Proc. Camb. Phil. Soc. (1964), **60**, 115
Printed in Great Britain

A Monte Carlo method applied to the Heisenberg ferromagnet

BY D. C. HANDSCOMB
Oxford University Computing Laboratory, 9 South Parks Road, Oxford

(*Received* 6 *May* 1963)

Abstract. Following on from a previous paper (5), we apply the new Monte Carlo method described there to the estimation of order parameters of a simple Heisenberg ferromagnet. By way of illustration, we include some results on the simple cubic lattice, comparing them with results obtained by conventional methods.

1. *Introduction.* In another paper (5), we have introduced a new Monte Carlo process, extending from classical to quantum-mechanical systems the principles underlying the powerful method first described by Metropolis and his collaborators (9). Here we discuss its application to a specific problem in statistical mechanics, the Heisenberg ferromagnet.

Fosdick and others (3), (4) have used Metropolis's method to estimate the order parameters of a two- or three-dimensional Ising lattice (8). When one is studying the phenomenon of ferromagnetism, especially the transition from the ordered to the disordered phase at the Curie temperature, one naturally would like to use the same method on the Heisenberg lattice (7) to which Ising's model is an approximation; much work has gone into the computation of series expansions for these parameters (2), (10), (11), and for checking purposes it is desirable to have independent estimates of them. Since Heisenberg's Hamiltonian is not scalar, however, one cannot use Metropolis's method as it stands.

In this paper we give all the necessary details of our process, and at the end we quote some results for the simple cubic lattice. They agree fairly closely with the series, except for the long-range order at low temperatures. The method will, however, need to be improved considerably before it can give useful results on its own.

2. *Modifications of the original method.* Before we apply our method to the Heisenberg model, we must make it slightly more general than it was in (5). When we express the Hamiltonian as

$$H = H_0 + \sum_{i=1}^{N} H_i, \tag{1}$$

we shall allow H_0 to be any operator that commutes with each H_i, instead of insisting that it be constant. We then replace the product $H_{i_1}\ldots H_{i_r}$ by $H_{i_1}\ldots H_{i_r}\exp(-\beta H_0)$, each time that it appears in (5), and we must replace H_0 by its expectation $\langle H_0\rangle$ in equations (10) and (11) of (5).

We mention in passing a second way in which the method may be generalized. We previously defined the probabilities f_r, for arbitrary λ, to be

$$f_r = e^{-1/\lambda}\lambda^r \int_0^{1/\lambda} e^t t^r\, dt, \tag{2}$$

satisfying the recurrence relation

$$f_r = 1 - \lambda r f_{r-1} \quad (r > 0). \tag{3}$$

We notice that (3) is numerically unstable in the direction of increasing r when $\lambda r > 1$, and in the reverse direction when $\lambda r < 1$. In order to use it to generate the f_r, therefore, we must first calculate the values at *two* starting points, namely

$$f_0 = 1 - e^{-1/\lambda} \tag{4}$$

and

$$f_r = \sum_{n=1}^{\infty} \frac{(-1)^{n+1} r!}{(r+n)!\,\lambda^n}, \tag{5}$$

for some large value of r. This is a slight inconvenience. Suppose, however, that we take for f_r any sequence of probabilities whatever ($0 < f_r < 1$), define

$$\lambda_r = \frac{1 - f_r}{r f_{r-1}}, \tag{6}$$

and replace equation (13) of (5) by

$$\tau(C_r) = \frac{(-\beta)^r \left(\prod_{s=1}^{r} \lambda_s \right) \operatorname{trace} \{H_{i_1} \ldots H_{i_r} \exp(-\beta H_0)\}}{p(i_1) \ldots p(i_r)}. \tag{7}$$

The method will then go through as before. In particular, if we take

$$f_r = \frac{1}{1 + \lambda(r+1)}, \quad \lambda_r = \lambda(1 + f_r/r), \tag{8}$$

these probabilities f_r do not differ greatly from those defined by equation (2), and are much easier to compute.

In this paper, however, we shall continue to define f_r by equation (2).

3. *The simple Heisenberg model.* We are not yet able to deal with the full generality of the Heisenberg model, notably with antiferromagnetic interactions, and the nature of the method precludes us from working on an infinite lattice.

