PHYSICAL REVIEW VOLUME 142, NUMBER 2 FEBRUARY 1966

Correlation Functions of Disordered Binary Alloys. I

PHILIP C. CLAPP
Ledgemont Laboratory, Kennecott Copper Corporation, Lexington, Massachusetts
AND
SIMON C. MOSS*
Department of Metallurgy, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received 9 August 1965; revised manuscript received 4 October 1965)

An exact expression for the correlation functions (or Warren short-range order parameters) of an alloy of arbitrary composition and range of interaction is derived in terms of the Flinn operators. An approximate method of solution is found by replacing certain operators by their averages. One such choice leads to the equations formerly deduced by Cowley, and constitutes an alternative derivation of Cowley's theory. A second choice yields Zernike's equations and provides a method of comparing the approximations inherent in these two theories. Finally, a third solution is suggested which appears to have the same order of accuracy as the former two, but which has the advantage of being exactly soluble. It also provides an inversion formula which, in theory, can be used to infer the form of the two-particle interaction energy as a function of distance in the crystal. The theory is valid only above the ordering temperature of an alloy, but may be used with equal facility for a variety of alloys with differing compositions and ranges of interaction.

I. INTRODUCTION

THE central problem considered in this paper is that of determining theoretically the correlation functions of a binary alloy. The more familiar question of finding the long- and short-range order parameters of an alloy is a special case of this general problem. We confine ourselves to the assumption that the configurational energy of the alloy is adequately approximated by the following pair Hamiltonian

$$H=+\tfrac{1}{2}\sum_{i,j}\left[V_{ij}{}^{AA}\sigma_i{}^A\sigma_j{}^A+V_{ij}{}^{BB}\sigma_i{}^B\sigma_j{}^B+V_{ij}{}^{AB}(\sigma_i{}^A\sigma_j{}^B+\sigma_i{}^B\sigma_j{}^A)\right],\quad(1)$$

where i, j refer to the sites of a regular rigid lattice and $\sigma_i{}^A$, $\sigma_i{}^B$ are the occupation numbers for each site. If, in a particular state of the crystal, site i is occupied by an A atom, $\sigma_i{}^A=+1$, $\sigma_i{}^B=0$ and if by a B atom, $\sigma_i{}^A=0$, $\sigma_i{}^B=+1$. $V_{ij}{}^{AB}$ is the energy of interaction between an A atom on site i and a B atom on site j, and is assumed symmetric with respect to the interchange of site labels.

The correlation functions we seek are the thermodynamic averages $\langle\sigma_i{}^A\sigma_j{}^B\rangle$, $\langle\sigma_i{}^A\sigma_j{}^A\rangle$ etc., and are defined as follows:

$$\langle\sigma_i{}^A\sigma_j{}^B\rangle=\sum_\gamma\langle\gamma|\sigma_i{}^A\sigma_j{}^Be^{-\beta H}|\gamma\rangle/\sum_\gamma\langle\gamma|e^{-\beta H}|\gamma\rangle,\quad(2)$$

where $\beta=1/kT$ and k and T are the Boltzmann constant and temperature. The states of the system $\langle\gamma|$ correspond to all possible configurations of N_A A atoms and N_B B atoms arranged on N sites. The total number of different states is therefore $N!/N_A!N_B!$ with $N_A+N_B=N$.

The theory of binary alloys based on the Hamiltonian given above is mathematically equivalent in most respects to the Ising theory of magnetic systems so that many of the results obtained for the spin-spin correlation functions in magnetic systems can be translated into predictions for the alloy correlation functions. Since an exact solution for the Ising linear chain with magnetic field has been obtained for the case of nearest-neighbor interactions only, this constitutes an exact solution for a one-dimensional binary alloy of any composition with the same restriction on range of interaction. Similarly, Onsager's exact solution[1] of the two-dimensional Ising lattice with nearest-neighbor interactions in zero magnetic field is equivalent to a complete description of a two-dimensional binary alloy of 50–50 composition.

There are no complete solutions for three-dimensional lattices and so one must resort to either so-called cluster approximations which give a solution in closed form at all temperatures or else some form of high- or low-temperature expansion[2] which is valid only in a limited temperature range. Although these expansions have the satisfying feature of exactness (apart from truncation errors) they are very tedious to carry out for interactions which have a range greater than nearest-neighbor distances. Furthermore, a new expansion must be computed for each correlation function, for each lattice type, and for each composition. The only high-temperature expansions for more than first neighbor correlation functions that we are aware of are contained in the work of Oguchi.[3] Oguchi gives expansions for the correlation functions in the 100 direction of an AB alloy with a simple cubic lattice. It is clear that a single expression for the correlation functions would be extremely useful for experimental analysis if it were reasonably accurate. For this reason our development will be more akin to the cluster approach.

Work in this direction on the correlation functions has previously been done by Zernike[4] and Cowley.[5] The

---

* Supported by a grant from the U. S. Atomic Energy Commission.

[1] L. Onsager, Phys. Rev. **65**, 117 (1944).

[2] See for example C. Domb, Advan. Phys. **9**, 149 (1960).

[3] J. Phys. Soc. Japan **6**, 31 (1951).

[4] F. Zernike, Physica **7**, 565 (1940).

[5] J. M. Cowley, Phys. Rev. **77**, 669 (1950); **120**, 1648 (1960); **138**, A1384 (1965).

recent work of Christy and Hall[6] is closely related to Cowley's theory, and the theory of Elliott and Marshall[7] as it applies to alloy systems is directly related to Zernike's work. Although Zernike and Cowley appear to have approached the problem from quite different directions, it will be part of our purpose to show that their approximations are quite closely related.

The above fact will emerge in the mathematical development presented in Sec. III of a general expression for the correlation functions of a binary alloy with arbitrary composition and range of interaction. This expression is exact and may be solved approximately by replacing certain operators by their averages, the choice being guided to some extent by physical arguments. It will be shown that one choice leads to Cowley's theory and a second more complex choice to Zernike's theory. Also, it will be possible to judge qualitatively the conditions under which each of these two theories constitutes a reasonable approximation.

In Sec. IV a third approximation to the exact correlation function equation is proposed which leads to a new theory of correlations in alloy systems, and which appears to have the same order of accuracy as the approaches of Cowley and Zernike. Furthermore, this theory gives a linear set of equations which may be solved in closed form, unlike the equations of the two preceding theories. It has, however, the disadvantage that the theory is valid only above the ordering temperature.

Section II introduces the operators, the statistics, and the transformations necessary for Secs. III and IV.

In a succeeding paper[8] (Part II) it is intended to apply the theory outlined in Sec. IV of this paper to current experimental results for Cu_3Au and, possibly, CuZn.

