PHYSICAL REVIEW VOLUME 145, NUMBER 1 6 MAY 1966

Application of the Cluster Variation Method to the Heisenberg Model with Arbitrary Spin and Range of Exchange*

Tohru Morita†
The Catholic University of America, Washington, D. C.

AND

Tomoyasu Tanaka
The Catholic University of America, Washington, D. C.
and
U. S. Naval Ordnance Laboratory, Silver Spring, Maryland

(Received 22 October 1965; revised manuscript received 29 November 1965)

The cluster variation method for the cooperative phenomena proposed by Kikuchi and reformulated and generalized by Morita, is applied to the Heisenberg model with arbitrary spin and range of exchange. A general expression for the two-body reduced density matrix is obtained in the approximation in which the clusters of pairs of lattice sites are retained correctly. The constant-coupling approximation for the Heisenberg model of $S \geqq 1$ is shown to be derived by satisfying the reducibility conditions $\mathrm{tr}_k \rho^{(2)}(j,k) = \rho^{(1)}(j)$ only partly, requiring the consistency for the zeroth and first moments of S_{jz} and ignoring the consistency for the second to $2S$th moments. A natural method of extending the constant-coupling approximation for the Heisenberg model to the cases with arbitrary spin and range of exchange is suggested.

I. INTRODUCTION

RECENTLY Strieb, Callen, and Horwitz[1] tried to generalize Horwitz and Callen's[2] investigation on the Ising model to the Heisenberg model and gave a cluster expansion for the Heisenberg model. In their expression for the free energy, a parameter $\tilde{S}$ is introduced in such a way that each term depends on $\tilde{S}$ but the entire sum does not. As soon as an approximation is introduced, the resulting expression becomes dependent on $\tilde{S}$. They determined $\tilde{S}$ such that the approximate expression for the free energy is a minimum. Their lowest two approximations were shown to be equivalent to the Weiss molecular-field approximation and the constant-coupling approximation proposed by Kasteleijn and Van Kranendonk.[3]

Callen and Callen[4] generalized the constant-coupling approximation on the basis of the analysis of Streib, Callen, and Horwitz, to the Heisenberg ferromagnet with first- and second-neighbor exchanges, for the purpose of investigating the properties of europium chalcogenides. In this case also, they started from an identity which is valid for any value of $\tilde{S}$ and determined the value of $\tilde{S}$, after an approximation was introduced, by a stationary condition.

Quite recently and independently Fujishiro, Takano, and Oguchi[5] have developed a theory which is also a generalization of the constant-coupling approximation. Their treatment is somewhat similar to that of Callen and Callen in that they also take the second-neighbor interaction into account. Based on some physical intuitions, guided by the constant-coupling approximation, they guess an effective Hamiltonian for two- or three-spin cluster. This Hamiltonian contains several unknown local-field parameters, and they determine these parameters by some sort of consistency conditions. One of the interesting conclusions obtained in FTO theory as a result of three-spin cluster calculation is that the two-dimensional triangular lattice cannot be ferromagnetic whereas the simple cubic lattice becomes ferromagnetic in spite of the fact that both lattices have the same number of nearest neighbors.

In order to determine an arbitrarily introduced parameter $\tilde{S}$, CC used a method which is different from the method used by FTO, therefore, they obtained different results. Neither of these methods has been justified from the basic principle of statistical mechanics. As a matter of fact, FTO state in their paper that they cannot verify the method from the first principle and there are some ambiguities of the method.

On the other hand, we have the cluster-variation method for the cooperative phenomena proposed by Kikuchi.[6] This method is formulated in the form of a variational principle and is known to give the Weiss molecular-field approximation for $S \geqq \frac{1}{2}$ and the constant-coupling approximation for $S = \frac{1}{2}$. Hence, it is expected that there exists some relation between the cluster-variation method and the CC method or the FTO method.

It is the purpose of this paper to analyze the nature of approximation involved in the constant-coupling ap-

---

[*] Work supported in part by the U. S. Air Force Office of Scientific Research Grant AF-AFOSR-445-66.

[†] On leave of absence from Shizuoka University, Shizuoka, Japan.

[1] B. Strieb, H. B. Callen, and G. Horwitz, Phys. Rev. **130**, 1798 (1963).

[2] G. Horwitz and H. B. Callen, Phys. Rev. **124**, 1757 (1961).

[3] P. W. Kasteleijn and J. Van Kranendonk, Physica **22**, 317 (1956).

[4] H. B. Callen and E. Callen, Phys. Rev. **136**, A1675 (1964), to be referred to as CC in the following.

[5] T. Fujishiro, F. Takano, and T. Oguchi, J. Phys. Soc. Japan **19**, 1666 (1964), to be referred to as FTO in the following.

[6] R. Kikuchi, Phys. Rev. **81**, 988 (1951); J. Chem. Phys. **19**, 1230 (1951); M. Kurata, R. Kikuchi, and T. Watari, *ibid.* **21**, 434 (1953). See also T. Morita, J. Phys. Soc. Japan **12**, 753 (1957).

proximation from the point of view of the cluster variation method and to find a natural generalization of the constant-coupling approximation to the Heisenberg model with arbitrary range of exchange interaction.

In Sec. II the cluster variation method due to one of the present authors (T. M.)[7] is reviewed. In Sec. III the application of the method to the Heisenberg model is discussed. The constant-coupling approximation and its generalization are discussed in Sec. IV. In Sec. V a comparison with the previous formulations is made. In Sec. VI the results of numerical calculations are discussed.

II. CLUSTER VARIATION METHOD

The cluster variation method for the cooperative phenomena was first proposed by Kikuchi[6] and later reformulated and generalized by one of the present authors (T. M.).[7] The method may be summarized as follows.

It is well known that the free energy F of a system of which the Hamiltonian is $\mathcal{H}$ is calculated by the variational principle:

$$F=\min\mathfrak{F}, \tag{2.1}$$

$$\mathfrak{F}=\mathrm{tr}\rho_t[\mathcal{H}+kT\ln\rho_t], \tag{2.2}$$

where the minimum is taken with respect to the trial density matrix ρ_t under the normalization condition that

$$\mathrm{tr}\rho_t=1. \tag{2.3}$$

The ρ_t which minimizes $\mathfrak{F}$ is the density matrix ρ in thermal equilibrium, i.e., $\rho=[\exp-\beta\mathcal{H}]/\mathrm{tr}[\exp-\beta\mathcal{H}]$.

