

# Computer-Oriented Approach to the Ensemble Theory

Z. Alexandrowicz<sup>1</sup>

Received July 14, 1971

---

Computer simulation of the ensemble behavior of large interacting systems is hampered by the relative paucity of low-energy configurations. The stochastic sampling must, therefore, be guided toward the capture of configurations which, though rare, are *typical*, in the sense of contributing the maximum term of the partition function,  $\text{Max log } \zeta$ . In a previous Monte Carlo study by the author, such a guidance was applied to an *ab initio* construction of lattice configurations with a stochastic chain of steps, the transition probability at each step depending of both near and far neighbors. This study is now put on a more systematic and broader basis. It is argued that any computer study of the emergence of long-range order in large systems must be guided by  $\text{Max log } \zeta$ . Thus, lattice configurations are often constructed by a cyclic variation, which, guided by nearest-neighbor interaction, tries to relax an arbitrary initial configuration toward equilibrium (Metropolis). Such a relaxation may be also guided by  $\text{Max log } \zeta$ , with transition probabilities dependent on both near and far neighbors. Of main interest is the relevance of such a calculation for the actual relaxation behavior of the system, and this is discussed at some length.

---

**KEY WORDS:** Computer study; Monte Carlo; Metropolis method; stochastic models; ensemble theory; maximum term of partition function; Ising lattice; long-range order; phase transition; relaxation process.

## 1. INTRODUCTION

The power of modern calculating machines has drawn attention to the possibility of applying, as it were, statistical mechanics from first principles:

---

<sup>1</sup> Polymer Department, Weizmann Institute of Science, Rehovot, Israel.

the machines make it possible to construct samples of model configurations for a given system, providing thereby a direct simulation of its ensemble average behavior. As of now, only highly idealized systems can be studied in this manner, which hampers the comparison to experimental data, but the limitation will hopefully be overcome with the help of future computing facilities. One thing should, however, be very clearly realized. Advanced facilities will enable the construction of model configurations for systems consisting of many particles interacting with a great deal of complexity. Yet it will never be possible to carry out the astronomically large numbers of constructions for the ensemble of all the configurations, or even for any reasonably large part of it. One should therefore essay to find only those configurations that are typical of the ensemble as a whole. From this point of view, the situation is not anymore desperate for, as the system becomes macroscopically large, it appears that a single "typical" configuration embodies the ensemble behavior in its entirety. The computer work must, however, be guided explicitly toward finding these typical configurations; in other words, future possibilities and limitations call for a computer-oriented approach to the ensemble theory. In this context, some previous work of the author, dealing with the computer description of large systems of interacting particles involving long-range order, is briefly reviewed here and put (it is hoped) on a more rational basis. It is then argued that the method of singling out the typical equilibrium configuration for a system should also enable a computer study of its relaxation behaviour.

## 2. THE PRINCIPLE OF MAXIMUM TERM (TYPICAL) CONFIGURATIONS

Consider the configurational partition function for a system of  $n$  interacting particles at thermal equilibrium

$$Z = N \sum_i \nu_i \exp(-\beta E_i) \quad (1)$$

where  $\beta = 1/kT$ ,  $N$  is the number of configurations in the ensemble, and  $\nu_i$  is their frequency at energy levels  $E_i$ . Of particular usefulness here will be the well-known<sup>(1)</sup> approximation valid for the macroscopic limit

$$Z \simeq N \text{Max}[\nu_i \exp(-\beta E_i)] = N \nu_i^* \exp(-\beta E_i^*) \quad (1a)$$

The astronomically large number of distinct configurations for an even moderately large system (e.g.,  $N = 2^n$  for the Ising lattice) rules out the possibility of constructing the ensemble with any imaginable computing

machine. An attempt to construct a representative statistical sample with the direct, or purely random, Monte Carlo (MC) computer technique encounters the following difficulty. Configurations belonging to high energy levels are vastly more numerous. The maximum term of Eq. (1a) exemplifies such a behavior:  $\nu_i$  increases at the same rate that  $\exp(-\beta E_i)$  decreases. To a significant interval of  $\beta E_i$  (which is of the order of  $n$ ) corresponds therefore an exceedingly rapid exponential variation of  $\nu_i$ . Thus a sample picked at random will consist entirely of configurations  $I$  belonging to the highest, or group of highest, energy levels. In effect, their frequency is

$$\nu_I = \text{Max}(\nu_i) \simeq 1 \quad (2)$$

Hence the “pseudo-partition” function  $z$  that sums  $M$  configurations constructed with the direct MC method reduces to merely

$$z_{\text{direct}} = M \sum_i \nu_i \exp(-\beta E_i) \simeq M \exp(-\beta E_I) \quad (3)$$

