha\gamma} \sum_{\tilde{\gamma}} W(\alpha \to \tilde{\gamma}) \tag{29}$$

と定義すれば，(29) 式は

$$\frac{dP(t)}{dt} = LP(t) \tag{30}$$

と書ける．ただし，$P(t)$ は状態 α に系が存在する確率を成分とするベクトル $\{P_\alpha(t)\}$ をあらわす．また，$\delta_{\alpha\gamma}$ はクロネッカーのデルタ

$$\delta_{\alpha\gamma} = 1 \quad (\alpha = \gamma) \tag{31}$$

$$= 0 \quad (\alpha \neq \gamma) \tag{32}$$

である．

この解は L の固有値 $\{\lambda_i\}$ と対応する固有ベクトル $\{V_i\}$ によって定まり，

$$P_\alpha(t) = P_\alpha^\infty + \sum_{i=1,2\ldots} a_i V_i^\alpha \exp(\lambda_i t) \tag{33}$$

となる．$\{a_i\}$ は初期分布 $P_\alpha(0)$ に依存する定数である．

ここで，$\{\lambda_i\}$ のなかで最大のものはギブス分布 P^∞ を固有ベクトルとして持つ $\lambda_0 = 0$ である．また，詳細釣合条件を考慮すると，λ_0 以外の $\{\lambda_i\}$ はすべて負の実数であることがわかる (あとの注参照)．これは，定常分布が一意で，ギブス分布になることの数学的表現になっている．

そこで，緩和時間スペクトル $\{\tau_i\} = \{-1/\lambda_i\}$ を定義すると，$\{\tau_i\}$ はすべて正の実数で，そのうち最大のもの τ_1 が系が記憶を失う (緩和する) 時間を決める．これは，系の状態がマルコフ連鎖に従って変化するとき，初期状態から τ_1 以上時間がたったのちに，τ_1 より十分大きい時間間隔でサンプルをぬきだせば，これらは定常分布 (ここではギブス分布) からランダムに選んだものと見なせるということを意味している．

$P_\alpha(t)$ は状態 α に系が存在する確率 であることを注意しておく．アンサンブル的に解釈すれば，(33) 式の意味は，多数の系を用意して，各々を独立に，与えられたダイナミクスに従って発展させたときに，$\{\tau_1\} = \{-1/\lambda_1\}$ より十分長い時間がたてば，多数の系の分布が P_α^∞ に近付くということである．1 個の系を追いかけた場合に，出発した状態から τ_1 以上の時間がたてば，そこで 1 回だけ観測した物理量がその期待値の良い近似になるという意味ではない．これが成り立つのは，系が良く定義された (単一の) 熱力学的極限をもち，系の大きさが十分大きく，測定する量が巨視的変数と見なせる場合に限る．この条件は，あとで述べるマルコフ連鎖モンテカルロ法の場合，とくに統計的情報処理への応用の場合には，満たされないと考えた方が無難である．1 個の系の場合，緩和時間はあくまで “記憶を失う” までの時間の目安と考えるのが正しい．

$\{\lambda_i\}$ の性質は次のようにして分かる．

まず $\exp(tL)$ が対角要素が正の確率行列であることと遷移のグラフの強連結性から，$\exp(tL)$ の固有値は絶対値が 1 以下であり，とくに定常状態に対応する $\exp(\lambda_0 t) = 1$ は単純根である．そこで，リウビル演算子 L の固有値 $\{\lambda_i\}$ は，$\lambda_0 = 0$ を除いてすべて実部が負になる．

詳細釣合の式 (26) は確率ベクトルの空間に定義された内積

$$(Q^1, Q^2)_G = \sum_\alpha Q^1_\alpha Q^2_\alpha (P^\infty_\alpha)^{-1} \tag{34}$$

のもとで，L がエルミートであることを示している．そこで，これがみたされる場合は，$\{\lambda_i\}$ はすべて実数であることがわかる．これは物理的には，詳細釣合条件をみたす連続時間の系の緩和は振動的ではありえないことを意味する．

離散時間の場合もほとんど同じ議論ができるが，偶数周期の振動的な緩和が原理的にはありうる．また，定常状態の一意性をいうためには付加的な条件がいる．たとえば，最低 1 つの α について $W(\alpha \to \alpha) \neq 0$ となれば十分である．以下であげる例はすべてこの条件を満たしている．まったく条件なしでは，詳細釣合の条件のもとでも周期 2 の非減衰振動がおこりうる.(式 (38-41) で定義されるダイナミクスで離散時間，$h = 0$ の場合が，簡単な例になっている．)

次に，系の大きさ N を変えたとき，緩和時間がどう変わるかを考える．以下では，局所的なダイナミクスを仮定し，系の大きさを変えた時には、単位時間の遷移確率を，物理現象として自然なようにスケールするものとする．単位時間の遷移確率の定義によっては，以下の τ_1 は τ_1/N に読みかえなくてはならない．この違いは，離散時間の場合には，マルコフ連鎖のステップ数を N で割ったものと割らないものの違いに対応している．あとで述べるマルコフ連鎖モンテカルロ法では，前者をモンテカルロステップ数 (MCS 数)，後者をスピンフリップ数または試行回数として区別している．

普通，系の大きさ N が大きくなると，系のとりうる状態の数 (あるいは状態空間の体積) は N の指数関数で増えるが，緩和時間は指数関数的に増えるとは限らない．これは系の各部分がすべて密接に関連しているとは限らないことに関連している．たとえば，格子上のモデルの場合，ある程度以上はなれた部分の相関は指数関数的に失われるのが普通である．このような場合，格子の 1 辺の大きさ $L(N = L^d$，d は次元) が相関距離以上になると，

$$\tau_1 \sim const. \tag{35}$$

となると思われる．

格子上のモデルでも例外的な場合には，もっと緩和が遅いことがある．たとえば，あとでででてくる正方格子上の動的イジング模型 (磁場なし，固定境界条件) の場合，“無秩序相” では (35) 式のようになるが，相転移点直上および “秩序相” では，

$$\tau_1 \sim L^z \tag{36}$$

のようになる．べき z は前者では $z = 2.13..$，後者では $z = 2$ となると信じられている．

さらに遅い場合としては，

$$\tau_1 \sim \exp(L^z) \tag{37}$$

も考えられる．このような場合の例はあとで示そう．

スピンがネットワーク上に配置されているモデルのように，要素同士が非局所的に結合されているモデルの場合も，全部の状態をまわるのに要する時間より局所的なダイナミクスのもとでの緩和時間の方がはるかに短いことは珍しくない．

2.2 例

2.2.1 2 状態遷移模型

もっとも簡単な場合として，1 個のイジングスピンを考えよう．ギブス分布 (13)，(14) を実現するためのひとつの方法は，

$$W(+1 \to -1) = \exp(-2h/T) \tag{38}$$

$$W(+1 \to +1) = 1 - \exp(-2h/T) \tag{39}$$

$$W(-1 \to +1) = 1 \tag{40}$$

$$W(-1 \to -1) = 0 \tag{41}$$

とすることである ($h > 0$ の場合)．これが (13)，(14) の分布に関して詳細釣合の条件 (26) をみたしていることはすぐわかる．W の別の選びかたとしては，

$$W(+1 \to -1) = \frac{\exp(-h/T)}{\exp(+h/T) + \exp(-h/T)} \tag{42}$$

$$W(+1 \to +1) = \frac{\exp(+h/T)}{\exp(+h/T) + \exp(-h/T)} \tag{43}$$

$$W(-1 \to +1) = \frac{\exp(+h/T)}{\exp(+h/T) + \exp(-h/T)} \tag{44}$$

$$W(-1 \to -1) = \frac{\exp(-h/T)}{\exp(+h/T) + \exp(-h/T)} \tag{45}$$

というのも可能である．

2.2.2 動的イジング模型

正方格子上のイジング模型の場合，いちばん簡単なダイナミクスの与え方は状態 α，γ がひとつのスピンを除いて一致する場合以外は $W(\alpha \to \gamma) = 0$ とすることである (1 spin flip)．これだけではまだ任意性があるが，先に考えた 1 個のイジングスピンの場合を一般化すると，最も簡単な例として，次の 2 つが考えられる．(あとの節との関連で，離散時間の場合を手続き的に説明した)．

- ダイナミクス A

  1. 乱数を用いてスピンの番号 j を選ぶ．
  2. 選んだスピン S_j が符号を変えたときのエネルギー変化 (符号に注意．普通，変化というときは変化後の値から前の値を引くのだが，以下では逆である．将来の版では系統的に直すかもしれない)．

     $$\triangle E = -2JS_j \sum_{i \in nn(j)} S_i \tag{46}$$

     を計算する．ただし，$nn(j)$ は j の 4 個の隣接格子点を表わす．

3. (0,1) の一様乱数 RND を発生し，

$$RND < \exp(\triangle E/T)$$

なら，注目しているスピンの符号を反転する．

これを繰り返す．

- ダイナミクス B

1. 乱数を用いてスピンの番号 j を選ぶ．
2. 選んだスピン S_j が $+1$ のときと -1 のときの系のエネルギー E_+，E_- をそれぞれ計算する．
3. (0,1) の一様乱数 RND を発生し，

$$RND < \frac{\exp(-E_+/T)}{\exp(-E_+/T) + \exp(-E_-/T)}$$

なら，注目しているスピンを $+1$ で，そうでなければ -1 で置き換える．

これを繰り返す．

実際の計算では，

$$\frac{\exp(-E_+/T)}{\exp(-E_+/T) + \exp(-E_-/T)} = \frac{1}{1 + \exp(\tilde{\triangle E}/T)} = \frac{1}{2}(1 - \tanh(\tilde{\triangle E}/2T)) \quad (47)$$

$$\tilde{\triangle E} = E_+ - E_- = -2J \sum_{i \in nn(j)} S_i \quad (48)$$

を利用すればダイナミクス A と同じ程度の計算量で済む．

なお，前にイジング模型はマルコフ場モデルの一種だと書いたが，これは，格子上の確率場の空間的性質について述べたものであり，ここで論じているマルコフ連鎖は時間方向の性質なので，混同してはならない．

2.3 マルコフ連鎖モンテカルロ法

2.3.1 計算手段としてのマルコフ連鎖

以上では，どちらかというと，実際の物理系のダイナミクスのモデルとしてマルコフ連鎖 (マルコフ過程) を考えた．以下では，逆に，ギブス分布からのサンプルを得るための手段として，定常分布がギブス分布になるマルコフ連鎖をシミュレートすることを考える．このように考えた場合，もとの分布が統計物理と関係ある必要はないので，確率分布からのサンプリングのための普遍的なアルゴリズム群が得られたことになる．一般には，マルコフ連鎖の “時間” は仮想的なものと考えられる．

物性物理の分野では，単にモンテカルロ法というとほぼこの方法を意味する．とくに区別する場合は，動的モンテカルロというようだが，それほど一般的な名称ではない．統計学や工学では他の種類のモンテカルロ法が広く使われているので，区別するための名称が必要である．最近，統計学の分野ではマルコフ連鎖モンテカルロ法 (Markov chain Monte

Carlo algorithms , $(\mathrm{MC})^2$) という名称が定着してきているので，本稿ではそれを用いる．モデルとしてのマルコフ場モデルと混同しないよう注意して欲しい．

なお，神経回路網関係では，ボルツマンマシンという用語があるが，この言葉は，モデルの名前なのか，学習法 (推定法) の名前なのか，マルコフ連鎖モンテカルロ法の工学的応用全体を指すのか，全く混乱している．できれば使わない方が良いだろう．

マルコフ連鎖モンテカルロ法を用いると，緩和時間より長い間隔でマルコフ連鎖から抜き出したサンプルが，与えられた分布からのランダムサンプルと見なせることを利用して，任意の量の期待値，周辺分布を計算することができる．物理量 (統計量) $A(\alpha)$ に対して，マルコフ連鎖から抜き出したサンプルを $\{\alpha_m\}$ とすると，与えられた分布 P_α^∞ での $A(\alpha)$ の期待値

$$\langle A(\alpha)\rangle = \sum_\alpha A(\alpha) P_\alpha^\infty \tag{49}$$

が，

$$\langle A(\alpha)\rangle = \lim_{M\to\infty} \frac{1}{M} \sum_{m=1}^{M} A(\alpha_m) \tag{50}$$

と表せるわけである．

抜き出したサンプルの間隔がマルコフ連鎖の緩和時間より長いという条件は実際は不必要である．サンプル間隔が緩和時間より短い長さ M の列を考える．この列は，緩和時間より長い間隔 (R とする) ごとにとびとびにサンプルしなおすことで，R 個の長さ M/R の列に分割することができる．M/R が十分大きければ，この部分列のそれぞれについて物理量の平均値が求める期待値に収束するので，全体の平均もまた正しい値に収束する．したがって，物理量の計算に要する手間が無視できるなら，すべてのサンプルを計算に使って差し支えない．ただし、分布のキュムラントなどの精密計算で不偏推定値を求めたい場合には，緩和時間とサンプル間隔に応じて補正が必要である (Kikuchi et al (1994))．

サンプルの列の長さ M は緩和時間に対して十分長くなければならない．理論的には系の緩和は任意の物理量に関してマルコフ連鎖の緩和時間のうち最長のもの (τ_1) に支配されるはずである．しかし，定量的な緩和の速さは物理量によって異なる．これは (33) 式の V_i に対しての物理量 $A(\alpha)$ の “射影” $\sum_\alpha V_i^\alpha A(\alpha)$ によって決まる．系を動かしているマルコフ連鎖と測定する物理量がともに特別な対称性をもっている場合には，V_1 への射影が零になるために，その物理量についてだけ緩和時間がとくに短くなることがある (ダイナミクス A や B におけるエネルギー E はその例である)．このような量の緩和時間で他の量の緩和時間を見積もってはいけない．

マルコフ連鎖モンテカルロ法の難点は，緩和時間の推定が (事前にも，事後的にも) 難しいことである．すでにみたように，局所的で自然なダイナミクスのもとでの緩和時間はきわめて長くなりうる．そのような場合には，全く間違った結果に導かれることもある．