Let us consider a system which consists of L localized parts (e.g., localized spin-magnetic moments, particles, oscillators, etc.) which have some internal degree of freedom. For convenience we may call such a localized part a spin in the following. We then assume that the Hamiltonian of the system is given in the form:

$$\mathcal{H}=\sum_i h(i)+\sum_{i>j} h(i,j)+\cdots, \tag{2.4}$$

where $h(i)$ is the energy of the spin localized at the ith lattice site due to an external field and $h(i,j)$ is the energy of interaction between two spins localized at ith and jth lattice sites, and so forth.

In order to work with the variational principle (2.1) and (2.2) for such a system, it is convenient to introduce the trial n-lattice site reduced density matrix which is defined by

$$\rho_t^{(n)}(i_1,\cdots,i_n)=\mathrm{tr}_{\{L\}-i_1-\cdots-i_n}\rho_t, \quad n=1,\,2,\,\cdots, \tag{2.5}$$

and satisfies the recurrence relation

$$\begin{aligned}\rho_t^{(n-1)}(i_1,\cdots,i_{n-1})=\mathrm{tr}_{i_n}\rho_t^{(n)}(i_1,\cdots,i_{n-1},i_n),\\ n=1,\,\cdots,\,L,\,\rho_t^{(0)}=1.\end{aligned} \tag{2.6}$$

Here tr_{i_n} and $\mathrm{tr}_{\{L\}-i_1-\cdots-i_n}$ mean, respectively, the trace over the degree of freedom of the i_nth spin and the trace over all the degrees of freedom of L spins except those of i_1th, $\cdots$, i_nth spins. The tr in Eqs. (2.2) and (2.3) means the trace over all the degrees of freedom of L spins in the system and it will alternatively be expressed as $\mathrm{tr}_{\{L\}}$. Similarly, a notation $\mathrm{tr}_{i_1,\cdots,i_n}$, which indicates the trace over the degrees of freedom of the i_1th, $\cdots$, i_nth spins, will be introduced.

One now introduces

$$\Gamma^{(1)}(j),\quad \Gamma^{(2)}(j,k),\ \cdots,\ \Gamma^{(L)}(1,\cdots,L);$$

$\gamma^{(1)}(j)$, $\gamma^{(2)}(j,k)$, $\cdots$, $\gamma^{(L)}(1,\cdots,L)$ successively by means of the following relations:

$$\begin{aligned}\Gamma^{(n)}(i_1,\cdots,i_n)=\mathrm{tr}_{i_1,\cdots,i_n}\rho_t^{(n)}(i_1,\cdots,i_n)\\ \times\ln\rho_t^{(n)}(i_1,\cdots,i_n),\end{aligned} \tag{2.7}$$

$$\begin{aligned}\Gamma^{(1)}(j)&=\gamma^{(1)}(j),\\ \Gamma^{(2)}(j,k)&=\gamma^{(1)}(j)+\gamma^{(1)}(k)+\gamma^{(2)}(j,k),\\ \Gamma^{(3)}(j,k,l)&=\gamma^{(1)}(j)+\gamma^{(1)}(k)+\gamma^{(1)}(l)\\ &\quad+\gamma^{(2)}(j,k)+\gamma^{(2)}(k,l)+\gamma^{(2)}(j,l)\\ &\quad+\gamma^{(3)}(j,k,l),\end{aligned} \tag{2.8}$$

$$\begin{aligned}&\cdots\\ \Gamma^{(L)}(1,\cdots,L)&=\sum_j\gamma^{(1)}(j)+\sum_{j>k}\gamma^{(2)}(j,k)+\cdots\\ &\quad+\gamma^{(L)}(1,\cdots,L).\end{aligned}$$

The function $\mathfrak{F}$ given by Eq. (2.2) is then expressed as

$$\begin{aligned}\mathfrak{F}=\sum_j\mathrm{tr}_j h^{(1)}(j)\rho_t^{(1)}(j)+\sum_{j>k}\mathrm{tr}_{j,k}h^{(2)}(j,k)\rho_t^{(2)}(j,k)+\cdots\\ +kT\{\textstyle\sum_j\gamma^{(1)}(j)+\sum_{j>k}\gamma^{(2)}(j,k)+\cdots\}.\end{aligned} \tag{2.9}$$

The variational principle is now stated in terms of $\rho_t^{(n)}$'s:

$$0=\delta\mathfrak{F}/\delta\rho_t^{(n)}(i_1,\cdots,i_n),\quad n=1,\,2,\,\cdots,\,L, \tag{2.10}$$

with restricting conditions (2.6).

The above formulation is rigorous and the $\rho_t^{(n)}$'s which make the expression (2.9) a minimum are the reduced density matrices, $\rho^{(n)}$'s, in thermal equilibrium, defined by

$$\rho^{(n)}(i_1,\cdots,i_n)=\mathrm{tr}_{\{L\}-i_1-\cdots-i_n}\rho.$$

In practical applications, however, one introduces an approximation where the series in Eq. (2.9) is truncated at a point: i.e., only the first few $\gamma^{(n)}$'s are included.

III. APPLICATION TO THE HEISENBERG MODEL WITH ARBITRARY SPIN AND RANGE OF EXCHANGE

The variational principle (2.1), (2.9), and (2.10) with Eq. (2.6) is applied to the Heisenberg model with arbitrary spin and range of exchange. The Hamiltonian for the system is

$$\mathcal{H}=-\sum_j HS_{jz}-\sum_{j>k}J_{jk}\mathbf{S}_j\cdot\mathbf{S}_k, \tag{3.1}$$

where H is the external magnetic field in a suitable unit and J_{jk} is the exchange integral between jth and

---

[7] T. Morita, J. Phys. Soc. Japan **12**, 1060 (1957).

kth spins. J_{jk} is not necessarily restricted to nearest neighbors.