Because of the relative smallness of  $\exp(-\beta E_I)$ , the contribution of such configurations to the true partition function is, however, incomparably smaller than the all-important maximum term of Eq. (1a). This implies that the purely random MC method does not enable us to capture those configurations that *alone* are typical of the system’s behavior. The difficulty was recognized long ago and hence a biased MC procedure has been adopted,<sup>(2,3)</sup> picking configurations with the help of a suitable probability  $P_i$ . In this case,

$$z_{\text{biased}} = M \sum_i \nu_i' [\exp(-\beta E_i)] / P_i \quad (4)$$

the frequencies  $\nu_i'$  in the biased sample evidently obeying

$$\nu_i' \propto \nu_i P_i \quad (5)$$

What is perhaps not clearly realized is the following. Unless the biasing procedure actually manages to offset the fatal decrease of  $\nu_i$  with increasing  $\exp(-\beta E_i)$  then, similar to Eqs. (2) and (3),

$$z_{\text{biased}} \simeq M [\exp(-\beta E_I)] / P_I ; \quad \nu_I' \simeq 1 \quad (6)$$

Here,  $z_{\text{biased}}$ , though much larger than  $z_{\text{direct}}$ , very possibly falls short of the maximum term of Eq. (1a) and in that case, the all-important typical configurations still cannot be captured.

To conclude: It seems that the attempt to get the largest possible value of  $z$  with a MC construction should constitute the guiding principle of the computer-simulated description of interacting systems. The following remarks should be made in this connection.

An invariant character of an MC sample, say with respect to  $E_i$  of the configurations, is due to the sharp predominance of one  $\nu_i'$  over all the others in a large system (except when near phase transition); hence it is doubtful whether it can be considered to be indicative of the reliability of a sampling procedure.<sup>(2)</sup> Moreover, with a large system, whatever be the size of a sample, it will still constitute a ridiculously small fraction of the ensemble. The effort spent in trying to construct large samples can therefore be spared altogether. In fact, searching for a largest  $z$ , or  $\log z$ , it is sufficient to construct only a *single configuration*; i.e., one assays to maximize

$$\begin{aligned}\log \zeta &= -\beta E_I - \log P_I \\ &= \log z_{\text{biased}}, \quad \text{for } M = 1\end{aligned}\quad (6a)$$

whereas a moderately-sized sample enables us to study fluctuations. This does not relate to the fact that to enable the search for  $\text{Max } \log \zeta$  many configurations should possibly be constructed in correspondance to different  $P_i$  and  $\nu_i'$  of variously biased MC. One last observation: suppose  $\log Z$  [Eq. (1a)] is indeed satisfactorily approximated by  $\log \zeta$  [Eq. (6a), the substitution of  $N$  by  $M = 1$  adding a constant]; in that case,

$$E_i^* = E_I$$

and

$$\log \nu_i^* = -\log P_I \quad \text{for } \log \zeta \sim Z \quad (7)$$

Statistical thermodynamics, however, relates  $\log Z$  to the ensemble average energy, which is the same as  $E_i^*$ , and to the entropy:

$$\log Z = -\beta E_i^* + (S/k) + \text{const} \quad (8)$$

Hence

$$S/k = -\log P_I + \text{const} \quad \text{for } \log \zeta \sim \log Z \quad (9)$$

Equation (9) provides an operational example on the equivalence of entropy and “bias of sampling.”<sup>(4)</sup>

### 3. DETAILED DISCUSSION OF BIASED MC METHODS

(a) To begin with, let the widely used “Metropolis”<sup>(5)</sup> method be recalled. One starts from an arbitrary initial configuration and derives from it a sequence of other configurations with a Markov chain of steps, each step corresponding to flipping the spin at a lattice site, for the Ising model, or, displacing a molecule in a volume, for the lattice gas, etc. A step of the sequence is illustrated in Fig. 1 for an Ising lattice in two dimensions. The transition probability from an  $i$  to a  $j$  configuration corresponds to flipping the spin of the  $k$ th site and is proportional to the Boltzmann factor

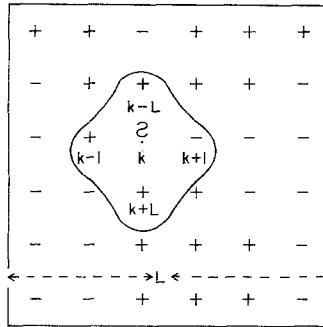


Fig. 1. A schematic drawing of the Metropolis cyclic relaxation process for a lattice of  $L$  sites in a row. The transition probability for site  $k$  depends on charges of the four nearest neighbors,  $k - 1$ ,  $k - L$ ,  $k + 1$ , and  $k + L$ .

