

## On the Functional Form of the Transition Probabilities in the Monte Carlo Method as used in Statistical Mechanics

C. H. J. Johnson

Division of Chemical Physics, CSIRO, P.O. Box 160, Clayton, Vic. 3168.

### Abstract

The transition probabilities  $W$  in the traditional Monte Carlo simulation process used in statistical mechanics are shown to satisfy a linear functional equation. General classes of solution to this equation are presented. A simple one-particle mean-field Ising model of a ferromagnet is used in an analytical comparison of the various possible forms of  $W$ .

### Introduction

The computer simulation of the behaviour of a statistical mechanical system is essentially one of simulated sampling over the set of configurations of a finite cluster or sample of particles taken in the infinite system. For Ising systems with spin  $\frac{1}{2}$  particles set at the node points of a lattice, this simulation is often taken to be a realization of the basic stochastic process describing the evolution of the system as it passes from one equilibrium state to another. As implemented in computer terms, this process is discrete in time and is defined over the discrete state space of all configurations of the finite sample of spin particles. The physical origin of the stochastic process may be considered to lie in the interaction of the spin system with a reservoir system, each system being able to induce state transitions in the other. Each system is thus open and, if we assume the reservoir to be large and constantly in thermodynamic equilibrium at some temperature  $T$ , we may proceed from the Liouville-von Neumann equation and set up for the spin particle system a master equation satisfied by the appropriate density operator  $\rho_s(t)$ . This master equation necessarily involves memory effects, expressed as integration over the past history, so that the associated stochastic process is generally nonMarkovian. However, if the spin system is sufficiently close to equilibrium at the reservoir temperature, the memory effects may be approximated in such a way that the process becomes Markovian. Allowing the spin state transitions to be single-spin reversals, and interpreting the diagonal elements of  $\rho_s(t)$  as the probabilities  $\rho(s; t)$  of the various states (i.e. configurations of the spin particles), the master equation may be written as the linear system (Argyres and Kelly 1964):

$$d\rho(s; t)/dt = - \sum_j W_j(s) \rho(s; t) + \sum_j W_j(s^j) \rho(s^j; t). \quad (1)$$

Here  $s = \{s_k\}$  is the vector of spin states of the particles of the sample; the vector  $s$  thus represents a configuration of the sample particles. The notation  $s^j$  means that the  $j$ th spin in  $s$  has been reversed, the other spins being unchanged. The transition

probability  $W_j(s)$  is the probability rate that the  $j$ th spin in  $s$  will reverse, changing  $s$  into  $s^j$ , while  $W_j(s^j)$  refers to the inverse transition.

In the computer simulation of the Markov process associated with equation (1), the transition events occur at equal intervals of 'computer time', with the value of the probability of transition determining whether or not the transition is to be 'accepted'. At each transition event, the sample particle correlations are computed, so that the simulation in effect solves equation (1) in a moment sense. In general, for practical reasons, only a few of the low-order moments are computed, the maximum possible order being that for which the correlation length is equal to the sample diameter.

As generally defined, the transition probabilities  $W_j(s)$  depend only on the instantaneous configuration of the spin particles and are not explicit functions of the time, i.e. the Markov process is time homogeneous and, considered in terms of the master equation (1), is a process with discrete states in continuous time. It follows from this that the underlying random variables which generate the process—in the present case those random variables that, through particle–reservoir coupling, generate the transition events—must be exponentially distributed (see e.g. Chapter 6 of Cox and Miller 1977). It is, of course, this underlying distribution function that sets the time scale for the transition process as expressed through the  $W_j(s)$ . If we are concerned only with the estimation of equilibrium properties, it probably does not matter how the equilibrium sequence of configurations is reached, so long as the Markov process in the equilibrium state is time homogeneous. However, if we are concerned with the approach to equilibrium and wish to estimate relaxation times for the moments, the underlying random variables may not be exponentially distributed and so a quite general stochastic process may be needed to describe the approach to equilibrium.

