PHYSICAL REVIEW VOLUME 116, NUMBER 3 NOVEMBER 1, 1959

Calculation of Order Parameters in a Binary Alloy by the Monte Carlo Method*

LLOYD D. FOSDICK
Digital Computer Laboratory, University of Illinois, Urbana, Illinois

(Received May 4, 1959)

A Monte Carlo sampling scheme similar to that used by Metropolis, Wood, and others in equations of state computations for gases has been used to investigate order-disorder phenomena in a face-centered cubic A_3B alloy. The model of the alloy assumes that the structure of the lattice is fixed and that interactions exist between first neighbors and second neighbors only. In most of the calculations detailed consideration is given to an array consisting of five unit cells on an edge (five hundred sites) with periodic boundary conditions. The long-range order and short-range order for first and second neighbors has been computed above and below the critical temperature. Using the energy parameter, $v_n=[(V_{AA}^{(r_n)}+V_{BB}^{(r_n)})/2]-V_{AB}^{(r_n)}$, for nth neighbors it is found that $v_2/v_1=-0.25$ and $v_1=816$ cal/mole gives the best agreement with experiments on $\mathrm{Cu_3Au}$. The critical temperature appears to vary linearly with the ratio v_2/v_1.

I. INTRODUCTION

With the high speed and large memory capacity available in modern computing machines there has been considerable interest in using them as a tool to attack certain problems in statistical mechanics by a rather direct Monte Carlo approach. This approach consists of generating a sample ensemble by Monte Carlo techniques in which the distribution of the systems approximates the correct statistical mechanical ensemble. The sample ensemble is then used to compute mean values of the desired thermodynamic functions.

Several years ago Metropolis and others[1] used this kind of a method to treat the two-dimensional hard sphere gas. Rosenbluth and Rosenbluth[2] extended these calculations to the three-dimensional system and obtained some results for a two-dimensional system of molecules with a Lennard-Jones interaction. Wood and Parker[3] continued these calculations to three-dimensional systems of molecules interacting with a Lennard-Jones potential. Wood and Jacobson[4] have also repeated some of the earlier work on the hard sphere system. This same method has also been applied to the lattice-gas model[5] and the two-dimensional Ising model.[6] In the latter two cases, where the exact results[7] are available for comparison, the Monte Carlo method gives estimates of the configuration energy which are accurate to 1% or better, except in the immediate neighborhood of the phase transition, in a reasonable amount of computing time.

This paper discusses an investigation of order-disorder phenomena in a binary alloy, A_3B, with face-centered cubic structure, using a Monte Carlo sampling technique similar to that used in the work mentioned above. The calculations were performed on an IBM 704 computer using a program known as Monte Crysto.

II. GENERAL DESCRIPTION OF THE METHOD

We consider a system of N atoms, A and B in the ratio 3:1, and assume them to be situated on the sites of a rigid, face-centered cubic lattice. The allowed states of this assembly of atoms correspond to the different arrangements of atoms on the sites, a specific state, say i, of the system being represented by the N component vector $\boldsymbol{\mu}_i$. The components $\mu_i(0)$, $\mu_i(1)$, $\cdots$, $\mu_i(N-1)$, are the two valued occupation numbers for the N sites. For the purpose of the discussion in the next section it will be assumed that $\mu_i(k)$ equals $+1$ or -1 corresponding to the kth site in the ith configuration state being occupied by an A or B atom, respectively. The energy, E_i, of the ith configuration is assumed to be due to the interactions between first neighbors and second neighbors only:

$$\begin{aligned}E_i=(N_{AA}^{(1)})_iV_{AA}^{(1)}+(N_{BB}^{(1)})_iV_{BB}^{(1)}\\+(N_{AB}^{(1)})_iV_{AB}^{(1)}+(N_{AA}^{(2)})_iV_{AA}^{(2)}\\+(N_{BB}^{(2)})_iV_{BB}^{(2)}+(N_{AB}^{(2)})_iV_{AB}^{(2)},\end{aligned}\qquad(1)$$

where the V's are the interaction energies for a pair of atoms indicated by the subscript AA, BB, or AB which are first neighbors, indicated by the superscript 1, or second neighbors, indicated by the superscript 2; the N's are the total numbers of pairs of atoms of the type indicated by the subscript and superscript in the ith configuration. For this system it is desired to compute the mean equilibrium values of the long- and short-range order parameters, which describe the interatomic correlations over large and small distances, as functions of temperature and the interaction energies.

The computation proceeds by generating a Markov chain of configurations $\boldsymbol{\mu}_0$, $\boldsymbol{\mu}_i$, $\boldsymbol{\mu}_m\cdots\boldsymbol{\mu}_n\cdots$ using the one-step transition probabilities p_{ij}, which give the probability that the state i will be immediately followed by the state j. These probabilities are the elements of

---

* This work was supported in part by the Office of Naval Research.

[1] Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller, J. Chem. Phys. **21**, 1087 (1953).

[2] M. N. Rosenbluth and A. W. Rosenbluth, J. Chem. Phys. **22**, 881 (1954).

[3] W. W. Wood and F. R. Parker, J. Chem. Phys. **27**, 720 (1957).

[4] W. W. Wood and J. D. Jacobson, J. Chem. Phys. **27**, 1207 (1957).

[5] Salsburg, Jacobson, Fichett, and Wood (to be published).

[6] L. D. Fosdick, Bull. Am. Phys. Soc. **2**, 239 (1957).