$\exp(-\beta \Delta E_{ij})$ . Here,  $\Delta E_{ij}$  represents the energy difference for the two states of the  $k$ th site due to its different interaction with nearest neighbors  $k - 1$ ,  $k + 1$ ,  $k - L$ , and  $k + L$  ( $L$  being the lattice edge length). It has been shown<sup>(2,3)</sup> that in the limit of ergodic behavior of long Markov chains, the initial configuration relaxes toward the typical configurations of Eq. (1a). Yet, doubts have been raised with regard to the actual efficiency of the relaxation and whether, under unlucky circumstances, it might not become entrapped in semi stable states for very long durations of computing time. The foregoing discussion of the extremely rapid variation of  $\nu_i'$  with a small difference of  $E_i$  suggests the adoption of a more extreme attitude. One should expect such "unlucky circumstances" to almost certainly arise while applying the Metropolis relaxation to large enough systems with long-range interactions, for, apparently, the method's bias does not enable us to control the variation of  $\nu_i'$  with  $E_i$  in what concerns long order. Thus, consider, for example, the configuration of the two-dimensional Ising lattice depicted in Fig. 2, which might very well arise in the course of relaxation from a less to a more ordered state, at a temperature below the critical one. The increased order dictates the dissolution of the large negative droplet immersed in the predominantly positive lattice. The MC relaxation, however, will proceed almost entirely at the droplet's surface due to the random exchange of spins at sites with two positive and two negative neighbors ( $\Delta E_{ij} = 0$ ). The random process becomes oriented toward dissolution as it arrives at sites for which a flip to positive is associated with  $\Delta E_{ij} < 0$ . But take the peeling of the droplet's top row of  $l$  sites, from left to right corner, corresponding to the consecutive flipping of  $l$  spins to positive. The process requires that  $l - 1$  random flips,  $\Delta E_{ij} = 0$ , happen to fall in the correct sense followed by the oriented step at the right corner site, where three positive neighbors are now

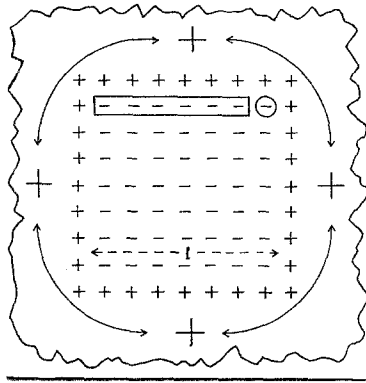


Fig. 2. A schematic representation of dissolution of the top row of a large ( $l^2$  sites) negative droplet immersed in a positive lattice. The Metropolis method requires that the  $l - 1$  negative spins (enclosed in the rectangle) flip to positive in succession, each with two positive and two negative neighbors,  $\Delta E_{ij} = 0$ . Only then is the succession followed by the (encircled) corner site flipping to positive with three positive and one negative neighbor,  $E_{ij} < 0$ .

adjacent to a negative spin,  $E_{ij} < 0$ . (Other peeling mechanisms are similar in this respect.) The chance of hitting it right, along the row and for all rows in turn, becomes increasingly slim as larger lattices and droplets are being treated. Clearly, the stepwise process lacks a “feeling of purpose,” or a guiding principle which would “tell” the *individual* steps that in the long run, in view of the balance, at  $\beta$ , of entropy versus energy for the *entire* lattice, it is “worthwhile” for the droplet’s spins to flip to positive. In the absence of such guidance, the unwanted droplets will just stay on “forever.” It seems, moreover, that other stepwise numerical calculations, molecular dynamics, for example, are open to a similar critique.

(b) A recently described *ab initio* construction of lattice configurations, with the help of “stochastic models”,<sup>(6-8)</sup> constitutes an attempt to utilize this guidance by  $\text{Max log } \zeta$  [Eq. (6a)]. To circumvent relaxation, one starts from an empty lattice and, with a Markov chain of steps, allots definite states to lattice sites in some ordered succession. The description detailed here is limited to one particular model which has been used for the square Ising lattice,<sup>(6)</sup> as illustrated in Fig. 3. The terminology is of two alternative states, or charges,  $\sigma = \pm 1$ , and the Markov chain proceeds for sites in a row from left to the right, and for rows from top to bottom (the rows being interconnected to a spiral). The transition probability is made dependent on both nearest and far neighbors. Thus at a  $k$ th step of the process, one computes the reduced charge of the two nearest neighbors,

$$\sigma' = (\sigma_{k-1} + \sigma_{k-L})/2 \quad (10)$$

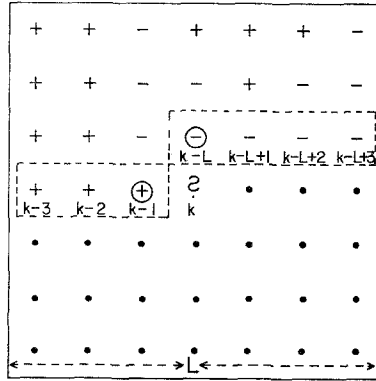


Fig. 3. A schematic drawing of the  $k$ th step in the *ab initio* construction of a square lattice of  $L$  sites in a row, guided by  $\text{Max log } \zeta$ . The transition probability for  $k$  depends on the past nearest neighbors  $k - 1$  and  $k - L$  (encircled) and of the more distant past neighbors  $k - L, k - L + 1, k - L + 2, k - L + 3, k - 1, k - 2,$  and  $k - 3$  (enclosed by the rectangle).

