The Spin-Wave Theory as a Variational Method and Its Application to Antiferromagnetism

RYOGO KUBO*

Institute for the Study of Metals, The University of Chicago, Chicago, Illinois

The spin-wave theory in a modified form is applied to antiferromagnetism. The interaction between different modes of spin-waves, which is due to the higher order terms omitted in the customary spin-wave treatment, is taken into account in a self-consistent manner. It is shown that the interaction terms cause a change of the frequency spectrum of spin-waves which is equivalent to the existence of an anisotropic field. The upper limits of the ground levels are calculated. The variational method can also be applied to the discussions of the stable configurations of magnetization vectors in anisotropic and external fields, and it is consistent with the molecular field theory.

INTRODUCTION

THE spin-wave theory first invented by Bloch[1] as an approximate method of treating ferromagnets at low temperatures has recently been applied by several authors to antiferromagnets.[2-8] The main difference between the two cases lies in the fact that the theory remains approximate for antiferromagnets even at absolute zero, where it becomes rigorous for ferromagnets. The reason is simple; it is because the spin-wave theory applied to antiferromagnets necessarily starts from the assumption that there exists an ordered pattern of the sublattices at low temperatures, whereas the ground state of an antiferromagnet should be a singlet which is described by a very complicated combination of the patterns of the spin ordering.

The inverse pattern of an ordered pattern, that is, the spin configuration in which the directions of spins are just reversed, is also to be included in the linear combination. The interaction between the opposite orderings, however, will be so indirect that it can be neglected, and a fairly good approximation is obtained by taking the configurations near one of the ordered states.[2] The transition from such a state to its opposite pattern should be so slow that neutron diffraction experiments[9] can actually prove the ordering. Such a spin-wave treatment of antiferromagnets is a reasonable approximation to the ground states as has been discussed in detail by P. W. Anderson,[2] and it can be used to derive the thermodynamic properties[3] as well as the microwave absorption of antiferromagnets at low temperatures.[5-7]

It should be remembered here that the spin-wave

theory customarily omits the higher order terms in the original Hamiltonian, which in the spin-wave language give the interactions between the different modes of the spin-waves and cause the broadening of resonance absorptions. Very little is known, however, concerning how these interactions effect the energy levels and the thermodynamic behavior of ferromagnets or antiferromagnets.[3] Obviously the contribution from the interaction terms is zero for ferromagnets at absolute zero, but it will be considerable at higher temperatures. It will be even greater for antiferromagnets, for which it is not zero at absolute zero. This is one of the reasons that one might feel somewhat skeptical about the spin-wave theory of antiferromagnets.

It should be also noticed that the neglect of interaction terms gives rise to a certain difficulty in the spin-wave theory. This is revealed, for example, in the fact that the fluctuation of the magnetization (of the sublattice in the case of antiferromagnets) diverges abnormally for both ferromagnets and antiferromagnets, or, more exactly, it is of the order of $N^{4/D}$, N being the total number of magnetic atoms or ions, and D the dimensionality of the lattice, if there exists no anisotropic field acting on the spins.[3] This abnormal fluctuation might, at first sight, be correlated with the free rotation of the resultant spin-moment in the absence of the anisotropic field. However, such an interpretation is inconsistent with the current picture that describes the spins primarily aligned by strong exchange forces, the resultant moment moving around in a relatively weak anisotropic field and also in an external field, if any. This picture requires the resultant moment to be a well-defined quantity with a normal fluctuation of the order of N. The fact that the spin-wave theory does not give such a picture may be regarded as a difficulty of the theory. Corresponding to this we have to suffer from many divergences if we apply simple perturbation methods to the interaction terms starting from the usual spin-wave approximation.