The detailed form of the  $W_j(s)$  depends on the nature of the coupling between the spin system and the reservoir, and may be determined in specific cases, such as spin–phonon coupling (Bolton *et al.* 1977). However, in the absence of any particular model for the coupling, the most that can be said about the form of the  $W_j(s)$  is that at equilibrium, when the Markov process has converged in time and where only fluctuations about the equilibrium state occur, the  $W_j(s)$  should satisfy the condition of detailed balance for the particular type of transition involved (i.e. single-spin reversals). In mathematical terms this condition becomes a linear functional equation for  $W_j(s)$  which simply depends on the energy  $\Delta E$  of the transition alone, and so does not depend explicitly on the configuration immediately prior to the transition (cf. the definition of a Markov process (Feller 1957)). Detailed balance now becomes a linear functional equation involving the transition probabilities for an event and its inverse, and with argument  $\Delta E/k_B T$ , where  $k_B$  is Boltzmann's constant. In the present paper we consider solutions to this equation and indicate how general classes of transition probabilities may be constructed. Since the transition process is the same for equilibrium and non-equilibrium states, we may assume the same form of transition probability for the two cases. By considering a one-particle mean-field Ising model of a ferromagnet, we are able to compare the rates of convergence of the Markov process for different transition probabilities.

### Functional Equation

For transitions based on single-spin reversals, the condition of detailed balance at equilibrium implies

$$W_j(s) P_{\text{eq}}(s) = W_j(s^j) P_{\text{eq}}(s^j), \quad (2)$$

where  $P_{\text{eq}}(s)$  is the equilibrium probability distribution function of the state  $s$ . If we define the Hamiltonian for the spin system in the absence of any applied magnetic field,

$$H(s) = - \sum_{\langle ij \rangle} J_{ij} s_i s_j, \quad (3)$$

and assume for  $P_{\text{eq}}(s)$  the Boltzmann distribution

$$P_{\text{eq}}(s) \propto \exp(-\beta H(s)), \quad \text{with} \quad \beta = 1/k_{\text{B}}T, \quad (4)$$

then detailed balance yields

$$\frac{W_j(s)}{W_j(s^j)} = \frac{\exp(-\beta H(s^j))}{\exp(-\beta H(s))} = \frac{\exp(-\beta E_j s_j)}{\exp(\beta E_j s_j)}, \quad (5)$$

where

$$E_j = \sum_k J_{jk} s_k, \quad (6)$$

the summation being over nearest neighbours.

To set up the functional equation we first define

$$\Delta H = H(s^j) - H(s) = 2E_j s_j, \quad (7)$$

and assume the transition probability to depend on the energy of transition  $\Delta H$  alone, so that on introducing the variable

$$\xi = \exp(\beta \Delta H) > 0 \quad (8)$$

we may write

$$W_j(s) = \phi(\xi), \quad (9)$$

and, by symmetry, for the inverse transition

$$W_j(s^j) = \phi(1/\xi). \quad (10)$$

We observe that the dependence of the transition probability on the single variable implies that  $W(s)$  depends not on the whole configuration  $s$  but only on the local state generated by the transition, i.e. the transition process is spatially Markovian. Using the forms (9) and (10), the condition of detailed balance becomes

$$\xi \phi(\xi) = \phi(1/\xi) \quad \text{for} \quad \xi > 0. \quad (11)$$

This linear functional equation is quite general within statistical mechanics and holds whenever the present Markov process applies. For example, it holds for molecular systems where the variables are continuous rather than discrete (Metropolis *et al.* 1953).

### Solutions to Functional Equation

To solve equation (11) we first proceed by inspection, and note that there are several classes of solutions. One class of solutions comprises those that are rational functions in  $\xi$ , of which the simplest may be written

$$\phi(\xi) = A(1 + 1/\xi). \quad (12)$$

Here  $A$  is a parameter which depends on the initial state and which must be proportional to the average number of transitions per unit time. This solution may be generalized to higher-degree functions to obtain, for example,

$$\phi(\xi) = A(1 + \xi)/(1 + \alpha\xi + \xi^2). \quad (13)$$

The solution (13) does not reduce to (12) for  $\alpha = 0$ .