and a smeared-out reduced charge for a larger group of  $2s + 1$  next to nearest neighbors, irrespective of their sequential ordering,

$$\sigma'' = (\sigma_{k-1} + \dots + \sigma_{k-s} + \sigma_{k-L} + \dots + \sigma_{k-L+s}) / (2s + 1) \quad (11)$$

Now, for a random sampling, the transition probabilities  $f$  and  $1 - f$  for a positive and negative charge, respectively, are both equal to  $1/2$ . For a biased sampling,  $f$  should be increased in the presence of positive neighbors,  $\sigma > 0$ , while, in complete symmetry,  $1 - f$  should be increased for  $\sigma < 0$ . The separate effects of far and near neighbors,  $\sigma'$  and  $\sigma''$ , are represented with the help of two adjustable parameters,  $a$  and  $b$ . Specifically, taking  $a$  and  $b$  in the interval from zero to one, the bias is conveniently given by the expression

$$f = 1 / (1 + a^{\sigma'} b^{\sigma''}) \quad (12)$$

Equation (12) defines a set of  $3(2s + 1)$  different probabilities, for nearest-neighbor charge  $\sigma' = 1, 0,$  or  $-1$  and for far-neighbor cloud charge increasing in  $(2s + 1)$  steps from  $-1$  to  $1$ . A moment's reflection shows that the equations allows for the symmetric behavior required of  $(1 - f)$ .

A choice of particular values for  $a$  and  $b$  fixes the set of transition probabilities, henceforth referred to as a given "stochastic model." This enables us to actually carry out the biased MC construction, whereas definite states are allotted to  $n = L^2$  lattice sites in succession. Let  $p_k$  be the probability (equal to one of the set of  $f$  or  $1 - f$ ) with which the choice of a particular state has been made at the  $k$ th step of the process. In the limit of large  $n$  the ergodic

behavior of long Markov chains assures the probability of the various states at any step to tend to a constant value independent of  $k$ . Hence the product of all  $p_k$  gives (to within a multiplicative factor) the probability of a set of  $n$  steps, in which the states appear at a certain relative frequency. That set will be chosen “almost invariably” by a particular stochastic model. The choice of a state defines, however, the incremental negative and positive interaction energy per lattice site,  $\mp \epsilon_k$ , for pairs of similar and dissimilar neighbor charges, respectively. The set of  $n$  steps therefore fixes a certain value of the interaction energy for the lattice,  $E_I$ . At correspondance, the product of all  $p_k$  gives the probability with which configurations of that definite energy are picked by the process; in brief; it is equal to  $P_I$  of Eq. (6a),

$$\sum_{k=1}^n \log p_k \sim \log P_I \quad \text{for large } n \quad (13)$$

Hence, for the stepwise constructed configurations,

$$\log \zeta = -\beta \sum_{k=1}^n \epsilon_k - \sum_{k=1}^n \log p_k \quad (14)$$

where  $\log \zeta$  is maximum ( $\sim \log Z$ ) for a *typical* configuration. Equation (14) enables us to calculate  $\log \zeta$  at given  $\beta$ , for a particular stochastic model (one lattice configuration only needs to be constructed), notably, for a defined parametric dependence of the transition probabilities and specified values of the parameters. The search for  $\text{Max } \log \zeta$  is then conducted by varying the parameter values and the nature of the dependence; for example, the dependence of  $\sigma'$  and  $\sigma''$  in Eq. (12) could be refined to a dependence of three groups of neighbors of a different proximity.