II. MATHEMATICAL DEVELOPMENT

A. The Grand Canonical Ensemble

The Hamiltonian given in Eq. (1) can be rewritten in a more compact and useful form by introducing operators $(\bar{\sigma}_i)$ which are called spin-deviation operators in magnetism and Flinn operators[6,9] in alloy theory. They are defined here as

$$\bar{\sigma}_i \equiv 2(\sigma_i{}^A - m_A) \equiv 2(m_B - \sigma_i{}^B)\,, \tag{3}$$

where $m_A = N_A/N$ and $m_B = N_B/N$. These operators have the property that $\bar{\sigma}_i = 2m_B$ if there is an A atom at i and $\bar{\sigma}_i = -2m_A$ if a B atom is at site i.[10] Introduction of these operators in Eq. (1) yields

$$H = +\tfrac{1}{4} \sum_{i,j}^{N} \left[V_{ij}{}^0 + V_{ij}{}^1(\bar{\sigma}_i + \bar{\sigma}_j) + V_{ij}\bar{\sigma}_i\bar{\sigma}_j \right] \tag{4}$$

with

$$V_{ij}{}^0 = \tfrac{1}{2}(V_{ij}{}^{AA} + V_{ij}{}^{BB} + 2V_{ij}{}^{AB})\,,\ \tfrac{1}{2}V_{ij}{}^1 = (V_{ij}{}^{AA} - V_{ij}{}^{BB})$$

and

$$V_{ij} = \tfrac{1}{2}(V_{ij}{}^{AA} + V_{ij}{}^{BB} - 2V_{ij}{}^{AB})\,.$$

Since $\sum_{i=1}^{N} \bar{\sigma}_i$ has the same value for all states of the system (in this case equal to zero), the terms in (4) containing $V_{ij}{}^1$ can be dropped. The $V_{ij}{}^0$ term may be eliminated for the same reason and the resulting Hamiltonian (5) will give all the essentials of the problem.

$$H = +\tfrac{1}{4} \sum_{i,j}^{N} V_{ij}\bar{\sigma}_i\bar{\sigma}_j\,. \tag{5}$$

If V_{ij} is positive the system prefers unlike atoms on sites i and j, and like atoms if V_{ij} is negative.

Before we attempt to evaluate thermodynamic averages such as (2), the task will be made simpler by changing from a canonical ensemble to a grand canonical ensemble. This will mean changing the set of allowable states from all configurations of N_A A atoms and N_B B atoms on N sites to the larger set of all configurations in which each site has either an A atom or a B atom with probability m_A and m_B, respectively. An average in the grand canonical ensemble of some arbitrary function, $F(\bar{\sigma}_i)$, of the $\bar{\sigma}_i$'s, is defined as

$$\langle F(\bar{\sigma}_i)\rangle = \sum_{\bar{\sigma}_k=\pm} F(\bar{\sigma}_i)\cdot\omega(\bar{\sigma}_1)\omega(\bar{\sigma}_2)\cdots\omega(\bar{\sigma}_N)e^{-\beta H} \Big/ \sum_{\bar{\sigma}_k=\pm} \omega(\bar{\sigma}_i)\cdots\omega(\bar{\sigma}_N)e^{-\beta H}\,, \tag{6}$$

where $\sum_{\bar{\sigma}_k=\pm}$ implies summing each of the $\bar{\sigma}_k$'s over the two possible values $+2m_B$ and $-2m_A$, and the weighting factors $\omega(\bar{\sigma}_i)$ account for the fact that the + and − values should occur with the relative frequency m_A and m_B, respectively. Hence, $\omega(\bar{\sigma}_i = +2m_B) = m_A$ and $\omega(\bar{\sigma}_i = -2m_A) = m_B$. We may write $\omega(\bar{\sigma}_i)$ as $(m_A)^{\sigma_i{}^A}(m_B)^{\sigma_i{}^B}$ and

$$\prod_{i=1}^{N} \omega(\bar{\sigma}_i) = (m_A)^{\Sigma\sigma_i{}^A}(m_B)^{\Sigma\sigma_i{}^B} \tag{7a}$$

or

$$\prod_{i=1}^{N} \omega(\bar{\sigma}_i) = (m_A)^{m_A N}(m_B)^{m_B N}\left[\left(\frac{m_A}{m_B}\right)^{1/2}\right]^{\Sigma\sigma_i}. \tag{7b}$$

Equation (7b) follows from Eq. (3). Henceforth $\sum$ without summation indices will imply a summation over the N sites of the crystal. Finally, we may write Eq. (6) as

$$\langle F(\bar{\sigma}_i)\rangle = \sum_{\bar{\sigma}_k=\pm} F(\bar{\sigma}_i)e^{\lambda\Sigma\bar{\sigma}_i}e^{-\beta H} \Big/ \sum_{\bar{\sigma}_k=\pm} e^{\lambda\Sigma\bar{\sigma}_i}e^{-\beta H}\,, \tag{8}$$

where $e^{\lambda} \equiv (m_A/m_B)^{1/2}$ and the factor $e^{\lambda\Sigma\bar{\sigma}_i}$ is analogous to the term $e^{-\mu N}$ which often appears in grand canonical

---

[6] D. O. Christy and G. L. Hall, Phys. Rev. **132**, 1959 (1963).

[7] R. J. Elliott and W. Marshall, Rev. Mod. Phys. **30**, 75 (1958).

[8] S. C. Moss and P. C. Clapp, Phys. Rev. (to be published).

[9] P. A. Flinn, Phys. Rev. **104**, 350 (1956).

[10] Flinn's definition is $\bar{\sigma}_i = +m_B,\ -m_A$.

averages and is responsible for making the grand canonical and canonical averages equivalent in the limit of large N.

B. Correlations

We shall now be concerned with finding the correlation between the occupation of some site j and a second site O, which can be thought of as the origin of the lattice without loss of generality. This will be expressed in terms of statistical probabilities such as: the probability that site O is occupied by a B atom, given that site j is occupied by an A atom. This conditional probability is written $P_{Oj}{}^{BA}$ and the second subscript and superscript will always refer to the given condition. In general, $P_{Oj}{}^{BA}$ is not equal to $P_{jO}{}^{AB}$. $P_{Oj}{}^{BA}$ may be written exactly as

$$P_{Oj}{}^{BA}=\sum_{\bar{\sigma}_k=\pm}\sigma_O{}^B\sigma_j{}^A e^{-\beta H}e^{\lambda\Sigma\bar{\sigma}_i}\Big/\sum_{\bar{\sigma}_k=\pm}\sigma_j{}^A e^{-\beta H}e^{\lambda\Sigma\bar{\sigma}_i}. \tag{9}$$

$\sigma_j{}^A$ is, in effect, a projection operator that allows only those states for which j is occupied by an A atom to count in the summation. Similarly, $\sigma_O{}^B$ selects those states having a B atom on O.

We may write expressions analogous to Eq. (9) for each of the other three binary probabilities $P_{Oj}{}^{AA}$, $P_{Oj}{}^{BB}$, and $P_{Oj}{}^{AB}$. It is apparent that these four probabilities are not independent since the following relations must hold:

$$P_{Oj}{}^{BA}+P_{Oj}{}^{AA}=1 \tag{10a}$$

$$P_{Oj}{}^{AB}+P_{Oj}{}^{BB}=1 \tag{10b}$$

$$m_BP_{Oj}{}^{AB}=m_AP_{Oj}{}^{BA}. \tag{10c}$$

Equation (10c) follows from the fact that $m_BP_{Oj}{}^{AB}$ is the probability of the simultaneous occurrence of an A on O and a B on j, and $m_AP_{Oj}{}^{BA}$ is the probability of B on O and A on j. Thus these weighted probabilities differ only in an interchange of indices which, because of the assumed symmetry of the crystal, cannot affect their value.

The three relations among four unknowns allow us to write all four probabilities in terms of one arbitrary parameter, which is customarily taken to be the Warren short-range order parameter, α_{Oj}.