従来から統計学などで使われてきたモンテカルロ法の多くは，欲しい分布を大域的に近似するような重み関数を与えて，それからのサンプリングを行なうものである．この方法では，問題が高次元，大規模になると，効果的な重み関数を与えるのがきわめて難しくなる．たとえば，1 次元で $(1-\epsilon)$ 位の近似度の場合，そのまま 1000 次元に拡張すれば $(1-\epsilon)^{1000}$ の程度の近似度になる恐れがある．これが統計物理の分野でマルコフ連鎖モンテカルロ法が愛用されてきた主な理由である．次元がひくく，目的の分布が適当な変数変換に

よって正規分布などで良く近似されるなら，従来の方法を使うべきである．

2.3.2 狭義のメトロポリス法

狭義のメトロポリス法は，動的イジングモデルの場合のダイナミクス A を一般化したものと考えられる．一般に，$W(\alpha \to \gamma) \neq 0$ で $\alpha \neq \gamma$ のとき $W(\alpha \to \gamma)$ を

if $\triangle E_{\alpha\gamma} \geq 0$

$$W(\alpha \to \gamma) = \frac{1}{c_{\alpha\gamma}} \tag{51}$$

if $\triangle E_{\alpha\gamma} < 0$

$$W(\alpha \to \gamma) = \frac{1}{c_{\alpha\gamma}} \exp(\triangle E_{\alpha\gamma}/T) \tag{52}$$

$$\triangle E_{\alpha\gamma} = E_\alpha - E_\gamma \tag{53}$$

のように決める．$W(\alpha \to \gamma) = 0$ のときは，$W(\gamma \to \alpha) = 0$ とする．また，$c_{\alpha\gamma} = c_{\gamma\alpha}$ は定数で，離散時間の場合，確率の和が 1 になるようにとる．詳細釣合の条件が満たされることがすぐ確かめられるので，マルコフ連鎖モンテカルロ法のひとつの族が得られたことになる．

以下，$\alpha = \{\alpha_i\}(i = 1, I)$ というように状態がベクトル的にあらわせる場合を考える．要素 α_i としては，たとえば，各々のスピン (局所的変数)S_i を考えればよい．

$\{\alpha_i\}$ が離散変数の場合，もっとも基本的なアルゴリズムは，以下の手順を繰り返すことで得られる．動的イジング模型 (ダイナミクス A) の場合は，スピンが 2 状態しかなかったので，ステップ (2) に相当するものがなかった．

1. “動かすことを試みる変数”α_j を決める．

   最も基本的な方法は，乱数を用いて j を選ぶことである．任意の i について選ばれる確率が零でなければ，選び方が等確率である必要はない．

2. 変数 α_j をどう変更するか，“動く先の候補”γ_j を決める．

   候補の選び方は等確率にする．すなわち，α_j のとりうる値を $\{g_j^k, k = 1..K_j\}$ とするとき，乱数を用いて確率 $1/(M_j - 1)$ で $\gamma_j = g_j^k$ ($g_j^k \neq \alpha_j$の現在の値) とする．

3. もとの状態 $\alpha = \{\alpha_1..\alpha_j..\alpha_I\}$ と動く先 $\gamma = \{\alpha_1..\gamma_j..\alpha_I\}$ について，動かした場合のエネルギーの変化

   $$\triangle E = E_\alpha - E_\gamma \tag{54}$$

   を計算する．

4. (0,1) の一様乱数 RND を発生し，$RND < \exp(\triangle E/T)$ なら，α_j を γ_j で置き換える．

ステップ (1) で乱数を用いずに一定の順番で j を選んでも厳密に正しい結果が得られる．このことは，複数のマルコフ連鎖の合成という視点からみると理解しやすい．各 $i(i = 1..I)$ について，α_i のみを動かす過程を考え，全体のマルコフ連鎖をそれら I 個の過程の合成と見るわけである．たとえば，正方格子の場合，格子を市松模様に塗り分けて，ひとつの色の格子点上のスピンを一通り動かしたのち，別の色の格子点にうつる方法 (checkerboard

updation) が可能である．これによって，アルゴリズムのベクトル化が可能になる．特別な場合には，規則的に選んだために緩和時間が長くなることもあるが，普通はあまり問題にならないようである．全部のスピンを “同時に” 動かすのはもちろん誤りである．

$\{\alpha_i\}$ が連続変数の場合は，次のようにすればよい．与えられた確率分布が

$$P(\alpha)d\alpha = \frac{\exp(-E(\alpha)/T)\Pi d\alpha_i}{\int_{-\infty}^{+\infty}\exp(-E(\alpha)/T)\Pi d\alpha_i} \tag{55}$$

の形で書かれている場合を考える ($\alpha_i \in (-\infty, \infty)$ とする)．

1. 適当な方法で j を選ぶ．
2. $\triangle\alpha$ を $(-\epsilon_j, \epsilon_j)$ の一様乱数として，$\gamma_j = \alpha_j + \triangle\alpha$ とする．
3. もとの状態 $\alpha = \{\alpha_1..\alpha_j..\alpha_I\}$ と動く先 $\gamma = \{\alpha_1..\gamma_j..\alpha_I\}$ について，動かした場合のエネルギーの変化
$$\triangle E = E_\alpha - E_\gamma \tag{56}$$
を計算する．
4. (0,1) の一様乱数 RND を発生し，$RND < \exp(\triangle E/T)$ なら，α_j を γ_j で置き換える．

ステップ (2) 以外は離散変数の場合と同じである．ステップ (2) で用いる $\triangle\alpha$ は必ずしも一様乱数である必要はないが，ステップ (2) では，区間 $(\alpha_j, \alpha_j+d\alpha_j)$ にいる変数を区間 $(\gamma_j, \gamma_j+d\gamma_j)$ に動かそうと試みる確率を，その逆，すなわち，$(\gamma_j, \gamma_j+d\gamma_j)$ から $(\alpha_j, \alpha_j+d\alpha_j)$ に動かそうと試みる確率と一致させるようにする必要がある．これらが一致しないと，ステップ (3),(4) をうまく調節しない限り，詳細釣合条件が破れてしまう．一様乱数の場合には，ϵ_j は α_j の値に依存しない定数でなければならない．

$\{\epsilon_i\}$ の大きさは，緩和時間を決める重要な要素である．経験的には (ステップ (4) で) “動く割合”(acceptance ratio) が５０％前後となるようにするのが良いといわれている．これは経験則であり，条件付き分布の裾が長い場合など，特別な場合に当てはまるかどうかは疑問である．ただ，acceptance ratio が大きすぎても緩和が悪くなることは覚えておく必要がある．

2.3.3 Gibbs Sampler(熱浴法)

動的イジング模型の場合のダイナミクス A の一般化が狭義のメトロポリス法であったのに対し，ダイナミクス B の一般化が Gibbs Sampler である．Gibbs Sampler というのは統計学者の用語であるが，物理でいう熱浴法 (heat bath algorithm) とほぼ同じ意味ではないかと思われる。今度は，α がベクトル的に $\{\alpha_i\}\,(i = 1..I)$ と書けることをはじめから仮定しよう．Gibbs Sampler のアルゴリズムは，$\{\alpha_i\}$ が離散変数の場合，以下の手順を繰り返すことで得られる．

1. 適当な方法で，“動かす変数” の番号 j を決める．

2. 変数 α_j の “とりうる値” にそれぞれ確率を割り当てる．

   α_j のとりうる k 個の値を $\{g_j^k, k = 1..K_j\}$ とする．着目している以外の成分 $\alpha_i (i \neq j)$ はそのままで，j 番目の成分 α_j が値 g_j^k をとった状態を，$\gamma^k = \{\alpha_1..g_j^k..\alpha_I\}$ とするとき，確率 q^k を，

$$q^k = \frac{\exp(-E(\gamma^k)/T)}{\sum_{\bar{k}} \exp(-E(\gamma^{\bar{k}})/T)} \tag{57}$$

   と定める．

3. 乱数を用いて，α_j を確率 q^k で状態 g_j^k に遷移させる．

   アルゴリズム的には，

   - $Q_k = \sum_{\bar{k}=1}^{k} q^{\bar{k}}$, $Q_0 = 0$ とする．
   - $(0,1)$ の一様乱数 RND を発生する．
   - $Q_{k-1} \leq RND < Q_k$ なら，α_j を g_j^k で置き換える．

   とすれば良い．

ステップ (2),(3) は，着目している成分以外を固定した条件つき分布 $P(\gamma_j \mid \alpha_1..\alpha_{j-1}, \alpha_{j+1}..\alpha_I)$ から γ_j をサンプルすることを意味する．狭義のメトロポリス法の場合と異なって，新しい α_j がどうなるかは，以前の α_j の状態に依存しない．詳細釣合条件が満たされることは簡単に確かめられるが，求める分布が不変なことを直接理解した方が早いかもしれない．

このアルゴリズムをそのまま連続変数の場合に拡張した場合，ステップ (2) の条件つき確率の計算が大変で，実用的でないことが多い．ただし，統計的情報処理の問題などでは，エネルギーの形が特別なために (あるいは，特別な形に設定したおかげで)，ステップ (2) が少ない計算量で実行できる場合がある．たとえば，離散変数と連続変数の両方を含んでいて，離散変数を固定したときに，連続変数の部分がガウス分布になるモデルはその典型である．また，ある分布の分散や平均値などがまた確率的に分布しているモデル (いわゆる混合分布モデル) のなかにも，Gibbs Sampler が簡単にインプリメントできるものがある．ステップ (2) の条件つき確率の計算が完全にできない場合のひとつの方策は，近似的な重み関数を導入することであるが，その方法については付録で述べる．

2.3.4 積分それ自体の計算

近似的な重み関数を与えて行なうモンテカルロ法は積分の近似計算法とみなすことができる．これに対して，マルコフ連鎖モンテカルロ法は，確率分布からのサンプルを得る方法である．別な表現をすれば，期待値 (積分の比) の計算法であるともいえる．

統計物理でも，自由エネルギー $F(T) = -T \log Z$ は積分 (の対数) そのものであって，積分の比ではないから，マルコフ連鎖モンテカルロ法では直接計算できない．自由エネルギーを計算する単純な方法は，(7) 式を用いて，

$$\frac{F(T)}{T} = (\frac{F}{T})_{T=\infty} + \int_0^{1/T} \langle E(T') \rangle d(\frac{1}{T'}) \tag{58}$$

とし，各 T での $\langle E(T) \rangle$ の計算に帰着させることである．大部分のモデルで，$(F/T)_{T=\infty}$ は簡単に計算できる．この方法は統計学での mariginal likelihood や分布の規格化定数の

計算にも利用できる．より進んだ方法としては付録で説明したマルチカノニカル法や交換法の応用が考えられる．

分配関数 Z を無理やりギブス分布での期待値の形に書くと

$$Z = \frac{\text{全状態数}}{\langle \exp(E/T) \rangle} \tag{59}$$

と書けるが，この表式は役に立たない．

“マルコフ連鎖モンテカルロ法によれば，どんな量の期待値でも計算できる” というのは正確な表現ではない．状態空間の一部で異常に大きくなる量の期待値はたとえ緩和時間が 1 のオーダーでもうまく計算できない．これは，母分散が異常に大きくなる場合に相当する．(59) 式の右辺の $\exp(E/T)$ はこのような量である．

2.3.5 simulated annealing

本稿の最初に述べたとおり，温度 T が小さいときには，ギブス分布は最低エネルギーの状態 (最大確率の状態) の上に集中する．実際，$T \to 0$ では，狭義のメトロポリス法のアルゴリズムは最急降下法ないし局所的な逐次改良法と見なせる．これを利用して，マルコフ連鎖モンテカルロ法において，T の大きいところから出発して，ゆっくりと $T \to 0$ とすることで，組合せ的最適化の問題を解くことができる．ゆっくりと $T \to 0$ とする意味は，はじめから $T = 0$ とするのに比べて，温度が有限のときに，エネルギー関数の局所的な極小のまわりから抜け出すチャンスが多いということにある．この方法を，材料が歪んだ状態 (局所的極小) にあるのを加熱と除冷 (anneal) によって除く方法にたとえて，simulated annealing 法 (模擬やきなまし法) と呼ぶ．

simulated annealing 法がうまくいくためには，有限温度で出現しやすい状態の集合 (大きな確率をもつ準安定状態) に大域的な極小が含まれていることが望ましい．また，厳密に最適解が必要とされる場合より，近似的な解で良い場合に向いた方法である．実際の計算は，組合せ的最適化の用語でいえば，確率的に繰り返しバックトラックを行なっているのに相当するので，きわめて時間がかかる．人間が複雑なアルゴリズムを考える手間の分も計算機が努力していると考えれば，ある程度は容認できるかもしれない．

数学的には，極めてゆっくりと温度 T を下げた場合に確率 1 で大域的な極小に到達することが証明できるらしいが，そのような証明に実際的な意味があるかどうかは疑わしい．また，物理との類似が強調されるが，これも成功を約束するものではない．マルコフ連鎖モンテカルロ法の “時間” を実際の物理的時間に対応させる方法は自明ではないが，直観的にいって，simulated annealing 法で利用できる計算時間は，実際の物理系で結晶化が起こる時間に比して極めて短いと思われる．

simulated annealing 法に関連した最適化手法として，マルコフ連鎖モンテカルロ法で有限温度の分布を生成する代わりに，適当な近似法，たとえば後述の平均場近似によって近似的な (周辺) 分布を計算し，$T \to 0$ とするという方法がある．いわゆる Hopfield-Tank 法による最適化はこの例として解釈できる．

simulated annealing 法の意義は，ランダムな揺らぎが工学的な応用で肯定的な意味を持つことを示したことである．逆に良くない影響は，マルコフ連鎖モンテカルロ法の本来の用途は分布からのサンプリングであるのに，最適化という文脈でのみ考える傾向が一部

で生じたことである．もちろん，多くの工学の問題では，(近似的) 最適解のみが意味を持つ．しかし，少なくとも 2 つの分野，統計物理と統計的情報処理，では高次元の確率分布からのサンプリングそれ自体が重要な課題なのである．

3 相転移

ここでは相転移 (phase transition) について入門的解説をする．相転移は統計物理で最も重要なトピックのひとつである．統計的情報処理における相転移の役割はいまのところはっきりしないが，非ガウス的な高次元の分布を考える場合，相転移の知識は背景として必要と思われる．また，統計物理の諸概念や近似法には相転移の問題をめぐって発達したものも多いので，それらの説明をするのも目的のひとつである．