It is well known that if one retains only $\sum_j \gamma^{(1)}(j)$ in the second line of Eq. (2.9) and approximates $\rho_t^{(2)}(j,k)$ by $\rho_t^{(1)}(j)\,\rho_t^{(1)}(k)$, then one obtains the Weiss molecular-field approximation. In this paper, we shall investigate the approximation in which $\gamma^{(2)}$ is also retained. In this approximation, the free energy is given by

$$F=\min\mathfrak{F},$$

$$\begin{aligned}\mathfrak{F}=&-\sum_j H\,\mathrm{tr}_j S_{jz}\rho_t^{(1)}(j)-\sum_{j>k} J_{jk}\,\mathrm{tr}_{j,k}\mathbf{S}_j\cdot\mathbf{S}_k\rho_t^{(2)}(j,k)+kT\sum_j \mathrm{tr}_j\rho_t^{(1)}(j)\ln\rho_t^{(1)}(j)\\&+kT\sum_{j>k}\left[\mathrm{tr}_{j,k}\rho_t^{(2)}(j,k)\ln\rho_t^{(2)}(j,k)-\mathrm{tr}_j\rho_t^{(1)}(j)\ln\rho_t^{(1)}(j)-\mathrm{tr}_k\rho_t^{(1)}(k)\ln\rho_t^{(1)}(k)\right].\end{aligned}\tag{3.2}$$

The recurrence relations for $\rho_t^{(1)}$ and $\rho_t^{(2)}$ are

$$\mathrm{tr}_j\rho_t^{(1)}(j)-1=0.\tag{3.3}$$

$$\mathrm{tr}_k\rho_t^{(2)}(j,k)-\rho_t^{(1)}(j)=0.\tag{3.4}$$

Since the recurrence relation (3.4) is a matrix equation it is equivalent to a set of $(2S+1)^2$ scalar relations. If one introduces a complete set of $(2S+1)^2$ matrices, 1 and $M_n(j)$ $[n=1,\,2,\,\cdots,\,(2S+1)^2-1]$ such that any matrix A of order $2S+1$ can be expanded:

$$A=a_0+\sum_n a_nM_n(j).\tag{3.5}$$

Then the subsidiary conditions (3.4) for the variation are replaced by

$$\mathrm{tr}_jM_n(j)[\mathrm{tr}_k\rho_t^{(2)}(j,k)-\rho_t^{(1)}(j)]=0,\qquad n=1,\,\cdots,\,(2S+1)^2-1,\tag{3.6}$$

and

$$\mathrm{tr}_{j,k}\rho_t^{(2)}(j,k)-1=0.\tag{3.7}$$

Denoting the Lagrange multipliers for the subsidiary conditions (3.3), (3.7), and (3.6), as

$$kT-kT\sum_k 1+f(j),\quad kT-f(j)-f(k)+f(j,k)$$

and

$$kT\ln\rho^{(1)}(j)-\lambda_{k,j}^{(n)},$$

the variational function is written as

$$\begin{aligned}\mathfrak{F}=&-\sum_j H\,\mathrm{tr}_j S_{jz}\rho_t^{(1)}(j)-\sum_{j>k} J_{jk}\,\mathrm{tr}_{j,k}\mathbf{S}_j\cdot\mathbf{S}_k\rho_t^{(2)}(j,k)+kT\sum_j \mathrm{tr}_j\rho_t^{(1)}(j)\ln\rho_t^{(1)}(j)\\&+kT\sum_{j>k}\left[\mathrm{tr}_{j,k}\rho_t^{(2)}(j,k)\ln\rho_t^{(2)}(j,k)-\mathrm{tr}_j\rho_t^{(1)}(j)\ln\rho_t^{(1)}(j)-\mathrm{tr}_k\rho_t^{(1)}(k)\ln\rho_t^{(1)}(k)\right]\\&-\sum_j\left[kT-kT\sum_k 1+f(j)\right]\left[\mathrm{tr}_j\rho_t^{(1)}(j)-1\right]-\sum_{j>k}\left[kT-f(j)-f(k)+f(j,k)\right]\left[\mathrm{tr}_{j,k}\rho_t^{(2)}(j,k)-1\right]\\&-\sum_j\sum_k \mathrm{tr}_j\left[kT\ln\rho^{(1)}(j)-\lambda_k(j)\right]\left[\mathrm{tr}_k\rho_t^{(2)}(j,k)-\rho_t^{(1)}(j)\right],\end{aligned}\tag{3.8}$$

where

$$\lambda_k(j)=\sum_n \lambda_{k,j}^{(n)}M_n(j).\tag{3.9}$$

The Lagrange multipliers must be determined such that $\rho_t^{(1)}(j)$ and $\rho_t^{(2)}(j,k)$ which minimize this expression satisfy the subsidiary conditions (3.3), (3.7), and (3.6). It is noted that the differentiations of the expression (3.8) with respect to $f(j)$, $f(j,k)$, and $\lambda_{k,j}^{(n)}$ give the subsidiary conditions. This implies that one can determine these Lagrange multipliers by the condition that $\mathfrak{F}$ is stationary, instead of the subsidiary conditions.[8]

As a result of the variation with respect to $\rho_t^{(1)}(j)$, $\rho_t^{(2)}(j,k)$, $f(j)$, and $f(j,k)$, one obtains

$$\rho^{(1)}(j)=\exp\beta[f(j)+HS_{jz}+\sum_k \lambda_k(j)],\tag{3.10}$$

$$\begin{aligned}\rho^{(2)}(j,k)=\exp\beta[&f(j,k)+H(S_{jz}+S_{kz})+\sum_{k'(\neq k)}\lambda_{k'}(j)\\&+\sum_{j'(\neq j)}\lambda_{j'}(k)+J_{jk}\mathbf{S}_j\cdot\mathbf{S}_k],\end{aligned}\tag{3.11}$$

$$F=\sum_j f(j)+\sum_{j>k}[f(j,k)-f(j)-f(k)],\tag{3.12}$$

where

$$\exp-\beta f(j)=\mathrm{tr}_j\exp\beta[HS_{jz}+\sum_k\lambda_k(j)],\tag{3.13}$$

$$\begin{aligned}\exp-\beta f(j,k)=\mathrm{tr}_{j,k}\exp\beta[&H(S_{jz}+S_{kz})+\sum_{k'(\neq k)}\lambda_{k'}(j)\\&+\sum_{j'(\neq j)}\lambda_{j'}(k)+J_{jk}\mathbf{S}_j\cdot\mathbf{S}_k],\end{aligned}\tag{3.14}$$

and $\lambda_k(j)$ is given by Eq. (3.9). $\lambda_{k,j}^{(n)}$ must be determined by the consistency condition (3.7) or by

$$\partial F/\partial\lambda_{k,j}^{(n)}=0.\tag{3.15}$$

From Eqs. (3.10), (3.11), and (3.4), one concludes that

$$\lambda_k(j)=0\ \text{and}\ \rho^{(2)}(j,k)=\rho^{(1)}(j)\rho^{(1)}(k)\ \text{if}\ J_{jk}=0.\tag{3.16}$$