The four probabilities in terms of α_{Oj} are given by Cowley[5] as

$$P_{Oj}{}^{AA}=m_A+m_B\alpha_{Oj}, \tag{11a}$$

$$P_{Oj}{}^{BA}=m_B-m_B\alpha_{Oj}, \tag{11b}$$

$$P_{Oj}{}^{AB}=m_A-m_A\alpha_{Oj}, \tag{11c}$$

$$P_{Oj}{}^{BB}=m_B+m_A\alpha_{Oj}, \tag{11d}$$

and by a simple computation, it may be discovered that

$$\langle\bar{\sigma}_O\bar{\sigma}_j\rangle=4m_Am_B\alpha_{Oj}. \tag{12}$$

Since only α_{Oj} is experimentally measurable it will be important to have our final results in terms of α_{Oj} alone, although it will turn out to be simpler to calculate each of the binary probabilities separately.

Equation (9) may be recast in a form which will allow us to obtain solutions in varying degrees of approximation. This is done by replacing $\sigma_O{}^B$ by its conditional average (i.e., the average of $\sigma_O{}^B$ when all other sites in the system are fixed and $\bar{\sigma}_O$ alone is allowed to vary). This may be seen to be

$$\langle\sigma_O{}^B\rangle_{\text{cond}}\equiv\frac{\displaystyle\sum_{\bar{\sigma}_O=\pm}\sigma_O{}^B\exp\left[\left(-\frac{\beta}{2}\sum_f V_{Of}\bar{\sigma}_f\right)\bar{\sigma}_O\right]e^{\lambda\bar{\sigma}_O}}{\displaystyle\sum_{\bar{\sigma}_O=\pm}\exp\left[\left(-\frac{\beta}{2}\sum_f V_{Of}\bar{\sigma}_f\right)\bar{\sigma}_O\right]e^{\lambda\bar{\sigma}_O}} \tag{13}$$

and performing the sum gives

$$\langle\sigma_O{}^B\rangle_{\text{cond}}=\frac{e^{-2m_A(\beta E_O+\lambda)}}{e^{-2m_A(\beta E_O+\lambda)}+e^{+2m_B(\beta E_O+\lambda)}}\equiv f_B(E_O), \tag{14}$$

where $E_O\equiv-\frac{1}{2}\sum_f V_{Of}\bar{\sigma}_f$ and may be thought of as the field which the atom at the origin sees as a result of some particular configuration of atoms around O (specified by the $\bar{\sigma}_f$'s).

We shall also need $\langle\sigma_O{}^A\rangle_{\text{cond}}$ and it is calculated to be

$$\langle\sigma_O{}^A\rangle_{\text{cond}}=\frac{e^{+2m_B(\beta E_O+\lambda)}}{e^{-2m_A(\beta E_O+\lambda)}+e^{+2m_B(\beta E_O+\lambda)}}\equiv f_A(E_O). \tag{15}$$

Using a theorem of statistics (which will be proven in the Appendix for our particular case), $\sigma_O{}^B$ may be replaced in Eq. (9) by $\langle\sigma_O{}^B\rangle_{\text{cond}}$ without altering the exactness of the relation. Therefore, we have

$$P_{Oj}{}^{BA}=\sum_{\bar{\sigma}_k=\pm}f_B(E_O)\sigma_j{}^A\rho\Big/\sum_{\bar{\sigma}_k=\pm}\sigma_j{}^A\rho, \tag{16}$$

where $\rho\equiv e^{-\beta H}e^{\lambda\Sigma\bar{\sigma}_i}$, the grand canonical density operator.

It is important to realize what has been accomplished by this transformation. The correlation between site j and the origin has been written in terms of the correlation between j and the shells of sites around O which have a direct interaction (V_{Of}) with O, and the operator $\bar{\sigma}_O$ no longer appears in the expression. In essence, the effect of an A atom at site j on the occupancy of the origin is now represented in terms of the effect it has on the shells of sites around O, which, in turn, pass this information on to O via their direct interactions with

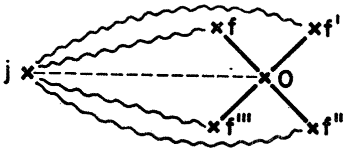

FIG. 1. The basic cluster for calculating the correlation between site j and the origin in terms of the correlations between site j and the neighbors of the origin.

the origin site. This is represented diagrammatically in Fig. 1 where the correlation between j and the origin (shown by a dashed line) has been replaced by the correlations between j and sites f, f', etc. (shown by wavy lines) and the direct interactions V_{Of} (indicated by solid lines) which these sites have with the origin site.

III. APPROXIMATE EQUATIONS FOR THE CORRELATION FUNCTIONS

In this section two methods will be developed for obtaining approximate solutions to the exact binary probability equation [Eq. (16)] which was given in Sec. II.

The first method involves the use of a frequency distribution function for E_O and by making a simple assumption about the behavior of this distribution function it is shown that the general equations previously derived by Cowley[5] result. Furthermore, if a somewhat more complex approximation is used for the distribution function, the theory of Zernike[4] is recovered. It then becomes possible to compare directly the basic approximations of these two theories and to determine under what conditions each should be a good approximation.

The second method for solving Eq. (16) approximately uses an exact expansion of the distribution function in terms of its moments. The theories of Zernike and Cowley are shown to be equivalent to certain approximations for the successive moments of this expansion. In addition, this expansion suggests a third approximation which would appear to have the same order of accuracy as the two other theories at temperatures above the ordering temperature, but has the advantage of yielding a set of linear equations for the correlation functions which are amenable to exact solution. This theory for correlation functions in alloy systems is discussed in detail in Sec. IV.

A. The Distribution Function

Returning to Eq. (16) we may write the right-hand side in the form of a sum over the set of discrete values $\{E_O\}$ which E_O is able to assume:

$$P_{Oj}{}^{BA}=\sum_{\{E_O\}} g_{j=A}(E_O)f_B(E_O), \tag{17}$$

where

$$g_{j=A}(E_O)=\sum_{\bar{\sigma}_k=\pm}\delta(E_O+\tfrac{1}{2}\sum_f V_{Of}\bar{\sigma}_f)\sigma_j{}^A\rho/\sum_{\bar{\sigma}_k=\pm}\sigma_j{}^A\rho$$

and $\delta(E_O+\frac{1}{2}\sum_f V_{Of}\bar{\sigma}_f)=1$ if $E_O=-\frac{1}{2}\sum_f V_{Of}\bar{\sigma}_f$, and $=0$ otherwise.

$g_{j=A}(E_O)$ is simply the frequency distribution of the random variable E_O under the condition that site j is occupied by an A atom, and satisfies the normalization requirement of $\sum_{\{E_O\}} g_{j=A}(E_O)=1$. Complete knowledge of this frequency distribution would be tantamount to an exact solution for the correlation functions. This, in three dimensions, is one of the unsolved problems of lattice statistics and so we content ourselves with different possible approximations to $g_{j=A}(E_O)$.