[7] G. Newell and E. Montroll, Revs. Modern Phys. **25**, 353 (1953).

the stochastic matrix P which operates on the 2^N component[8] probability vector $\boldsymbol{\psi}_k$. The components of $\boldsymbol{\psi}_k$ are just the probabilities for each of the 2^N states of the system at the start of the kth step in the chain. The transformation $\boldsymbol{\psi}_k \rightarrow \boldsymbol{\psi}_{k+1}$, resulting from one step, is described by the equation

$$\boldsymbol{\psi}_k P = \boldsymbol{\psi}_{k+1}, \tag{2}$$

and for n steps we clearly have

$$\boldsymbol{\psi}_k P^n = \boldsymbol{\psi}_{k+n}. \tag{3}$$

Provided that the elements of P satisfy the usual normalization condition

$$\sum_{\text{all } j} p_{ij} = 1 \quad (\text{all } i), \tag{4}$$

and that the states of the chain are ergodic it can be shown[9,10] that a limit for P^n exists. In particular

$$\lim_{n\to\infty} [P^n]_{ij} = u_j > 0, \tag{5}$$

and the vector $\boldsymbol{\Psi}$, having as components the limit probabilities u_j, is an eigenvector of P with eigenvalue 1, hence

$$\boldsymbol{\Psi} P = \boldsymbol{\Psi}. \tag{6}$$

This equation and the normalization condition

$$\sum_{\text{all } j} u_j = 1, \tag{7}$$

uniquely determine $\boldsymbol{\Psi}$.

The essence of the computational procedure lies in selecting P such that the elements of $\boldsymbol{\Psi}$ are just the Boltzmann weight factors

$$u_j = e^{-E_j/kT} \Big/ \sum_{\text{all } j} e^{-E_j/kT}. \tag{8}$$

With this choice of P, the existence of the limit of P^n implies that after a "sufficiently long time" the states of the chain $\boldsymbol{\mu}_0$, $\boldsymbol{\mu}_i$, $\boldsymbol{\mu}_j \cdots \boldsymbol{\mu}_n \cdots$ form an ensemble approximating the distribution defined by Eq. (8). The average value $\langle F \rangle$ of some function F_j of the system parameters

$$\langle F \rangle = \sum_{\text{all } j} \left[F_j e^{-E_j/kT} \Big/ \textstyle\sum_j e^{-E_j/kT} \right], \tag{9}$$

can then be approximated by

$$\langle F \rangle_{\text{approx}} = (1/M) \sum_{t=\delta}^{\delta+M-1} F(t), \tag{10}$$

where $F(t)$ is the value of F for the system at the tth step in the chain. The error in $\langle F \rangle_{\text{approx}}$ will tend to zero as δ and M tend to infinity but the practicality of this method, of course, rests on the possibility that the error can be made small enough without an excessive amount of computing.

It is clear from the conditions expressed by Eqs. (6), (7), and (8) that P is not uniquely defined. Ideally the freedom remaining in the choice of P should be constrained by a condition expressing the desire for rapid convergence of the approximations of the average values. From the standpoint of over-all computing time rapid convergence implies that both δ, denoting the point in the chain at which the averaging process started, and M, denoting the number of samples over which the average is performed, are small. The choice of δ depends on the rate of convergence of P^n which in turn depends primarily on the second largest eigenvalue of P. This is immediately evident if one expresses the probability vector $\boldsymbol{\psi}_0$ for the initial state of the chain (which contains 1 for one of its elements and zero for all the rest) as a linear combination of the eigenvectors of P. The problem of picking P such that the magnitude of the second largest eigenvalue and hence δ is minimized has not been solved. It should be pointed out that minimization of δ does not necessarily imply minimization of the computing time since the computing process thus defined might be so complicated as to require more over-all time than a simpler method requiring a larger δ. Nevertheless, the solution to the above problem would be of considerable interest. In its absence the choice of P is determined largely by the desire for simplicity and hence speed in carrying out the calculations which correspond to making a single step in the chain. Without knowledge of the second largest eigenvalue a good criterion for picking δ is lacking. Previous applications of this method have generally resolved this problem by directly considering the chain of sample values of some parameter of the system and setting δ at that point where the results for this parameter appear to become "steady." Experience has shown that this is not always a safe practice since the system may become partially trapped in a metastable state, exhibiting a steady behavior, and only after a long (on the computational time scale) time does it leave the metastable state. The system may become trapped for such a long time that states which should make a significant contribution to the ensemble average are completely omitted; this happened in the early hard sphere equation of state calculations.[4] The present computation scheme uses a quite different criterion for fixing δ which considerably reduces the chance that a metastable situation will be overlooked. This procedure involves the simultaneous generation of two or more independent chains with different initial states with the requirement that the mean values of certain system parameters agree among the different chains to within

---

[8] For simplicity we will ignore here the restrictions imposed by the 3:1 composition.

[9] This discussion follows closely that of Wood and Parker, reference 3.

[10] W. Feller, *Probability Theory and Its Applications* (John Wiley & Sons, Inc., New York, 1950), Chap. 15.

a preset margin of error. Details are given in the next section.

Referring again to Eq. (10), the number of terms, M, used for computing $\langle F\rangle_{\text{approx}}$ depends in detail on the physical system and the accuracy desired. When a relatively large volume of configuration space makes a significant contribution to the ensemble averages, as near a phase transition, it is necessary to take correspondingly more terms for a given accuracy. In the present computations the standard deviations of the averages are computed to verify that M is large enough for the desired accuracy. If M is held fixed, large variations in the standard deviations may be used to indicate the possible presence of a phase transition.

III. DETAILS OF THE COMPUTATION

A. The Model

The size of the system is a parameter of the Monte Crysto program which permits systems of up to fifteen unit cells on an edge, however, almost all of the computations were done with a system containing five unit cells on an edge, and consequently five hundred lattice sites. Indices i, j, k are used to label the sites as illustrated in Fig. 1(a) and Fig. 1(b) for the $5\times5\times5$ model. To simulate a lattice of infinite extent periodic boundary conditions are imposed, hence in the $5\times5\times5$ model the indices are interpreted modulo 10.

Every site of the model is occupied by either an A or B atom. The ratio of A atoms to B atoms, 3:1, is held fixed throughout the calculation, new configuration states being generated by the interchange of pairs of atoms. A configuration state of the model is conveniently realized in a binary computer since a single binary digit can be used to represent each site; the digit is one if the site is occupied by an A atom, while it is zero if the site is occupied by a B atom.