The above result has been obtained without specifying the set of $M_n(j)$'s. The result is simplified by the following choice[9] of $M_n(j)$'s; that is,

$$S_{jz}^n,\quad n=1,\,2,\,\cdots,\,2S$$

---

[8] This property that the Lagrangian multipliers are determined by the variation principle is quite general. For instance, the variational principle (2.1)–(2.3) is written as follows: The free energy F is equal to the value of the functional defined by

$$\mathfrak{F}\{\rho_t,\lambda\}=\mathrm{tr}\rho_t[\mathcal{H}+kT\ln\rho_t]+\lambda[\mathrm{tr}\rho_t-1]$$

for the values of ρ_t and λ which make this expression simultaneously stationary. This implies

$$\frac{\delta\mathfrak{F}\{\rho_t(\lambda),\lambda\}}{\delta\lambda}=\left[\frac{\delta\mathfrak{F}\{\rho_t(\lambda),\lambda\}}{\delta\rho_t(\lambda)}\right]_\lambda\frac{\delta\rho_t(\lambda)}{\delta\lambda}+\left[\frac{\delta\mathfrak{F}\{\rho_t(\lambda),\lambda\}}{\delta\lambda}\right]_{\rho_t(\lambda)}=0,$$

where $\rho_t(\lambda)$ is the ρ_t which makes $\mathfrak{F}(\rho_t,\lambda)$ stationary for fixed λ. Consequently one can calculate F in two steps: Minimize $\mathfrak{F}\{\rho_t,\lambda\}$ with respect to ρ_t for fixed λ, introduce the resulting $\rho_t(\lambda)$ into $\mathfrak{F}$, and then calculate the stationary value of $\mathfrak{F}\{\rho_t(\lambda),\lambda\}$ with respect to λ.

[9] *Note added in proof.* What is essential here is to choose a set of completely diagonal matrices and completely off-diagonal matrices for $Mn(j)$'s. The present choice is one of such possibilities.

and

$$[(S_j^+)^n+(S_j^-)^n]S_{jz}^m, \quad n=1, 2, \cdots, 2S$$

$$[(S_j^+)^n-(S_j^-)^n]S_{jz}^m, \quad m=0, 1, \cdots, 2S-n.$$

Now the subsidiary conditions (3.6) read

$$\mathrm{tr}_{j,k}S_{jz}^n\rho^{(2)}(j,k)-\mathrm{tr}_jS_{jz}^n\rho^{(1)}(j)=0, \qquad n=1, 2, \cdots, 2S, \tag{3.17}$$

and

$$\begin{aligned}\mathrm{tr}_{j,k}[(S_j^+)^n\pm(S_j^-)^n]S_{jz}^m\rho^{(2)}(j,k)&\\ -\mathrm{tr}_j[(S_j^+)^n\pm(S_j^-)^n]S_{jz}^m\rho^{(1)}(j)&=0,\\ n=1, 2, \cdots, 2S;\quad m=0, 1, \cdots, 2S-n.&\end{aligned} \tag{3.17a}$$

Equation (3.9) reads

$$\begin{aligned}\lambda_k(j)=\sum_{n=1}^{2S}\lambda_{k,j}^{(n)}S_{jz}^n+\sum_{n,m}\lambda_{k,j}^{(+n,m)}[(S_j^+)^n+(S_j^-)^n]S_{jz}^m&\\ +\sum_{n,m}\lambda_{k,j}^{(-n,m)}[(S_j^+)^n-(S_j^-)^n]S_{jz}^n.&\end{aligned}$$

The Lagrangian multipliers, $\lambda_{k,j}^{(n)}$ and $\lambda_{k,j}^{(\pm n,m)}$, must be determined to secure the consistency conditions given by Eqs. (3.17) and (3.17a). Now $\lambda_{k,j}^{(\pm n,m)}$ are set equal to zero, because then both terms in Eq. (3.17a) become individually zero and the consistency conditions given by Eq. (3.17a) are automatically secured. Hence, $\lambda_k(j)$ has been reduced to

$$\lambda_k(j)=\sum_{n=1}^{2S}\lambda_{k,j}^{(n)}S_{jz}^n, \tag{3.18}$$

where $\lambda_{k,j}^{(n)}$ must be determined by the conditions given by Eq. (3.17).[10] The implication of the fact that both terms in Eq. (3.17a) are individually zero will be discussed in the following paragraph.

In our present problem of the Heisenberg ferromagnet, the total Hamiltonian commutes with the total magnetization and hence the average value of an operator, which does not conserve the total magnetization, such as S_j^-, $S_j^-S_{jz}$, $(S_j^-)^2S_{jz}$, $S_j^-S_{kz}$, $S_j^-(S_k^+)^2S_{kz}$, etc., is zero. It follows that $\rho^{(1)}(j)$, $\rho^{(2)}(j,k)$, $\cdots$, commute with S_{jz}, $S_{jz}+S_{kz}$, $\cdots$, respectively. It is noted here that the $\rho^{(1)}(j)$ and $\rho^{(2)}(j,k)$ determined in the present approximation, where $\rho^{(2)}(j,k)$ is taken account of, satisfy these properties automatically.

IV. CONSTANT-COUPLING APPROXIMATION AND ITS GENERALIZATION

Let us introduce an approximation where the consistency relations (3.17) for $n=2, \cdots, 2S$ are neglected. Hence the Lagrangian multipliers $\lambda_{k,j}^{(2)}$, $\cdots$, and $\lambda_{k,j}^{(2S)}$ for these conditions are dropped. As a result, Eqs. (3.10)–(3.15) reduce to

$$\rho^{(1)}(j)=\exp\beta[f(j)+(H+\textstyle\sum_k\lambda_{k,j}^{(1)})S_{jz}], \tag{4.1}$$

$$\rho^{(2)}(j,k)=\exp\beta[f(j,k)+(H+\textstyle\sum_{k'(\neq k)}\lambda_{k',j}^{(1)})S_{jz}+(H+\sum_{j'(\neq j)}\lambda_{j',k}^{(1)})S_{kz}+J_{jk}\mathbf{S}_j\cdot\mathbf{S}_k], \tag{4.2}$$

$$F=\textstyle\sum_j f(j)+\sum_{j>k}[f(j,k)-f(j)-f(k)], \tag{4.3}$$

where

$$\exp-\beta f(j)=\mathrm{tr}_j\exp\beta(H+\textstyle\sum_k\lambda_{k,j}^{(1)})S_{jz}, \tag{4.4}$$

$$\exp-\beta f(j,k)=\mathrm{tr}_{j,k}\exp\beta[(H+\textstyle\sum_{k'(\neq k)}\lambda_{k',j}^{(1)})S_{jz}+(H+\sum_{j'(\neq j)}\lambda_{j',k}^{(1)})S_{kz}+J_{jk}\mathbf{S}_j\cdot\mathbf{S}_k] \tag{4.5}$$