A second class of solutions consists of those that are rational in  $\xi^{\frac{1}{2}}$ , of which the simplest is

$$\phi(\xi) = A/\xi^{\frac{1}{2}}, \quad (14)$$

and which generalizes to

$$\phi(\xi) = A/(1 + \alpha\xi^{\frac{1}{2}} + \xi), \quad (15)$$

of which a particular form ( $\alpha = 0$ ) is

$$\phi(\xi) = A/(1 + \xi). \quad (16)$$

Note that equation (16) may be written

$$\phi(\xi) = A\xi^{-\frac{1}{2}}/(\xi^{\frac{1}{2}} + \xi^{-\frac{1}{2}}).$$

A third class of solutions arises from the substitution

$$\phi(\xi) = \xi^{-\frac{1}{2}}\psi(\xi) \quad (17)$$

so that  $\psi$  satisfies

$$\psi(\xi) = \psi(1/\xi). \quad (18)$$

In the first instance this equation is satisfied by any function  $f(\xi, \xi^{-1})$  which is symmetric in its arguments. The solution

$$\phi(\xi) = \xi^{-\frac{1}{2}}f(\xi, \xi^{-1}) \quad (19)$$

clearly contains some of the earlier solutions.

On inspecting equation (18) we see that it is a particular case of Schroeder's equation (Hille 1972)

$$f(g(\xi)) = \lambda f(\xi) \quad \text{for} \quad \xi > 0, \quad (20)$$

with  $g(\xi) = 1/\xi$  and  $\lambda = 1$ . It would appear that the solutions of equation (18) belong to a function algebra defined over the field of rational functions symmetric in  $\xi$  and that this field may be extended to include  $|\log \xi|$ .

Another general class of solutions to equation (18) may be constructed by a division of the interval  $(0, 1)$  into dyadic subintervals. Thus, if we define

$$\psi(\xi) = h(\xi) \quad \text{for} \quad 0 < \xi < 1, \quad (21a)$$

then we have

$$\psi(\xi) = h(1/\xi) \quad \text{for} \quad \xi > 1. \quad (21b)$$

More generally, if  $N$  be a positive integer, and we define

$$\begin{aligned} \psi(\xi) &= h_0(\xi) & \text{for} & \quad 0 < \xi < 2^{-N} \\ &= h_k(\xi) & & \quad 2^{-k} < \xi < 2^{-k+1}, \end{aligned}$$

with  $k = 1, 2, \dots, N$ , then we obtain

$$\begin{aligned}\psi(\xi) &= h_0(1/\xi) & \text{for } \xi > 2^N \\ &= h_k(1/\xi) & 2^{k-1} < \xi < 2^k,\end{aligned}$$

with  $k = 1, 2, \dots, N$ . The function  $\psi$  is then defined for all  $\xi > 0$ .

### Identification of Solutions to Functional Equation

Some of the solutions (listed in the previous section) to the detailed-balance functional equation (11) may readily be identified with the various transition probabilities used in computer simulations, although most of the results given here are new. Thus, for the Ising system we have from equation (14) the simple exponential form

$$W_j(s) = A \exp(-\beta E_j s_j), \quad (22)$$

which in the molecular context is called the 'symmetric' transition probability (Wood 1968). From equation (22) we derive

$$W_j(s) = A \cosh(\beta E_j) \{1 - \tanh(\beta E_j s_j)\}, \quad (23)$$

or simply

$$W_j(s) = \frac{1}{2} A \{1 - s_j \tanh(\beta E_j)\} \quad (24)$$

on redefining the constant  $A$ . This is the Suzuki-Kubo form.

From equation (16) we have

$$W_j(s) = A \{1 + \exp(2\beta E_j s_j)\}^{-1}, \quad (25)$$

from which we may also derive the Suzuki-Kubo form (24). From the equations (21) the simplest solution is obtained by taking  $h(\xi) \equiv 1$ , and this recovers equation (22). If we take  $h(\xi) = A\xi^{\pm 1}$  we have

$$\begin{aligned}\phi(\xi) &= A & \text{for } 0 < \xi < 1 \\ &= A/\xi & \xi > 1,\end{aligned} \quad (26)$$

which yields

$$W_j(s) = A \quad \text{for } E_j s_j < 0 \quad (27a)$$

$$= A \exp(-2\beta E_j s_j) \quad E_j s_j > 0, \quad (27b)$$

which in the molecular context is the 'asymmetric' form originally used by Metropolis *et al.* (1953).