True, even a greatly increased  $\log \zeta$  may be still far removed from the maximum value ( $\sim \log Z$ ), which is unknown. However, it becomes possible to objectively compare the results for different models, the “best results” corresponding to those of a highest value of  $\log \zeta$  at given  $\beta$ . (This is important when studying highly idealized systems, for which no experimental data are available.) It appears that such best results derived with the help of very simple models provide quite a good description of the main features of cooperative behavior. This is illustrated, for example, by the results that have been obtained<sup>(6)</sup> for a relatively large square Ising lattice of  $L^2 = 360^2$  sites. Figure 4 describes the dependence of the fraction of antiparallel neighbor pairs  $q$  (related to the interaction energy) upon the exponential “Ising temperature” factor  $\exp(-2K)$ . Figure 5 describes the long-range parameter  $R$  (related to the spontaneous magnetization) as a function of the reduced temperature  $T/T_c$ . The best results are described by the filled-in points (different types of points represent various stochastic models) and are seen



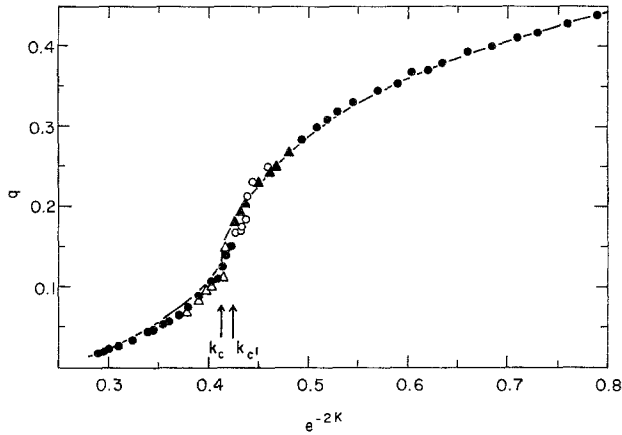


Fig. 4. The dependence of the fraction of antiparallel pairs  $q$  on  $K = J/kT$ , for the Ising square lattice. The solid line represents Onsager's exact solution; the circles and triangles, the results<sup>(6)</sup> for various Monte Carlo model *ab initio* constructions, those filled in black signifying the overall "best" results, i.e., of highest  $\log \zeta$  at given  $K$ .

to agree quite well with the line describing Onsager's exact solution. Without going into details, which have been discussed elsewhere, the following aspects may be briefly reviewed.

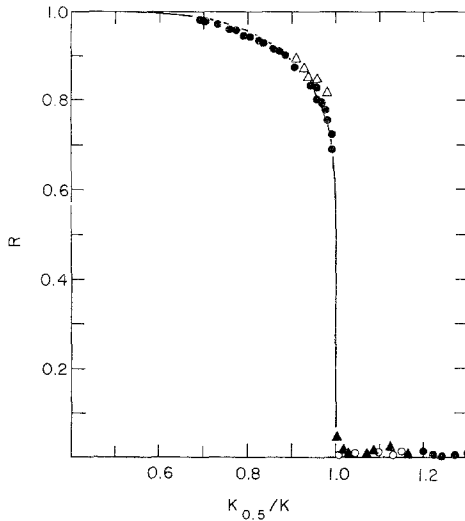


Fig. 5. The long-range parameter  $R$  versus  $K_{0.5}/K (\approx T/T_c)$  for the Ising square lattice. Here,  $K_{0.5}$  corresponds to  $R = 0.5$  (or to  $K_c$ , in effect) and equals 0.4407 and 0.428, respectively, for the exact theory and for the Monte Carlo model's "best" results.<sup>(6)</sup> The designation of the solid line and of the Monte Carlo points is as in Fig. 4.

The extension of the method to other lattice systems is almost immediate. Thus the application to a cubic Ising lattice involves only a redefinition of neighbor sites. The same intuitive concepts guide the choice of the model transition probabilities, namely a joint dependence of nearest and far neighbors, aided by symmetry considerations. Indeed, the results obtained in three dimensions seem to be more accurate (for example, critical point  $K_c$  was 0.223 versus 0.2217 of series expansions). An attempt to include the presence of an external field, other types of interactions, etc. should also involve a nonessential modification only of the stochastic models.

Furthermore, the method need not be limited to systems for which the lattice corresponds to any tangible physical property. The "lattization" is a matter of convenience, enabling one to construct the configurations of a system with any ordered sequence of steps. As an example, the critical behavior of a gas consisting of hard cubes with a square-well potential has been treated fairly successfully in the following manner.<sup>(6)</sup> A cubic volume is cut by a large number of parallel planes into thin square slices (in actual execution, the thickness of a slice was 1/180 that of the hard-cube molecules). The guided stochastic process considers one slice after another in succession, deciding whether a molecule should be placed on the slice and, if so, where it should be located on its surface. The transition probability of a present choice depends on the location of molecules in several neighboring "past" slices. Thus the occupancy of regions of a present slice is either excluded or favored (to a varying degree) depending on whether a region is overlapped by the hard core, or by the attractive potential, of one or several molecules belonging to the past slices. The term  $\log \zeta$  is computed very much as for the Ising lattice [Eq. (14)], except for the inclusion of the  $-\beta pV$  term for a gas at a given temperature and pressure. The difference worth noting is that the use of an arbitrary lattice (here, cutting the volume into slices) implies that different stochastic models might describe the same system of  $n$  particles with a different number of steps. In that case,  $\log P_I$  has to be merely renormalized for the purpose of objective comparison. In yet another application of the method, a stochastic model has been used<sup>(7)</sup> for the approximate description of random chains with "excluded volume" interactions. Here, the model transition probability favors the successive linking of the chain segments in a forward direction: This tends to decrease the interaction energy (due to self-intersections) at the expense of decreased  $-\log P_I$  (entropy of random linking). Once again, the typical configurations are picked by maximizing  $\log \zeta$ .