and $\lambda_{k,j}^{(1)}$ must be determined by $\mathrm{tr}_jS_{jz}[\mathrm{tr}_k\rho^{(2)}(j,k)-\rho^{(1)}(j)]=0$ or by

$$\partial F/\partial\lambda_{k,j}^{(1)}=0. \tag{4.6}$$

For the Heisenberg ferromagnet where

$$\begin{aligned}J_{jk}&=J \text{ if } j \text{ and } k \text{ are nearest neighbors},\\ &=0 \text{ otherwise},\end{aligned}$$

one finds

$$\rho^{(1)}(j)=\exp\beta[f_1+(H+z\lambda)S_{jz}], \tag{4.7}$$

$$\rho^{(2)}(j,k)=\rho_{nn}^{(2)}(j,k)=\exp\beta\{f_2+[H+(z-1)\lambda](S_{jz}+S_{kz})+J\mathbf{S}_j\cdot\mathbf{S}_k\} \text{ when } j \text{ and } k \text{ are nearest neighbors}, \tag{4.8a}$$

$$=\rho^{(1)}(j)\rho^{(1)}(k), \text{ otherwise}. \tag{4.8b}$$

$$F/L=(1/2)zf_2-(z-1)f_1, \tag{4.9}$$

where z is the number of nearest-neighbor lattice sites and

$$\exp-\beta f_1=\mathrm{tr}_j\exp\beta(H+z\lambda)S_{jz}, \tag{4.10}$$

$$\exp-\beta f_2=\mathrm{tr}_{j,k}\exp\beta\{[H+(z-1)\lambda](S_{jz}+S_{kz})+J\mathbf{S}_j\cdot\mathbf{S}_k\}, \tag{4.11}$$

---

[10] In fact, $\lambda_k(j)$ given by Eq. (3.18) and the set of Eqs. (3.10)–(3.15) and Eq. (3.17) are noticed to be the result of the variation, when the subsidiary conditions given by Eq. (3.17a) are ignored and only those given by Eq. (3.17) are considered. In general, the free energy given by the variation in which some of subsidiary conditions are ignored is a lower bound to the free energy which would be obtained when all the subsidiary conditions were considered. Since $\rho^{(1)}(j)$ and $\rho^{(2)}(j,k)$ thus obtained satisfy all the ignored subsidiary conditions as mentioned above and give the lower bound for the free energy, they must be the $\rho^{(1)}(j)$ and $\rho^{(2)}(j,k)$ which make the free energy a minimum and satisfy all the subsidiary conditions.

and λ must be determined either by the consistency condition

$$\mathrm{tr}_j S_{jz}{}^{(1)} = \mathrm{tr}_{j,k} S_{jz} \rho_{nn}{}^{(2)}(j,k)\,, \tag{4.12}$$

or

$$\partial(F/L)/\partial\lambda = 0\,. \tag{4.13}$$

This set is seen to be equivalent to the constant-coupling approximation proposed by Kasteleijn and Van Kranendonk, in their original form, where Eq. (4.12) is used to determine λ or in the form derived by Strieb, Callen, and Horwitz, where Eq. (4.13) is used to determine λ. In the above we have neglected the subsidiary conditions (3.17) for $n=2, \cdots, 2S$, which is exact for $S=\frac{1}{2}$. For this reason, Kasteleijn and Van Kranendonk found that their constant-coupling approximation had the same result as the one by Kikuchi[11] who considered the Heisenberg ferromagnet only of $S=\frac{1}{2}$.

The above formulation has been set forth without restricting the range of exchange and hence it is applicable to the ferromagnets with first- and second-neighbor exchanges:

$$\begin{aligned} J_{jk} &= J \text{ if } j \text{ and } k \text{ are nearest neighbors},\\ &= J' \text{ if } j \text{ and } k \text{ are next-nearest neighbors},\\ &= 0 \text{ otherwise}. \end{aligned}$$

For this case, one has

$$\rho^{(1)}(j) = \exp\beta[f_1 + (H + z\lambda + z'\lambda')S_{jz}]\,, \tag{4.14}$$

$$\rho^{(2)}(j,k) = \rho_{nn}{}^{(2)}(j,k) = \exp\beta\{f_2 + [H + (z-1)\lambda + z'\lambda'](S_{jz} + S_{kz}) + J\mathbf{S}_j\cdot\mathbf{S}_k\}\,,$$

when j and k are nearest neighbors; (4.15a)

$$= \rho_{nnn}{}^{(2)}(j,k) = \exp\beta\{f_2' + [H + z\lambda + (z'-1)\lambda'](S_{jz} + S_{kz}) + J'\mathbf{S}_j\cdot\mathbf{S}_k\}\,,$$

when j and k are next-nearest neighbors; (4.15b)

$$= \rho^{(1)}(j)\rho^{(1)}(k)\,, \text{ otherwise, and} \tag{4.15c}$$

$$F/L = \tfrac{1}{2}zf_2 + \tfrac{1}{2}z'f_2' - (z + z' - 1)f_1, \tag{4.16}$$

where z and z' are the numbers of nearest- and next-nearest-neighbor lattice sites, respectively, and

$$\exp -\beta f_1 = \mathrm{tr}_j \exp\beta(H + z\lambda + z'\lambda')S_{jz}\,, \tag{4.17}$$

$$\exp -\beta f_2 = \mathrm{tr}_{j,k} \exp\beta[(H + (z-1)\lambda + z'\lambda') \times (S_{jz} + S_{kz}) + J\mathbf{S}_j\cdot\mathbf{S}_k]\,, \tag{4.18a}$$

$$\exp -\beta f_2' = \mathrm{tr}_{j,k} \exp\beta[(H + z\lambda + (z'-1)\lambda') \times (S_{jz} + S_{kz}) + J'\mathbf{S}_j\cdot\mathbf{S}_k]\,. \tag{4.18b}$$

The conditions determining λ and λ′ are either

$$\mathrm{tr}_j S_{jz} \rho^{(1)}(j) = \mathrm{tr}_{j,k} S_{jz} \rho_{nn}{}^{(2)}(j,k) \tag{4.19a}$$

and

$$\mathrm{tr}_j S_{jz} \rho^{(1)}(j) = \mathrm{tr}_{j,k} S_{jz} \rho_{nnn}{}^{(2)}(j,k) \tag{4.19b}$$

or

$$\partial(F/L)/\partial\lambda = 0 \tag{4.20a}$$

and

$$\partial(F/L)/\partial\lambda' = 0\,. \tag{4.20b}$$

V. COMPARISON WITH PREVIOUS FORMULATIONS

One can now examine the Callen-Callen formulation and the Fujishiro-Takano-Oguchi formulation under the light of the cluster variation method. For definiteness the comparison is limited to the two-spin cluster approximations for the ferromagnetic case with two exchange integrals J and J'.