One fact that can be immediately determined about $g_{j=A}(E_O)$ is the mean of $E_O(\langle E_O\rangle_{j=A})$ with respect to $g_{j=A}(E_O)$:

$$\begin{aligned}\langle E_O\rangle_{j=A}&=\sum_{\{E_O\}} E_O g_{j=A}(E_O)\equiv-\tfrac{1}{2}\sum_f V_{Of}\langle\bar{\sigma}_f\rangle_{j=A}\\&=-\tfrac{1}{2}\sum_f V_{Of}(+2m_BP_{fj}{}^{AA}-2m_AP_{fj}{}^{BA})\\&=-m_B\sum_f V_{Of}\alpha_{fj},\end{aligned} \tag{18}$$

where the last equality follows from Eq. (11). Similarly the mean with the complementary distribution function $g_{j=B}(E_O)$ is found to be $\langle E_O\rangle_{j=B}=+m_A\sum_f V_{Of}\alpha_{fj}$. Our treatment to this point has been exact and we now introduce an approximation which will yield a closed form solution to the problem and will be mathematically equivalent to Cowley's theory.

The approximation is

$$\langle(E_O)^n\rangle_{j=A}\equiv\sum_{\{E_O\}}(E_O)^n g_{j=A}(E_O)\cong(\langle E_O\rangle_{j=A})^n \tag{19a}$$

and

$$\langle(E_O)^n\rangle_{j=B}\equiv\sum_{\{E_O\}}(E_O)^n g_{j=B}(E_O)\cong(\langle E_O\rangle_{j=B})^n. \tag{19b}$$

Since a knowledge of all the moments of a distribution function is sufficient to determine the distribution function it can be shown that the form of $g(E_O)$ which satisfies Eq. (19) is a delta function, i.e.,

$$g_{j=A}(E_O)=\delta(E_O-\langle E_O\rangle_{j=A}) \tag{20a}$$

$$g_{j=B}(E_O)=\delta(E_o-\langle E_O\rangle_{j=B}). \tag{20b}$$

Equation (17) is now soluble and we obtain

$$P_{Oj}{}^{AA}=f_A(\langle E_O\rangle_{j=A}), \tag{21a}$$

$$P_{Oj}{}^{AB}=f_A(\langle E_O\rangle_{j=B}), \tag{21b}$$

$$P_{Oj}{}^{BA}=f_B(\langle E_O\rangle_{j=A}), \tag{21c}$$

$$P_{Oj}{}^{BB}=f_B(\langle E_O\rangle_{j=B}). \tag{21d}$$

Remembering the functional forms of f_A and f_B from Eqs. (14) and (15) and inserting the values of $\langle E_O\rangle_{j=A}$ and $\langle E_O\rangle_{j=B}$ from Eq. (18), the results of the calculation can be put in the following convenient combination

$$\frac{P_{Oj}{}^{AA}P_{Oj}{}^{BB}}{P_{Oj}{}^{AB}P_{Oj}{}^{BA}}\equiv\frac{(m_A+m_B\alpha_{Oj})(m_B+m_A\alpha_{Oj})}{(1-\alpha_{Oj})^2m_Am_B}=\exp(-2\beta\sum_f V_{Of}\alpha_{fj}), \tag{22}$$

which is the result of Cowley[5] with the difference that the site indices O, j are interchanged. Cowley arrived at Eq. (22) by arguments which, on the surface, appear to be quite different from those used here although the two arrive at the same mathematical endpoint.

We shall now use the present derivation of Cowley's expression to determine under what conditions Eq. (22) should give a good fit to experimental results. In physical terms we have approached the problem of calculating the correlation between the origin and j by fixing the occupancy of site j (to be an A atom, say) and trying to ascertain how this affects the occupancy of the origin site. In terms of the operator $\bar{\sigma}_O$, before the occupancy of site j was fixed, the average of $\bar{\sigma}_O$ was zero, but now this average will be shifted slightly (by an amount $\langle\sigma_O\rangle_{j=A}=2m_B\alpha_{Oj}$). We have shown that we are able to calculate the shift in $\langle\bar{\sigma}_O\rangle$ by another route (see Fig. 1) since $\bar{\sigma}_O$ is directly affected only by the sites f with which it has an interaction V_{Of} so that the effect of j on O must be passed through these sites f. Consequently, fixing an A atom at j can be imagined to "polarize" slightly the sites around O which then causes a shift in $\langle\bar{\sigma}_O\rangle$ from its random value. This second way of calculating $\langle\bar{\sigma}_O\rangle_{j=A}$ was expressed by Eq. (16). The condition that the value of $\langle\bar{\sigma}_O\rangle_{j=A}$ be the same in both cases gives the set of equations for the α_{Oj}'s. The approximation of replacing $g(E_O)$ by a delta function means in physical terms that we have assumed that the origin "sees" its neighbors frozen in their *average* polarized configuration and does not follow their fluctuations about this average. There are two cases in which this makes good sense. The first is at temperatures well below an ordering temperature in which we know that the atoms really are frozen into a few highly ordered states. In the case of an AB alloy in a simple cubic lattice two states predominate, A on α sites with B on β sites or A on β sites with B on α sites. Choosing j to be an A site resolves this twofold degeneracy and makes the occupancy of the neighbors of O determinate. Consequently site O will see on the average essentially one configuration of its neighbors with decreasing fluctuations about that configuration as the temperature decreases.

The second case in which the distribution function can be well represented by a delta function is at temperatures above the ordering temperature when a large number of sites have an interaction with the origin. This will be the situation when V_{Of} is a long-range interaction and the dimensionality of the lattice is high. This may be seen by regarding the operators $\bar{\sigma}_f$ in $E_O=-\frac{1}{2}\sum_f V_{Of}\bar{\sigma}_f$ as random variables capable of taking each of two values with a certain probability. At temperatures well above the ordering temperature these variables will be only weakly correlated and it is known from the theory of independent random variables that a distribution function of a sum of such variables peaks more and more sharply about $\langle E_O\rangle$ as the number of variables increases and reaches a delta function form in the limit of an indefinitely large number of variables. Consequently, we suggest that Cowley's expression will be worst for lattices of low coordination numbers and short-range interactions and best for the converse.

In view of the foregoing, one of the chief criticisms that can be made of Cowley's approximation is that it takes no account of the fluctuations in E_O about its average value, which can be expected to cause appreciable effects near the ordering temperature. It is natural, therefore, to try to incorporate these fluctuations into the theory and give a finite width to $g(E_O)$. Zernike's theory[4] and the closely related theory of Elliott and Marshall[7] do just this by using a different approximation to $g(E_O)$ which we will now discuss.

Zernike makes the assumption that the operators occurring in E_O may be regarded as stochastically independent of each other so that if the relative probability of the two values which each operator is able to assume were known the probability of any configuration of the neighbors would be given by a simple product of the individual probabilities. In this way he is able to obtain a first approximation for the fluctuations in E_O about its average value and arrives at a distribution function with a finite width.