### Rates of Convergence to Equilibrium

To compare the effects of different forms of transition probability on the rate of convergence to equilibrium of the simulation process, it is convenient to consider the analytical behaviour of a particular physical model which uses these different probability functions. Accordingly, we consider the behaviour of a simple one-particle mean-field Ising model of a ferromagnetic system of particles. Let the spin particles be located at the node points of a lattice and consider the behaviour of the particle

at node point  $P$ . We regard this particle as a one-particle sample from the infinite system. If  $z$  is the coordination number of the lattice then the sample particle will interact with its  $z$  nearest neighbours. However, the states of these nearest neighbours are unknown and must be estimated using the statistics of the sample. If the state variable for the sample particle is  $s$  with values of  $\pm 1$  then the simplest estimator for the states of the external particles is the site mean  $\langle s \rangle$ . The values of this estimator lie in the range  $[-1, 1]$  and so its direct use is perhaps somewhat unphysical. However, it does have the advantage of making this one-particle sample exhibit a critical behaviour, i.e. there exists a temperature  $T_c$  such that  $\langle s \rangle = 0$  for  $T \geq T_c$  and  $\langle s \rangle \neq 0$  for  $T < T_c$ . Thus, the model qualitatively at least contains the right kind of physics. By assigning  $\langle s \rangle$  to the external particles, we have in effect taken the thermodynamic limit by imposing a closure on the sequence of site correlation functions. The closure here occurs at the first term  $\langle s \rangle$ , and it follows that all higher correlations exist and are simply powers of  $\langle s \rangle$ . The existence of all correlations implies the existence of a critical temperature, although it might be zero (cf. the one-dimensional chain).

In the absence of a magnetic field, the Hamiltonian may be written

$$H = -zJs\langle s \rangle, \quad (28)$$

so that we have for the partition function

$$Z = \sum_{s=\pm 1} \exp(Kzs\langle s \rangle), \quad (29)$$

where  $K = J/k_B T$ . From equation (29) we define  $\langle s \rangle$  (self-consistently) by

$$\langle s \rangle = \tanh(Kz\langle s \rangle). \quad (30)$$

From equation (30) we have

$$\begin{aligned} \langle s \rangle &= 0 & \text{for} & & 0 < zK \leq 1 \\ &\neq 0 & & & zK > 1. \end{aligned}$$

That is, a critical temperature for the site mean  $\langle s \rangle$  (i.e. long-range order) exists and is given by  $T_c = z$ , in units of  $J/k_B$ . For the specific heat we find

$$C_u = \langle (H - \langle H \rangle)^2 \rangle = J^2 z^2 \langle s \rangle^2 (1 - \langle s \rangle^2),$$

which has a maximum at  $\langle s \rangle = \frac{1}{2}$ , for which  $Kz = 1.2463\dots$

To discuss the time-dependent behaviour of the one-particle sample, and hence of the infinite system, we proceed from the master equation (1). Thus, for  $s = \pm 1$  we have

$$dP(s, t)/dt = -W(s \rightarrow -s)P(s, t) + W(-s \rightarrow s)P(-s, t), \quad (31)$$

where  $P(s, t)$  is the probability that the sample is in state  $s$  and  $W(s \rightarrow -s)$  is the transition probability rate out of that state. Expressing  $P(s, t)$  in terms of  $\langle s \rangle$ ,

$$P(s, t) = \frac{1}{2}(1 + s\langle s \rangle),$$

and specializing to the case  $s = 1$  we have

$$d\langle s \rangle / dt = F^{(1)} + F^{(2)} \langle s \rangle, \quad (32)$$

where

$$F^{(1)} = -W(s \rightarrow -s) + W(-s \rightarrow s) \quad \text{and} \quad F^{(2)} = -W(s \rightarrow -s) - W(-s \rightarrow s). \quad (33)$$