The purpose of the present paper is, in the first place, to examine the importance of these interactions terms, and secondly, to remove the difficulty mentioned above,

[1] F. Bloch, Z. Physik **61**, 206 (1930).
[2] P. W. Anderson, Phys. Rev. **86**, 694 (1952).
[3] R. Kubo, Phys. Rev. **86**, 929 (1952).
[4] I. Shoji, Busseiron-Kenkyu (in Japanese) **39**, 55 (1951).
[5] T. Nakamura, Prog. Theoret. Phys. **7**, 539 (1952).
[6] Keffer, Kaplan, and Yafet, Am. J. Phys. (to be published).
[7] J. M. Ziman, Proc. Phys. Soc. (London) **A65**, 540, 548 (1952).
[8] J. R. Tessman, Phys. Rev. **88**, 1132 (1952).
[9] C. G. Shull and J. S. Smart, Phys. Rev. **76**, 1256 (1949); Shull, Wollan, and Strausser, Phys. Rev. **83**, 333 (1951).

applying the spin-wave theory in a different form from the usual one, namely as a variational method. To do this, we choose a system of spin-waves with a certain spectrum of proper frequencies representing the best approximation of the spin-system. These spin-waves are different from those obtained in the usual spin-wave theory. They are determined in a self-consistent way, taking the interaction terms into account. Thus, for instance, the ground levels obtained are the best upper limit of the true values as far as we confine ourselves in the picture of spin-waves. This variational method will give a finite frequency for the spin-wave of infinite wavelength in the absence of an anisotropic field. This follows because, in antiferromagnets, where the spin-waves collide with each other even at absolute zero, the wave trains must be of finite length and accordingly the frequency of long waves cannot be exactly zero. The method proposed here can be applied to ferromagnets as well, but here we shall confine ourselves to antiferromagnets.

1. VARIATIONAL METHOD IN QUANTUM STATISTICS

First we describe in a general way the variational method employed. Let the Hamiltonian of the system under consideration be H. We define the variational functions for this Hamiltonian by the eigenfunctions of a suitable operator H_{tr}, called the trial Hamiltonian, which is assumed to be of the form

$$H_{tr} = \sum_r c_r X_r, \qquad (1.1)$$

where X_r's are some known operators, and the sum should include the adjoint operators such as $c_r^* X_r^*$ if X_r is not Hermitian. The trial Hamiltonian should be selected of solvable form, and in most cases must be separable into components having only a few degrees of freedom, so that the eigenvalue problem of H_{tr} will be solved explicitly in the form

$$H_{tr}\Psi_n = \epsilon_n \Psi_n, \qquad (1.2)$$

where Ψ_n and ϵ_n are naturally functions of the constants $\{c_r\}$. If such a trial Hamiltonian is properly chosen, the smallest diagonal element of H in the representation Γ diagonalizing H_{tr}, will be an approximation of the lowest level of H, the best approximation of which will be given by minimizing it with respect to the variational parameters $\{c_r\}$: As to the excited levels, there is a theorem proved by Peierls[10] which states that the approximate partition function

$$Z = \sum_n \exp(-H_{nn}/kT),$$

constructed by the diagonal elements of H in any representation, can never be larger than the true partition function of the system $Z_0 = \text{trace } \exp(-H/kT)$. Therefore, at finite temperatures, the approximate free

[10] R. Peierls, Phys. Rev. 54, 918 (1938).

energy

$$F = -kT \log Z \qquad (1.3)$$

should be minimized with respect to $\{c_r\}$. The calculated minimum of the free energy is the best approximation as far as we confine ourselves in the frame of the trial system set up by Eq. (1.1). The trial Hamiltonian itself might then be considered as a certain representation of the system at the temperature T.

The variational procedure of this program is simplified if the diagonal elements H_{nn} can be given in the form

$$H_{nn} = \bar{H}(\cdots, \bar{X}_r, \bar{X}_r^*, \cdots), \qquad (1.4)$$

where $\bar{X}_r$ means the diagonal element of the operator X_r in the representation Γ. $\bar{X}_r$ is dependent on the quantum number n and on the parameters $\{c_r\}$. The diagonal elements of H_{tr}, or the eigenvalues, are expressed as

$$\epsilon_n = H_{tr}(n) = \sum_r c_r \bar{X}_r(n). \qquad (1.5)$$

Then it can easily be proved that the extremum condition of F, Eq. (1.3), is reduced to a set of equations

$$\frac{\partial \bar{H}}{\partial \bar{X}_r} = c_r \qquad r = 1, 2, \cdots,$$

$$\bar{X}_r = \frac{\partial F_{tr}}{\partial c_r} \qquad r = 1, 2, \qquad (1.6)$$