It has been also noted<sup>(6,7)</sup> that in some cases the MC construction, represented by a stochastic model, need not be carried out at all. Instead,  $E_I$  and  $P_I$  can be computed from an analytical formulation of the process. This possibility seems to be limited to the more approximate models (except

for a one-dimensional system<sup>(6)</sup>), but is important in relating the guided MC method to theoretical treatments. For the present purpose, however, these and other details need not be described any further.

#### 4. STOCHASTIC MODELS AND RELAXATION

At present, we would like to consider the broad possibilities of guiding computer study by the maximum term principle. The following features need to be underlined in this connection.

The maximization of  $\log \zeta$  can be carried out to advantage only with a quite small number of adjustable parameters, and thus the choice of different transition probabilities is necessarily limited. For that reason, a stochastic model description of a large interacting system in more than one dimension is always approximate. For example, the number of potentially different transition probabilities for the Markov chain describing the square Ising lattice in Fig. 3 is  $2^L$  (all the  $L$  sites in the last, "exposed" row originate chains of successive interactions that reach a present site). These, however, are grouped together into the set of  $3(2s + 1)$  probabilities, defined with the help of only two parameters  $\sigma'$  and  $\sigma''$  [Eq. (12)].

Notwithstanding this fact, quite shrewd guesses can be made in general with regard to the desired definition of the model set of transition probabilities.<sup>(6-8)</sup> Such guesses are aided by the intuitive understanding of what "goes on" at the microscopic level of a particular system. Notably, of great help are: the fact that a description needs to differentiate in detail various states of the nearest neighbors whereas far-off neighbors may be treated in terms of broader groupings, considerations of symmetry in space and with respect to an inverted sense of the interaction, and, lastly, the expected dependence of the probability upon multiple overlapping of a similar interaction. Indeed, this shrewd definition of the parametric dependence of transition probabilities terminates the physicist's workday, while the further evolution of the "best" MC description ( $\text{Max log } \zeta$ ) is left to the brute force of the calculating machine over the weekend....

In view of these observations, the *ab initio* MC construction that has been described at some length in the preceding section does not have to constitute the sole alternative enabling the computer study to be guided by  $\text{Max log } \zeta$ . True, the *ab initio* technique enjoys the advantage of having the desired configuration constructed right from the beginning of the process (except the starting boundary, which is irregular due to lack of past neighbors), with no need for any subsequent cycling. This greatly shortens the computing time and, for certain cases, facilitates the formulation of an analytical description. Yet, there are also important disadvantages. The shrewd guess

of a transition probability can rely on information with respect to only half of the environment of a present site. Thus, in the example of Fig. 3, one knows the charge of the "past" neighbors,  $k - 1, k - 2, \dots, k - L + 1$ , and  $K - L$ , but not of the "future" ones,  $k + 1, k + 2, \dots, k + L - 1$ , and  $k + L$ . This lack of definite information is made up by the method's parametric maximization, which tries to allow for the *average* effect of past upon future. Still, the method remains necessarily approximate, even for a short-range order. Moreover, the dependence of the transition probability on only the past half of the environment destroys certain symmetry properties of the description; for example, with the model of Fig. 3, the charge correlation becomes unequal along the right and left diagonals. On the other hand, in the course of the cyclic construction (Fig. 1), the entire environment of a present site is known (i.e., the charge of all lattice sites, although with the Metropolis method, the transition probability actually depends only on the four, directly interacting, neighbors,  $k - 1, k - L$ , and  $k + L$ ). Admittedly, the information is incomplete in another way: it is not final, as in the *ab initio* construction, but becomes modified at the subsequent steps of the process. The information that has been employed in the course of past steps is therefore rendered obsolete in the course of the transition of the *entire* lattice from one configuration to another (see the example of Fig. 2). With regard to short-range order, however, not too much importance should be attached to this disadvantage. Hence, an alternative approach worth trying is to combine what seems to be the best features of the two techniques, notably, carry out the cyclic relaxation of the Metropolis method, but define transition probabilities in terms of adjustable parameters depending on near and far neighbors [somewhat in the manner of Eq. (12)], thereafter, of course, adopting the guidance of  $\text{Max log } \zeta$ . The investigation of possible advantage of this approach with particular systems is now underway.