3.1 2 次元イジング模型の相転移

正方格子上で最隣接のみ相互作用するイジング模型を考える ((15)) . いま磁場 h が零の場合を考えると外部からかえられる変数は J/T だけである．

この場合，ギブス分布からランダムに選んだパターン (snapshot) を観察すると，J/T が小さい場合はパターンは確率 $1/2$ で黒と白をでたらめにばらまいたものに近いが，J/T が大きくなるにしたがって黒同士，白同士集まる傾向が強まり，同色のスピンのクラスター (空間的かたまり) ができる．さらに結合 J を強く，あるいは温度 T を低くしていくと，$J/T = 0.4407..$ のあたりでクラスターの大きさが急激に大きくなり，系全体に広がるようになる (図 2) . この点を相転移点または臨界点 (critical point)，これより弱結合 (高温) の側を無秩序相 (disordered phase)，強結合 (低温) の側を秩序相 (ordered phase) という．温度 T を変えたときのこのような変化は非ガウス性のあらわれである．ガウス分布の場合，相関は温度 T に単純に比例して強くなるだけで，相転移のような現象は存在しない．

以下では，相転移をもっときちんと定義する方法をを 2 つ紹介する．

ひとつの方法では，スピン 1 個当りの磁化 m を考える．この場合，磁場 h が零の有限系では，黒と白の対称性から，m のギブス分布での期待値 $\langle m \rangle$ は常に零である．ただし，境界条件が黒と白の対称性を破っていれば別である．いま，境界条件を周囲が全部黒 $(+1)$ の固定境界条件としよう．すると，無秩序相では，系の一辺の長さ L を無限大にした時，$\langle m \rangle \to 0$ となるが，秩序相ではこれが有限値にとどまる，すなわち $\langle m \rangle \to m_\infty > 0$ となることがわかる．まわりを白 (-1) に固定すれば，当然 $\langle m \rangle \to -m_\infty < 0$ となる．定量的にいうと，$\mid m_\infty \mid$ は J/T が大きくなると図 3 のように増大し，J/T が無限大のとき 1 になる．秩序相では熱力学的極限が 2 つあるとでもいうべき状況が起きているわけで，これを自発的な対称性の破れ (spontaneous symmetry breaking) という．また，$\pm m_0$ を自発磁化という．いまの場合，秩序相は自発磁化のある相 (強磁性相，ferromagnetic phase) として特徴付けられるわけである．いったん，固定境界条件を与えてしまえば，$\langle m \rangle$ が $\pm m_\infty$ のまわりで揺らぐ幅は $1/\sqrt{N}$ の程度になるので，この意味では m を巨視的変数と呼ぶことができる．一般に，相転移に際して磁化 m のような役割を演ずる巨視的変数を秩序変数 (order parameter) という．

別の見方として，スピン間の空間相関関数

$$\tilde{C}(i,j) = \langle S_i S_j \rangle \tag{60}$$

をもとに議論することもできる．$\tilde{C}(i,j)$ の値は無秩序相ではスピン i とスピン j の距離 d を無限に大きくすると指数関数的に零に収束する．

$$\tilde{C}(d) \sim \exp(-d/\xi) \tag{61}$$

このときの ξ を相関距離といい，臨界点に近付くと大きくなり，臨界点で発散する．ξ は大ざっぱにいえばクラスターの大きさの目安となる量である．これに対して，秩序相では $\tilde{C}(i,j)$ は有限の値に収束する．

$$\tilde{C}(d) \sim const. \tag{62}$$

臨界点は異常な点である．たとえば，相関関数を

$$C(i,j) = \langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle \tag{63}$$

と定義しなおすと，これは無秩序相でも秩序相でも，(熱力学的極限の取り方に注意すれば) 指数関数的に零に収束する．しかし，臨界点の直上は例外で，零への収束はべき的になる．これに関連して，ゆらぎがガウス分布に漸近しない，磁場に対する感受率 (帯磁率，susceptibility) が発散する， snapshot がある意味でフラクタル的になる，などいろいろ特異なことがおきる．

ここで， J/T がどんなに小さくても，分布 (15) において “もっとも確率の高い状態” は S_i が全部黒 (+1) の状態か，全部白 (-1) の状態の 2 つであるということを注意しておく．これは，式 (15) から自明のことである．しかし，図 2 をみればわかるように，この 2 つの状態は， J/T がきわめて大きい場合を除いて，決して “典型的な状態” ではない．図 2 に示した snapshot そのものの出現確率は低くても，それに “性質が似た状態” は多数あり，それらの出現確率の和は大きい値になるのである．

“典型的な状態” という表現をきちんと定義するのは，“乱数” という概念を定義するのと似ていて，簡単ではない．熱力学的極限が一意的に定義されていて，系の大きさが十分大きい場合には，ある状態のエネルギーがギブス分布でのエネルギーの期待値に十分近い値をとるとき，その状態を “典型的” と定義するのがひとつの方法である．

この定義によると，典型的状態の数の対数 S は，エネルギーの期待値を $\langle E \rangle$ とすると，次の式で評価される．

$$\exp(S) = \left(\frac{\exp(-\langle E \rangle / T)}{Z} \right)^{-1} \tag{64}$$

これは，自由エネルギー $F = -T \log Z$ を使えば，

$$S = (\langle E \rangle - F)/T \tag{65}$$

と書ける． S が統計物理におけるエントロピーである． $P_\alpha = \exp(-E_\alpha/T)/Z$ とするとき

$$S = -\sum_\alpha P_\alpha \log P_\alpha \tag{66}$$

とも書ける．エントロピーという言葉は情報処理の分野でさまざまなレベルで使われているため，本稿ではこの言葉をなるべく使わないようにしたが，そのために多少持って回った表現になった部分もある．

3.2 相転移とダイナミクス

動的イジング模型の場合に相転移はどう見えるだろうか．周期境界条件ないし自由境界条件の有限系を考えよう．ダイナミクスとしては局所的なものを考える．

無秩序相の側から臨界点に近付くと，緩和時間は急激に長くなる (critical slowing down)．snapshot でみると，クラスターの大きさが大きくなり，それらが時間的にゆっくりと変化する．さらに結合を強くして，臨界点の直上から秩序相にはいるあたりでは，1 つのクラスターが系全体をおおう程度に大きくなる．系が主として黒のスピンでおおわれる時期と白でおおわれる時期が短い過渡期をはさんで交代で出現する．このときの磁化 m の時間的振舞いを (図 4) に示す．磁化の値が永遠に大きく揺らぎ続けることがわかる．

秩序相における緩和時間 τ_1 と系の一辺の長さ L との関係は，周期境界条件，自由境界条件の場合，

$$\tau_1 \sim \exp(L^{d-1}) \tag{67}$$

となると考えられている．これは，黒っぽい状態と白っぽい状態との間を動くのに要する時間が局所的ダイナミクスのもとでは非常に長くなることに対応しており，自発的対称性の破れのダイナミクスへの反映である．

固定境界条件 (周囲を黒に固定) の場合，

$$\tau_1 \sim L^2 \tag{68}$$

となると考えられている．こんどは，“行ったり来たり” はないが，系の真中に大きな白いクラスタができる可能性があり，これが消えるのにある程度時間を要するのでべきの発散があらわれる．

前に述べた通り，無秩序相では

$$\tau_1 \sim const. \tag{69}$$

である．

これらのこと (及びあとの注) からわかるように，秩序相ではマルコフ連鎖モンテカルロ法の緩和が遅くなるので，注意しなければならない．いまの場合には，ランダムな初期条件を避けて，黒または白に揃った状態を初期条件として用いることでほぼうまくいくが，秩序相の様子が未知の場合やあとで述べる 1 次転移の場合にはそう簡単に対処できない．

周期境界条件，自由境界条件の場合，固定境界条件の場合からの類推で，3 番目の固有値について，

$$\tau_2 \sim L^2 \tag{70}$$

となるように思われるが，実際は，

$$\tau_2 \sim L^3 \tag{71}$$

となると考えられている．これは “橋” ないし “帯” のようなスピン配置 (図 5) の影響であるが，この状態の存在は実際上も緩和を妨げる最大の原因になっている．

3.3 平均場近似

イジング模型の振舞い，とくに相転移を理解するために簡単な近似的取扱いを試みる．スピン S_i の磁化の期待値を m_i とおく．スピン間の相互作用 J が場所によらないイジング模型 (15) では，m_i は i によらないので，以下これを m と書く．

$$\langle S_i \rangle = m \tag{72}$$

さて，

$$E(\{S_i\}) = -J\sum_{nn} S_i S_j - h\sum_i S_i \tag{73}$$

において，S_j を期待値で置き換え，

$$E(\{S_i\}) = -J\sum_{nn} S_i m - h\sum_i S_i = -\sum_i (4Jm+h)S_i \tag{74}$$

と近似することを考える．S_i が黒 (白) なら，S_i の近くのスピンは普通より黒 (白) である確率が高いはずだが，いまの近似はこのような空間相関を無視して，すべてを 1 個のスピンについての確率 (1 体の周辺分布) で表わすことにあたっている．

(74) 式は S_i が，外場とまわりの 4 個のスピンの影響を加えた “平均場”$4m+h$ のもとにある単独のスピンとして振舞うことを意味しているから，(22) 式より，

$$\langle S_i \rangle = \tanh(\frac{4Jm+h}{T}) \tag{75}$$

となる．

ここで，(72) 式より $\langle S_i \rangle = m$ でなければならないことを思い出すと，自己無撞着方程式 (self consistent equation)

$$m = \tanh(\frac{4Jm+h}{T}) \tag{76}$$

を得る．以上のような近似法を，平均場近似 (mean field approximation) または分子場近似 (molecular field approximation) という．

$h = 0$ の場合は (76) 式は，

$$m = \tanh(\frac{4J}{T}m) \tag{77}$$

となる．この式は，

$$\frac{J}{T} > \frac{1}{4} \tag{78}$$

のとき，自明な解 $m = 0$ 以外に非零の 2 解 $m = \pm m_\infty (\neq 0)$ を持つ．

$$\frac{J}{T} \leq \frac{1}{4} \tag{79}$$

のときは解は $m = 0$ のみである．

前者 $(J/T > 1/4)$ の場合，$m = 0$ は (77) 式を逐次代入でとく場合，不安定解となる．あとで示すように，この解は自由エネルギーの極大値にあたる非物理的な解である．

これを除くと，自己無撞着方程式の解は図 3 と似た様子を示し，$J/T > 1/4$ が秩序相，$J/T < 1/4$ が無秩序相に対応する．臨界点は $J/T = 1/4 = 0.25$ である．この臨界点の値は真の値 $J/T = 0.4407..$ に比して，かなり弱結合側 (高温側) にかたよっており，2 次元系では平均場近似による相転移の記述があまり正確でないことを示している．一般に平均場近似は次元が高いほど良い．

平均場近似は自由エネルギーに対する近似としても定式化できる．方程式 (76) は次の式の極値をあたえる．

$$\frac{F_{MF}}{N} = -\frac{4Jm^2}{2} - hm + T\{p\log p + (1-p)\log(1-p)\} \tag{80}$$

但し，p はスピンが黒である確率で

$$m = p - (1-p) = 2p - 1 \tag{81}$$

である．

(80) 式を理解する最もよい方法は，ギブス分布を特徴付ける (8) 式の近似式と考えることである．(80) 式と (8) 式の違いは，2 番目の項 (“エントロピー項”) にある．平均場近似は，(8) 式の 2 番目の項で P_α をスピン 1 個に関する確率の積で近似することに当たっている (相関の無視)．

(8) 式にそれを最小にする分布であるギブス分布を代入すると，自由エネルギーになることから，(80) 式に自己無撞着方程式の解を代入したものは自由エネルギーの近似値を与えると考えられる．(80) 式をグラフにしたものを図 6 にしめす．秩序相では，図 6 の右のように，2 つの最小と 1 つの最大 (不安定解に対応) があり，無秩序相では図 6 の左のように最小が 1 つだけある．このあいだにある分岐点が相転移点に対応するわけである．

平均場近似は (8) 式を独立性の制限をつけた部分空間内で最小化しているのに対し，真のギブス分布は無制限の最小化で得られると考えることもできる．このことからすぐ分かるように，平均場近似によって求めた自由エネルギーは真のそれより常に大きい．

平均場近似はスピン間の空間相関を無視した近似なので，画像生成とくにテクスチャー生成のモデルとしては使えない．実際，無秩序相で $h = 0$ の場合，(大域的極小としては) 空間的に一様な解しか得られない．しかし，このことは平均場近似が画像復元に役立たないことを意味するわけではない．これについてはあとでまた触れる．

> 平均場理論の定性的な結果，たとえば，臨界点の近くで秩序変数が (上の例でいえば)
>
> $$m \sim ((J/T) - (J/T)_{critical})^{1/2} \tag{82}$$
>
> のように立ち上がることは，系が少数の秩序変数で記述されることと，秩序変数のみたす方程式が黒と白の対称性を満たす範囲で十分一般である (生成的である) ことだけから理解できる．これをランダウの現象論という．“カタストロフィー理論” の祖型のような話である．
>
> しかし，詳しい研究によると臨界点の非常に近くでは，平均場理論あるいは (素朴な) ランダウの現象論で予言される振舞いは一般には正しくないことが分かっている．たとえば，2 次元イジング模型では (82) の代わりに，
>
> $$m \sim ((J/T) - (J/T)_{critical})^{1/8} \tag{83}$$
>
> となる．これは，臨界点の近傍では現象が本質的に無限次元となることと関連している．これらの振舞いを理解するためには，臨界点でのフラクタル的な性質をとりこんだ理論 (スケーリング理論，繰り込み群の理論) が必要である．

3.4 1 次転移と 2 次転移

いままで $h = 0$ の 2 次元イジング模型で J/T を変えたときの相転移を例にして説明をしてきたが，これは 2 次の相転移 (second order transition) の例であって，相転移にはほかに 1 次の相転移 (first order transition) もある．

1 次の相転移の最も簡単な例は，2 次元イジング模型の秩序相 ($J/T > 0.4407..$) で J/T を固定して磁場 h の符号を変えたときに起きる転移である．これについて，平均場近似を用いて説明しよう．

平均場近似の自由エネルギー (80) で，J/T を秩序相 $(J/T\ >\ 0.25)$ の値に固定して h を負 → 零 → 正と動かすと，グラフの概形は (図 7 の左)→(図 7 の中)→(図 7 の右) のように変化する．磁場 h が零を横切るとき，自由エネルギー最小の点 (•) は左の谷から右の谷へと跳躍する．(図 7) は 1 スピンあたりの自由エネルギーであり，系の大きさ N が非常に大きい場合を考えていることに注意すると，左右の谷の高さが僅かに違っても，非常に鋭い変化が起きることがわかる．