For the case of an AB alloy with nearest-neighbor interactions (which is the only case given by Zernike), any operator $\bar{\sigma}_k$ can take on the values $+1$ or -1. Let us further assume we are dealing with a square lattice (as in Fig. 1) and the neighbors f, f', f'', and f''' of O are labeled from 1 to 4. If there were no constraints $\bar{\sigma}_f$ would assume the values $+1$ and -1 with equal probability, but if the constraint that j be occupied by an A atom (i.e., $\bar{\sigma}_j=+1$) is introduced then these probabilities are no longer equal. From the definitions of Eq. (11) we have

$$P_{fj}{}^{AA}\equiv P(\bar{\sigma}_f=+1\,|\,\bar{\sigma}_j=+1)=\tfrac{1}{2}(1+\alpha_{fj}), \quad \text{(23a)}$$

$$P_{fj}{}^{BA}\equiv P(\bar{\sigma}_f=-1\,|\,\bar{\sigma}_j=+1)=\tfrac{1}{2}(1-\alpha_{fj}), \quad \text{(23b)}$$

which may be condensed to

$$P(\bar{\sigma}_f\,|\,\bar{\sigma}_j=+1)=\tfrac{1}{2}(1+\bar{\sigma}_f\alpha_{fj}), \quad \text{(23c)}$$

and similarly

$$P(\bar{\sigma}_f\,|\,\bar{\sigma}_j=-1)=\tfrac{1}{2}(1-\bar{\sigma}_f\alpha_{fj}). \quad \text{(23d)}$$

Using Zernike's assumption that each neighbor f of the origin is independently correlated with j, one may write that the probability of any of the 2^4 configurations of the neighbors of O is

$$P(\bar{\sigma}_1\bar{\sigma}_2\bar{\sigma}_3\bar{\sigma}_4\,|\,\bar{\sigma}_j=\pm1)=\frac{1}{2^4}\prod_{f=1,4}(1\pm\bar{\sigma}_f\alpha_{fj}). \quad \text{(24)}$$

Equation (24) provides the necessary information to calculate the frequency distribution

$$g_{j=A}(E_O=-\tfrac{1}{2}\textstyle\sum_f V_{Of}\bar{\sigma}_f)$$

in terms of the α_{fj}'s and it is apparent from Eq. (24) that since products of the α_{fj}'s occur, $g_{j=A}(E_O)$ will be nonlinear in the α_{fj}'s. We have drawn the distribution functions $g_{j=A}(E_O)$ and $g_{j=B}(E_O)$ obtained from the Cowley and Zernike expressions in Fig. 2 for purposes of comparison. α_{fj} has been taken as 0.1 independent of f for the sake of illustration although it is unrealistic that the α_{fj}'s would be equal at unequal distances. The last step is to use Eq. (17) to obtain Zernike's equation relating α_{Oj} to the α_{fj}'s and V_{Of} by performing the sum over E_O. The final result is

$$\begin{aligned}\alpha_{Oj}=&\frac{1}{2^4}(t_4+2t_2-2t_{-2}-t_{-4})(\alpha_{1j}+\alpha_{2j}+\alpha_{3j}+\alpha_{4j})\\&+\frac{1}{2^4}(t_4-2t_2+2t_{-2}-t_{-4})\\&\quad\times(\alpha_{1j}\alpha_{2j}\alpha_{3j}+\alpha_{1j}\alpha_{2j}\alpha_{4j}+\alpha_{1j}\alpha_{3j}\alpha_{4j}+\alpha_{2j}\alpha_{3j}\alpha_{4j}),\end{aligned}\tag{25}$$

where $t_n=\tanh(-n\beta V/2)$ and V is the common value of the nearest-neighbor interaction (i.e., $V=V_{O1}=V_{O2}$, etc.).

Since j in Eq. (25) may be taken as any of the N sites of the lattice (except the origin), Eq. (25) is, in fact, $N-1$ nonlinear equations in $N-1$ unknowns which must be solved by some further approximation. Zernike treats this problem in detail for the cubic simple, body-centered and face-centered lattices in his original paper.[4]

B. Moment Expansion for the Correlation Function

We will now give an expansion for the correlation functions which will put the theories of Cowley and Zernike in clearer perspective with respect to each other and will indicate where possible improvements can be made in the calculations.

Equation (17) can be written exactly in terms of the successive moments of the distribution function by using a Maclaurin expansion for $f_B(E_O)$ in the variable βE_O as follows:

$$\begin{aligned}P_{Oj}{}^{BA}&=\sum_{\{E_O\}}g_{j=A}(E_O)\left[f_B(0)+\beta E_Of_B'(0)+\frac{\beta^2E_O{}^2}{2!}f_B''(0)+\cdots\right]\\&=f_B(0)+\beta f_B'(0)\langle E_O\rangle_{j=A}+\frac{\beta^2}{2!}f_B''(0)\langle E_O{}^2\rangle_{j=A}+\cdots+\frac{\beta^n}{n!}f_B{}^{(n)}(0)\langle E_O{}^n\rangle_{j=A}+\cdots,\end{aligned}\tag{26}$$

where $f_B{}^{(n)}(0)$ is $\partial^nf_B(E_0)/\partial^n(\beta E_O)$ evaluated at $\beta E_O=0$.

We may write analogous expansions for the other binary probabilities and by grouping terms obtain:

$$\begin{aligned}2\alpha_{Oj}=&P_{Oj}{}^{AA}+P_{Oj}{}^{BB}-P_{Oj}{}^{AB}-P_{Oj}{}^{BA}\\=&\beta[f_A'(0)-f_B'(0)][\langle E_O\rangle_{j=A}-\langle E_O\rangle_{j=B}]+\frac{\beta^2}{2!}[f_A''(0)-f_B''(0)][\langle E_O{}^2\rangle_{j=A}-\langle E_O{}^2\rangle_{j=B}]+\cdots\\&+\frac{\beta^n}{n!}[f_A{}^{(n)}(0)-f_B{}^{(n)}(0)][\langle(E_O)^n\rangle_{j=A}-\langle(E_O)^n\rangle_{j=B}]+\cdots.\end{aligned}\tag{27}$$

From Eqs. (14) and (15) we have

$$\begin{aligned}f_A(E_O)-f_B(E_O)&=\frac{e^{2m_B(\beta E_O+\lambda)}-e^{-2m_A(\beta E_O+\lambda)}}{e^{2m_B(\beta E_O+\lambda)}+e^{-2m_A(\beta E_O+\lambda)}}\\&=\tanh(\beta E_O+\lambda)\end{aligned}\tag{28}$$

and so we may write Eq. (27) as

$$2\alpha_{Oj}=\sum_{n=1}^{\infty}\frac{\beta^n}{n!}T_n(\lambda)[\langle E_O{}^n\rangle_{j=A}-\langle E_O{}^n\rangle_{j=B}],\tag{29}$$

where we define

$$T_n(\lambda)\equiv\left.\frac{\partial^n\tanh(X+\lambda)}{\partial X^n}\right|_{X=0}.$$

This again is an exact expression for the correlation

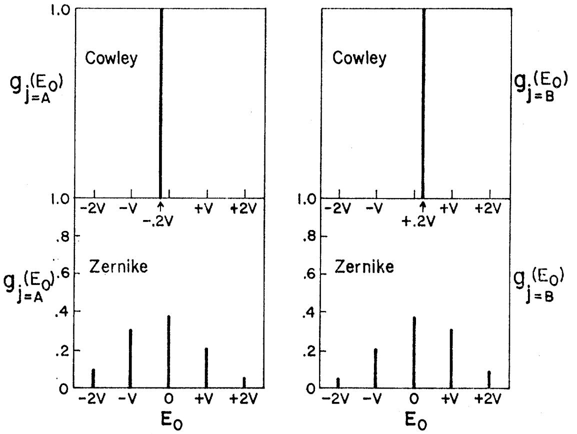

FIG. 2. Probability distribution of the field at the origin site as predicted by the Zernike and Cowley theories for the two cases of an A atom or a B atom at site j. The correlation between j and the neighbors of the origin has been taken as 0.1.