The CC formulation is equivalent to setting

$$\lambda = J\bar{S}\,, \quad \lambda' = J'\bar{S} \tag{5.1}$$

in Eqs. (4.16), (4.17), (4.18a), and (4.18b). The parameter $\bar{S}$ is then determined by minimizing the free energy (4.16); this results in the condition [cf. Eq. (25) in Ref. 4].

$$\mathrm{tr}_j S_{jz} \rho^{(1)}(j) = w\ \mathrm{tr}_{j,k} S_{jz} \rho_{nn}{}^{(2)}(j,k) + (1-w)\ \mathrm{tr}_{j,k} S_{jz} \rho_{nnn}{}^{(2)}(j,k)\,, \tag{5.2}$$

where

$$w = [(z-1)zJ + zz'J']/(z + z' - 1)(zJ + z'J')\,. \tag{5.3}$$

One can see here that neither of the consistency conditions (4.19a) and (4.19b) are satisfied. Equation (5.2) is a weighted average of two consistency conditions with the weight (5.3).

In the FTO formulation they assume the effective Hamiltonians as follows:

$$\mathcal{H}^{(1)}(j) = -[H + zJ + z'J']\bar{S}S_{jz}\,, \tag{5.4}$$

$$\mathcal{H}^{(2)}(j,k) = -[H + (z-1)J + z'J'] \times \bar{S}(S_{jz} + S_{kz}) - J\mathbf{S}_j\cdot\mathbf{S}_k\,, \tag{5.5}$$

for the nearest neighbors only. This is again equivalent to setting

$$\lambda = J\bar{S}\,, \quad \lambda' = J'\bar{S}$$

in $\rho^{(1)}(j)$ and $\rho_{nn}{}^{(2)}(j,k)$ given by Eqs. (4.14) and (4.15a). In the FTO theory, however, there is no pre-

---

[11] R. Kikuchi, Ann. Phys. (N. Y.) 4, 1 (1958).

scription given to find the free energy. In order to determine the parameter $\bar{S}$, the consistency condition (4.19a) for the first moment is employed.

First of all, one should remember the fact that, in the case of either the CC or FTO theory, there exists an assumption which should be questioned: Why must $\bar{S}$ be determined by minimizing the truncated free energy in the CC theory? Why must the effective Hamiltonians be given by Eqs. (5.4) and (5.5) for $\rho^{(1)}(j)$ and $\rho_{nn}{}^{(2)}(j,k)$ in the FTO theory? Why must $\rho_{nnn}{}^{(2)}(j,k)$ not play any role in the FTO theory?

Secondly, Eq. (5.1) may be the most natural choice of the relation between λ and λ', if one has to introduce it with just the aid of intuition. If we accept Eq. (5.1), our equations reduce to their equations in some sense. However, the choice (5.1) results in an ambiguity for the condition determining $\bar{S}$; FTO used Eq. (4.19a); CC used a weighted average of Eqs. (4.19a) and (4.19b). Different weighted averages of Eqs. (4.19a) and (4.19b) would also be possible. Moreover, there appears another inconsistency in the CC theory. Sould the magnetization be given by Eq. (5.2) or should it be given as the derivative of the free energy with respect to the external magnetic field? CC used the latter form, which is found to be another weighted average of Eqs. (4.19a) and (4.19b).

One can now say that the CC and FTO formulations are justified, in some sense, from the cluster variation method, and one can understand the nature of approximations in these formulations. However, the justification is just for the pair approximation combined with the consistency condition for the first moment.[12]

The cluster variation method is free from all these inconsistencies. The minimization condition for the free energy and the consistency condition for the first moment are completely identical as stated in Sec. III. Furthermore, it is possible to introduce many more local-field parameters depending upon the desired degree of accuracy of the treatment.

VI. NUMERICAL CALCULATION FOR EuS

It is not attempted to carry out extensive numerical calculations in this paper for the following reasons. First of all the two-spin cluster approximation would still not be satisfactory to treat cooperative phenomena like ferromagnetic phase transitions, although it is a one-step improvement over the molecular-field approximation which does not take two-spin correlations into account at all. This unsatisfactoriness becomes prominent especially when there exists a nonzero antiferromagnetic exchange integral between next-neighbor pairs in addition to the nearest-neighbor ferromagnetic exchange integral. Secondly, even with an improvement in the present formulation which is a significant improvement from statistical mechanics theory point of view, over the CC and the FTO theories, the behavior of the spontaneous magnetication near the Curie temperature may not be modified very much compared with the previous calculations. Nevertheless the numerical calculations are discussed in this section simply because the procedures by which various physical quantities are calculated are slightly different from those in the previous treatments.

Following Callen and Callen, we shall calculate the dependence of the Curie temperature on the ratio of the nearest-neighbor and next-nearest-neighbor exchange integrals for a face-centered lattice of $S=\frac{1}{2}$ and $\frac{7}{2}$ and then calculate the temperature dependence of the spontaneous magnetization for a special ratio of $J'/J=-0.4$. for a face-centered-cubic lattice of $S=\frac{1}{2}$ and $S=\frac{7}{2}$. This ratio is the one which has been estimated for EuS by Charap and Boyd,[13] who gave

$$J/k=0.20°\text{K},\quad J'/k=-0.08°\text{K}.$$

Callen and Callen used these values in their investigation of EuS.