一般に熱力学的極限で自由エネルギーの 1 階微分である秩序変数が不連続に変化するような転移を 1 次転移という (式 (6) 参照)．これに対して $h\ =\ 0$ の 2 次元イジング模型で J/T を変えたときの転移のように，自由エネルギーの 2 階微分ではじめて不連続性が表れる場合を 2 次転移という．この例は外場によって対称性が破れる場合であったが，1 次転移は自発的対称性の破れをともなって起きることもある．あとで説明するポッツ模型の 1 次転移はその例である．

図 6 や図 7 をみると，1 次転移は “準安定状態 $(\mathrm{metastable\ state})$ と別の準安定状態の交換”，2 次転移は “準安定状態の不安定化および別の準安定状態の出現” であると定義したくなる．ここで，“準安定状態” という言葉は，分布が局所的なダイナミクスでは移動しにくい “峰” を複数個持つときに，ひとつの峰に含まれる状態の全体を意味する．この場合，ある準安定状態の “重要度” は準安定状態に含まれる状態の確率やエネルギーの 平均値 で決まるのではなく，確率の 和 で決まる．

> ここで用いた，“分布の峰” という概念はダイナミクス，あるいは，状態間の近さをどう定義するかに依存する．また，あとで述べるように，秩序変数の周辺分布の峰には対応しても，もとの分布の “峰” とは解釈しがたい “準安定状態” もある．これらを考えると，自由エネルギーの n 階微分の不連続性にもとずく次数の定義の方が曖昧さがない．しかし，ダイナミクスに基づく直観的な理解も重要である．実際，上の例での左の谷と右の谷の間の変化には，局所的なダイナミクスのもとでは長い時間がかかり，いわゆる履歴現象 (ヒステリシス) が観察される．

4　格子スピン模型 (格子上のマルコフ場モデル) の例

以下では統計物理における格子スピン模型 (格子上のマルコフ場モデル) について，典型的なものをいくつか紹介する．また，それらの性質や相転移について述べる．以下のモデルで外部から変えられる変数がひとつしかない場合，その外部変数を変えることと温度を変えることはしばしば等価である．このような場合，たとえば，弱結合は高温と同じ意味になる．

4.1　ポッツ模型

イジング模型はスピンが ± 1 の 2 状態のみをとるいわば 2 色のモデルであったが，その多色版がポッツ模型として知られているものである．ポッツは人名である．

q 状態のスピン (ポッツスピン)

$$S_i \in \{1, 2.., q\} \tag{84}$$

を用いて，ポッツ模型のエネルギーは，

$$E = -J \sum_{nn} \delta_{S_i S_j} \tag{85}$$

のように書かれる．ここで，$\delta_{S_i S_j}$ はクロネッカーのデルタで，

$$\delta_{S_i S_j} = 1 \quad (S_i = S_j) \tag{86}$$

$$= 0 \quad (S_i \neq S_j) \tag{87}$$

で定義される．

ポッツ模型で興味深いのは，2 次元の場合，スピンのとりうる値の数 q が $q \leq 4$ か $q > 4$ で相転移の性質に違いがあって前者 (イジング模型を含む) では 2 次転移，後者では 1 次転移になることである．3 次元以上あるいは平均場近似では，$q > 2$ で 1 次転移になる．

ポッツ模型の 1 次転移は，自発的な対称性の破れを伴っている．J/T の値が 1 次転移の転移点の近傍にあるとき，ランダムな状態を初期状態としてマルコフ連鎖モンテカルロ法のシミュレーションを行なうと，次のような様子が見られる．

- 長い間，見かけ上乱雑な状態が続く，ときどき，いろいろな色のスピンのクラスター (空間的なかたまり) ができるが，すぐに消えてしまう．
- 何回もこれを繰り返すうち，q 個のうちどれかある色 (偶然できまる) のスピンのクラスターが “臨界核” の大きさに達する．
- クラスターは成長し，系の大半を占める．特定の色のスピンの “地” の上に他の $q-1$ 色のスピンがばらばらと出現する様相が長い間続く．
- 長い時間の後に，この状態が崩れ，また乱雑な状態に戻る．(以下，これを永遠に繰り返す)．

自発的な対称性の破れている場合でも, イジング模型で見られるような 2 次転移のときは，乱雑な状態は相対的に短い時間しか出現しない．これに対して，1 次転移の転移点近傍では,“乱雑な状態” そのものが準安定状態になっているところに違いがある．

上記のような現象の時間スケールは，J/T の値と系の大きさによって，ワークステーションで簡単に観察できる程度かもしれないし，宇宙の年齢以上かもしれない．

> 1 次転移のあるポッツモデルでは，どの “色” もほぼ同じだけ含んだ状態で構成された “乱雑な状態の集合” 1 個と，q 通りの “色” のうち特定のものが多い “規則的な状態の集合” q 個の合計 $q+1$ 個が競合する．“乱雑な状態の集合” は，含む状態それぞれの確率は低いが，含む状態数は多い．“規則的な状態の集合” は，含む状態の個々の確率は高いが，含む状態数は少ない．両方の要素の兼ね合いで，J/T のある範囲では両者の重みが同じ程度になる訳である．“規則的な状態の集合” でのエネルギーの期待値は，“乱雑な状態の集合” でのエネルギーの期待値よりいつも低い．状態の数の効果 (“エントロピー”) の重要性がよく分かる例である．
>
> “乱雑な状態の集合” については，もともとのギブス分布の “峰” という表現は適切ではないかもしれない．この場合も，適当な秩序変数の組，たとえば各色のスピンの割

合を取り，その空間 ($q-1$ 次元の射影空間) での周辺分布をみれば，確かに $q+1$ 個の“峰” が存在する．

このような型の 1 次転移は珍しいものではなく，身近な例としては液体と固体の相転移がある．この場合，液体 =“乱雑な状態の集合”，固体 =“規則的な状態の集合” であり，前者は 1 通り，後者は結晶のできる位置によって連続無限通りある．

4.2 反強磁性イジング模型

いままでイジング模型で $J > 0$ の場合のみを考えてきた．磁性体理論ではこれを強磁性 (ferromagnetic) という．$J < 0$ の場合 (反強磁性 (antiferromagnetic) という) の場合はどうなるだろうか．

正方格子で磁場 $h = 0$ の場合，格子を市松模様 (奇格子と偶格子) にわけて偶格子のスピンを反転し，同時に $J \to -J$ としてもエネルギーが変わらないことに注意すると，J/T の絶対値が同じ強磁性模型の場合と等価であることが分かる．たとえば，基底状態の対応は図 8 のようになる．もちろん，$h \neq 0$ のときは等価でない．

3 角格子の場合はこのような変換がないので，$J > 0$ の場合と $J < 0$ の場合は本質的に違うモデルになる．たとえば強磁性の場合に存在する $h = 0$ での 2 次相転移は，反強磁性の場合にはない．また，強磁性の場合の基底状態は全部黒と全部白の 2 つだけだが，反強磁性の場合は基底状態が無限個 ($\sim \exp(0.338L^2)$ 個) ある．これらの異常性は，格子が 2 部グラフになっていないことから生ずる．3 角形の面のまわりにスピンを配置する場合，四角形の場合と違って，すべての隣合う頂点が違う色になるように黒と白を配置することができないのである (図 9)．

一般に，ある面をかこむすべての相互作用を同時に “満足” させる (局所的に最小エネルギーにする) ことができないとき，その面 (plaquette) はフラストレート (frustrate) しているという．3 角格子上の反強磁性イジング模型は完全にフラストレートしたモデルの例である．

4.3 ANNNI 模型

こんどは，最隣接相互作用のみでなく，その次 (隣の隣) まで相互作用のあるイジング模型を考える．とくにここでは，正方格子ないし立方格子上で 1 方向にのみ隣の隣まで相互作用 $-J_2$ があり，その符号が最隣接相互作用 J_1 と反対であるようなモデルを考える．2 次元の場合の相互作用を図 1 0 にしめす．このモデルを ANNNI 模型 (anisotropic next nearest neighbor Ising model) という．

一見このようなモデルでは，パターンの定性的特徴は J_2/J_1 のみに依存し，温度によらないように思われる．ところが，これは正しくない．いまの場合，基底状態のパターン ($T = 0$ のパターン) は，J_2/J_1 の値によって図 1 1 に示す 2 種類のどちらか (もしくはこれらを黒白逆転，平行移動したもの) になる．ところが有限温度ではこのどちらとも違った周期のパターンがあらわれる．

簡単のために，平均場近似の場合を考える．隣の隣の相互作用があるのと直交する方向に層にわけて，各層の 1 スピン当り磁化を m_l とすると，自己無撞着方程式は，

$$m_l = \tanh\left(\frac{zJ_1 m_l + J_1 m_{l+1} + J_1 m_{l-1} - J_2 m_{l+2} - J_2 m_{l-2}}{T}\right) \tag{88}$$

となる．z は 2 次元で 2 ，3 次元で 4 である．(88) 式の解を数値的にもとめ，対応する自由エネルギーを最小にするものを拾うと図１２のような極めて複雑な相図が得られる．図１２は $z = 4$ の場合で，数字は周期構造の波数，FM は自発磁化が零でない相 (強磁性相) を示す．

平均場近似の結果 (図１２) がどの程度正しいかは難しい問題である．3 次元では図１２の構造のうち大きな部分は存在するが，細かい部分はゆらぎによって消されてしまうと考えられている．2 次元ではゆらぎが強すぎて，非自明な周期パターンは相としては存在し得ない．しかし，それらは依然，短距離相関 (局部的なパターン) のかたちで見られるかもしれない．

このようなモデルの相転移を記述するには，非常に多くの秩序変数を用意しなくてはならない．しかし，その場合でも分布を定めるには，たとえば最隣接の相関と隣の隣との相関の 2 つがあれば十分である．いわゆる十分統計量と秩序変数は全く違う概念であることを注意しておく．

4.4 離散ガウス模型・ＳＯＳ模型

離散ガウス模型は無限個の離散状態をとる模型である．

$$S_i \in \mathrm{Z} \quad (\text{整数}) \tag{89}$$

を用いて，離散ガウス模型のエネルギーは，

$$E = -J \sum_{nn} (S_i - S_j)^2 \ \square \tag{90}$$

と書かれる．

ＳＯＳ模型 (solid on solid model) はこれと似たモデルで，エネルギーが，

$$E = -J \sum_{nn} \mid S_i - S_j \mid \tag{91}$$

となるものである．

これらのモデル (2 次元) は，結晶の表面の性質の研究などに使われる．この場合 $\{S_i\}$ は表面の高さを表わす．

これらの特徴は，ラフニング転移と呼ばれる独特の相転移がみられることである．これは物理的には，高温側ではなめらかな結晶表面が存在しなくなることにあたっている．この転移の変わったところは，弱結合相全体が臨界点のような状況になっていることである．ＳＯＳ模型についてのモンテカルロ法の結果を図１３に示す．

4.5 平面回転子模型 (古典 $\mathbf{XY}$ 模型)

離散変数のモデルばかり取り上げたので，連続変数のモデルについても触れることにする．一般に応用の広い平面回転子模型 (planar model，古典 XY 模型ともいう) を取り上げることにしよう．このモデルのスピンは位相変数

$$S_i \in [0, 2\pi) \tag{92}$$

である．平面内を回転する矢印のようなものが，正方格子や立方格子上に配列していると考えればよい (図１４)．

エネルギーは，

$$E = -J \sum_{nn} \cos(S_i - S_j) \tag{93}$$

という内積の形をとり，対応するギブス分布は，

$$P(\{S_i\}) d\{S_i\} = \frac{\exp(-E(\{S_i\})/T) \Pi dS_i}{\int_0^{2\pi} \exp(-E(\{S_i\})/T) \Pi dS_i} \tag{94}$$

となる．やはり，$J > 0$ を強磁性，$J < 0$ を反強磁性という．

古典 XY 模型はある種の磁性体のモデルの古典極限として考えられたが，S_i を波動関数の位相と考えることで，超伝導や超流動の問題にも応用できる．また，厳密には平衡統計物理の範疇をはみだす例であるが，リミットサイクルを持つ振動子群の雑音中での振舞いもこれに似たモデルで論じることができる．この場合，S_i はリミットサイクルの位相と解釈される．

３次元 (立方格子) の場合，このモデルは普通の 2 次転移を示す．ただ，離散変数の場合と違って，スピンがそろう方向は任意なので，秩序相は無限種類ある．２次元 (正方格子) の場合はもっと奇妙で，ＫＴ転移 (Kosterlitz Thouless transition) と呼ばれる相転移が起きる．ＫＴ転移の場合，強結合側全体が臨界点のような性質をもっていると考えられている．この転移は前に述べたラフニング転移と深い関係があるが，物理的には “渦対の解離” として記述できる (図１４)．以上は強磁性の場合である．反強磁性でフラストレーションがある場合にも興味深い現象が報告されているが，ここでは触れない．

5　ランダム系の統計物理

ここではいわゆるランダム系の統計物理について解説する．ランダム系の統計物理は比較的最近発達した分野であるが，統計物理の統計的情報処理への応用を考える上では非常に重要と思われる．一般にベイズ的な復元問題では，対応するランダム系の問題でランダムに決められる部分に，雑音を含んだデータがはいることが多い (たとえば画像復元問題での “磁場”)．

5.1　ランダム系の物理

いままで扱ったのは，いわば均質な系の統計物理であった．それに対して，ランダム系とここで呼ぶのは，モデルにはじめから不純物などの不均一性があって，その不均一性が確率的に定められている場合である．

この場合，それらの乱れは最初に一度定めたら動かさないとする．それらの乱れも熱ゆらぎで動くとすれば，乱れを定める変数を含めたギブス分布を考えることに帰着されてしまい，本質的に均質な系の統計物理と変わらなくなってしまう．この点を強調する場合，“凍結された乱れ (quenched randomness) を持つ系” という言葉を使うことがある．以下では凍結された乱れについての平均を “ランダム平均” と呼び，$\langle\ \rangle_{random}$ と書く．これはギブス分布による平均 (“熱平均”) を表わす $\langle\ \rangle$ と区別するためである．