Although one can distinguish the six coupling energies, $V_{AA}^{(1)}$, $V_{BB}^{(1)}$, $V_{AB}^{(1)}$ for first neighbors and $V_{AA}^{(2)}$, $V_{BB}^{(2)}$, $V_{AB}^{(2)}$ for second neighbors, only the two energy parameters

$$v_1=[(V_{AA}^{(1)}+V_{BB}^{(1)})/2]-V_{AB}^{(1)}, \tag{11}$$

$$v_2=[(V_{AA}^{(2)}+V_{BB}^{(2)})/2]-V_{AB}^{(2)}, \tag{12}$$

arise explicitly in the computation. The computation is concerned with the change in the energy, ΔE, corresponding to an interchange of a pair of first neighbor atoms and it is easily seen that if a pair of such atoms, A and B, located on sites S and S', respectively, are interchanged, then

$$\Delta E=2(N_B^{(1)}-N_B^{(1)\prime}-1)v_1+2(N_B^{(2)}-N_B^{(2)\prime})v_2, \tag{13}$$

where $N_B^{(1)}$ and $N_B^{(1)\prime}$ are the numbers of first neighbor B atoms of S and S', respectively, before the interchange and $N_B^{(2)}$, $N_B^{(2)\prime}$ are like quantities for second neighbors. It is to be noticed that positive v_1 tends to make the ordered state have lower energy, while negative v_1 would energetically favor a complete separation of A and B atoms. Since second neighbors tend to be alike in the ordered state, a negative v_2 enhances the tendency to order produced by a positive v_1. For the later discussion we introduce a parameter λ, the ratio of v_2 to v_1,

$$v_2=\lambda v_1. \tag{14}$$

B. Generation of the Chain of Sample Configurations

As mentioned above, new configurations are generated by interchanging pairs of first neighbor atoms. This process will be described first and then it will be shown that it satisfies the convergence conditions. Let us fix attention on a particular site of the lattice which is to be called the central site. The manner in which this site is selected will be described shortly. One of the twelve first neighbors of the central site is selected at random[11] and if the atom on this first neighbor site is unlike the atom on the central site, then the change in energy ΔE, which would result from an interchange of their positions is computed. If $\Delta E\leq 0$, the positions of the two atoms are interchanged and the resultant configuration is taken to be the new configuration of the system. On the other hand if $\Delta E>0$ then the positions of the two atoms are interchanged with probability $e^{-\Delta E/kT}$; that is, a random number, RN, taken from a uniform distribution on the interval (0,1) is generated and if $RN<e^{-\Delta E/kT}$, then the interchange is made, otherwise it is not made. In either case the resultant configuration is taken to be the new configuration of the system. If the first neighbor site and the central site are occupied by like atoms, the new configuration is the same as the present configuration. With the completion of these considerations relative to a particular central site a new central site is selected and the process described above is repeated.

The face-centered cubic lattice can be pictured as being composed of the interpenetration of four simple cubic sublattices. Referring to Figs. 1(a) and 1(b),

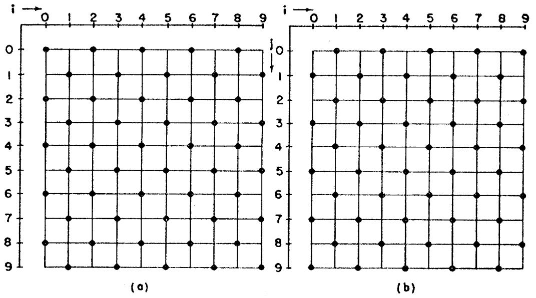

FIG. 1. The model showing coordinates of the sites. Layers shown in (a) and (b) alternate; (a) corresponds to even k, and (b) corresponds to odd k.

---

[11] Random numbers are generated by the multiplicative congruent method: see *Symposium on Monte Carlo Methods*, edited by H. A. Meyer (John Wiley & Sons, Inc., New York, 1956), p. 17.

sublattice I is composed of sites with i, j, and k even; sublattice II is composed of sites with i and j odd and k even; sublattice III is composed of sites with i odd, j even and k odd; sublattice IV is composed of sites with i even, j and k odd. The central sites are picked methodically, first picking all sites in sublattice I, next all sites in sublattice II, and so forth. In sublattice I the site $(0, 0, 0)$ is chosen first as a central site and new central sites are selected by advancing the i coordinate in steps of 2 until all sites in sublattice I, row $(i, 0, 0)$ are selected; next, sites in row $(i, 2, 0)$ of sublattice I are similarly selected, this continues until all sites in this sublattice with $k=0$ are selected; next, the sites in plane $k=2$ are similarly selected, etc., until finally all of the sites of sublattice I have been processed as central sites. Following this the sites in sublattices II, III, and IV are similarly selected. One iteration of the calculation consists of one pass through the entire lattice, processing each site as just described.

It is now to be shown that this process does produce a chain of states which, in the limit of a long chain, tend to be distributed according to the Boltzmann probabilities. That the ergodicity condition is satisfied follows immediately from the following observations. There is a nonzero probability for interchanging any pair of unlike first neighbor atoms. Any two configuration states differing only by the interchange of a pair of unlike atoms may be linked by many different chains corresponding to successive interchanges of first neighbor pairs. Finally, since there is nonzero probability for the interchange of any such pair of atoms there is a nonzero n-step transition probability linking any two configuration states.

Next, we will verify that the stochastic matrix, P, for this process does satisfy Eq. (6), where the elements of $\boldsymbol{\Psi}$ are given by Eq. (8). Let $P(k)$ be the stochastic matrix associated with consideration of the kth site as the central site. The product matrix

$$P=P(0)P(1)\cdots P(N-1) \tag{15}$$

is the stochastic matrix for one iteration of the calculation, where the N sites have been numbered in the order in which they are considered as central sites. The elements of $P(k)$ are defined precisely as follows:

(a) $i\neq j$:

$$\begin{aligned}[][P(k)]_{ij}&=\tfrac{1}{12}e^{-(E_j-E_i)/kT}, && (E_j>E_i)\\ &=\tfrac{1}{12}, && (E_j\leq E_i)\end{aligned}$$

if and only if

$$\begin{aligned}\mu_i(k)+\mu_j(k)&=0,\\ \mu_i(k')+\mu_j(k')&=0,\\ \mu_i(m)+\mu_j(m)&=\pm 2,\end{aligned}$$

where k and k' are first neighbor sites and the third condition is true for all m except $m=k$ or $m=k'$. If these conditions are not satisfied then this matrix element is zero.