functions, the only unknowns being the moments $\langle E_O{}^n\rangle_{j=A,B}$ and we can gain some understanding of the theories of Zernike and Cowley by comparing their moment expansion with the exact series. First, it should be realized that since $\langle E_O\rangle_{j=A,B}$ is a sum of two particle correlations, the higher moments $\langle E_O{}^n\rangle_{j=A,B}$ are, in general, combinations of $n+1$ particle correlations. In order to obtain a closed form solution to the problem these higher moments which occur on the right-hand side of Eq. (29) must be broken up in some approximate way into products of two particle correlations. We have already seen that Cowley accomplished this by the breakup scheme $\langle E_O{}^n\rangle_{j=A,B}\cong\langle E_O\rangle^n{}_{j=A,B}$. Zernike, on the other hand, allows the individual operators in Eq. (29) to fluctuate about their mean value but assumes they move independently of each other. This leads to the following type of decomposition which is equivalent to Zernike's approximation.

$$\begin{aligned}\langle E_O{}^n\rangle_{j=A,B}&=\langle(-\tfrac{1}{2}\sum_f^{1,k} V_{Of}\bar{\sigma}_f)^n\rangle_{j=A,B}\\&=\sum_{p_1\cdots p_k}\frac{n!(-\frac{1}{2})^n}{p_1!p_2!\cdots p_k!}\langle(V_{O1}\bar{\sigma}_1)^{p_1}\cdots(V_{Ok}\bar{\sigma}_k)^{p_k}\rangle_{j=A,B}\\&\cong\sum_{p_1\cdots p_k}\frac{n!(-\frac{1}{2})^n}{p_1!\cdots p_k!}\langle(V_{O1}\bar{\sigma}_1)^{p_1}\rangle_{j=A,B}\cdots\langle(V_{Ok}\bar{\sigma}_k)^{p_k}\rangle_{j=A,B}\,,\end{aligned}\tag{30}$$

where the sum $\sum_{p_1\cdots p_k}$ is taken over all combinations of the p_i's which satisfy the condition

$$\sum_{i=1}^{k} p_i=n\,.$$

k is the number of sites which directly interact with O. This is the mathematical statement that the correlations with j of each of the f sites surrounding O are to be retained and the correlations between sites f and f' are to be discarded, which is Zernike's assumption.

It may be seen that both Zernike and Cowley give the first term in Eq. (29) correctly but use different approximations for the second and higher terms. In breaking up a multiparticle correlation function $\langle\bar{\sigma}_1\cdots\bar{\sigma}_p\rangle$, it is reasonable to expect that the best approximation would be obtained by retaining the strongest pair correlations. That is, if $\bar{\sigma}_1$ is strongly coupled to $\bar{\sigma}_2$, but not to the rest, and $\bar{\sigma}_3$ is strongly coupled to $\bar{\sigma}_4$, but not to the rest, etc., then one could guess that a sensible decomposition of the multiparticle correlation would be $\langle\bar{\sigma}_1\bar{\sigma}_2\rangle\langle\bar{\sigma}_3\bar{\sigma}_4\rangle\cdots\langle\bar{\sigma}_{p-1}\bar{\sigma}_p\rangle$. Let us then consider the multiparticle correlation which occurs in the second term of Eq. (29), $\langle(\sum_f V_{Of}\bar{\sigma}_f)^2\rangle_{j=A,B}$. A particular operator $\bar{\sigma}_f$ will be most strongly correlated with itself and next most strongly with operators $\bar{\sigma}_{f'}$ of those sites which are closest to f. Cowley ignores all these correlations and effectively writes $\langle(\bar{\sigma}_f)^2\rangle_{j=A,B}\cong\langle\bar{\sigma}_f\rangle_{j=A,B}\langle\bar{\sigma}_f\rangle_{j=A,B}$ and $\langle\bar{\sigma}_f\bar{\sigma}_{f'}\rangle\cong\langle\bar{\sigma}_f\rangle_{j=A,B}\langle\bar{\sigma}_{f'}\rangle_{j=A,B}$. Zernike also ignores the second type of correlation (i.e., $\langle\bar{\sigma}_f\bar{\sigma}_{f'}\rangle\rightarrow\langle\bar{\sigma}_f\rangle\langle\bar{\sigma}_{f'}\rangle$) but does keep the self-correlation so that $\langle(\bar{\sigma}_f)^2\rangle$ is retained intact. Similar comparisons can be made between these two theories for higher moments as well.

We may now draw some qualitative conclusions about the usefulness of Zernike's and Cowley's approximations and the possibilities for improved closed form calculations. At first glance it would appear that since Zernike retains the self-correlation and so is able to include a rough estimate of the fluctuations of the distribution function $g(E_O)$ about its mean value, it should be superior to Cowley's theory in all cases. This conclusion is true when the number of operators in E_O is small, the most striking case being the linear chain with nearest-neighbor interactions. For this problem, Zernike's result is very close to the exact expression for the correlation functions and Zernike correctly predicts no cooperative transition. Cowley's theory, on the other hand, is significantly poorer and incorrectly predicts a finite-ordering temperature. This defect may be traced directly to the fact that fluctuations have been ignored. However, as the dimension of the lattice increases, Ashkin and Lamb[11] have pointed out that Zernike's result becomes steadily poorer. This can be understood from the fact that the neighbors of the central site O are becoming more and more correlated with each other via the increased number of coupling paths present in lattices of higher coordination. Increasing the range of interaction has the same effect, and this undoubtedly places Zernike's independent neighbor calculation of $g(E_O)$ in increasingly greater error. On the other hand, the increased correlation between neighbors and the inclusion of more neighbors with a longer range of interaction tends to narrow $g(E_O)$ and so Cowley's estimate of $g(E_O)$ becomes more accurate.

In addition, there is the pragmatic feature that Zernike's calculation becomes extremely difficult to carry out as the coordination number increases and to our knowledge has never been attempted for more than nearest-neighbor interactions. This places it in the same

---

[11] J. Ashkin and W. E. Lamb, Phys. Rev. 64, 159 (1943).

area of problems which it is feasible to treat with exact high- and low-temperature expansions. If one could give a good fit at the ordering temperature, it would still have an advantage over exact expansions, but it is clear that this is just the region in which the independent neighbor assumption will be worst since short-range order builds up quite rapidly in the neighborhood of the transition point.

Cowley's theory is best, however, for those problems for which exact expansions become increasingly tedious, if not impossible to do with any accuracy (i.e., increasing coordination of the lattice and increasing range of interaction). In addition, the mathematical difficulty in solving Cowley's infinite set of nonlinear equations to some approximation does not increase appreciably for these more complex cases, so that it is likely that such a theory will continue to remain useful for some time. Fortunately, it is possible to find a set of simultaneous linear equations for the correlation functions which have the same order of accuracy as Cowley's nonlinear set *above* the ordering temperature and are capable of being solved exactly for an arbitrary range of interaction and arbitrary coordination.

IV. LINEAR EQUATIONS FOR THE CORRELATION FUNCTIONS ABOVE THE ORDERING TEMPERATURE

Both Zernike and Cowley obtain an infinite set of coupled nonlinear equations in the α_{Oj}'s which must be solved in some approximate way to obtain values for the separate α_{Oj}'s. It will be the purpose of this section to show that at temperatures above the ordering temperature a much more practical set of linear equations for the α_{Oj}'s may be used which have the same degree of accuracy as either of the nonlinear forms of Zernike and Cowley. Moreover, this linear set of equations may be solved exactly for each of the α_{Oj}'s and for any range of interaction. The solution is such that the pairwise interaction energy can be derived as a function of interatomic distance directly from the experimental data.