多くの場合，実際に実験の対象になるのは 1 個の大きな系である．このような系に関して，ランダム平均が正しい答えを与えるという保証はない．ある量について，ランダム平均が正しい答えをだす場合，その量は self averaging であるという．多くの場合，物理量は self averaging であるが，常にそうだとはいえない．

解析的な手法で研究する際にとくに問題となることであるが，分配関数 Z のランダム平均 $\langle Z\rangle_{random}$ を計算してもあまり意味がないことを注意しておく．物理量は自由エネルギーの微分で書けるので，必要なのは $\log Z$ のランダム平均 $\langle \log Z\rangle_{random}$ である．前者をアニール平均，後者をクエンチ平均と呼ぶ場合がある．

> これは，次のように考えてもわかることである．ほとんどお互いに無関係な多数の部分からなる大きな系を考えよう．このとき，Z は各部分の寄与の積になるのに対し，$\log Z$ は各部分の寄与の和になる．系を大きくしたとき，Z の分布は対数正規分布に近付くのに対し，$\log Z$ の分布はガウス分布に近くなる．前者の場合に平均値で全体を代表させるのはあきらかに問題がある．

以下では，ランダム系の例として，最隣接相互作用をするイジング模型のランダム版を 2 種類あげる．いずれもその特異な性質に興味がもたれ，詳しい研究の対象になっている．

5.2　スピングラス模型

まずはじめに，結合定数が $\{J_{ij}\}$ が場所によってランダムに定められる模型を考えよう．

このような模型のうちとくに興味がもたれているのは，$\{J_{ij}\}$ が正負両方の値をとる場合で，イジングスピンではない場合も含めて，一般にスピングラス (spin glass) 模型といわれている．このような物理系はたとえば，化学的性質の近い強磁性体と反強磁性体を混

ぜ合わせることによって実現できる．グラス (glass) といっても，本当のガラスは不純物のために結晶化が妨げられているわけではないので，スピングラスとはあまり似ていない．

もっとも簡単な場合として，イジング模型で結合定数の絶対値を一定として，符号を確率 p で正，$1-p$ で負となるようにとった場合を考えよう．式で書くと，イジングスピン $\{S_i\}$ を用いて，

$$E(\{S_i\}) = -\sum_{nn} J_{ij} S_i S_j - h \sum_i S_i \tag{95}$$

$$P_r(J) = p\delta(J - J_0) + (1-p)\delta(J + J_0) \tag{96}$$

$$J_{ij} \leftarrow i.i.d.P_r(J) \tag{97}$$

と書ける．ただし，(97) 式の意味は J_{ij} を各隣接対 (i,j) に対してそれぞれ独立にえらぶという意味である．このモデルはしばしば $\pm J$ スピングラス模型と呼ばれる．$P_r(J)$ をガウス分布にとったモデルも考えられている．

これらのモデルの特徴は，図１５のようにフラストレートしている面とそうでない面が混在するため，$\{J_{ij}\}$ に依存するきわめて複雑な基底状態を持つことである．$h = 0$ の場合，基底状態をもとめる問題に対しては，2 次元 (正方格子) では多項式算法が存在するが，3 次元 (立方格子) では存在しない (いわゆるＮＰ完全問題になる)．

いま，どの 1 個のスピンを動かしてもエネルギーが大きくなるような状態を局所的極小として定義すると，$\pm J$ スピングラス模型 $(p \sim 0.5, h \sim 0)$ には非常に多くの局所的極小が存在する．この様子はしばしば図１６のように表現されるが，横軸は本来無限次元の空間であることに注意しなくてはならない．

局所的なダイナミクスのもとでの緩和時間は J/T が十分大きいときはそれほど長くないが，J/T の大きいところ (低温領域) では非常に大きくなる．これは，物理的にみると，低温では，熱雑音の大きさが図１６の谷と谷の間を乗り越えるのに十分でなくなるため，$T = 0$ での局所的極小に対応した準安定状態が生じてくるためである．

$\pm J$ スピングラス模型 $(p = 0.5, h = 0)$ でなんらかの相転移が起きるか，起きるとしたらどういう相転移か，という問題は，かって盛んに議論された．現在では，2 次元 (正方格子) の場合には相転移がなく，3 次元 (立方格子) の場合には $T/J \sim 1.2$ である種の相転移があると信じられている．この場合の秩序変数としては，たとえば，

$$q = \frac{1}{N} \sum_i \langle S_i \rangle^2 \tag{98}$$

(あるいはそのランダム平均) が提案されている．この量は基底状態がわからなくても定義できる．

> $T \neq 0$ でのパターンを特徴付けるために，共分散行列 $\{\langle S_i S_j \rangle\}$ の固有値問題を調べた研究がある (主成分分析あるいは KL 法に対応する)．温度 T が大きい極限では $\{\langle S_i S_j \rangle\}$ の固有ベクトルは $\{J_{ij}\}$ の固有ベクトルと一致するが，温度を下げて行くとそれからずれてゆく．

5.3 ランダム磁場模型

次に磁場 h_i が場所によってランダムな場合を考える．式で書くと，たとえば，

$$E(\{S_i\}) = -J \sum_{nn} S_i S_j - \sum_i h_i S_i \tag{99}$$

$$P_r(h) = p\delta(h - h_0) + (1-p)\delta(h + h_0) \tag{100}$$

$$h_i \leftarrow i.i.d.P_r(h) \tag{101}$$

となる．これをランダム磁場イジング模型 (random field Ising model) という．以下では $J > 0$ の場合のみ考える．

このモデルも多くの準安定状態を持つ場合がある．その場合，低温での緩和時間が非常に長くなる．基底状態を求める問題は決して自明ではないが，$J > 0$ なら，任意の次元で多項式算法が存在することが知られている．これを示すには，基底状態を求める問題をグラフの最小カットを求める問題に帰着させる．後者は多項式算法で解けることが知られている．この対応付けは次のようにして行う．まず，与えられた格子上の点以外に U と D の 2 点をとる．ランダム磁場が $+h$ の点と U，$-h$ の点と D を結び，これと元の格子を合わせた全体をひとつのグラフと見る．次に，このグラフの辺上に流量を割り付ける．格子の各辺にはすべて J を与え，U，D と格子の各点を結ぶ辺には h を与える．$J < 0$ だと，流量に正負両方の符号が生じてうまくいかない．

いま，$p = 0.5$ の場合を考える．$J = 1$ とおいて，h と温度 T を独立な外部変数と考える．このとき，平均場近似は，十分弱い有限の h について，T を十分小さくすれば，強磁性相 (固定境界条件のもとで 1 スピン当りの磁化 m の熱力学的極限が有限) があることを示唆する (5.4.1参照)．これによれば，十分弱いランダム磁場は強磁性相を壊さないということになる．これが事実かどうかは難しい問題であるが，2 次元の場合は誤りである (どんな弱いランダム磁場も長距離秩序を壊す) ことがわかっている．3 次元の場合は正しいと思われている．

5.4 ランダム系の平均場理論

簡単な例ではひとつの概念と見えたものが，問題が難しくなると “縮退” がとけて，いくつもの概念に分裂することは珍しくない．平均場近似の場合も，ランダム系に適用した場合，解釈によっていろいろな近似になりうる．ここではその例をしめす．

5.4.1 空間ゆらぎを完全に無視した場合

ランダム磁場イジング模型の平均場近似を考える．

まず，ギブス分布に関する平均だけでなく，ランダム磁場に関する平均についても空間的ゆらぎを無視する場合を考えよう．各スピンの磁化のギブス分布による期待値をさらにランダム平均したもの $\bar{m}$ を導入する．ランダム平均を取っているので，この値はどのスピンについても同じである．

$$\bar{m} = \langle\langle S_i \rangle\rangle_{random} \tag{102}$$

すると，モデル (99,100,101) の “平均場近似” は次の自己無撞着方程式をあたえる．

$$\bar{m} = p\tanh(\frac{zJ\bar{m} + h}{T}) + (1-p)\tanh(\frac{zJ\bar{m} - h}{T}) \tag{103}$$

ここで z は最隣接格子点の数．

$p = 0.5$ の場合にこの式及び対応する自由エネルギーから導かれる相図は図１７のようになる ($J = 1$ とおいている)．

5.4.2　ランダムネスの空間ゆらぎを残した場合

同じランダム磁場イジング模型の平均場近似でも，ランダム磁場のゆらぎを残した近似も可能である．

この場合，磁化の期待値としてスピンごとに違う値 m_i を割り当てることになる．自己無撞着方程式は，N(スピン数) 個の連立方程式となる．

$$m_i = \tanh(\frac{J\sum_{nn} m_j + h_i}{T}) \tag{104}$$

対応する自由エネルギーは，

$$F_{MF} = -J\sum_{nn} m_i m_j - \sum_i h_i m_i + T\{p_i \log p_i + (1-p_i)\log(1-p_i)\} \tag{105}$$

但し，p_i はスピン S_i が黒である確率で

$$m_i = p_i - (1-p_i) = 2p_i - 1 \tag{106}$$

となる．

自己無撞着方程式 (104) の解として得られる m_i は空間相関を持つが，ギブス分布という観点からみると，これらはあくまで一体の確率であり，熱ゆらぎに関してはスピンの空間相関は無視されている．

実際にこの近似で求められたパターン $(T=0)$ を図１８に示す．この場合，磁場 $\{h_i\}$，スピン間の結合 J，１体近似の範囲の熱ゆらぎが協調して図１８のようなパターンができているわけで，画像修復に平均場近似がある程度有効なことが示唆されている．

また，磁場について空間平均してしまう場合と違って，この近似は $T \to 0$ で原理的に正しい基底状態を与える．原理的といったのは，自己無撞着方程式の解が沢山ある場合があり，この場合には解は初期条件に依存するからである．基底状態はこれらの多くの解のひとつにすぎない．たとえば，図１８の (a)，(c) はどちらも同じ系 ($\{h_i\}$ がまったく同じ) についての自己無撞着方程式の解であるが，異なる部分がある．

３次元の立方格子の場合にこの近似による相図が作られているが，結果は定性的に図１７に似ている．

5.4.3　相互作用の範囲を無限大にした場合

最後に相互作用の範囲を無限大にしたモデル (いわゆる “全部つながっている” モデル) を考えよう．

“全部つながった” 強磁性のイジング模型

$$E(\{S_i\}) = -\frac{J}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} S_i S_j \tag{107}$$

の場合，熱力学的極限で平均場近似が厳密であることが鞍点法を用いて示せる．

ただしこの場合，J をスピンの数に反比例して減少させつつ熱力学的極限をとらなくてはならない．すなわち，J^* を一定として，

$$J = J^*/N \tag{108}$$

とする．

この模型のスピングラス版がいわゆるＳＫ模型 (Sherrington Kirkpatrick model) である．

$$E(\{S_i\}) = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} J_{ij}S_iS_j \tag{109}$$

$$P_r(J) = \frac{1}{\sqrt{2\pi\tilde{J}^2}}\exp(\frac{-(J-J_0)^2}{2\tilde{J}^2}) \tag{110}$$

$$J_{ij} \leftarrow i.i.d.P_r(J) \quad (i>j) \tag{111}$$

$$J_{ji} = J_{ij} \tag{112}$$

ただし，こんどは

$$J_0 = J_0^*/N \tag{113}$$

$$\tilde{J} = \tilde{J}^*/\sqrt{N} \tag{114}$$

とする．

このＳＫ模型を解析的に扱う場合に計算すべきものは $\langle \log Z\rangle_{random}$，すなわち，

$$\int\{\log\sum_{config.}\exp(-\sum_{ij}\frac{J_{ij}}{2}S_iS_j)\}\ \frac{1}{\sqrt{2\pi\tilde{J}^2}}\Pi_{i>j}\exp(\frac{(J_{ij}-J_0)^2}{2\tilde{J}^2})\Pi_{i>j}dJ_{ij} \tag{115}$$

である．$\sum_{config.}$ は 2^N 通りのスピン配位についての和であり，$\int\Pi_{i>j}dJ_{ij}$ は $\frac{N(N-1)}{2}$ 個の結合定数 J_{ij} についての積分である．

(115) 式の計算を難しくしているのは和と積分のあいだに入っている log であるが，レプリカ法 (replica method) では，

$$\langle \log Z\rangle_{random} = \lim_{n\to 0}\frac{\langle Z^n\rangle_{random}-1}{n} \tag{116}$$

と書きかえることで，この困難を回避する．

ＳＫ模型では，n が正の整数の場合，$\langle Z^n\rangle_{random}$ の計算はスピンが nN 個あるランダムでない系の分配関数の計算に帰着される．さらに，相互作用の範囲が無限大であることを利用すると，この量の N が大きいときの漸近形を鞍点法によって求めることができる．これで計算できるのは，あくまで n が正の整数の場合に限られるが，それを強引に $n\to 0$ に “解析接続” することで，(116) の右辺を計算しようというわけである．

この場合，もっとも議論されたのは n 個の鞍点の選び方である．これらをすべて同じとして得られた解 (SK 解，レプリカ対称解) は低温で非物理的になることがわかり，非常に複雑な鞍点の選び方 (replica symmetry breaking) を行った解 (Parisi 解) が提案された．この解は全く違う (同じぐらい複雑な) 方法で得られた解やシミュレーションの結果と一致するので，正しいと思われている．鞍点の選び方の問題は，図１６のようにたくさんの準安定状態が存在して，それらの間の関係に階層的 (ultrametric) な性質があることと関係がある．

解析的研究の結果としてわかったことは，この模型の解は，(104) のような型の方程式をみたさないことである．直接結合しているスピンの影響を表わす項でけでなく，もうひとつ高次の効果 (反作用場， reaction field) まで考慮しなくてはならない．これは相互作用の範囲が無限大のモデルについての従来の常識とは異なっている．

最終的な相図としては図１９のようなものが得られている．P は無秩序相 (常磁性相)，F,F' は強磁性相であり，SG がこのモデル特有の凍結した相 (スピングラス相) である．また，準安定状態の分布についても詳しく調べられている．それによれば，系の大きさ N が大きいとき，“重要でない” 準安定状態の数は非常に多い ($\exp(aN)$ の程度) が，“重要” なものは非常に少ない (平均して 1 個から数個の程度) という結果が得られている．

ＳＫ模型の性質は短距離相互作用をするモデルとはかなり違っているらしい．むしろ，工学にあらわれる組合せ的最適化問題 (ランダム問題の場合) にＳＫ模型に似たものがあると考えられている．また，いわゆる連想記憶モデルとの関連も論じられている．