(b) $i=j$:

$$[P(k)]_{jj}=\tfrac{1}{12}\textstyle\sum_m[1-e^{-(E_m-E_j)/kT}]+\tfrac{1}{12}n_j(k),$$

where the summation over m is a sum over all states for which

$$[P(k)]_{jm}=\tfrac{1}{12}e^{-(E_m-E_j)/kT},$$

and $n_j(k)$ is the number of first neighbor sites of k, in the jth configuration, which are occupied by the same type of atom as site k.

Consider the vector $\boldsymbol{\Phi}$ given by

$$\boldsymbol{\Psi}P(k)=\boldsymbol{\Phi}, \tag{16}$$

where the elements of $\boldsymbol{\Psi}$ are given in Eq. (8). The jth element of $\boldsymbol{\Phi}$,

$$[\boldsymbol{\Phi}]_j=\textstyle\sum_i u_i[P(k)]_{ij}, \tag{17}$$

will consist of a sum of terms of the following three types:

$$\text{(i)}\quad \frac{e^{-E_i/kT}}{\sum_i e^{-E_i/kT}}\times\frac{1}{12}e^{-(E_j-E_i)/kT}=\frac{1}{12}\frac{e^{-E_j/kT}}{\sum_i e^{-E_i/kT}},$$

$$\text{(ii)}\quad \frac{e^{-E_m/kT}}{\sum_i e^{-E_i/kT}}\times\frac{1}{12},$$

$$\begin{aligned}\text{(iii)}\quad &\frac{e^{-E_j/kT}}{\sum_i e^{-E_i/kT}}\times\frac{1}{12}\Big\{\sum_m(1-e^{-(E_m-E_j)/kT})+n_j(k)\Big\}\\ &=\frac{1}{12}\left\{\sum_m\frac{(e^{-E_j/kT}-e^{-E_m/kT})}{\sum_i e^{-E_i/kT}}+\frac{n_j(k)e^{-E_j/kT}}{\sum_i e^{-E_i/kT}}\right\},\end{aligned}$$

where the types (i) and (ii) arise from the off-diagonal matrix elements. In particular, the type (i) terms arise from transitions into the configuration j which result in an increase in configuration energy and the type (ii) terms arise from transitions into the configuration j which result in a decrease in configuration energy. The type (iii) term arises from the diagonal matrix element. It should be observed that there will be precisely 12-$n_j(k)$ terms of type (i) and (ii) altogether, since there will be one such term for every unlike first neighbor of the atom on site k. For each term of type (ii) there is a term of equal magnitude and opposite sign in (iii), resulting in a cancellation. Combining the terms which remain, it will be found that one obtains

$$[\boldsymbol{\Phi}]_j=e^{-E_j/kT}/\textstyle\sum_i e^{-E_i/kT}=u_j. \tag{18}$$

Therefore, we have the result

$$\boldsymbol{\Psi}=\boldsymbol{\Psi}P(k) \tag{19}$$

and since this is true for all k, we have

$$\boldsymbol{\Psi}=\boldsymbol{\Psi}P(0)P(1)\cdots P(N-1)=\boldsymbol{\Psi}P. \tag{20}$$

It has thus been shown that P does satisfy the conditions for the existence of the limit, Eq. (5), and that the limit elements are the desired ones. Furthermore it is to be observed from Eq. (19) that the result $\mathbf{\Psi}=\mathbf{\Psi}P$ is independent of the order in which the central sites are chosen.

C. Estimation of Convergence

The Monte Crysto program treats eight independent lattices simultaneously, generating from them eight statistically independent chains of sample configurations. Four of the lattices are given an initial configuration of perfect order; these lattices will be referred to as the low-temperature (LT) lattices since they begin from the equilibrium configuration at absolute zero. In the remaining four lattices the A and B atoms are distributed on the sites at random in the initial configuration and these lattices are correspondingly called the high-temperature (HT) lattices.

At the end of each iteration the current values of six configuration parameters for each of the eight lattices are recorded: $N[A/\mathrm{I}]$, $N[A/\mathrm{II}]$, $N[A/\mathrm{III}]$, $N[A/\mathrm{IV}]$, the number of A atoms on each sublattice; $N_{AB}{}^{(1)}$, the number of first neighbor AB pairs of atoms; $N_{AB}{}^{(2)}$, the number of second neighbor AB pairs of atoms. All four numbers $N[A/\mathrm{I}]$, $N[A/\mathrm{II}]$, $N[A/\mathrm{III}]$, and $N[A/\mathrm{IV}]$ are recorded for checking purposes. Storage space in the computer is provided for recording thirty-two values of each parameter for each lattice. The first phase of the computation consists of performing thirty-two iterations on the eight lattices, thus filling up this space.

Following this the second phase begins, during which a "convergence" test is made at the end of each iteration. In this phase each new set of configuration parameters replaces the oldest set (32 iterations old) held in the storage. Let $\langle N_{AB}{}^{(1)}\rangle_{\mathrm{LT},32}$ be the average value of $N_{AB}{}^{(1)}$, where the average is taken over the last thirty-two iterations for all four low-temperature lattices and let $\langle N_{AB}{}^{(1)}\rangle_{\mathrm{HT},32}$ be the same average over the high-temperature lattices. Likewise $\langle N_{AB}{}^{(2)}\rangle_{\mathrm{LT},32}$ and $\langle N_{AB}{}^{(2)}\rangle_{\mathrm{HT},32}$ are defined for second neighbors. The convergence test is passed when the following two conditions are both satisfied:

$$\frac{|\langle N_{AB}{}^{(1)}\rangle_{\mathrm{LT},32}-\langle N_{AB}{}^{(1)}\rangle_{\mathrm{HT},32}|}{\langle N_{AB}{}^{(1)}\rangle_{\mathrm{LT},32}+\langle N_{AB}{}^{(1)}\rangle_{\mathrm{HT},32}}<\epsilon_1, \tag{21}$$

and

$$\frac{|\langle N_{AB}{}^{(2)}\rangle_{\mathrm{LT},32}-\langle N_{AB}{}^{(2)}\rangle_{\mathrm{HT},32}|}{\langle N_{AB}{}^{(2)}\rangle_{\mathrm{LT},32}+\langle N_{AB}{}^{(2)}\rangle_{\mathrm{HT},32}}<\epsilon_2, \tag{22}$$

where ϵ_1 and ϵ_2 are small positive numbers.