where $\bar{H}$ is considered as a function of $\{\bar{X}_r\}$ in the form of Eq. (1.4) and F_{tr} is the "trial free energy" of the system represented by H_{tr}, that is,

$$F_{tr} = -kT \log \sum_n \exp(-\epsilon_n/kT). \qquad (1.7)$$

The solution of Eq. (1.6) gives the most favorable values of the variational parameters and the averages of the operators X_r's at the temperature T. Inserting these values, the function $\bar{H}$ of Eq. (1.4) gives the average of the energy, and the equation

$$F = F_{tr} - \bar{H}_{tr} + \bar{H} \qquad (1.8)$$

gives the free energy, which is an upper limit of the rigorous one. At absolute zero, the equations are reduced to

$$\frac{\partial \bar{H}}{\partial \bar{X}_r} = c_r \qquad r = 1, 2, \cdots,$$

$$\bar{X}_r = \frac{\partial \bar{H}_{tr}{}^0}{\partial c_r} \qquad r = 1, 2, \cdots, \qquad (1.9)$$

where $\bar{H}_{tr}{}^0$ means the lowest eigenvalue of H_{tr}. Of course, Eq. (1.9) is easily proved directly.

Applied to a many-electron system, Eq. (1.9) is nothing but the Hartree-Fock equation, provided that the set of our trial Hamiltonians covers the whole set of Hamiltonians corresponding to the one-body approxi-

mation. Then Eq. (1.6) is the extension of the Hartree-Fock equation at finite temperatures. We shall not go further in this, but we see that the variational method used in the following just corresponds to the Hartree-Fock method in the sense that the coupled oscillators are represented in a self-consistent manner by a set of independent oscillators.

2. SPIN OPERATORS

The spin-wave theory is most conveniently formulated by using the "*spin deviation*" operators introduced by Holstein and Primakoff.[11] The spin deviation is defined by

$$n = S - S_z, \tag{2.1}$$

S being the magnitude of the spin and S_z its component along a given axis of quantization. In terms of the annihilation and creation operators a and a^*, which will be called the *spin-deviation operators* in the following, the other components are written as

$$S_x + iS_y = (2S)^{\frac{1}{2}}\left(1 - \frac{a^*a}{2S}\right)^{\frac{1}{2}} a,$$

$$S_x - iS_y = (2S)^{\frac{1}{2}} a^*\left(1 - \frac{a^*a}{2S}\right)^{\frac{1}{2}}, \tag{2.2}$$

where a and a^* satisfy the commutation rule

$$aa^* - a^*a = 1 \tag{2.3}$$

and

$$a^*a = n. \tag{2.4}$$

The merit of this formulation lies primarily in this commutation rule, which is much simpler than those for the spin operators. The spin-wave method consists, essentially, of writing the Hamiltonian in terms of the Fourier components of the spin-deviation operators defined over the whole crystal, and diagonalizing the expression by retaining only quadratic terms in the Fourier components. Naturally the equivalent approximation is obtained in a semiclassical way,[12] which has been adopted for antiferromagnets by Anderson.[2] But we prefer the Holstein-Primakoff expression, since we want to give rigorous expressions for the higher terms.

However, there is an inherent difficulty in this expression. Due to the commutation rule of Eq. (2.3), the spin-deviation operators have to be matrices of infinite dimensions in the representation diagonalizing n, Eq. (2.4). By definitions, (2.1) and (2.2), the parts where $n \geqq 2S+1$ (or $|S_z| > S$) is separated from the part $|S_z| \leqq S$. This separability no longer holds in the spin-wave formulation, so that the unwanted contributions will necessarily come in from the part $n \geqq 2S+1$.

Therefore, if we wish to go a step further than the simple spin-wave theory, we have to supplement the definitions (2.1) and (2.2), which are primarily valid for $0 \leqq n \leqq 2S$, by some definitions for $n \geqq 2S+1$.