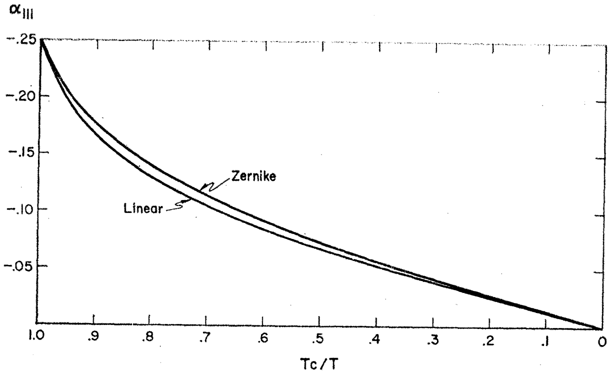

FIG. 3. First-neighbor order parameters as a function of temperature above the transition point, computed by the linear approximation and by Zernike's approximation for an AB alloy on a body-centered cubic lattice (Cu-Zn). Only nearest-neighbor interactions are assumed to be nonzero.

Returning to the exact expansion for α_{Oj} [Eq. (29)], we have already seen that Cowley and Zernike have accounted for the first term on the right-hand side exactly but their expressions correspond to some approximation for the higher terms. Cowley's approximation for these higher terms is probably quite bad since the dominant contributions in these higher order correlations can be expected to come from self-correlations of site f, and correlations of site f with sites near to f. These terms are ignored by Cowley's theory and only the correlation between f and j (which is usually distant from f) is retained. However, Cowley's theory does give quite good results for the α_{Oj}'s when compared with experiment[12,13] from which it can probably be concluded that the higher order terms in the expansion are negligible at most temperatures above the ordering temperature. Consequently, the particular approximation used for these higher terms will likely have little effect on the final result. It is then reasonable to suggest that the higher order terms be entirely neglected and Eq. (29) be approximated as follows

$$2\alpha_{Oj}=\beta T_1(\lambda)[\langle E_O\rangle_{j=A}-\langle E_O\rangle_{j=B}]. \tag{31}$$

Using Eq. (18) and the definition of $\lambda\equiv\frac{1}{2}\ln(m_A/m_B)$ gives

$$\alpha_{Oj}=-2m_Am_B\beta\sum_f V_{Of}\alpha_{fj}. \tag{32}$$

A particular case of this general formula has previously been used by Walker and Keating[14] and Paskin[12] in a calculation for β Cu–Zn, and Clapp[15] has indicated the justification of this linear approximation for A-B alloys generally, using arguments similar to those presented here. Walker and Keating found that a calculation of the nearest-neighbor correlation in β Cu–Zn at 10% above T_c gave -0.182, -0.180, and -0.171 using the Elliott-Marshall theory, the Zernike theory, and the linear approximation, respectively.

For a more detailed quantitative comparison of the linear approximation, Fig. 3 gives the nearest-neighbor correlation in β Cu–Zn above the ordering temperature as calculated by the Zernike and linear formulas. It may be noted that the two calculations differ at most by about 8% at $T_c/T\simeq 0.8$. They are in perfect agreement at $T=T_c$ and have the same asymptotic behavior as $T\rightarrow\infty$.

As a test of an A_3B alloy we have compared the linear theory with the Monte Carlo calculations of Fosdick[16] for two choices of the second-neighbor interaction in Fig. 4. The Monte Carlo calculations were performed on a lattice containing $5\times5\times5$ unit cells and periodic

---

[12] A. Paskin, Phys. Rev. **134**, A246 (1964).

[13] S. C. Moss, J. Appl. Phys. **35**, 3547 (1964).

[14] C. B. Walker and D. T. Keating, Phys. Rev. **130**, 1726 (1963).

[15] P. C. Clapp, Phys. Letters **13**, 305 (1964).

[16] L. D. Fosdick, Phys. Rev. **116**, 565 (1959).

boundary conditions were assumed. At temperatures well above the critical point, Fosdick's results are probably quite close to the exact correlations of an infinite lattice, but at temperatures near the critical point (say $T_c/T > 0.9$) his values will be a much cruder approximation because of the long-range cooperative effects which are now becoming important in an infinite lattice. Since this is also the region of greatest inaccuracy for the linear theory it is a moot point as to which calculation is closer to the truth here. For $T_c/T < 0.9$, the two calculations do not differ seriously.

We have not made a quantitative comparison between the Cowley and linear approximations because of the inherent difficulty in solving Cowley's nonlinear equations to sufficient accuracy to make a close numerical comparison meaningful.

The set of equations (32) can be diagonalized by introducing the Fourier transforms of α_{Oj} and V_{Of}, i.e.,

$$\alpha_{Oj} = \frac{1}{v_k}\int d^3k\, \alpha(k) e^{-i2\pi \mathbf{k}\cdot\mathbf{r}_j} \tag{33a}$$

$$V_{Of} = \frac{1}{v_k}\int d^3k\, V(k) e^{-i2\pi \mathbf{k}\cdot\mathbf{r}_f}, \tag{33b}$$

where $\mathbf{r}_j$ is the vector from the origin to site j.

The integration is taken over one unit cell of the reciprocal lattice, the volume of which is v_k. Inserting (33) in (32) yields equations in terms of $\alpha(k)$ and $V(k)$ which give the following solution for any value of k:

$$\alpha(k) = \frac{C}{1 + 2m_A m_B \beta V(k)}, \tag{34a}$$

$$C = \left[\frac{1}{v_k}\int d^3k \frac{1}{1 + 2m_A m_B \beta V(k)}\right]^{-1}, \tag{34b}$$

where C has been chosen to satisfy the condition

$$\alpha_{OO} = \frac{1}{v_k}\int d^3k\, \alpha(k) = 1.$$

Finally, one has as the general solution for α_{Oj}:

$$\alpha_{Oj} = \frac{C}{v_k}\int d^3k \left(\frac{e^{-i2\pi \mathbf{k}\cdot\mathbf{r}_j}}{1 + 2m_A m_B \beta V(k)}\right). \tag{35}$$

Equations (34) and (35) can be used for experimental analyses in a number of ways. The most direct way in principle, since $\alpha(k)$ (which is actually the appropriately reduced x-ray intensity) rather than α_{Oj} is the quantity directly measured in the x-ray diffuse scattering by the sample (see Moss[13]), would be to insert the experimental values of $\alpha(k)$ in Eq. (34) and thus calculate $V(k)$. The nature of $V(r)$ would then be obtained by Fourier inversion. This inversion, however, requires a complete map of $\alpha(k)$ in k space which, in general, is not possible since strong Bragg reflections completely obscure the x-ray diffuse scattering in certain regions of k space. To make matters worse these regions usually occur where $\alpha(k)$ has its minima and hence where the maximum values of $V(k)$ would lie [v.i. Eq. (34)].

One can proceed in another direction by making some specific guess as to the form of $V(r)$, either assuming that $V(r)$ is essentially short range and has a finite value only for the first few neighbors or assuming a particular long-range form such as the $V(r)$ proposed by Harrison and Paskin[12] for Cu–Zn. The undetermined constants and the validity of the assumed form of $V(r)$ are then determined by fitting Eq. (34) to the data.