The third phase of the computations begins as soon as the convergence test is passed. In this phase the results for the eight lattices are no longer distinguished and they are combined into a grand average. Thirty-two additional iterations are performed on each lattice to provide a total of five hundred and twelve samples for the calculation of the final averages. At the beginning of the third phase the sites of the lattice are classified as α or β in the eight systems; in each system the sites of the sublattice currently having a minority of A atoms are called the β sites and the sites of the other three sublattices are then called α sites. This is necessary in order to properly combine the results of the eight systems for the estimation of the long-range order. Aside from the bias inherent to the initial configuration (presumably there is none for the high-temperature lattices) the four alternate patterns of long-range order are equally likely since nothing in the program prejudices a particular sublattice to be a β sublattice.

IV. RESULTS

The primary limiting factor on the size of the model was the computing time, which for the 5×5×5 model amounted to 7.5 seconds per iteration per lattice. The computing time is essentially proportional to the number of lattice sites. In two cases the computations were repeated on a model having ten unit cells on an edge.

Computations were performed for a variety of values of kT/v_1, representing points above and below the critical temperature. Three different values of the second neighbor interaction parameter $\lambda=0$, $-\frac{1}{4}$, $-\frac{1}{2}$ were used. The convergence test constants ϵ_1 and ϵ_2, were each equal to 0.01.

A summary of the results is presented in Table I. The quantities α_1 and α_2 are Cowley's[12] short-range order parameters. The Cowley order parameter α_i, describing the order among atoms separated by a distance r_i (i.e., ith neighbors) is defined by

$$\alpha_i=1-P_{AB}(r_i)/2X_AX_B, \tag{23}$$

where $P_{AB}(r_i)$ is the probability of finding a pair of atoms AB separated by a distance r_i, X_A is the fraction of A atoms in the system, and X_B is the fraction of B atoms in the system. In the estimates obtained here for α_1 and α_2 the probabilities $P_{AB}(r_1)$ and $P_{AB}(r_2)$ have been replaced by

$$\langle N_{AB}{}^{(1)}\rangle/6N \quad\text{and}\quad \langle N_{AB}{}^{(2)}\rangle/3N,$$

respectively, where the averages have been taken over the ensemble of five hundred and twelve configurations obtained as described above. The indicated errors are the standard deviations. It is to be noted that a perfectly ordered face-centered cubic lattice, with $X_A=\frac{3}{4}$, $X_B=\frac{1}{4}$ yields $\alpha_1=-\frac{1}{3}$, $\alpha_2=1$, and a completely random arrangement of atoms yields $\alpha_1=0$, $\alpha_2=0$. The quantity S in the third column of results in Table I is the usual long-range order parameter defined by

$$S=1-\frac{P(A/\beta)}{X_A}, \tag{24}$$

where $P(A/\beta)$ is the probability of finding an A atom

---

[12] J. M. Cowley, Phys. Rev. 77, 669 (1950).

TABLE I. Summary of results.

kT/v_1	λ	$-\alpha_1$	α_2	S	Iterations to converge
0.95[a]	0	0.230 ±0.001	0.411 ±0.006	0.370 ±0.010	178
1.0	0	0.188 ±0.001	0.230 ±0.003	0.078 ±0.005	125
1.25	0	0.156 ±0.001	0.136 ±0.002	0.017 ±0.003	82
1.25[b]	0	0.1547±0.0002	0.1354±0.0005	0.0108±0.0013	83
1.5	0	0.138 ±0.001	0.104 ±0.002	0.004 ±0.004	35
1.75	0	0.125 ±0.001	0.085 ±0.002	0.015 ±0.004	33
2.0	0	0.114 ±0.001	0.069 ±0.002	0.000 ±0.003	32
10.0	0	0.034 ±0.001	0.004 ±0.001	0.000 ±0.002	32
10.0[b]	0	0.0325±0.0003	0.0043±0.0004	0.0013±0.0008	46
0.8[a]	−0.25	0.3080±0.0004	0.913 ±0.002	0.955 ±0.001	56
1.25[a]	−0.25	0.295 ±0.001	0.867 ±0.002	0.928 ±0.001	51
1.5	−0.25	0.240 ±0.001	0.640 ±0.004	0.771 ±0.004	131
1.6[a]	−0.25	0.194 ±0.001	0.409 ±0.007	0.445 ±0.013	82
1.75	−0.25	0.145 ±0.001	0.182 ±0.002	0.022 ±0.005	68
2.0	−0.25	0.129 ±0.001	0.140 ±0.002	0.009 ±0.004	40
2.25	−0.25	0.117 ±0.001	0.119 ±0.002	0.002 ±0.004	37
0.8[a]	−0.5	0.3062±0.0004	0.909 ±0.002	0.953 ±0.001	57
1.25	−0.5	0.2982±0.0004	0.883 ±0.002	0.940 ±0.001	136
1.75	−0.5	0.284 ±0.001	0.835 ±0.003	0.907 ±0.003	274
2.0	−0.5	0.245 ±0.001	0.688 ±0.003	0.812 ±0.002	99
2.25	−0.5	0.146 ±0.001	0.251 ±0.004	0.185 ±0.010	220
2.4	−0.5	0.134 ±0.001	0.210 ±0.004	0.037 ±0.008	50
2.5	−0.5	0.126 ±0.001	0.182 ±0.002	−0.0019±0.005	46
3.2	−0.5	0.096 ±0.001	0.108 ±0.002	0.011 ±0.004	35
10.0	−0.5	0.036 ±0.001	0.024 ±0.001	−0.003 ±0.003	32

[a] Initial conditions modified to partially order HT lattices.
[b] Model with ten unit cells on an edge.

on a β site. In the estimates obtained here for S the probability $P(A/\beta)$ has been replaced by

$$4\langle N(A/\beta)\rangle/N,$$

where again the average has been taken over the ensemble of five hundred and twelve configurations. In the fourth column of results the number of iterations performed before the convergence condition was satisfied is given.