6　付録: マルコフ連鎖モンテカルロ法の技術的側面

6.1　補足的注意

ここで述べるのは文字通り補足的・初歩的なことである．とくに計算のベクトル化などには全く触れていない．これは筆者がその方面に暗いためであるが，本や論文に書いてないような常識的なことを書いておくのも，それなりの役に立つと思ったからでもある．

緩和時間の判断

前にも述べたが，マルコフ連鎖モンテカルロ法の緩和時間を見つもることは困難なことが多い．緩和が多少遅い場合よりも，きわめて遅い場合の方がかえって気づきにくいので注意する必要がある．緩和時間を長くさせる要因としては，たとえば次のような場合がある。

- いわゆる悪条件 (ill-condition) の場合．すなわち，状態空間に細長い “山” (エネルギーでいえば “谷”) がある場合．狭義のメトロポリス法は $T \to 0$ とすると一種の最急降下法になる．このことからもわかるように，局所的なダイナミクスを持つマルコフ連鎖モンテカルロ法は悪条件に弱い．

  相転移とダイナミクスの項 (3.2節) で触れた “critical slowing down” や空間パターンが違う性質の領域に分かれてその境界が乱歩する場合なども、一種の悪条件ともいえるが、ガウス分布での悪条件の場合とは本質的に違う面を含んでいる。

  悪条件に陥っていないかどうか試すには，長いシミュレーションで測定量がトレンドを持たないかどうか調べるのがひとつの方法である．

- 準安定状態が複数あって抜けでるのに長時間を要する場合．

  普通の数値的最適化法が局所的極値にとらえられてしまうような場合でも，マルコフ連鎖モンテカルロ法は確率的要素を含むため，そこから抜けだせる可能性がある．しかし，有利なことばかりではない．たとえば，局所的極値は単に “抜け” れば良いのではなく，もしそのまわりの準安定状態がある程度重要であれば，またその近くに戻ってこなければならない．準安定状態 A,B の重要度 (含む状態の確率の和) が同じなら, 詳細釣合条件により， A を抜けて B に移るのに要する時間の期待値と A にまた戻ってくるのに要する時間の期待値は同じになる．また， A の重要度が B の $1/10$ なら，戻ってくるのに要する時間は抜けるのに要する時間の 10 倍程度になる．ここで，エネルギーの値は準安定状態の重要度の目安にはならないことに注意する必要がある．各準安定状態に含まれる状態の数は著しく違うかもしれない．

  ポッツ模型 (4.1節) の項で説明したように，最適化の場合の局所的極小に対応しない型の準安定状態もありうる．この場合，確率的要素のためにかえって狭く高い “峰” が見つけにくくなっているともいえる．非一様な系でそれに類した現象が起きた場合，大変厄介である．

  複数の準安定状態のひとつだけに捉えられていないことを確かめるには，初期状態を変えて何回もシミュレーションを行ない，結果が一致するかどうか調べるしかない．

この場合，ランダムな初期状態だけでなく，極端な初期状態も試みるとよい．もし，分布のモード (基底状態) がわかれば，モードを初期状態にとったときとランダムに初期状態をとった場合を比較するのも有益である．“良い初期状態” を選んだときだけ，良い (と思う) 答がでれば満足するという考え方は，問題の構造を完全に把握していない限り危険である．

マルコフ連鎖の最初の部分をどの程度 “捨てる” (物理量の測定に使用しない) のがよいかという問題は，実際上重要であるが，さほど簡単ではない．理論的にいえば，よい結果を得るためには，最長の緩和時間 τ_1 よりずっと長い間マルコフ連鎖上の平均をとる必要がある．これに対し，はじめの部分はせいぜい緩和時間の数倍の長さ捨てれば十分なはずである．この考えに従えば，捨てる部分は使う部分の $1/50$ とか，せいぜい $1/10$ で十分なはずである．しかし，多くの場合，初期状態はギブス分布からの “典型的なサンプル” ではない．これは式 (33) で，係数 a_1 や a_2 がギブス分布からの “典型的なサンプル” を初期状態にとったときと比べて非常に大きい値をとる可能性があることを意味する．このように考えると，最初の方を沢山捨てるのも理由がないわけではない．単一の熱力学的極限が存在して，系が十分大きく，測定量が巨視的とみなせる場合には，最初の方を沢山捨てるのが確かに良いが，統計的情報処理への応用では，これらの条件が満たされない場合も多い．筆者は，生成したマルコフ連鎖のはじめの $1/2$ から $1/4$ 程度を捨てて，残りを期待値の計算に使用しているが，これで良いのかどうかはわからない．

近似的な条件つき確率を用いた Gibbs Sampler

ステップ (2) の条件つき確率の計算が完全にできない場合は，近似的な重み関数 $h_i(\alpha_i)$ を用いて，次のようなアルゴリズムを構成することができる．これは，条件つき確率の部分に “従来型” のモンテカルロ法の考えを利用しているとも解釈できる．

はじめに，α_j のとりうる値 x に対して，“仮の確率”$h_j(x)$ を定める．$h_j(x)$ は簡単にサンプルできるものを選ぶ．$q_j(x) \neq 0$ ならば $h_j(x) \neq 0$ を満たす必要がある．この $h_j(x)$ を利用して以下の手続きを構成する．

1. 適当な方法で，“動かす変数” の番号 j を決める．
2. 確率分布 $h_j(x)$ からのランダムサンプルを γ_j とする．
3. もとの状態 $\alpha = \{\alpha_1..\alpha_j..\alpha_I\}$ と動く先の候補 $\gamma = \{\alpha_1..\gamma_j..\alpha_I\}$ について，

$$\triangle E = E_\alpha - E_\gamma \tag{117}$$

$$w = \exp(\triangle E/T)\frac{h_j(\alpha_j)}{h_j(\gamma_j)} \tag{118}$$

を計算する．

4. (0,1) の一様乱数 RND を発生し，$RND < w$ なら，α_j を γ_j で置き換える．

アルゴリズム構成上の常識

- 複数のマルコフ連鎖の合成

  複数のマルコフ連鎖を組み合わせる方法は，複雑な系のマルコフ連鎖モンテカルロ法を構成するために有益である．たとえば，系が 2 種類のスピン $\{\psi_i\}$，$\{\phi_i\}$ からなるときは，$\{\psi_i\}$ を動かすサブルーチンと $\{\phi_i\}$ を動かすサブルーチンを別々に作って，交互に呼べばよい．いくつかのマルコフ連鎖を合成したときに，ひとつひとつの過程が詳細釣合条件をみたせば，合成された過程もみたす．遷移グラフの強連結性の条件の方は合成された過程がみたせば十分である．

- 不等号による制約の処理

  狭義のメトロポリス法で変数の動く範囲が不等号で制約されている場合は誤りやすい．もっとも簡単な方法は，制約のない場合のマルコフ連鎖を構成した上で，単純に “はみ出す場合は試行を受理しない” とすることである．制約のない場合に詳細釣合条件がみたされていれば，こうして作ったマルコフ連鎖も詳細釣合条件をみたす．

プログラム上の常識

- 場合分け

  狭義のメトロポリス法の場合，$\triangle E_{\alpha\gamma}$ の計算では，$\triangle E_{\alpha\gamma} \geq 0$ の場合を分けるのが普通である．条件文がひとつ増えるかわり，該当する場合の乱数と $\exp$ の計算が回避できる．また，これによって，無用の $\mathrm{overflow}$ が $\exp$ の計算で生じるのが防げる場合がある．計算機によっては，$\mathrm{underflow}$ を防止するために，$\exp$ の中身 $\triangle E_{\alpha\gamma}/T$ が小さ過ぎるときの場合わけも必要になる．

- 乱数

  乱数を複数の手続きに使用するので，乱数に系列相関があるとおかしな結果になることがある．実際には，乱数関係でおこるひどい失敗は，初歩的なプログラムミスによるものが多い．

  よくある間違いの例．
  以下では，IR を $\{1..M\}$ の整数乱数，RND を $(0,1)$ の実数乱数とする．

  - IR を下位ビットがランダムでない整数乱数 IRND から IR=MOD(IRND,M)+1 として作ってしまった．
  - IR を RND から IR=INT(RND*M)+1 として作ったところ，IR=M+1 となる場合が生じてしまった．これはプログラムの移植のさいに問題になることがある．
  - 単精度実数を供給する乱数ルーチンに，間違って倍精度の配列を渡してそれに乱数を取り込んでしまった.

  乱数ルーチンは最も内側のループで呼ばれるから，呼びだしのオーバーヘッドが大きい．乱数をまとめてとるか，乱数ルーチンをプログラム中に展開すると早くなる．

- バグ

  マルコフ連鎖モンテカルロ法は，プログラムが簡単な代わり，バグがわかりにくいという特性がある．これは乱数を使うためもあるが，ひとつひとつは簡単なステップがより集まって複雑な効果を生ずるというアルゴリズムの特性によるところが大きい．もし，小さい系での厳密な計算との比較テストが可能なら，はじめにやっておいた方が良い．

- プログラムの移植

  移植して結果が変わる場合には本当に間違っている場合とそうでない場合がある．本当に間違っている場合は，疑似乱数ルーチンの移植性がないためであることが多いので，まずそれをテストするのが良い．

  本当は間違っていない場合も多い．これは，マルコフ連鎖モンテカルロ法のアルゴリズムが実数比較を含むことから生ずる．とくに，場合分けによって乱数を節約した場合は，乱数列がひとつ分ずれることで，それ以降まったく違う状態列を生じて驚かされることがある．他に問題がなければ，どちらの状態列もほぼおなじ期待値を与える．違うけれども，どちらも正しいわけである．

6.2 新しいアルゴリズム

いままで，説明してきたアルゴリズムは物理的なアナロジーにもとづくものと，それに多少手を加えたものであった．マルコフ連鎖モンテカルロ法を純粋に計算手法としてとらえるならば、実際の物理系の緩和との対応がつく必要はないとも考えられる。最近，この考えに従って ，さまざまな非局所的ダイナミクスが提案されている．また，学習によって効率のよくなるアルゴリズムや本来必要とされている確率分布を拡張したものをシミュレートする方法なども考えられる．以下ではそれらについて簡単に解説する．

> “新しい” 方法ではないが， ギブス分布からのサンプルをつくり出す方法として，ランジュバン方程式によるものがある．ランジュバン方程式による数値計算は，素粒子物理などでは有力な手段であり，統計的情報処理に応用する試みもいくつかある．これについては，Neal(1996), Kronfeld(1993), Amit et al.(1991) などを参照されたい．

6.2.1 非局所的ダイナミクス

非局所的なダイナミクス，すなわち，一度に沢山の局所的変数を変化させるアルゴリズムをうまく設計するのは難しい．狭義のメトロポリス法の場合一度に動かす局所的変数の数を単純に増やすと， acceptance ratio が下がるため，効率が悪くなることが多い．また，Gibbs Sampler の場合，条件つき確率の計算がどんどん面倒になっていく．個々の問題に応じて上手な方法を考えないと駄目である．

イジング模型，ポッツ模型など，本稿で論じた大部分の模型については，効率の良い非局所的ダイナミクスは知られていなかった．比較的最近になって， Swendsen と Wang が格子上のイジング模型 (一般にポッツ模型) について，臨界点近傍での緩和を著しく早くするような，非局所的なアルゴリズムを与えた (Swendsen and Wang (1987), Wang and Swendsen(1990))．この方法では，変化させるスピンのクラスターをランダムに選ぶのではなく，現在の状態 (スピン配置) を参照して適応的に作る．適応的な作り方をして，かつ，詳細釣合条件を満たすのは容易ではないが，彼らはイジング模型の特別な表現を利用するとそれが可能になることを示した．この結果は Wolfe によって平面回転子モデルなどに拡張された (Wolff(1989), Wang and Swendsen(1990), 鏡映変換を利用するのがポイント)．

Swendsen-Wang の方法は，その後もさまざまな拡張がなされているが，ランダム系・非一様な系への拡張は必ずしも容易ではない．ランダム磁場の問題やそれに関連した画像処理の問題に適用する試みがよっていくつか報告されている (Marinari and Parisi(1992), Gray(1993)) が，単純な適用はうまくいかないようである (少し拡張した方法が Higdon(1993) にあるが，詳細不明)．より興味深いのは， Swendsen と Wang 自身による格子上のスピングラス模型への適用である (Swendsen and Wang(1986,1988))．要点は，さまざまな温度の系を同時に結合させてシミュレートするところにある．これを巧妙に行なうことで，どの温度の系もその温度でのギブス分布を実現し，かつ高温の系で検出された同時に動きやすい部分の情報が自然に低温の側に伝わっていくようにするのである．この発想は，あとで述べる拡張アンサンブルの考えとも共通性がある．しかし、この方法も、 2 次元のスピングラス模型ではうまくいくが、 3 次元では十分に機能しないようである。

非局所的アルゴリズムへの別のアプローチとしては，階層的な方法 (multi-grid 的な方法) がある (いわゆるモンテカルロ繰り込み群とは別物)．これについては Goodman and Sokal(1989)

を参照されたい．さらに，階層的な方法を Swendsen-Wang の方法と組み合わせてイジング模型の臨界現象を扱った研究もある (Kandel et al.(1989))．Kandel らのアルゴリズムが実用的かどうかは不明であるが，単純な系に対するマルコフ連鎖モンテカルロ法のプログラムとして最も複雑なものであることは確かである (一段階層を上がるごとに格子のネットワークの形が変わり，またステップごとにも変わる)．

6.2.2 ダイナミクスの学習

シミュレーションを進めていくにしたがって，ダイナミクス自体が変わっていき，しだいに緩和時間が短くなるような方法を設計できれば有用であろう．この場合，途中では厳密にマルコフ連鎖モンテカルロ法の満たすべき条件が成立しなくてもよいだろう．十分に “学習” させてから，ダイナミクスを固定して，サンプルをとりはじめれば良いからである．

この方向の仕事としては，タンパク質のシミュレーションにおいて，狭義のメトロポリス法におけるステップ幅 (ϵ) の学習を扱った仕事がある (Bouzida et al.(1992))．この場合，ϵ は 1∼3 次元程度のベクトルであるが，対応する状態変化は非局所的なものである. しかし，状態変化の種類そのものを学習によって獲得する試みはなされていないようである．また，Bouzida et al. は，学習を行ないながら (ダイナミクスを固定しないで) サンプルを取ることを提唱しているが，これは誤った結果を招く可能性がある．逆にダイナミクスを固定した場合にこの方法でどの程度の効率の上昇があるかは不明である．