The model that we shall refer to consists of a section of some regular lattice, of n sites each occupied by a single particle of spin $\frac{1}{2}$. In order to approximate more closely to the infinite lattice we impose periodic boundary conditions whose effect is as if opposite sides of the section were in contact. Denoting the spin-operator for the tth site by $\mathbf{s}_t$ and its z-component by s_t^z, we assume the Hamiltonian H of the system to be of the form

$$H = -2\mu\mathscr{H} \sum_t s_t^z - \sum_{t>t'} \sum 2J_{tt'}(\mathbf{s}_t, \mathbf{s}_{t'}), \tag{9}$$

together with a constant term which we shall always omit. Here $\mathscr{H}$ represents the external magnetic field (acting in the z-direction), μ is the Bohr magneton, $J_{tt'}$ represents the exchange interaction between the particles on sites t and t', and $(,)$ denotes an inner product. We require that $J_{tt'} \geqslant 0$. The interaction need not be the same between every pair of sites, and need not be confined to nearest neighbours.

Our object will be to estimate the equilibrium values of the short-range order σ, long-range order S^2, polarization P, and magnetic susceptibility χ per particle of this

model, as functions of temperature T and magnetic field $\mathscr{H}$. We may conveniently represent the first three by the expectations of the operators

$$O_S = -(H - H_0)/NJ, \tag{10}$$

where H_0, N and J are defined in equations (16) and (19),

$$O_L = 4\left(\sum_t \mathbf{s}_t, \sum_t \mathbf{s}_t\right) \Big/ n(n+2), \tag{11}$$

and

$$O_P = \sum_t s_t^z/n, \tag{12}$$

while, denoting the equilibrium value of O by $\langle O \rangle$ as usual, the dimensionless susceptibility χ/K is derived from the relation

$$\langle O\chi \rangle = \frac{\partial}{\partial L}\langle O_P \rangle, \tag{13}$$

with L defined by equation (21).

4. *Transformation of the model.* Writing $E(t, t')$ for the operator that interchanges the spins on sites t and t', we know ((1), §58) that

$$E(t, t') = \tfrac{1}{2}\{1 + 4(\mathbf{s}_t, \mathbf{s}_{t'})\}, \tag{14}$$

so that equation (9) may be rewritten as

$$H = -2\mu\mathscr{H}\sum_t s_t^z - \sum_{t>t'}\sum J_{tt'} E(t, t'). \tag{15}$$

Equation (15) is of the form of equation (1), with

$$\left.\begin{aligned} H_0 &= -2\mu\mathscr{H}\sum_t s_t^z, \\ H_i &= -J_{t_i t_i'} E(t_i, t_i') \end{aligned}\right\} \quad (i = 1, 2, \ldots, N). \tag{16}$$

The number of components, N, is the number of pairs of sites tt' for which $J_{tt'}$ does not vanish, and we let $t_i t_i'$ be the ith in some ordering of such pairs.

If

$$P(C_r) = E(t_{i_1}, t_{i_1}') \ldots E(t_{i_r}, t_{i_r}'), \tag{17}$$

where C_r is any sequence $i_1, \ldots, i_r$ of indices selected from the range $1 \leqslant i \leqslant N$, then $P(C_r)$ is a permutation of the sites of the lattice. Let $P(C_r)$ be represented as the product of $k = k(C_r)$ cycles, of lengths $a_1, \ldots, a_k$. We may then show that

$$\begin{aligned} &\operatorname{trace}\{H_{i_1} \ldots H_{i_r} \exp(-\beta H_0)\} \\ &\quad = \prod_{i=i_1}^{i_r} (-J_{t_i t_i'}) \operatorname{trace}\left\{P(C_r) \exp\left(-2\beta\mu\mathscr{H}\sum_t s_t^z\right)\right\} \\ &\quad = \prod_{i=i_1}^{i_r} (-J_{t_i t_i'}) \prod_{j=1}^{k(C_r)} \{2\cosh(a_j \beta\mu\mathscr{H})\}. \end{aligned} \tag{18}$$

(The symbol β stands, as usual, for $(\mathbf{k}T)^{-1}$ where $\mathbf{k}$ is Boltzmann's constant.)
In order to satisfy equation (12) of (5), take

$$p(i) = J_{t_i t_i'} \Big/ \sum_{j=1}^{N} J_{t_j t_j'} = J_{t_i t_i}/NJ, \quad \text{say.} \tag{19}$$

We then have
$$\tau(C_r) = (NK\lambda)^r \prod_{j=1}^{k(C_r)} \{2\cosh(a_j L)\}, \tag{20}$$

where
$$K = \beta J \quad \text{and} \quad L = \beta\mu\mathscr{H}. \tag{21}$$

The quantities f_r, defined by equations (2) or (5), $p(i)$, defined by equation (19), and $\tau(C_r)$, defined by equation (20), are all that we require in order to set up the Markov chain described in §3 of (5); we shall show, in §5 below, what this implies in practice. We defer until then proof that this chain has a limit distribution.