From equations (8) and (28) we define

$$\xi = \exp(\beta \Delta H) = \exp(\gamma \langle s \rangle), \quad \text{with} \quad \gamma = 2zK, \quad (34)$$

and

$$W^+ \equiv W(s \rightarrow -s) = \phi(\xi) \quad \text{and} \quad W^- \equiv W(-s \rightarrow s) = \phi(\xi^{-1}). \quad (35)$$

In equilibrium  $d\langle s \rangle / dt = 0$ , so that we have

$$\begin{aligned} \langle s \rangle_{\text{eq}} &= -F^{(1)} / F^{(2)} = (W^- - W^+) / (W^- + W^+) \\ &= \frac{1 - \phi(\xi_{\text{eq}}) / \phi(\xi_{\text{eq}}^{-1})}{1 + \phi(\xi_{\text{eq}}) / \phi(\xi_{\text{eq}}^{-1})} \end{aligned}$$

and hence

$$\langle s \rangle_{\text{eq}} = (\xi_{\text{eq}} - 1) / (\xi_{\text{eq}} + 1) \quad (36)$$

on using (equation 11). Clearly,  $\langle s \rangle_{\text{eq}}$  is independent of  $\phi$ , that is, of  $W$ , as it must be. Using the definition (34) of  $\xi$  we recover the definition (30).

Using detailed balance equation (11) we may write the master equation (32) in the form

$$d\langle s \rangle / dt = -\{(1 - \xi) + (1 + \xi)\langle s \rangle\} \phi(\xi). \quad (37)$$

Setting  $\langle s \rangle = \langle s \rangle_{\text{eq}} + \Delta$  and substituting in equation (37) we obtain

$$d\Delta / dt = -\alpha \Delta,$$

where

$$\alpha = \frac{(1 + \xi_{\text{eq}})^2 - 2\gamma \xi_{\text{eq}}}{1 + \xi_{\text{eq}}} \phi(\xi_{\text{eq}}).$$

By definition  $\phi(\xi_{\text{eq}}) > 0$  and by inspection it follows that  $\alpha > 0$  for all  $T$ . Hence, the relaxation process must converge to the equilibrium state for all temperatures, and it follows further that the rate of convergence is a maximum when  $\phi(\xi_{\text{eq}})$  is a maximum. From this it appears, in particular, that the symmetric form (22) gives a rate of convergence greater than that of the asymmetric form (27) which in turn gives a greater rate than the form (25) used by Flinn and McManus (1961). This last result is in disagreement with the analysis of Cunningham and Meijer (1976). Of course, the present one-particle model is very special and, as Valleau and Whittington (1977) pointed out, the matter is quite involved. However, much of their analysis would appear invalid, since the transition probabilities in their examples do not appear to satisfy detailed balance, and it is difficult to see how the equilibrium state could ever be reached with a Markov process based on these probabilities. It must be emphasized that in the simulation of a statistical mechanical system the Markov property only holds with any accuracy when the stochastic process is close to convergence at equilibrium. After all, detailed balance is an expression of equilibrium.

## References

- Argyres, P. N., and Kelly, P. L. (1964). *Phys. Rev.* **134**, A98.
- Bolton, H. C., Fuller, B., Johnson, C. H. J., and Leng, C. A. (1977). *J. Phys. C* **5**, 791.
- Cox, D. R., and Miller, H. D. (1977). 'The Theory of Stochastic Processes' (Chapman & Hall: London).
- Cunningham, G. W., and Meijer, P. H. E. (1976). *J. Comput. Phys.* **20**, 50.
- Feller, W. (1957). 'An Introduction to Probability Theory and its Applications', Vol. 1 (Wiley: New York).
- Flinn, P. A., and McManus, G. M. (1961). *Phys. Rev.* **124**, 54.
- Hille, E. (1972). 'Methods in Classical and Functional Analysis' (Addison-Wesley: Reading, Massachusetts).
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., and Teller, E. (1953). *J. Chem. Phys.* **21**, 1087.
- Valleau, J. P., and Whittington, S. G. (1977). *J. Comput. Phys.* **24**, 150.
- Wood, W. W. (1968). In 'Physics of Simple Liquids' (Eds N. H. V. Temperley, J. S. Rowlinson, and G. S. Rushbrooke), p. 115 (North-Holland: Amsterdam).

Manuscript received 16 February 1979