The number of iterations that had to be performed before the convergence condition was satisfied varied considerably. Above the critical temperature the convergence condition was satisfied relatively fast; usually fifty iterations were sufficient. Below the critical temperature the number of iterations was sometimes very large. The $\lambda=0$ case presented the most difficulty in this respect and the lack of data for that case is due to the fact that in a number of the runs which were attempted the convergence condition was never satisfied; usually after five hundred iterations without convergence the run would be terminated. The long times to convergence were not as common in the runs corresponding to nonzero λ. The source of this difficulty was found to be due to the absence of a consistent pattern of long-range order in some of the high-temperature lattices. These lattices would reach fairly rapidly configurations of order consistent with the specified values of kT/v_1 and λ, except that in one plane, or sometimes two planes, the order would be out of phase. In the out of phase plane the positions of the A and B atoms are interchanged with respect to the pattern of order established elsewhere in the lattice. Figure 2 illustrates an out of phase plane in an otherwise perfectly ordered lattice. The low-temperature lattices, that is the ones starting from a configuration of perfect order, did not develop the out of phase planes. As a result the first condition, Eq. (21), relating to the short-range order of first neighbors would be satisfied but the second convergence condition, Eq. (22), relating to the short-range order of second neighbors would not be satisfied. It was observed that these out of phase planes were very stable and only rarely did one disappear, even after hundreds of iterations.

This behavior is easily explained.[13] Consider a perfectly ordered lattice, Fig. 2, and notice that the introduction of an out of phase plane does not change the number of first neighbor AB pairs. Consequently when $\lambda=0$, and therefore only first neighbor interactions

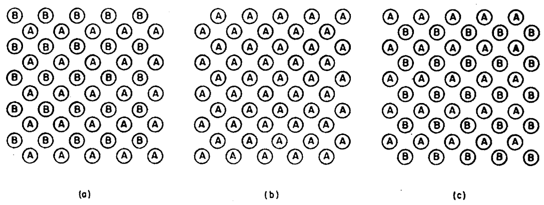

FIG. 2. A perfectly ordered lattice consists of successive planes of atoms alternating between the pattern shown in (a) and that shown in (b). A plane of atoms arranged as shown in (c) substituted for one of the planes of type (a) constitutes an out-of-phase plane.

---

[13] This phenomenon is also supported by experimental evidence. Chipman has inferred from his observations that out of phase regions, one layer thick, with about nine atoms per layer, were present; see D. R. Chipman, J. Appl. Phys. **27**, 739 (1956).

are present, configurations with out of phase planes have the same energy as one that has no out of phase planes. It is evident that this will remain approximately true even when the long-range order is not quite perfect. Finally, since configurations with out of phase planes clearly have higher entropy it follows that such configurations are to be expected when $\lambda=0$.

The stability of an out of phase plane is a natural consequence of the nature of the elementary steps which are made in configuration space by the present Monte Carlo method. The elementary transition probabilities relate to interchanges of just pairs of atoms and since reordering an out of phase plane by such interchanges requires going through a set of states of higher energy the corresponding transition probabilities are small. This "energy barrier" problem could be reduced by making an elementary step correspond to the change in position of a large group of atoms rather than just a single pair. On the other hand, this "large step" approach requires a longer computation time per step and it is therefore not clear that such an approach would really improve the method. There is another reason for not adopting the latter method. Although the kinetics of this computation do not necessarily represent the kinetics of the ordering process in the real system it seems that in the present scheme there is a closer correspondence than there would be if the "large step" approach were used.

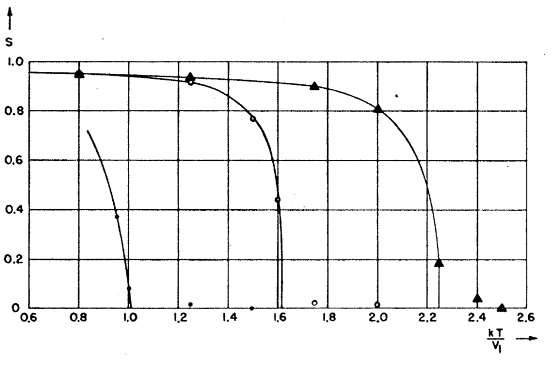

FIG. 3. Long-range order S vs kT/v_1: •,$\lambda=0$; ○,$\lambda=-0.25$; ▲,$\lambda=-0.5$.

Although it was felt that it would be desirable to obtain statistics on the appearance of the out of phase planes, it was evident that such an investigation would require excessive amounts of computing time with the present program. Therefore, attention was focused on the class of configurations having no out of phase planes and the results presented here apply to these configurations. In order to preclude appearance of out of phase planes in a few of the more difficult cases the initial conditions were altered. The four high-temperature lattices were given configurations with a well established pattern of long-range order. These starting configurations were, in fact, the resultant configurations of the four low-temperature lattices after ninety-seven iterations with $kT/v_1=2.0$ and $\lambda=-0.5$. This alteration of the starting conditions was made in just five of the computations; this is indicated in Table I by the superscript a.

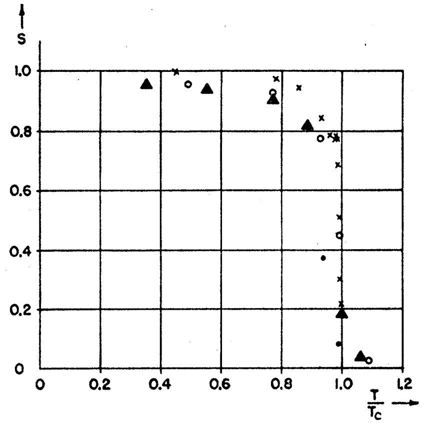

FIG. 4. Long-range order S vs T/T_c: •,$\lambda=0$; ○,$\lambda=-0.25$; ▲,$\lambda=-0.5$; ×, experimental results of Keating and Warren[14] for Cu$_3$Au.