Mathematically, the most convenient definition will be that which is periodic. Let us introduce a new operator $[n]$ refined by

$$(n'|[n]|n'') = n'\delta_{n'n''}, \quad \text{for } 0 \leqq n' \leqq 2S,$$
$$(n'|[n]|n') = (n''|[n]|n''),$$
$$\text{if } n' = n'', \mod (2S+1). \tag{2.5}$$

$[n]$ is commutable with n, and so is a periodic function of n. With this we define the spin operators by

$$S_z = S - [n],$$
$$S_x + iS_y = (2S)^{\frac{1}{2}} fa, \tag{2.6}$$
$$S_x - iS_y = (2S)^{\frac{1}{2}} a^* f,$$

where f is a function of n defined by

$$f = \left(1 - \frac{[n]}{2S}\right)^{\frac{1}{2}} ([n]+1)^{\frac{1}{2}} (n+1)^{-\frac{1}{2}}. \tag{2.7}$$

In the representation where all the n's are diagonal, the Hamiltonian of a spin-system will correspond to a direct product of matrices, each of which is defined for each of the spins and, because of the definition (2.6), each composed of an infinite repetition of a $(2S+1)$-dimensional matrix. This Hamiltonian, which has an infinite dimension with respect to the spin-deviations, is to be distinguished from the usual Hamiltonian of the spin-system with $(2S+1)$-dimensions for each of the spins. Yet the new Hamiltonian has the same spectrum of the eigenvalues as the usual one. In particular, the upper limit of the lowest eigenvalue is the same for both. Thus the new Hamiltonian treated by the variational method gives the answer for the spin-system. The variational method will also be applied at finite temperatures using Eq. (2.6), although in this case it is less rigorous than at absolute zero because the virtual multiplicity introduced by the periodic definition of (2.6) might result in some errors.

An alternative of the definition (2.6) would be to introduce projection operators such as

$$(n'|\epsilon|n') = 1, \quad 0 \leqq n' \leqq 2S$$
$$= 0, \quad 2S < n'$$

and to multiply the density matrix of the spin-system by the product of such operators defined for all the spins. But this introduces mathematical difficulties, because then one has either to treat the problem approximately as for the spherical model of ferromagnets, or to abandon the simple commutation rule of Eq. (2.5) introducing the operators $\epsilon a \epsilon$ and $\epsilon a^* \epsilon$ instead of a and a^*. The latter method may be practical only for the case of $S = \frac{1}{2}$.[5]

---

[11] T. Holstein and H. Primakoff, Phys. Rev. **58**, 1908 (1940).

[12] G. Heller and H. A. Kramers, Proc. Roy. Acad. Sci. Amsterdam, **37**, 378 (1934); L. Hulthén, *ibid.* **39**, 190 (1936); M. J. Klein and R. S. Smith, Phys. Rev. **80**, 1111 (1951).

3. DIAGONAL ELEMENTS OF THE HAMILTONIAN OF ANTIFERROMAGNETS IN SOME TRIAL REPRESENTATIONS

For simplicity we assume the Hamiltonian of an antiferromagnet to be

$$H_{ex}=J\sum_{\langle jk\rangle} \mathbf{S}_j\cdot\mathbf{S}_k \qquad (3.1)$$

in the absence of anisotropy and of external fields. Assuming two opposite directions as the quantization axes of the spins on each of the sublattices, into which the lattice is assumed to be divided in the ordered state, we introduce the expressions for the spin operators by

$$\begin{aligned} S_{jz}&=S-[n]_j,\\ S_{jx}+iS_{jy}&=(2S)^{\frac{1}{2}}f_ja_j,\\ S_{jx}-iS_{jy}&=(2S)^{\frac{1}{2}}a_j^*f_j, \end{aligned} \qquad (3.2)$$

and

$$\begin{aligned} S_{kz}&=-S+[n]_k,\\ S_{kx}+iS_{ky}&=(2S)^{\frac{1}{2}}b_k^*f_k,\\ S_{kx}-iS_{ky}&=(2S)^{\frac{1}{2}}f_kb_k, \end{aligned} \qquad (3.3)$$

with $[n]$ and f defined by (2.5) and (2.7). Inserting Eqs. (3.2) and (3.3) in (3.1) one obtains

$$\begin{aligned} H_{ex}=&-\tfrac{1}{2}NJzS^2\\ &+zJS(\sum_j [n]_j+\sum_k [n]_k)-J\sum_{jk}[n]_j[n]_k\\ &+JS\sum_{jk}\{f_ja_jf_kb_k+a_j^*f_jb_k^*f_k\} \end{aligned} \qquad (3.4)$$

which is, strictly speaking, different from (3.1) in that the dimensions of (3.4) is ∞^N, whereas that of (3.1) is $(2S+1)^N$.