The possibility to guide a cyclic relaxation by the  $\text{Max log } \zeta$  principle touches, however, upon a far more important question: Can the method of stochastic models also be employed to advantage for a computer study of nonequilibrium conditions? For that purpose, let us return to the previous example of a square Ising lattice (Fig. 2) undergoing a relaxation which begins at a configuration far removed from equilibrium. Thus, consider an actual relaxation process, which is due to a spontaneous flipping of spins at random sites of a lattice, while the whole system is in contact with a heat bath at  $\beta$ . Suppose a model stochastic process, which flips spins at random sites with a Markov chain of steps, is set up to describe the actual relaxation. The definition of the transition probabilities for the model process (generally dependent on time) will be, constantly adjusted through parametric variation, to give a maximum incremental increase,  $\text{Max}(\delta \log \zeta)$ , for the variation of  $\log \zeta$  at any stage of the process. Since  $\delta \log \zeta$  is to be computed for a number

number of steps  $n'$  which, on one hand, is sufficiently small to describe instantaneous increments while, on the other, is large enough to enable the ergodic behavior of the Markov chain to be effectively attained [justifying Eqs. (13) and (14) to be used for computing  $\log \zeta$ ], the question of incremental maximization has to be carefully analyzed. At present, however, let it be only said that although  $n'$  is required to be small relative to the lattice size  $n$ , it should still be large enough to cover homogeneously the microscopic variation of a lattice configuration, in the sense that averages over  $n'$  and over  $n$  are equal. (In terms of the droplet argument of Fig. 2, it is required that  $n'$  is large enough to include the rare corner sites,  $\Delta E_{ij} < 0$ .)

Now, consider the line describing the rise of  $\zeta$ , first for the model relaxation guided by  $\text{Max}(\delta \log \zeta)$ , from  $\zeta_{\text{init}}$  to  $\zeta_{\text{equil}}$  (to within a specified accuracy), in  $t$  steps. This line should lie above any other line describing in  $t$  steps the evolution from  $\zeta_{\text{init}}$  for an arbitrarily different model relaxation process, for the first line has the highest initial slope and it cannot be crossed over by the second line at any later stage; this would imply that at the crossing point,  $\delta \log \zeta$  for the second line is larger than for the first, which constitutes, however,  $\text{Max}(\delta \log \zeta)$ . Hence the top value  $\zeta_{\text{equil}}$  is reached fastest with the first line. This brings us to the conclusion that, among various model relaxation processes, the one guided by  $\text{Max}(\delta \log \zeta)$  provides the most efficient way to attain equilibrium. Is this, however, the actual relaxation pathway for our physical system? For there is an important difference between the model and the actual process. With the former, at each stage of the process, the computer is ordered (by varying parameters) to derive a set of alternative configurations, starting from a given initial configuration. One of these is picked in accordance to  $\text{Max}(\delta \log \zeta)$  as the new starting configuration, and the relaxation continues. But is there any counterpart to that artificial control in the actual relaxation process? The question can be subdivided into two: Is there a counterpart to having a choice of various alternatives? And, if so, is the alternative chosen in accordance with  $\text{Max}(\delta \log \zeta)$ ? To deal with the questions, let us paraphrase what has been earlier, with regard to incrementing  $\log \zeta$  for the model process: The actual relaxation is considered in terms of successive intervals of  $n'$  steps (at each step, one site is relaxed);  $n'$  is sufficiently large to homogeneously cover local fluctuations or microscopic variability; hence values of  $\log \zeta$  can be computed for the originally homogeneous configuration (in the sense that any microscopic variation is repeated many times over) as well as for the later configurations at intervals of  $n'$ . With large enough systems, the microscopic variability becomes so small compared to the lattice size that we are free to require (except possibly at phase transition) that  $n'$  be also sufficiently small to give rise to an almost reversible change. In that case, a sequence of increments of  $n'$  states constitutes in effect an ensemble of configurations,

which offers the requisite choice of alternatives for the relaxation process. But will the choice proceed in accordance to a largest value of  $\log \zeta$ ? Recall that the argument for equilibrium conditions was that the ensemble of all configurations can be represented by a single configuration of  $\text{Max} \log \zeta$ . It seems that at each instant of the relaxation, after  $t$  steps, we still deal with an ensemble of configurations, which, however, is restricted to only those configurations which could originate from the initial one (at  $t = 0$ ), in  $t$  steps. This restricted ensemble can be also represented by its proper set of typical configurations, which give a  $\text{Max}(\log \zeta)_t$  smaller than the equilibrium  $\text{Max}(\log \zeta)$ , but, the best that can be had at time  $t$ . The maximum increase of  $\log \zeta$  would therefore constitute the variation principle enabling us to characterize the actual relaxation at each instant of time.