In Sec. IV, the free energy is given as a sum of three terms as

$$F/L=\tfrac{1}{2}zf_2+\tfrac{1}{2}z'f_2'-(z+z'-1)f_1. \tag{6.1}$$

The f_1 is given by Eq. (4.1) and is easily calculated in the representation in which S_{jz} is diagonal. The result is

$$\exp-\beta f_1=\sinh[(2S+1)\beta H_1/2]/\sinh[\beta H_1/2], \tag{6.2}$$

where

$$H_1=H+z\lambda+z'\lambda'. \tag{6.3}$$

The f_2 is given by Eq. (4.11) and is calculated in the representation where $(\mathbf{S}_j+\mathbf{S}_k)^2=\mathbf{S}_j{}^2+\mathbf{S}_k{}^2+2\mathbf{S}_j\cdot\mathbf{S}_k$ and $S_{jz}+S_{kz}$ are diagonal. The result is

$$\exp(-\beta f_2)=\frac{\sum_{\mathcal{S}=0}^{2S}\sinh[(2\mathcal{S}+1)\beta H_2/2]\exp[\beta J\mathcal{S}(\mathcal{S}+1)/2]}{\sinh[\beta H_2/2]}e^{-\beta JS(S+1)}, \tag{6.4}$$

---

[12] *Note added in proof.* Recently papers (O. Nagai, J. Phys. Soc. Japan **18**, 510 (1963), and a subsequent private communication) came to the authors' attention in which a two-sublattice antiferromagnet and also a Heisenberg ferromagnet are treated in the two-spin cluster approximation. In these papers effective one-spin and two-spin Hamiltonians are introduced based on some intuitive considerations but without any attempt to derive them. These Hamiltonians are exactly the ones which would be obtained if the method of the present paper is applied to the problem in the approximation in which the two-spin cluster is included and the first moment of the magnetization is determined consistently as is done in this section. In Nagai's papers the local field parameters are also determined by the same self-consistency conditions. Therefore, the method used in these papers is completely justified from the point of view of the cluster variation method within the specified approximations. Nagai calculated the Curie temperature for the Heisenberg ferromagnet of spin $\frac{1}{2}$. This result is identical to the corresponding part given in the next section.

[13] S. H. Charap and E. L. Boyd, Phys. Rev. **133**, A811 (1964).

where

$$H_2=H+z\lambda+z'\lambda'-\lambda. \tag{6.5}$$

The f_2' which is given by Eq. (4.18b) is similarly calculated as

$$\exp(-\beta f_2')=\frac{\sum_{\mathcal{S}=0}^{2S}\sinh[(2\mathcal{S}+1)\beta H_2'/2]\exp[\beta J'\mathcal{S}(\mathcal{S}+1)/2]}{\sinh[\beta H_2'/2]}e^{-\beta J'S(S+1)}, \tag{6.6}$$

where

$$H_2'=H+z\lambda+z'\lambda'-\lambda'. \tag{6.7}$$

The equations to determine λ and λ' are obtained by the variation (4.20a), (4.20b) or the self-consistency conditions (4.19a), (4.19b). This means that λ and λ' must be determined such that all of three values:

$$2m_1=2\langle S_{jz}\rangle_1\equiv 2\frac{\partial(-\beta f_1)}{\partial\beta H}=(2S+1)\coth[(2S+1)\beta H_1/2]-\coth[\beta H_1/2], \tag{6.8}$$

$$2m_2=\langle S_{jz}+S_{kz}\rangle_2\equiv\frac{\partial(-\beta f_2)}{\partial(\beta H)}=\frac{1}{2}\left\{\frac{\sum_{\mathcal{S}=0}^{2S}(2\mathcal{S}+1)\cosh[(2\mathcal{S}+1)\beta H_2/2]\exp[\beta J\mathcal{S}(\mathcal{S}+1)/2]}{\sum_{\mathcal{S}=0}^{2S}\sinh[(2\mathcal{S}+1)\beta H_2/2]\exp[\beta J\mathcal{S}(\mathcal{S}+1)/2]}-\coth[\beta H_2/2]\right\}, \tag{6.9}$$

and

$$2m_2'=\langle S_{jz}+S_{kz}\rangle_2'\equiv\frac{\partial(-\beta f_2')}{\partial(\beta H)}=\frac{1}{2}\left\{\frac{\sum_{\mathcal{S}=0}^{2S}(2\mathcal{S}+1)\cosh[(2\mathcal{S}+1)\beta H_2'/2]\exp[\beta J'\mathcal{S}(\mathcal{S}+1)/2]}{\sum_{\mathcal{S}=0}^{2S}\sinh[(2\mathcal{S}+1)\beta H_2'/2]\exp[\beta J'\mathcal{S}(\mathcal{S}+1)/2]}-\coth[\beta H_2'/2]\right\} \tag{6.10}$$

are equal to each other.

As the first step, the relation between the Curie temperature and the value of the ratio J'/J will be discussed.

Just below the Curie temperature, the magnetization in the absence of the external field is very small. Since the local fields, λ and λ', are the Lagrangian multipliers for the consistency conditions for the magnetization, λ and λ' should vanish in the limit as the magentization vanishes. Therefore in Eqs. (6.3), (6.5), (6.7), (6.8), (6.9), and (6.10) H is set equal to zero and λ and λ' are regarded as very small parameters. Expanding the hyperbolic functions in Eqs. (6.8), (6.9), and (6.10) in powers of βH_1, βH_2, and $\beta H_2'$, respectively, and retainin the leading terms, one obtains

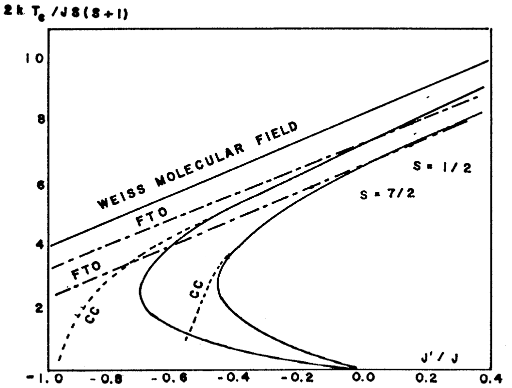


FIG. 1. Curie temperature as a function of exchange constants for spin values $\frac{1}{2}$ (the lower curve) and $\frac{7}{2}$ (the upper curve). The molecular-field result, the Callen-Callen result, and the Fujishiro-Takano-Oguchi result are also shown.

$$m_1=\beta(z\lambda+z'\lambda')\tfrac{1}{3}[(S+\tfrac{1}{2})^2-(\tfrac{1}{2})^2], \tag{6.11}$$

$$m_2=\beta(z\lambda+z'\lambda'-\lambda)\tfrac{1}{3}\times\tfrac{1}{8}\Phi_S(\beta J), \tag{6.12a}$$

$$m_2'=\beta(z\lambda+z'\lambda'-\lambda')\tfrac{1}{3}\times\tfrac{1}{8}\Phi_S(\beta J'), \tag{6.12b}$$

where

$$\Phi_S(\beta J)=\frac{\sum_{\mathcal{S}=0}^{2S}(2\mathcal{S}+1)^3\exp[\beta J\mathcal{S}(\mathcal{S}+1)/2]}{\sum_{\mathcal{S}=0}^{2S}(2\mathcal{S}+1)\exp[\beta J\mathcal{S}(\mathcal{S}+1)/2]}-1. \tag{6.13}$$