Now let us introduce the Fourier components of the spin-deviation operators

$$\begin{aligned} a_\lambda&=(2/N)^{\frac{1}{2}}\sum e^{ij\lambda}a_j\\ a_\lambda^*&=(2/N)^{\frac{1}{2}}\sum e^{-ij\lambda}a_j^*\\ b_\lambda&=(2/N)^{\frac{1}{2}}\sum e^{-ik\lambda}b_k\\ b_\lambda^*&=(2/N)^{\frac{1}{2}}\sum e^{ik\lambda}b_k^*. \end{aligned} \qquad (3.5)$$

Substituting these expressions into Eq. (3.4) we have a very complicated operator, which in itself may be beyond mathematical means. But we can give exact expressions for its diagonal elements in a certain family of representations.

The spin-wave approximation is obtained from Eq. (3.4) if we retain there only the quadratic terms of spin-deviation operators. This is given by

$$\begin{aligned} H_{\text{spin wave}}=&-\tfrac{1}{2}NJzS^2\\ &+zJS\sum_\lambda [a_\lambda^*a_\lambda+b_\lambda^*b_\lambda+\gamma_\lambda(a_\lambda b_\lambda+a_\lambda^*b_\lambda^*)], \end{aligned} \qquad (3.6)$$

where

$$\gamma_\lambda=\sum_\rho e^{i\rho\lambda}/z$$

with ρ denoting the vectors to the nearest neighbors

from a lattice point. The higher order terms of the spin-deviation operators omitted in Eq. (3.6) represent the interactions of the spin-wave modes. Now we define the trial Hamiltonians of the spin-waves by

$$H_{tr}=\sum_\lambda [c_{1\lambda}a_\lambda^*a_\lambda+c_{2\lambda}b_\lambda^*b_\lambda+c_{3\lambda}(a_\lambda b_\lambda+a_\lambda^*b_\lambda^*)], \qquad (3.7)$$

which is a generalized form of Eq. (3.6). The operator (3.7) can easily be diagonalized to

$$\begin{aligned} H_{tr}=\sum_\lambda [&(n_{1\lambda}+\tfrac{1}{2})(\omega_\lambda^0+\Delta\omega_\lambda)\\ &+(n_{2\lambda}+\tfrac{1}{2})(\omega_\lambda^0-\Delta\omega_\lambda)-\tfrac{1}{2}(c_{1\lambda}+c_{2\lambda})], \end{aligned} \qquad (3.8)$$

where

$$\omega_\lambda^0=[\{\tfrac{1}{2}(c_{1\lambda}+c_{2\lambda})\}^2-c_{3\lambda}^2]^{\frac{1}{2}}, \quad \Delta\omega_\lambda=\tfrac{1}{2}(c_{1\lambda}-c_{2\lambda}).$$

The two modes with the same wave number λ correspond to the precessional motions of spins in opposite sense. These classical pictures have been discussed by Keffer et al.[6] More general forms of the trial Hamiltonian will be any quadratic form of spin-deviation operators, but for simplicity we shall not go into such generalizations. Equation (3.7) means that we confine ourselves to the representations where

$$\operatorname{diag} a_\lambda=\operatorname{diag} a_\lambda^*=\operatorname{diag} b_\lambda=\operatorname{diag} b_\lambda^*=0,$$

$$\begin{aligned} &\operatorname{diag} a_\lambda a_{\lambda'}=\operatorname{diag} a_\lambda^*a_{\lambda'}=\operatorname{diag} a_\lambda^*a_{\lambda'}^*\\ &=\operatorname{diag} a_\lambda b_{\lambda'}=\cdots=\operatorname{diag} b_\lambda^*b_{\lambda'}^*=0 \text{ for } \lambda\neq\lambda'. \end{aligned} \qquad (3.9)$$