Swendsen and Wang とは別の方針でクラスターダイナミクスのスピングラス模型への適用を試みたものが Liang(1992) にあるが、そこでもある種の学習が試みられている. 近似的な重み確率を用いた Gibbs Sampler において，重み $h_j(x)$ を学習させる試みもあるらしい．

これらの試みの多くは，実用的には良い結果を生む場合があるとしても，どちらかというと小手先の改良という感を免れえない．これは，学習した結果を記憶するための良い表現がないことに起因している．数値的最適化における準ニュートン法も一種の自己改良アルゴリズムとみなせるが，この場合は，最終的に極値に到達するのが目的であるため，ヘッセ行列の逆行列 (ガウス分布の共分散行列に相当) に情報を貯蔵することで十分であった．マルコフ連鎖モンテカルロ法は分布からのサンプルを得るのが目的であって，真に大域的な面をもっているので，ガウス分布による表現では十分とはいえない．

すぐあとで説明する拡張アンサンブルの方法のいくつかも “学習” という側面を含むが，学習させるのがダイナミクスでなくシミュレートする確率分布自体である点が異なっている．

6.2.3 拡張アンサンブル

必要な確率分布をなんらかの意味で拡張ないし修正した確率分布に関して，それを不変にするマルコフ連鎖を構成する方法が注目されている．

1. “Metropolis-coupled Markov Chain Monte Carlo” algorithm
   “時間的一様な並列アニーリング” algorithm
   “(レプリカ) 交換” algorithm

もっとも簡単なのは，Geyer(1991) 及び木村・瀧 (1990) によって提案された方法である (Geyer and Thompson (1995)，小西ほか (1995) も参照)．Geyer はこれを “Metropolis-coupled Markov Chain Monte Carlo”(MCMCMC) と呼んで，統計学への応用を論じている．木村・瀧は simulated annealing の新しい方法として説明しているが，この方法の利点は与えられた分布からのサンプルを正しく生成できることにあるのだから，最適化への応用に限定するのはもったいない．同じ方法は，伊庭 (1993a,1993b)，及び，Hukushima and Nemoto(1996) によっても考えられた．後者では 3 次元のスピングラス模型への応用が論じられている．

“MCMCMC” あるいは交換法では，外部変数のことなる確率分布の族をまとめて，並列的にシミュレーションする．その際に同時分布関数を不変に保つような ”入れ替え ”を考えることで，準安定状態にとらえられることを防ごうとする．たとえば，外部変数のことなる確率分布として温度 T_m の違うギブス分布の族 $\{P_m(x_m)\}$

$$P_m(x_m) = \frac{\exp(-\frac{E(x_m)}{T_m})}{Z_m} \tag{119}$$

を考えた場合には，同時分布関数 $\Pi_m P_m(x_m)$ を不変に保つ入れ替えのアルゴリズムとして，

(a) $E(x_{m+1})$, $E(x_m)$ を計算する．

(b) $\triangle = -(\, E(x_{m+1}) - E(x_m)\,)(\, \frac{1}{T_{m+1}} - \frac{1}{T_m}\,)$ を計算する．

(c) 乱数 $rnd \in [0,1]$ を発生させ，$rnd < exp(-\triangle)$ なら x_m と x_{m+1} を “入れ替える”．
すなわち，いままで，温度 T_m で動いていた系の状態を初期状態として温度 T_{m+1} の計算をはじめ，温度 T_{m+1} で動いていた系の状態を初期状態として温度 T_m の計算をはじめる．

という方法が考えられる．各温度 T_m の系はそれぞれ独立に狭義のメトロポリス法もしくは Gibbs Sampler によって時間発展するとし，それに各 m についての上述の操作を組み合わせたものが，全体のアルゴリズムを構成する．各 m について T_m と T_{m+1} が十分近ければ，入れ替わりが頻繁に起こり，高温での速い緩和が，低温での緩和を促進する効果を持つことが期待される．

入れ替わりが頻繁に起こるために必要な $\beta_m = 1/T_m$ と $\beta_m = 1/T_{m+1}$ の間隔 $\delta\beta$ はどの程度になるだろうか．

$$\sigma_E^2 = \frac{d\langle E\rangle}{d\beta} = \langle (E - \langle E\rangle)^2\rangle \tag{120}$$

とすると，

$$\langle E\rangle_{T_m} - \langle E\rangle_{T_{m+1}} \sim \sigma_E^2 \delta\beta \tag{121}$$

であり，エネルギーの揺らぎは σ_E 程度である．エネルギー期待値の差と $\delta\beta$ の積が十分小さい条件は，

$$(\langle E\rangle_{T_m} - \langle E\rangle_{T_{m+1}})\delta\beta \sim 1 \tag{122}$$

エネルギー期待値の差とエネルギーの揺らぎが同じ程度になる条件は，

$$(\langle E\rangle_{T_m} - \langle E\rangle_{T_{m+1}}) \sim \sigma_E \tag{123}$$

となるが，両者はともに，

$$\delta\beta \sim 1/\sigma_E \tag{124}$$

に帰着される．系の大きさ N に対して σ_E^2 は $O(N)$ の量であるから，同じ温度範囲に必要な系の数は $O(\sqrt{N})$ で増大することになる．また，臨界点の近傍では σ_E が大きいので，多数の系が必要なこともわかる．

2. "simulated tempering" algorithm

類似の方法は， Marinari and Parisi(1992) によっても提案され，ランダム磁場のイジング模型に対して効果的であることが示されている ("simulated tempering")．この方法では温度のことなる多数の系を同時にシミュレートするかわりに，ひとつの系で温度が内部変数となっていると考える．温度を内部変数と見なして狭義のメトロポリス法によって変更しても， 温度に関する条件つき分布 はその温度でのギブス分布になっている．しかし，単に温度を変数としただけでは，高温の状態の方が圧倒的に起こりやすくなり，欲しい領域での情報はほとんど手に入らない (バイアスはないが分散が大きくなる)．これは，普通の温度を固定した方法では，欲しい領域の情報が十分に含まれている代わり，低温で局所的極値のまわりの準安定状態にトラップされやすくなるのと対照的である． simulated tempering では，温度に関して罰金項を設けて，

$$E_{ST} = E/T - g(T) \tag{125}$$

をシミュレートすることで，この中間を実現するようにする．とくに，適当な T の範囲で，罰金項 $g(T)$ の大きさをその温度での分配関数の対数 $\log Z_T$ と (定数を除いて) 等しくとれば，その範囲の各温度に系が存在する (周辺) 確率は一様になる．温度を離散化して，

$$g(T_{m+1}) = \langle E\rangle_{T_m}(1/T_{m+1} - 1/T_m) + g(T_m) \tag{126}$$

ととっても近似的には同じである．分配関数の値やエネルギーの期待値をあらかじめ知ることはできないが，何回も計算を繰り返しすことで罰金項を学習することは可能である．

simulated tempering で温度を変化させるステップ幅 dT には，理想的な罰金項を選んだとしても上限があり，それ以上では acceptance ratio が下がってしまう．普通は温度の大きい分布の方が，典型的な状態の総数 (エントロピー) が大きい．これに対して，温度の高い側から低い側への遷移はあくまで個々の状態の間のエネルギー差で決まる．したがって，温度についての周辺分布を一様にすると，温度の高い側から低い側への遷移がエントロピーの差の分だけ起こりにくくなることは避けられないのである．この限界は，逆温度の幅 $d\beta$ でみて，

$$d\beta \sim 1/\sigma_E \tag{127}$$

となる．これは，Metropolis-coupled chain で入れ替わりが頻繁に起こるために必要な温度の間隔と simulated tempering で可能なステップ幅が同じ程度であることを示している．

3. “multicanonical” algorithm

   さらに別の方法として，Berg らによる multicanonical algorithm がある ((Berg and Neuhaus(1992), Berg and Celik(1992a,1992b))．これも独立の発見であるが，ポッツ模型の 1 次転移や格子上のスピングラスなどに適用されており，現在のところもっとも実績がある．Berg らの方法では多数の系を同時にシミュレートするのではなく，ひとつの修正した系を考えるので，その点では simulated tempering に類似している．ただし，こちらの方法では温度を変化させる代わりに，エネルギー E を修正する．わざと揺らぎを大きくしてやることで，低温での凍結現象を防ぐという発想である．

   修正したエネルギー E_{MC} がもとのエネルギー $E(\alpha)$ だけの関数として $E_{MC} = f(E(\alpha))$ と書けるとする．すると，任意の温度 T での物理量 $A(\alpha)$ の期待値は，E_{MC} を用いたシミュレーションから得られたサンプル列 $\{\alpha_m\}$ から，

   $$\langle A \rangle_T = \frac{\sum_m A(\alpha_m) \exp(-E(\alpha_m)/T + f(E(\alpha_m))}{\sum_m \exp(-E(\alpha_m)/T + f(E(\alpha_m))} \tag{128}$$

   と求められる．これは一般には形式的な表現に過ぎず，$f(E)$ が温度 T で “十分情報がとれるように” 選ばれていないと，右辺の分散が大きくなり，意味のある答えは得られない．

   Berg らの方法では，もとのエネルギー $E(\alpha)$ の周辺分布が必要な範囲で一様になるように $f(E)$ を選ぶ．これは，何回もシミュレーションを繰り返して，最適な $f(E)$ を推定することで達成できる．Berg らはそのための近似的学習則を与えている．ポッツ模型のように一様な系では小さい系での情報を大きい系での計算に利用することもできる．状態空間全体から一様にサンプルしたのでは，低温での情報は全く得られなくなってしまい駄目である．エネルギー E の周辺分布を一様化することが本質的である．

   multicanonical algorithm の発想は，前の 2 つの方法に似ているが，本質的に違う点もある．たとえば，ポッツ模型の 1 次転移の問題 (Berg and Neuhaus(1992)) では multicanonical algorithm と前 2 者のアルゴリズムの振舞いは大きく違うと予想される．これは，multicanonical algorithm は温度でなく，それに “共役な” エネルギーの方を考えているためである．

これらの方法は，一組の外部変数での結果を同時に得たい場合に向いている．とくに，積分の計算，たとえば，自由エネルギーや分布の規格化定数の計算に有効である．

いったん，展望ができてしまうとさまざまなバリエーションが考えられる．ひとつの拡張の方向だけに触れておく．系全体を入れ替えるのではなく，一部を入れ替えるという考えである．これは，“heated Metropolis” algorithm (Lin and Thompson(1993)) 及び “unfair genetic” algorithm (高畠 (1992)) の名前で提案されている．後者では genetic algorithm における cross over が意識されている．これらの方法では入れ替える部分が大き

くなると acceptance ratio が急激に低下するのではないかと想像される．ランダムに選んだ部分を入れ替えるならば，全体をそっくり入れ替えるか，少数の変数のみを入れ替えるかのどちらかが良いのではないだろうか．前に触れたスピングラスに対する Swendsen-Wang のアルゴリズムも，系の一部を入れ替えるという文脈で理解できるが，入れ替えるクラスターを現在の状態を参照して適応的に作るという点で異なっている．

7　参考文献

(1 章)

統計物理の教科書は沢山ある．筆者は Wannier(1966) の日本語版を使っているが，客観的に良いといえるかどうかはわからない．情報処理への応用を意識して書かれた入門書としては，篠本 (1992) がある．立場としては本稿に近いが，内容はずっと易しく，応用面にも触れている．

(2 章)

統計物理でのマルコフ連鎖モンテカルロ法について，網羅的に成果を集めた本として，Binder (1986, 1987, 1995) がある．2 冊目の巻頭の総説 Binder and Stauffer(1984) は入門向きである．また，教科書として Binder and Heermann (1992) があるが，相転移の解析法などに重点がおかれているようである．

統計学者の書いたマルコフ連鎖モンテカルロ法の総合報告としては，Gilks et al. (1996), Neal (1996), Smith and Roberts(1993), Besag and Green(1993), Geyer(1992) などがある．伊庭 (1996) は本稿の姉妹編ともいえるが，よりアルゴリズム中心に書かれている．メトロポリス法の “原典” は Metropolis et al.(1953)．

simulated annealing については，Kirkpatrick et al.(1983) 及び Mezard et al.(1987), van Hemmen and Morgenstern(1987) の該当箇所を参照．また，画像再構成への応用は，Geman,S. and Geman,D.(1984), Geman,D. and Geman,S.(1986), Besag (1989), Winkler(1995)，Hopfield-Tank 法については，Hopfield and Tank(1985), Reger et al.(1984) をそれぞれ参照．

(3 章と 4 章)

イジング模型のダイナミクスについては Miyashita and Takano(1985)，ポッツ模型については Wu(1982), Kikuchi and Okabe (1992)，三角格子上の反強磁性イジング模型については Wannier(1950)，ANNNI 模型については Bak and von Boehm(1980)，離散ガウス模型・SOS 模型については Müller-Krumbhaar(1986)，平面回転子模型については Kosterlitz and Thouless(1973), Kogut(1979), Miyashita et al.(1978) をそれぞれ参照．

スピン模型に関しての教科書としては，少し古いが，小口武彦 (1970) がある．また，本報告では触れなかった，素粒子論での格子スピン模型の応用 (格子ゲージ理論) についてのやさしい解説が Kogut(1979) にある．

(5 章)

ランダム系の物理についてのまとまった教科書として，Ziman(1979) があるが，実は本文に書いたことはほとんどこれに載っていない．スピングラスを中心とした研究が 1980 年代に大きく進展したためである．これらを含んだ本や総合報告としては，Mezard et al.(1987), van Hemmen and Morgenstern(1987), Binder and Young(1986), Takayama(1989)，高山 (1991) がある．このうち Mezard et al.(1987) と高山 (1991) では SK 模型の周辺に重点が置かれている．ランダム磁場模型の平均場近似については Ro et al.(1985) を参照した．ランダム磁場模型の基底状態を求めるアルゴリズムに関しては Angles d'Auriac et al.(1985) を参照．神経回路網 (連想記憶) に関しては，Mezard et al.(1987), van Hemmen and Morgenstern(1987) の該当箇所を参照のこと．また，本報告では触れなかった神経回路網の学習の統計力学に関する総合報告が Watkin and Rau(1993) にある．

(統計的情報処理への応用)

本文では統計的情報処理への応用についてまったく述べなかった．その代わり，ここでいくつか参考文献をあげておく(既出のものもあり)．統計学におけるマルコフ連鎖モンテカルロ法の応用についてはプレプリントサーバーhttp://www.stats.bris.ac.uk/MCMC/がある．

- マルコフ場の画像再構成への応用
  Geman,S. and Geman,D.(1984)
  Geman,D. and Geman,S.(1986)
  Besag (1989)
  Winkler (1995)
  Marroquin et al.(1987)
  Chellappa and Jain (1993)
  Amit et al. (1991)
  Ogata (1990)
  伊庭 (1991a)

- 統計学一般へのマルコフ連鎖モンテカルロ法の応用
  Gilks et al. (1996)
  Smith and Roberts(1993)
  Geyer and Thompson(1992)
  Gelfand and Smith (1990a,1990b)
  伊庭 (1991b, 1996)

- モデルとしての “ボルツマンマシン”
  Hinton and Sejnowski (1986)

- パーセプトロンへのマルコフ連鎖モンテカルロ法の応用
  Neal (1996)

- ランダム系の平均場近似の応用
  (Hopfield-Tank 法のところで触れたもののほかに)
  Peterson and Hartman (1989)
  Geiger and Girosi (1991)

- 気体・液体論のモデル・手法の空間点過程への応用
  Ogata and Tanemura(1984)
  長谷川・種村 (1986)

Amit,Y.,Grenander,U. and Piccioni,M(1991)
Structural image restoration through deformable templates
Journal of the American Statistical Association 86 376-387.