Some preliminary restrictions on the form of $V(r)$ can be deduced from the observed distribution of $\alpha(k)$ because of the relation between the symmetry of $V(k)$ and $\alpha(k)$ implied by Eq. (34). This restriction gives a minimum range which $V(r)$ must have to account for the shape and symmetry of $\alpha(k)$ and it is possible to conclude for instance that current data on Cu_3Au[13] imply an interaction at least as far out as third nearest neighbors, so that the agreement which Moss obtained with the Cowley theory using V_1 and V_2 alone represented something more of a forced fit than had originally been thought. This is not a reflection on the accuracy or inaccuracy of the Cowley theory but means that V_3 is not negligible in Cu_3Au and should be included in any theoretical fit to the Cu_3Au data.

A second restriction on $V(r)$ may be obtained from a knowledge of the ordered state of the alloy since the maxima of $\alpha(k)$ fall at those points in k space which correspond to the wave vector of the ordered state even at temperatures above the critical temperature. This implies [via Eq. (34)] that $V(k)$ *must* have absolute minima at these points. For example, Cowley[5] fitted his data on Cu_3Au with values for the first three nearest-neighbor interactions of $V_1 = 358k$, $V_2 = -34k$, and $V_3 = -19k$. These values yield minima for $V(k)$ at positions in k space which are incompatible with the ordered structure of Cu_3Au by the above criterion.

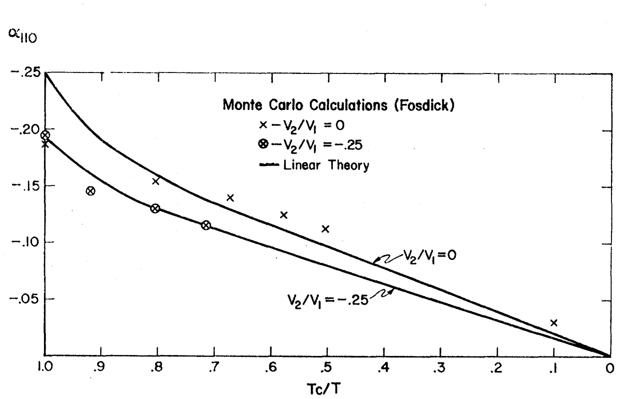

FIG. 4. The linear approximation compared with Fosdick's Monte Carlo calculations for the first-neighbor order parameter in Cu_3Au. The second-neighbor interaction V_2 was given the two values $-0.25\ V_1$ and zero.

In conclusion it may be said that the linear theory should prove useful in determining the strength and range of the pairwise interactions in alloys from the high-temperature data on short-range correlations, since the mathematics involved in making such a determination is quite tractable regardless of crystal symmetry, alloy composition, or range of the pairwise interaction. The theory should be particularly useful for testing the validity of long-range interaction models.

It should be remembered that the linear theory is valid only above the ordering temperature and that its accuracy improves as the temperature increases. A rough guess based on the few comparisons made with three dimensional alloys (to be reported upon in Part II) is that the linear theory probably is accurate enough for most purposes to within 10% of the ordering temperature. The range of validity for a particular system can be crudely estimated by comparing the temperature variation of the measured values of $\alpha(k)$ with the temperature behavior predicted by Eq. (34) [assuming $V(k)$ temperature-independent].

ACKNOWLEDGMENTS

One of us (S. C. M.) would like to express his appreciation to the U. S. Atomic Energy Commission for support of this general program of work.

APPENDIX

The theorem to be proven is that the operator $\sigma_O{}^B$ occurring in Eq. (9) of the text may be replaced by the conditional average, i.e., Eq. (A1) below is exact.

$$\sum_{\bar{\sigma}_k=\pm} \sigma_O{}^B\sigma_j{}^A\rho \Big/ \sum_{\bar{\sigma}_k=\pm} \sigma_j{}^A\rho = \sum_{\bar{\sigma}_k=\pm} \langle\sigma_O{}^B\rangle_{\text{cond}}\sigma_j{}^A\rho \Big/ \sum_{\bar{\sigma}_k=\pm} \sigma_j{}^A\rho, \tag{A1}$$

where

$$\langle\sigma_O{}^B\rangle_{\text{cond}} \equiv \sum_{\bar{\sigma}_O=\pm} \sigma_O{}^B\sigma_j{}^A\rho \Big/ \sum_{\bar{\sigma}_O=\pm} \sigma_j{}^A\rho$$

and $\rho = e^{-\beta(\frac{1}{4}\Sigma V_{ij}\bar{\sigma}_i\bar{\sigma}_j)}e^{\lambda\Sigma\bar{\sigma}_i}$.

The sum $\sum_{\bar{\sigma}_k=\pm}$ implies summing each of the $N\bar{\sigma}$ operators in the lattice (including $\bar{\sigma}_O$) over its two allowed values $+2m_B$ and $-2m_A$. We proceed by first imagining the sum over $\bar{\sigma}_O$ to be performed first in the numerator on the r.h.s. of Eq. (A1), i.e.,

$$\sum_{\bar{\sigma}_k=\pm} \langle\sigma_O{}^B\rangle_{\text{cond}}\sigma_j{}^A\rho = \sum_{\bar{\sigma}_k=\pm}{}' \langle\sigma_O{}^B\rangle_{\text{cond}}\sigma_j{}^A\rho, \tag{A2}$$

where $\sum_{\bar{\sigma}_k=\pm}{}'$ implies a summation for all operators other than $\bar{\sigma}_O$. It will be noticed from the definition of $\langle\sigma_O{}^B\rangle_{\text{cond}}$ above or Eq. (14) of the text that it contains no dependence on the operator $\bar{\sigma}_O$ since a summation over $\bar{\sigma}_O$ has removed this dependence. Therefore the order of $\sum_{\bar{\sigma}_O=\pm}$ and $\langle\sigma_O{}^B\rangle_{\text{cond}}$ in Eq. (A2) can be interchanged. This gives (putting in the definition of $\langle\sigma_O{}^B\rangle_{\text{cond}}$):

$$\sum_{\bar{\sigma}_k=\pm} \langle\sigma_O{}^B\rangle_{\text{cond}}\sigma_j{}^A\rho = \sum_{\bar{\sigma}_k=\pm}{}' \Big(\sum_{\bar{\sigma}_O=\pm} \sigma_O{}^B\sigma_j{}^A\rho \Big/ \sum_{\bar{\sigma}_O=\pm} \sigma_j{}^A\rho \Big) \sum_{\bar{\sigma}_O=\pm} \sigma_j{}^A\rho \tag{A3}$$

and now the factors $\sum_{\bar{\sigma}_O=\pm} \sigma_j{}^A\rho$ can be cancelled

$$\sum_{\bar{\sigma}_k=\pm} \langle\sigma_O{}^B\rangle_{\text{cond}}\sigma_j{}^A\rho = \sum_{\bar{\sigma}_k=\pm}{}' \Big(\sum_{\bar{\sigma}_O=\pm} \sigma_O{}^B\sigma_j{}^A\rho \Big) = \sum_{\bar{\sigma}_k=\pm} \sigma_O{}^B\sigma_j{}^A\rho \tag{A4}$$

which is the required equality.

The proof depends upon the use of a grand canonical ensemble and the property that the $\bar{\sigma}$ operators commute.