These conditions allow us to derive exact expressions for the diagonal elements of H_{ex}, Eq. (3.4), provided that we make an additive condition that

$$\begin{aligned} &\operatorname{diag} a_\lambda^*a_\lambda=0(1), \quad \operatorname{diag} b_\lambda^*b_\lambda=0(1),\\ &\operatorname{diag} a_\lambda b_\lambda=0(1), \quad \operatorname{diag} a_\lambda^*b_\lambda^*=0(1). \end{aligned} \qquad (3.10)$$

This means that we exclude those states where some modes of spin-waves are highly excited. One might suppose that such excitation may happen for the infinitely long waves, but we have to exclude such configurations because they correspond to the rotation of all the spins in phase.[3]

With the above conditions we can show the exact expression of the diagonal element of H_{ex} is given by

$$\begin{aligned} \operatorname{diag} H_{ex}=&H_{ex}(A, B, C_\rho, C_\rho^*)\\ =&-\frac{N}{2}zJS^2\Bigg[1-\frac{1}{S}AG(A)-\frac{1}{S}BG(B)\\ &+\frac{1}{z}\sum_\rho F_{0\rho}(A, B, C_\rho, C_\rho^*)\\ &+\frac{1}{z}\sum_\rho F_{1\rho}(A, B, C_\rho, C_\rho^*)\Bigg], \end{aligned} \qquad (3.11)$$

where

$$A = \frac{2}{N} \sum_\lambda \text{diag } a_\lambda{}^* a_\lambda,$$

$$B = \frac{2}{N} \sum_\lambda \text{diag } b_\lambda{}^* b_\lambda,$$

$$(3.12)$$

$$C_\rho = \frac{2}{N} \sum_\lambda e^{i\rho\lambda} \text{ diag } a_\lambda b_\lambda,$$

$$C_\rho{}^* = \frac{2}{N} \sum_\lambda e^{-i\rho\lambda} \text{ diag } a_\lambda{}^* b_\lambda{}^*,$$

and

$$G(A) = 1 - (2S+1)A^{2s}\{(1+A)^{2s+1} - A^{2s+1}\}^{-1}, \quad (3.13)$$

$$F_{0\rho} = \sum_{l=0}^\infty \sum_{m=0}^\infty \sum_{n=0}^\infty \binom{l+m}{l}\binom{l+n}{l}[l+m][l+n]$$

$$\times (A + AB - C_\rho C_\rho{}^*)^m (B + BA - C_\rho C_\rho{}^*)^n (C_\rho C_\rho{}^*)^l$$

$$\times [(1+A)(1+B) - C_\rho C_\rho{}^*]^{-1-2l-m-n}, \quad (3.14)$$

$$F_{1\rho} = C_\rho \sum_{l=0}^\infty \sum_{m=0}^\infty \sum_{n=0}^\infty (l+1)\binom{l+m+1}{l+1}\binom{l+n+1}{l+1}$$

$$\times f(l+m)f(l+n)(A+AB-C_\rho C_\rho{}^*)^m$$

$$\times (B+AB-C_\rho C_\rho{}^*)^n (C_\rho C_\rho{}^*)^l$$

$$\times [(1+A)(1+B) - C_\rho C_\rho{}^*]^{-2-2l-m-n}. \quad (3.15)$$

(See Appendix.) Since the diagonal elements of H_{ex} are expressed by Eq. (3.11) as was assumed in Eq. (1.4), we can apply the variational formulas given in Sec. 1.