Angles d'Auriac, Preissmann,M. and Rammal,R.(1985)
The random field Ising model : algorithmic complexity and phase transition
Journal de Physique Letters 46 L173-L180.

Bak,P. and von Boehm,J.(1980)
Ising Model with solitons, phasons, and "the devil's staircase"
Physical Review B 21 5297-5307.

Berg,B.A. and Neuhaus,T.(1992)
Multicanonical ensemble: a new approach to simulate first-order phase transitions
Physical Review Letters 68 9-12.

Berg,B.A. and Celik,T.(1992a)
New approach to spin-glass simulations
Physical Review Letters 69 2292-2295.

Berg,B.A. and Celik,T.(1992b)
Multicanonical spin glass simulations
Internatinal journal of modern physics C 3 1251-1274.

Besag,J.(1989)
Towards Bayesian image analysis
Journal of Applied Statistics 16 395-407.

Besag,J and Green,P.J.(1993)
Spatial statistics and Bayesian computation
Journal of Royal Statistical Society B 55 25-37.

Binder,K.(ed.)(1986)
Monte Carlo Methods in Statistical Physics (2nd ed.)
Topics in current physics vol.7, Springer-Verlag, Berlin.

Binder,K.(ed.)(1987)
Applications of the Monte Carlo Method in Statistical Physics (2nd ed.)
Topics in Current Physics vol.36, Springer-Verlag, Berlin.
(1st ed.(1984), 2nd ed. は未入手).

Binder,K.(ed.)(1995)
The Monte Carlo Method in Condensed Matter Physics (2nd ed.)
Topics in Applied Physics vol.71, Springer-Verlag, Berlin.

Binder,K. and Stauffer,D. (1984)
A simple introduction to Monte Carlo simulation and some specialized topics
in Applications of the Monte Carlo Method in statistical physics
Topics in current physics vol.36
Ed. K. Binder, Springer-Verlag, Berlin
(2nd ed.(1987) にも載っていると思うが未確認).

Binder,K. and Young,A.P. (1986)
Spin glasses: experimental facts, theoretical concepts, and open questions
Reviews of Modern Physics 58 801-976.

Binder,K. and Heermann,D.W. (1992)
Monte Carlo simulation in statistical physics An introduction (2nd ed.)
Springer series in solid-state sciences 80
Springer-Verlag, Berlin.

Bouzida,D., Kumar,S and Swendsen,R.H.(1992)
Efficient Momte Carlo methods for the computor simulation of biological molecules
Physical Review A 45 8894-8901.

Chellappa,R. and Jain,A. (Eds.)
Markov Random Fields: Theory and Application
Academic Press San Diego 1993.

Geiger,D. and Girosi,F.(1991)
Parallel and deterministic algorithms for MRF's : surface reconstruction and integration
IEEE Transactions on Pattern Analysis and Machine Intelligence 13 401-412.

Gelfand,A.E. and Smith,A.F.M.(1990a)
Sampling-based approaches to calculating marginal densities
Journal of the American Statistical Association 85 398-409.

Gelfand,A.E., Hills,S.E., Racine-poon,A. and Smith,A.F.M.(1990b)
Illustration of Bayesian inference in normal data models using Gibbs sampling
Journal of the American Statistical Association 85 972-985.

Geman,S. and Geman,D.(1984)
Stochastic relaxation, Gibbs distributions,
and the Bayesian restoration of images
IEEE Transactions on Pattern Analysis and Machine Intelligence 6 721-741.

Geman,D. and Geman,S.(1986)
Bayesian image analysis
NATO ASI series F20
Disordered Systems and Biological Organization
eds. Bienenstock,E et al.
Springer-Verlag.

Geyer,C.J.(1991)
Markov chain Monte Carlo maximum likelihood
In Computing Science and Statistics:
Proc. of the 23rd symposium on the interface(E.M.Keramides,ed.).

Geyer,C.J.(1992)
Practical Markov chain Monte Carlo
Statistical Science 7 473-511.

Geyer,C.J. and Thompson,E.A.(1992)
Constrained Monte Carlo maximum likelihood for dependent data
Journal of Royal Statistical Society B 54 657-699.

Geyer,C.J. and Thompson,E.A. (1995)
Annealing Markov chain Monte Carlo with applications to ancestral inference
Journal of the American Statistical Association 90 909-920.

Gilks,W., Spiegelhalter,D. and Richardson,S.(eds)(1996)
Markov Chain Monte Carlo in practice
Chapman and Hall, London New York.

Goodman,J. and Sokal,A.D.(1989)
Multigrid Monte Carlo method. Conceptual foundations
Physical Review D 40 2035-2071.

Gray,A.J.(1993)
In discussion on the meeting on the Gibbs sampler and other Markov chain Monte Carlo Methods
Journal of Royal Statistical Society B 55 58-61.

長谷川政美, 種村正美 (1986)
なわばりの生態学 生態のモデルと空間パターンの統計
東海大学出版会.

Higdon,D.(1993)
In discussion on the meeting on the Gibbs sampler and other Markov chain Monte Carlo Methods
Journal of Royal Statistical Society B 55 78-78.

Hinton,G.E. and Sejnowski,T.J.(1986)
Learning and relearning in Boltzmann machines
in Parallel Distributed Processing Vol.1
Eds. Rumelhart,E. and McClelland,J.L.
MIT press Cambridge.

Hopfield,J.J. and Tank,D.W.(1985)
"Neural" computation of decisions in optimization problems,
Biological Cybernetics 52 141-152.

Hukushima,K and Nemoto,K. (1996)
Exchange Monte Carlo method and application to spin glass simulations
Journal of Physical Society of Japan, 65, 1604-1608.

伊庭幸人 (1991a)
メトロポリス的モンテカルロ法の巨視的パラメータ推定への応用
– 2 次元イジング模型の場合 –
統計数理 39 No.1 1-21.

伊庭幸人 (1991b)
メトロポリス的モンテカルロ法の擬ベイズ法への応用
– 変化点問題を例として –
統計数理 39 No.2 225-244.

伊庭幸人 (1993a)
ベイズ統計と統計物理 - 有限温度での情報処理 -
物性研究 60-6 (1993 年 9 月号) 677-699.

伊庭幸人 (1993b)
統計数理研究所 年度末発表会 (1992 年度) 要旨
メトロポリス的モンテカルロ法の緩和について
統計数理 41 No.1 65-67.

伊庭幸人 (1996)
マルコフ連鎖モンテカルロ法とその統計学への応用
統計数理 44 No.1 49-84.

Kandel,D., Domany,E. and Brandt,A. (1989)
Simulation without critical slowing down: Ising and three-state Potts models
Physical Review B 40 330-344.

Kikuchi,M. and Okabe,Y.(1992)
Order-Parameter distribution function and order of the phase transition of the ferromagnetic Potts model
Journal of Physical Society of Japan 61 3503-3510.

Kikuchi,M, Ito,N. and Okabe,Y. (1994)
Statistical dependence and related topics
in Computer Simulation in Condensed Matter Physics VII
Eds. D.P.Landau, K.K.Mon, and H.B.Schüttler, Springer-Verlag, Berlin.

木村宏一, 瀧和男 (1990)
時間的一様な並列アニーリングアルゴリズム
電子情報通信学会 NC-90-1 1-8.

Kirkpatrick,S, Gellatt,C.D. and Vecchi,M.P.(1983)
Optimization by simulated annealing
Science 220 671-680.

Kogut,J.B.(1979)
An Introduction to lattice gauge theory and spin systems
Reviews of Modern Physics 51 659-713.

小西健三、瀧和男、木村宏一 (1995)
温度並列シミュレーテッド・アニーリング法とその評価
情報処理学会論文誌 Vol.36 No.4 797-807.

Kosterlitz,J.M. and Thouless,D.J.(1973)
Ordering, metastability and phase transition in two-dimensional systems
Journal of Physics C 6 1181-1203.

Kronfeld,A.(1993)
Dynamics of Lengevin simulations
Progress of Theoretical Physics Supplement No.111 293-311.

Liang,S. (1992)
Application of cluster algorithm to spin glasses
Physical Review Letters 69 2145-2148.

Lin,S. and Thompson,E.A.(1993)
In discussion on the meeting on the Gibbs sampler and other Markov chain Monte Carlo Methods
Journal of Royal Statistical Society B 55 81-82.

Marinari,E. and Parisi,G.(1992)
Simulated Tempering: a New Monte Carlo Scheme
Europhysics Letters 19 451-458.

Marroquin,J., Mitter,S. and Poggio,T. (1987)
Probabilistic solution of ill-posed problems in computational vision
Journal of the American Statistical Association Vol.82 76-89.

Metropolis.N, Rosenbluth,A.W., Rosenbluth,M.N.,
Teller,A.H. and Teller,E (1953)
Equation of state calculation by fast computing machines
Journal of the chemical physics 21 1087-1092.

Mezard,M., Parisi,G and Virasoro,M.A.(1987)
Spin glass theory and beyond
World Scientific Singapore.

Miyashita,S., Nishimori,H., Kuroda,A. and Suzuki,M.(1978)
Monte Carlo simulation and static and dynamic critical behavior of the plane rotator model
Progress of Theoretical Physics 60 1669-1683.

Miyashita,S. and Takano,H.(1985)
Dynamical nature of the phase transition of the two-dimensional kinetic Ising model
Progress of Theoretical Physics 73 1122-1140.

Müller-Krumbhaar,H.(1986)
Momte Carlo simulation of crystal growth
In Binder,K.(ed.) Monte Carlo Methods in statistical physics (2nd ed.)
Topics in current physics vol.7
Springer-Verlag Berlin.

Neal,R.M. (1996)
Bayesian Learning for Neural Networks
Lecture Notes in Statistics 118
Springer.

Ogata,Y.(1990)
A Monte Carlo method for objective Bayesian procedure
Annals of the Institute of Statistical Mathematics 42 403-433.

Ogata,Y. and Tanemura,M.(1984)
Likelihood analysis of spatial point patterns
The Journal of the Royal Statistical Society B, Vol.46 496-518.

小口武彦 (1970)
磁性体の統計理論 (**物理学選書** 12)
掌華房.

Peterson,C. and Hartman,E. (1989)
Explorations of the mean field thoery learning algorithm
Neural Networks 2 475-494.

Reger,J.D.,Binder,K. and Kinzel,W (1984)
Investigation of the validity of the "slow cooling" iterative mean-field method for the study of ground-state properties of spin-glasses
Physical Review B 30 4028-4030.

Ro,C., Grest,G.S., Soukoulis C.M. and Levin,K.(1985)
Irreversibility in random-field ferromagnets and diluted antiferromagnets
Physical Review B 31 1682-1685.

篠本滋 (1992)
情報の統計力学
丸善.

Smith, A.F.M. and Roberts,G.O.(1993)
Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods
Journal of Royal Statistical Society B 55 3-23.

Swendsen,R.H. and Wang,J.S.(1986)
Replica Monte Carlo simulation of spin-glasses
Physical Review Letters 57 2607-2609.

Swendsen,R.H. and Wang,J.S.(1987)
Nonuniversal critical dynamics in Monte Carlo simulation
Physical Review Letters 58 86-88.

Swendsen,R.H. and Wang,J.S.(1988)
Low-Temperature properties of the $\pm J$ Ising spin glass in two dimension
Physical Review B 38 4840-4844.

高畠一哉 (1992)
不公平な遺伝的アルゴリズム
1992 年 **電子情報通信学会秋期大会** A-172.

Takayama,H (ed.) (1989)
Cooperative dynamics in complex physical systems
Springer-Verlag Berlin.

高山一 (1991)
スピングラス
丸善.

van Hemmen,J.L. and Morgenstern,I. (eds.) (1987)
Heidelberg colloquium on glassy dynamics
Lecture note in physics 275, Springer-Verlag Berlin.

Wang,J.S. and Swendsen,R.H.(1990)
Cluster Monte Carlo algorithms
Physica A 167 565-579.

Wannier,G.H.(1966)
Statistical physics
(**邦訳 統計物理学** I ,II **紀伊国屋書店** (1974,1975)).

Wannier,G.H.(1950)
Antiferromagnetism. The triangular Ising net
Physical Review 79 357-364.

Watkin,T.L.H. and Rau,A.(1993)
The statistical mechanics of learning a rule
Revew of Modern Physics 65 498-556.

Winkler,G.(1995)
Image analysis, Random fields and Dynamic Monte Carlo Methods
A Mathematical Introduction
Springer-Verlag, Berlin.

Wolff,U.(1989)
Collective Monte Carlo updating for spin systems
Physical Review Letters 62 361-364.

Wu,F.Y.(1982)
The Potts model
Review of Modern Physics 54 235-268.

Ziman,J.M.(1979)
Models of Disorder
Cambridge University Press Cambridge.
(邦訳 乱れの物理学 丸善 (1982)).