The long-range order S is plotted as a function of kT/v_1 for different λ in Fig. 3. The present method of calculation will never yield zero long-range order because the system considered is necessarily finite in size and only in the limit of an infinite system can a true critical temperature with an associated vanishing of the long-range order make its appearance. Also, for purely statistical reasons, the results obtained from a small sample will exhibit fluctuations; an estimate of these statistical fluctuations is given by the standard deviations shown in Table I. However, it is clear from Fig. 3 that the long-range order becomes very close to zero after the sharp drop in S and it is reasonable to assume that S would tend to zero at these higher temperatures with increasing lattice size (though this is not apparent from the results of two runs made with the large model). Therefore, the curves have been extrapolated to zero long-range order as shown in Fig. 3, thus establishing an estimate of the critical temperature kT_c/v_1. For $\lambda=0$, -0.25, and -0.5 it is found that

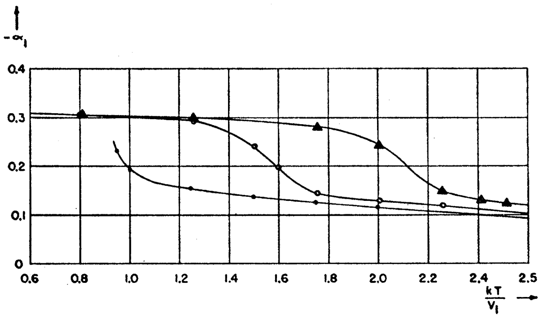

FIG. 5. Short-range order parameter $-\alpha_1$ vs kT/v_1: •,$\lambda=0$; ○,$\lambda=-0.25$; ▲,$\lambda=-0.5$.

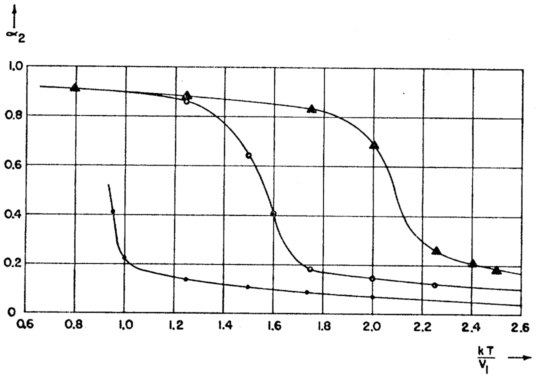

FIG. 6. Short-range order parameter α_2 *vs* kT/v_1: •,$\lambda=0$; ○,$\lambda=-0.25$; ▲,$\lambda=-0.5$.

$kT_c/v_1=1.01$, 1.61, and 2.25, respectively. The variation of kT_c/v_1 with λ is essentially linear. In Fig. 4 these results are compared with the experimental measurements of long-range order in Cu_3Au made by Keating and Warren.[14] This alloy is chosen for the comparison since the lattice structure remains face-centered cubic both above and below the critical temperature, and the lattice parameter changes only very slightly; the lattice parameter changes from 3.7478 A at 18°C to 3.7940 A at 620°C with a discontinuity of about 0.0044 A at 385°C.[15] The experimental results show a higher degree of order at low temperatures and a more abrupt drop of the long-range order at the critical temperature than the calculated results. The discrepancy cannot be explained by an error in location of the critical temperature from faulty extrapolation of the curves in Fig. 3. It seems likely that the small size of the model is responsible. This is inferred from computations on the Ising model where it has been found that the more accurate computation schemes produce a more abrupt vanishing of long-range order. In this figure no significant difference is apparent between results for $\lambda=-0.25$ and $\lambda=-0.5$.

The short-range order parameters are plotted as functions of kT/v_1 in Figs. 5 and 6. They are plotted as functions of T/T_c in Figs. 7 and 8. Cowley's measurements[16] of α_1 and α_2 in Cu_3Au are shown in the latter figures. Since the present calculations can be expected to give the most accurate results at higher temperatures, it appears from Fig. 7 that $\lambda=-0.25$ gives the best agreement with the experimental results. In Fig. 8 the curves for α_2 all fall so close together that an attempt to fit the experimental data by a choice of λ doesn't seem to be too meaningful. Furthermore, since the measurement of α_2 is less accurate than that of α_1 there is a good chance that the discrepancy indicated by the results in Fig. 8 may be due to experimental error.

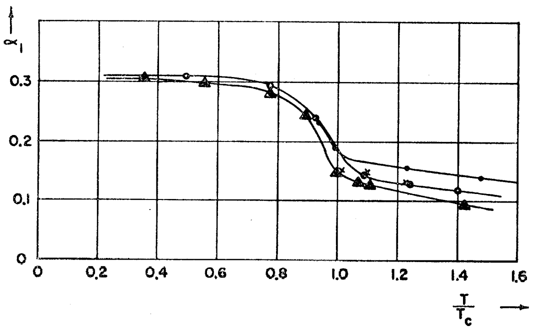

FIG. 7. Short-range order parameter $-\alpha_1$ *vs* T/T_c: •,$\lambda=0$; ○,$\lambda=-0.25$; ▲,$\lambda=-0.5$; ×, experimental results of Cowley[16] for Cu_3Au.