One can write the conditions $m_1=m_2=m_2'=0$, as $\beta\rightarrow\beta_c$, as follows:

$$\Phi_S(\beta_cJ)=2[(2S+1)^2-1]/(1-t), \tag{6.14a}$$

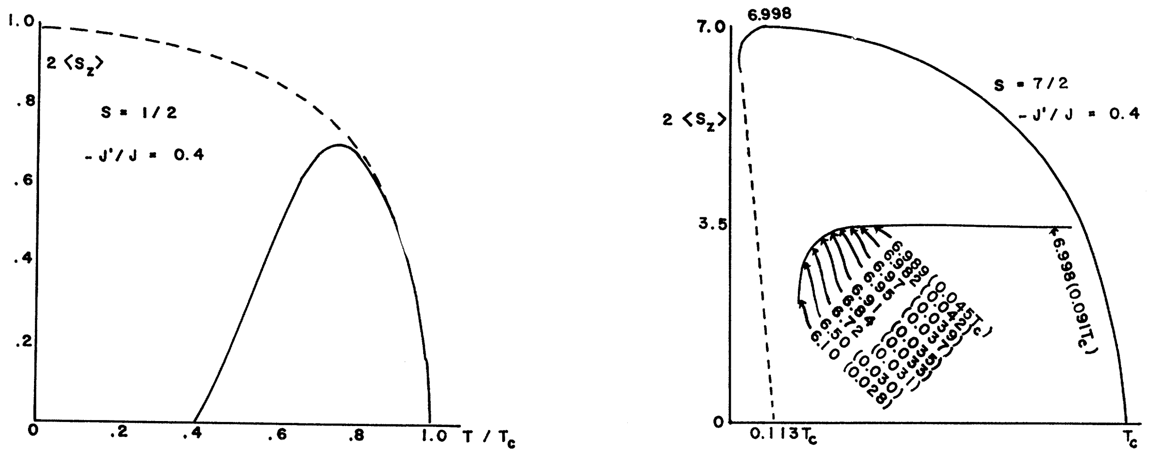
$$\Phi_S(\beta_cJ')=2[(2S+1)^2-1]/(1-t'), \tag{6.14b}$$

where

$$t=\lambda/(z\lambda+z'\lambda')=1/[z+z'(\lambda'/\lambda)], \tag{6.15a}$$

and

$$t'=\lambda'/(z\lambda+z'\lambda')=t(\lambda'/\lambda). \tag{6.15b}$$



(a) (b)

FIG. 2. (a) Magnetization as a function of temperature for spin $\frac{1}{2}$ and $J'/J=-0.4$. (b) Magnetization as a function of temperature for spin $\frac{7}{2}$ and $J'/J=-0.4$. A detailed behavior near the anti-Curie temperature is also shown.

Set $\beta_c J$ to an arbitrary value, and then calculate $\Phi_S(\beta_c J)$ by Eq. (6.13). This value of $\Phi_S(\beta_c J)$ is substituted into Eq. (6.14a) and (λ'/λ) will be calculated ($z=12$, $z'=6$ for fcc). With these t and (λ'/λ), t' is found by Eq. (6.15b), and then from the numerical table of $\Phi_S(\beta J)$ the value of $\beta_c J'$ is found. In this way the relation between $\beta_c J$ and J'/J is found and that is plotted in Fig. 1 for $S=\frac{1}{2}$ and $S=\frac{7}{2}$. One sees in Fig. 1 that the two-spin-cluster approximation gives an unphysical anti-Curie point for negative values of J'/J. The origin of the anti-Curie point is easily traced back to the general form of the density matrix for the second-nearest-neighbor pair. As long as J' is negative, $\mathcal{S}=0$ is the only contribution to the summations in Eq. (6.10) at $\beta=\infty$. Hence $m'=0$ is the only possible value at $\beta=\infty$. If the magnetization is determined consistently, then $m_1=m_2=m_2'=0$ is found to be the only possibility at $\beta=\infty$. In Fig. 1, the results of FTO and CC are also included. In their results the unphysical anti-Curie point does not appear. This is because in FTO theory the density matrix for the second-neighbor pair does not play any role at all. In the CC theory, the magnetization is calculated as a weighted average of magnetizations for the nearest-neighbor pair and for the second-neighbor pair as given by Eq. (5.2). Hence, even if the magnetization vanishes for the second-neighbor pair, the weighted magnetization does not become zero and the anti-Curie point does not appear. The CC curves go to zero very rapidly as J'/J becomes more negative than a certain value, as reported in their paper. We found, however, that their curves actually extend much further to negative J'/J values.

As the second step, the temperature dependence of the spontaneous magnetization will be discussed. In the following calculation, the value of J'/J is fixed to be -0.4.

We choose a temperature lower than the Curie temperature. This means, then, that the values of βJ and $\beta J'$ are fixed. Plot $2m_1$, $2m_2$, and $2m_2'$ as functions of βH_1, βH_2, and $\beta H_2'$, respectively. In this way, three curves are obtained. Those curves are now looked at as giving βH_1, βH_2, and $\beta H_2'$ as functions of $2m$. From these curves $\beta\lambda=\beta H_1-\beta H_2$, $\beta\lambda'=\beta H_1-\beta H_2'$, and $\beta H=\beta H_1-(z\beta\lambda+z'\beta\lambda')$ are evaluated as functions of $2m$ and then the value of $2m$ for which $\beta H=0$ is calculated. This $2m$ is the consistent magnetization in the absence of the external field for the temperature βJ.

The spontaneous magnetization computed in this way as a function of temperature is plotted in Fig. 2 for $S=\frac{1}{2}$ and $S=\frac{7}{2}$.

For the case of $S=\frac{7}{2}$, the anti-Curie point is found to be $0.113T_c$. The spontaneous magnetization approaches the saturation value very closely at low temperatures. The curve overshoots the anti-Curie point, reaches maximum value of $6.998/2$ at $0.091T_c$, and then suddenly drops to zero at about $0.028T_c$.

This appearance of the anti-Curie point for a system in which the second-neighbor interaction is antiferromagnetic is a characteristic of the two-spin-cluster approximation. This suggests strongly that at least the three-spin correlation should be taken into account. Application of the cluster variation method to this case will be a subject of future consideration.

ACKNOWLEDGMENT

The authors express their appreciation to Dr. Earl Callen. Discussion with him suggested the present investigation.