The operator O_L is transformed by the modified form of equation (5) of (5),
$$\Omega(C_r) = \frac{\text{trace}\{OH_{i_1}\ldots H_{i_r}\exp(-\beta H_0)\}}{\text{trace}\{H_{i_1}\ldots H_{i_r}\exp(-\beta H_0)\}}, \tag{22}$$
into the observation
$$\Omega_L = \left[\sum_i a_i^2 + 2\coth L\sum_i (a_i\tanh a_i L) + \left\{\sum_i (a_i\tanh a_i L)\right\}^2 - \sum_i (a_i^2\tanh^2 a_i L)\right] \Big/ n(n+2), \tag{23}$$
which, in the absence of a magnetic field, reduces to
$$\Omega_L(L=0) = 3\sum_i a_i^2/n(n+2), \tag{24}$$
and O_P becomes similarly
$$\Omega_P = \sum_i (a_i\tanh a_i L)/n. \tag{25}$$

The dimensionless susceptibility in zero fields may thus be derived from (13); the corresponding observation is
$$\Omega_\chi = \frac{\partial}{\partial L}\Omega_P(L=0) = \sum_i a_i^2/n. \tag{26}$$
By virtue of equation (11) of (5),
$$\mathscr{E}(r) = \beta\langle H_0 - H\rangle, \tag{27}$$
we may obtain the short-range order parameter from the observation
$$\Omega_S^* = r/NK, \tag{28}$$
which is simpler than the Ω_S of equation (22) with $O = O_S$.

As we explained in (5), the result of these transformations is that we can set up a Markov chain whose limit distribution is such that the expectation of each observation Ω is equal to the corresponding parameter $\langle O\rangle$ of the model. We use the fact that the states passed through by a realization of the Markov chain are eventually almost certain to be distributed according to this limit distribution, and take the average value of each Ω over the C_r's that occur.

5. *Realization of the Markov chain.* The sample space consists of all sequences $C_r = i_1, \ldots, i_r$ with $r \geqslant 0$ and $1 \leqslant i \leqslant N$. The transition probabilities depend on f_r, $p(i)$, and the ratios
$$\rho(C_r, i) = \tau(C_r i)/\tau(C_r). \tag{29}$$
It is convenient to obtain ρ directly, as follows.

When the permutation $P(C_r)$ is expressed as a product of cycles, the sites t_i, t_i' belong either to the same cycle or to distinct cycles. In the former case, let them divide their cycle into two segments of lengths a and a' (so that $a+a'$ is the length of the original cycle). Then

$$\rho(C_r, i) = 2NK\lambda/(1+\tanh aL \tanh a'L). \tag{30}$$

In the latter case, if the lengths of the two cycles concerned are b and b', we have

$$\rho(C_r, i) = \tfrac{1}{2}NK\lambda(1+\tanh bL \tanh b'L). \tag{31}$$

The Markov chain is now realized as follows.

At all times we keep a record of the current sequence C_r and of the induced permutation $P(C_r)$. At each step we generate two independent random numbers, ξ and ξ', uniformly distributed in the interval (0, 1). If $\xi < f_r$, we select the 'forward' direction; otherwise, the 'backward' direction.

If the forward direction is selected, we next generate a random index i according to the probability distribution $p(i)$, and compute $\rho(C_r, i)$. If $\rho(C_r, i) < \xi'$, the next sequence is a repetition of C_r; otherwise it is the extended sequence $C_r i$ and r is increased by 1.

If the backward direction is selected, and $r \neq 0$, let $C_r \equiv i' C_{r-1}$. We then compute $\rho(C_{r-1}, i')$. If $\xi'\rho(C_{r-1}, i') > 1$, the next sequence is the permuted sequence $C_{r-1}i'$; otherwise it is the truncated sequence C_{r-1} and r is decreased by 1. If $r = 0$, the current sequence C_0 is repeated.

It is easily shown that this process corresponds exactly to that described in (5) and that the limit distribution exists, since

$$\tau(C_r i)\, p(i)/\tau(C_r) = \rho(C_r, i)\, p(i) < 2KN\lambda \tag{32}$$

for all C_r and i, so that condition (14) of (5) is satisfied.