The value $\lambda=-0.25$ lies between that given by Cowley,[12] $\lambda=-0.1$, and that given by Fournet,[17] $\lambda=-0.5$. The value of v_1 can be estimated, taking for kT_c/v_1 the value 1.61 which was obtained for $\lambda=-0.25$ and $T_c=394$°C, measured by Cowley; this gives

$$v_1=816 \text{ cal/mole},$$

which is to be compared with

$$v_1=711 \text{ cal/mole} \quad \text{(Cowley)},$$
$$v_1=802 \text{ cal/mole} \quad \text{(Fournet)}.$$

Fournet's result applies to a "best fit" to experimental short-range order data; a lower value, 604 cal/mole, must be used for a best fit to experimental long-range order data.[17]

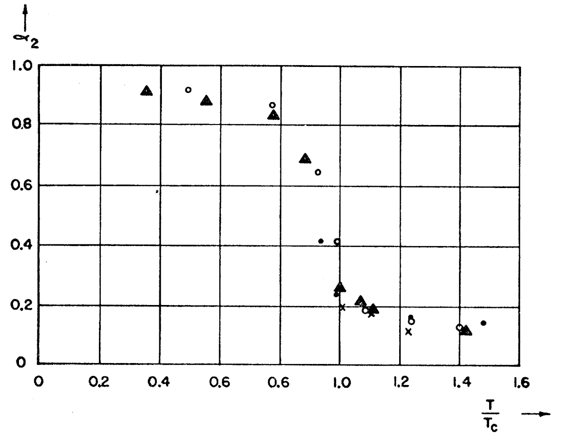

FIG. 8. Short-range order parameter α_2 *vs* T/T_c: •,$\lambda=0$; ○,$\lambda=-0.25$; ▲,$\lambda=-0.5$; ×, experimental results of Cowley[16] for Cu_3Au.

The results appear to indicate the presence of a phase change but they are not accurate enough for one to conclude with any definiteness whether the phase change is first or second order. If any success is to be had in resolving this point with some degree of certainty by the present method, it is clear that an amount of

---

[14] D. T. Keating and B. E. Warren, J. Appl. Phys. **22**, 286 (1951).

[15] E. A. Owen and Y. H. Liu, Phil. Mag. **38**, 342 (1939).

[16] J. M. Cowley, J. Appl. Phys. **21**, 24 (1950).

[17] Lester Guttman, *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press, Inc., New York, 1956), Vol. 3, p. 145.

computation is required which is beyond the practical limitations of presenting computing equipment. The system studied here represents about the upper limit in size and complexity that can be reasonably handled with present computing equipment.

ACKNOWLEDGMENTS

The author thanks the Digital Computer Laboratory of the Midwestern Universities Research Association, where the initial computations were performed, and the Applied Mathematics Division of the Argonne National Laboratory where the bulk of these computations were performed on their IBM 704 computer. The author is indebted to R. F. King of the latter Laboratory for supervising the running of the program. The author also thanks A. Nordsieck and A. H. Taub for their advice and encouragement in this work.

PHYSICAL REVIEW VOLUME 116, NUMBER 3 NOVEMBER 1, 1959

Exciton Spectrum of Cadmium Sulfide

D. G. THOMAS AND J. J. HOPFIELD
Bell Telephone Laboratories, Murray Hill, New Jersey

(Received June 10, 1959)

The reflectance and fluorescent spectra of hexagonal CdS crystals have been measured at 77° and 4.2°K using polarized light in the region of 5000 A. Structure not previously reported has been found in the reflectivity curves which leads to the identification of three exciton series. These can be understood in terms of the splitting of the valence band into three levels at $\mathbf{k}=0$. The observation of excited exciton states and the polarization properties of the excitons make possible: (1) The determination of two of the three exciton binding energies, (2) the determination of the energy splittings of the three valence bands, (3) the verification of the symmetry assignments of the valence and conduction bands, and (4) correlation of the work of others with the present work, showing that the definite intrinsic effects are consistent both with our observations and our interpretations. The fluorescent experiments strongly suggest that the radiative decay of excitons occurs not directly, but from localized impurity exciton states in agreement with theory.

I. INTRODUCTION

THERE exists in the literature much evidence of fine structure in various optical effects displayed by CdS at wavelengths near the onset of intrinsic absorption. In spite of these detailed studies the description of the band structure and exciton energy levels in CdS has remained in a state of confusion. The present work represents an attempt to correct this situation and was suggested by the reflectivity measurements of Dutton.[1] From theoretical considerations, together with observations made on ZnO, it was expected that there should occur in the reflectance spectrum two strong peaks in addition to the three already reported. We have therefore remeasured the reflectance at 77°K over a wider range of wavelengths, and have extended the observations to 4.2°K. In order to correlate the various effects, the fluorescence spectrum was also measured at both nitrogen and helium temperatures.

Gross and collaborators,[2–4] and others[5] have studied the absorption spectra of thin crystals at 4.2°K. Gross has pointed out that for wavelengths shorter than 4860 A the absorption lines observed are strong and are an intrinsic property of the crystal, and so are to be ascribed to exciton states. At longer wavelengths the lines are rather weak and vary from one crystal to another. Some lines are independent of the thickness of the crystal and so are connected with surface effects. The weak lines are associated with imperfections, although they cannot be attributed to excited impurity states in the sense that this term is normally understood. They apparently correspond to the formation of excitons in the field of an imperfection which has the effect of slightly lowering the exciton energy. These states may be called impurity excitons. Although Gross has described one set of lines active in light polarized perpendicular to the hexagonal c axis (hereafter denoted by $E\perp c$) between 4855 and 4806 A as being due to the $n=3$, 4, 5, 6 states of a hydrogen-like exciton series with the $n=1$ and 2 states missing, this description is not entirely satisfactory as there are other lines unaccounted for.

The uv-stimulated fluorescent spectra at 4.2°K may be divided into two regions. At wavelengths longer than about 5100 A there is a series of regularly spaced peaks having half-widths of about 10 A, with an energy difference between peaks equal to that of the longitudinal optical phonon.[5,6] This is referred to as the

---

[1] D. Dutton, Phys. Rev. **112**, 785 (1958).

[2] E. F. Gross, Suppl. Nuovo cimento **3**, 672 (1956).

[3] Gross, Razbirin, and Iakobson, J. Tech. Phys. U.S.S.R. **27**, 1149 (1957) [translation: Soviet Phys. (Tech. Phys.) **2**, 1043 (1957)].

[4] E. F. Gross and B. S. Razbirin, J. Tech. Phys. U.S.S.R. **27**, 2173 (1957) [translation: Soviet Phys. (Tech. Phys.) **2**, 2014 (1957)].

[5] L. R. Furlong and C. F. Ravilious, Phys. Rev. **98**, 954 (1955).

[6] Lambe, Klick, and Dexter, Phys. Rev. **103**, 1715 (1956).