To conclude, the present discussion tries to explain, though it certainly does not claim to prove, how a mere extension of the concepts describing equilibrium might possibly lead to the principle of relaxation of very large systems being guided by  $\text{Max}(\delta \log \zeta)$ . One could, of course, *postulate* this principle, which is attractive in view of the fastest attainment of equilibrium, and then apparently would find it related to the postulate of minimum production of entropy for the total system, lattice plus thermal bath. Here, however, we have tried to argue that a sweeping postulate might be avoided by a careful analysis of incremental variation for the process.

From a more practical standpoint, let it be said that the argument suggests (though, once again, does not claim to prove) that model stochastic processes guided by  $\text{Max}(\delta \log \zeta)$  should enable the description of the relaxation of large interacting systems. A computer method unguided by consideration of  $\log \zeta$  of the entire lattice, like the one of Metropolis or molecular dynamics, will satisfactorily describe the relaxation of systems with short-range order, but will apparently go astray in the presence of long-range interactions (unless impossibly large samples were to be computed and weighted for each stage anew). As for the practical execution of the guidance, the description is necessarily approximate, depending on the correctness of the shrewd guess of transition probabilities and on the power of the machine maximization. The increments for  $n'$  steps need then presumably not be taken too small; in fact,  $n'$  of the order of  $n$  itself appears to constitute a reasonable choice. Furthermore, finding the largest increment  $\Delta(\log \zeta)$  at each stage depends on the succession of all preceding stages; hence the approximate evaluation of the relaxation process might run astray much more easily than a corresponding description of an equilibrium state. As has been said already, the execution of the process for simpler systems is now being studied, in an attempt to find efficient ways to arrive at equilibrium. However, with regard to the study of relaxation itself, in view of the idealized simplicity of the system (dictated by present-day computer facilities), the author is at a

loss with regard to experimental data which might be compared with the results of the computer.

Finally, several general remarks, in the order of their increasingly speculative content, should be added.

The guidance by  $\text{Max log } \zeta$ , whether at equilibrium or for a relaxation, does not constitute a technique for saving computer time. Instead, it purports to enable the description of certain systems, notably large and with long-range interactions, which otherwise should not be amenable at all to a computer study. True, attaining a satisfactory description of the long-range interactions might be reflected by a relatively small correction of values of the thermodynamic functions, which are at any rate computed in an approximate manner. Except when studying phase transitions, one could doubt the worth of all such efforts; if, however, one is concerned with the morphology of the typical configurations rather than with their thermodynamic properties (for example, in relation to the evolution of biological, social, or cosmic organization), it is the emergence of long-range order from apparently chaotic interactions which is of utmost importance. In this connection, a thought-provoking aspect of the relaxation being guided by  $\text{Max log } \zeta$  should be made more explicit. It will be recalled that the informal style of discussing the dissolution of the droplet of Fig. 2 inevitably dipped to a teleological attitude of a "feeling of purpose," "telling" individual steps of the process, of what is "worthwhile," etc. But take an observer of this relaxation process who, far from being an ignoramus, knows very well all the direct interactions and short-range correlations for the system. Say he fixes his attention to the preferred sequential flipping of the spins at the surface of the droplet to positive, accompanied by  $\Delta E_{ij} = 0$ , which has been described before in connection with the detailed mechanism of the droplet dissolution. The only thing that will escape the observer's attention is the existence of the extremely rare corner sites, at which the sequential flipping is accompanied by  $\Delta E_{ij} < 0$ , which drives the flipping to positive, increasing  $\text{log } \zeta$  of the entire system. (The presence of such sites is, of course, easily inferred for the regular Ising lattice, but not so for some highly complex intracellular organization, for example.) What the observer will see is an intrinsically chaotic process being mysteriously guided toward a very definite goal. Depending on whether the "system" is considered to be intelligent or not, the observer will then attribute this guidance to either free will or to divine power.

## REFERENCES

1. E. Schrödinger, *Statistical Thermodynamics* (Cambridge, London, 1938).
2. L. D. Fosdick, in *Methods in Computational Physics*, Vol. 1 (Academic Press, New York, 1963), p. 245.

3. W. W. Wood, in *Physics of Simple Liquids* (North-Holland Publ. Co., Amsterdam, 1968), Chapter 5.
4. E. T. Jaynes, *J. Math. Phys.* **106**:620 (1957).
5. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**:1087 (1953).
6. Z. Alexandrowicz, *J. Chem. Phys.* **55** : 2765 (1971).
7. Z. Alexandrowicz, *J. Chem. Phys.* **55** : 2780 (1971).
8. Z. Alexandrowicz, *J. Chem. Phys.*, In press.