We have still to choose a value for λ. Now when $\lambda r = 1$ the value of f_r is approximately $\frac{1}{2}$, so that the forward and backward directions are selected equally often. If this is true near to the equilibrium mean value of r, then we may expect (although we cannot prove it) that the number of steps at which the sequence C_r is unaltered or merely permuted is smaller than would otherwise be the case, and that therefore the speed at which the process explores the sample space is the greater. For this kind of reason we have taken

$$\lambda^{-1} = KN\sigma^*, \tag{33}$$

where σ^* is a prior guess at the value of $\sigma = \langle O_S \rangle$. We have also arranged for the Markov chain to start from an initial sequence of approximately this length.

6. *Practical details and some results.* We have written a program to carry out this process on the Ferranti Mercury computer, restricting our attention for the time being to the case $L = 0$. The size of the fast-access store of this machine effectively limits us to lattices of not more than 1024 sites. Most of the running time of the program is taken up by the process of deciding, at each step, whether or not two sites belong to the same cycle of the current permutation. Since the average length of the cycle containing a given object in a random permutation on n objects is $\frac{1}{2}(n+1)$, this decision

takes an average time of the order of $150(n+1)$ microseconds per step, so that on a lattice of maximum size we cannot exceed a speed of about 6 steps a second. This is not affected significantly by the shape or dimensionality of the lattice, but we do improve on it for small values of K, when cycles tend to be short.

The values of the quantities Σa_i^2 and r (as also $\Sigma(a_i \tanh a_i L)$ and $\Sigma(a_i \tanh a_i L)^2$, when $L \neq 0$) are kept up to date simply by calculating their increments at every step, so that we have Ω_L, Ω_P, Ω_χ, and Ω_S^* readily available. We take our estimates from the running averages of these observations.

In order to show that these running-average estimators converge with probability one as the experiment continues, consider the following simplification of the process. Let a Markov chain on the set of non-negative integers have transition probabilities

$$\left.\begin{aligned} &p(r \to r+1) = f_r, \quad p(r+1 \to r) = 1-f_{r+1} \quad (r \geqslant 0) \\ &p(0 \to 0) = 1-f_0 \end{aligned}\right\}, \tag{34}$$

where $\{f_r\}$ is a monotone decreasing sequence of probabilities whose limit f is less than $\frac{1}{2}$. By comparing this with a similar chain in which f_r is replaced by p for all r, where $\frac{1}{2} > p > f$, we may show that the mean square recurrence time of any state, and the mean square sum of the values of r occurring between successive recurrences of any state, are both finite. It follows ((12), pp. 96 onwards) that the running average of r, as well as that of any bounded function, has the Central Limit property. We hope that this holds also for the more complicated actual process; assuming this to be so, we can say that the standard deviation of each of our estimators will ultimately decrease as the inverse square root of the number of steps taken. Unfortunately there is no way we can predict at what point this convergence will take effect.

Table 1. *Monte Carlo results for the simple cubic lattice*

K	n	χ/K	S^2	$K\sigma$
0·2	6	2·04 ± 0·07	0·012 ± 0·0004	0·131 ± 0·003
0·3	9	2·97 ± 0·09	0·017 ± 0·0005	0·199 ± 0·005
0·4	13	6·32 ± 0·25	0·037 ± 0·001	0·294 ± 0·003
0·5	13	18·1 ± 2	0·11 ± 0·01	0·403 ± 0·003
0·6	12	65·2 ± 3	0·38 ± 0·02	0·523 ± 0·003
0·7	9	100·7 ± 4	0·59 ± 0·03	0·631 ± 0·004
0·8	9	125 ± 3	0·73 ± 0·02	0·73 ± 0·009

The results in Table 1 were obtained on a simple cubic lattice of 512 sites, with nearest-neighbour interactions only. The progress of each experiment was divided into n runs of 1000 steps, where n is the number shown in the second column, and the various averages were computed for each run separately. The means and standard deviations quoted in the table are estimated from the sample distribution of these n averages. The last three columns come from the respective values of Ω_χ, Ω_L, and Ω_S^*. For comparison, Table 2 gives corresponding values (for the infinite lattice) computed from the series of Rushbrooke and Wood (10), (11), and of Dyson (2), respectively, on either side of the transition temperature. The series used are shown in Table 3.

Since the Monte Carlo computations involve a finite lattice, they will not show a phase-transition, however large a sample we take, so that the small values in the fourth column of Table 1 for $K < 0{\cdot}5$ and the large values in the third column for $K > 0{\cdot}5$ are meaningless but unavoidable. (Rushbrooke estimates the position of the phase-transition as at $K \cong 0{\cdot}56$.) There appears to be some disagreement between our estimates and Dyson's estimates of the long-range order, for which we offer no explanation, but in other respects we agree reasonably well. We cannot yet claim that our results are very useful; some reader may, however, be encouraged, by the demonstration that a Monte Carlo attack on this problem is possible, to develop more powerful methods on similar lines.

Table 2. *Series results for the same lattice*

High-temperature series			Low-temperature series		
K	χ/K	$K\sigma$	K	S^2	$K\sigma$
0·2	1·984	0·1265	0·5	0·2464	0·3907
0·3	3·049	0·2057	0·6	0·4269	0·5246
0·4	6·017	0·2926	0·7	0·5511	0·6441
0·5	16·21	0·3834	0·8	0·6375	0·7564

Table 3

High-temperature series

$$K/\chi = 1 - 3K + 3K^2 - 2K^3 + 3K^4/8 + 29K^5/40 + 1919K^6/480 + \ldots$$
$$\sigma = \tfrac{1}{2} + 3K/4 - 3K^2/8 - 9K^3/16 + 35K^4/32 + 461K^5/320 - 2359K^6/640 + \ldots$$

Low-temperature series

$$S = 1 - 5{\cdot}224\theta^{\frac{3}{2}} - 6{\cdot}30\theta^{\frac{5}{2}} - 23{\cdot}0\theta^{\frac{7}{2}} - 111{\cdot}2\theta^4 - \ldots$$
$$\sigma = 1 - 13{\cdot}375\theta^{\frac{3}{2}} - 26{\cdot}417\theta^{\frac{7}{2}} - 107{\cdot}30\theta^{\frac{9}{2}} - 132{\cdot}24\theta^5 - \ldots$$
$$(\theta = 0{\cdot}079575/K)$$

The work in this paper (and in (5)) is taken largely from the thesis submitted by the author to the University of Oxford (6) under the terms of a studentship from the United Kingdom Atomic Energy Authority. He would like to express his thanks to the Authority and to the University of Illinois for their financial support and for permission to use their respective computers; to his supervisor, Dr J. M. Hammersley, and to Drs L. D. Fosdick, W. A. Marshall, R. J. Elliott, and others for supplying him with ideas.

REFERENCES

(1) Dirac, P. A. M. *The principles of quantum mechanics* (3rd edition, Oxford, 1947).
(2) Dyson, F. J. Thermodynamic behavior of an ideal ferromagnet. *Phys. Rev.* **102** (1956), 1230–1244.
(3) Ehrman, J. R., Fosdick, L. D. and Handscomb, D. C. Computation of order parameters in an Ising lattice by the Monte Carlo method. *J. Math. Phys.* **1** (1960), 547–558.
(4) Fosdick, L. D. Calculation of order parameters in a binary alloy by the Monte Carlo method. *Phys. Rev.* **116** (1959), 565–573.
(5) Handscomb, D. C. The Monte Carlo method in quantum statistical mechanics. *Proc. Cambridge Philos. Soc.* **58** (1962), 594–598.

(6) HANDSCOMB, D. C. *Monte Carlo methods applicable to the Heisenberg ferromagnet* (Thesis, University of Oxford, 1962).
(7) HEISENBERG, W. Theorie des Ferromagnetismus. *Z. Physik*, **49** (1928), 619–636.
(8) ISING, E. Beitrag zur Theorie des Ferromagnetismus. *Z. Physik*, **3** (1925), 253–258.
(9) METROPOLIS, N., ROSENBLUTH, A. W. and M. N., TELLER, A. H. and E. Equations of state calculations by fast computing machines. *J. Chem. Phys.* **21** (1953), 1087–1092.
(10) RUSHBROOKE, G. S. and WOOD, P. J. On the high-temperature susceptibility for the Heisenberg model of a ferromagnet. *Proc. Phys. Soc.* **68** (1955), 1161–1169.
(11) RUSHBROOKE, G. S. and WOOD, P. J. On the Curie points and high-temperature susceptibilities of Heisenberg model ferromagnets. *Molecular Phys.* **1** (1958), 257–283.
(12) CHUNG, K. L. *Markov chains with stationary transition probabilities* (Springer; Berlin, 1960).