4. GROUND LEVELS OF ANTIFERROMAGNETS

Now we shall go to the discussion of ground levels. In the absence of any anisotropy and magnetic field, we can assume in Eq. (3.8) that $c_{1\lambda} = c_{2\lambda}$ by symmetry. Then we have $A = B$ in Eq. (1.2) and Eqs. (1.9) are reduced to

$$\alpha = \frac{\partial \bar{H}_{ex}}{\partial C} \bigg/ \frac{\partial \bar{H}_{ex}}{\partial A},$$

$$A = B = -\frac{1}{2} + \frac{1}{N}\sum_\lambda (1 - \alpha^2 \gamma_\lambda{}^2)^{-\frac{1}{2}}, \quad (4.1)$$

$$C = C^* = -\frac{1}{N}\sum_\lambda \alpha\gamma_\lambda{}^2(1-\alpha^2\gamma_\lambda{}^2)^{-\frac{1}{2}},$$

where we assume that C_ρ is independent of ρ, which is true for cubic and also tetragonal systems. A and C are functions of α, which is the only parameter now remaining in our variational problem, although we started from a general assumption, Eq. (3.8) involving $(3/2)N$ parameters. Thus Eq. (4.1) gives the most

favorable value of α. The functions A and C can be calculated from the expressions[2,3,13]

$$A_1 (\sin\psi) = \frac{1}{\pi}K (\sin\psi) - \frac{1}{2},$$

$$C_1 (\sin\psi) = -\frac{1}{\pi}\sin\psi D (\sin\psi),$$

$$A_2 (\sin\psi) = \frac{2}{\pi^2}K^2\left(\sin\frac{\psi}{2}\right) - \frac{1}{2},$$

$$C_2 (\sin\psi) = -\cosec\psi A_2 (\sin\psi)$$

$$\qquad (4.2)$$

$$\qquad - \int_0^\psi A_2 (\sin\theta) \sec\theta \cot^2\theta d\theta,$$

$$A_3 (\sin\psi) = \frac{2}{\pi}\int_0^{\pi/2} A_2 (\sin\psi \sin\theta)d\theta,$$

$$C_3 (\sin\psi) = -\cosec\psi A_3 (\sin\psi)$$

$$\qquad - \frac{2}{\pi}\int_0^{\pi/2} A_2 (\sin\psi \sin\theta) \cot^2\theta d\theta,$$

where K and D are the complete elliptic integrals, or from the power series like

$$A_3 (\sin\psi) = \frac{1}{2}\sum_{n=1}^\infty \left[2^{-2n}\binom{2n}{n}\right]^4 \sin^{2n}\psi,$$

$$C_3 (\sin\psi) = -\frac{1}{2}\sum_{n=1}^\infty (2n-1)^{-1}\left[2^{-2n}\binom{2n}{n}\right]^4 \sin^{2n}\psi.$$

In these equations the suffices refer to the dimensionality of the lattice and the three-dimensional lattice is assumed here to be the body-centered type (CsCl type arrangement). Some examples of the numerical values of A and C are given in Table I.

If A and C are small enough, an approximate solution of Eq. (4.1) is given by

$$\alpha \approx 1 + [S^{-1} + 4(1-1/2S)^{\frac{1}{2}} - 4](A(1) - C(1)) < 1,$$

which can be used for three-dimensional cases. For example, if $S = 1$, the most favorable value of α is around 0.975. In this case, the interaction between the spin-waves raises the frequency of long waves to a few percent of the frequency of the shortest waves. In effect, this is equivalent to a certain anisotropic field, which, if present, helps the ordering and produces a finite frequency for long waves. Therefore, if we start from the wave functions obtained by our variational method, all the divergent difficulties are removed and the fluctuations will always be normal.

Some numerical results of the preliminary calculations are shown in Table II. For convenience, we list

[13] G. N. Watson, Quart. J. Math. **9–10**, 269 (1938); Quart. J. Pure and Appl. Math. **39**, 27 (1908).

the values of c defined by the equation

$$E_{\text{ground}} = -\frac{N}{2}zJS^2(1+cS^{-1}).\qquad(4.3)$$

Anderson[14] has proved by a variational argument that c lies in the limits

$$0 < c < z^{-1}.\qquad(4.4)$$

Our calculation replaces the lower limits of c in (4.4) by more accurate values. The spin-wave theory,[2,3] neglecting the higher terms, gives the approximate values of c:

$$c_\infty = -2[A(1)+C(1)],\qquad(4.5)$$

which should be the limit of our values for $S = \infty$.

For one-dimensional cases, Eq. (4.5) gives surprizingly good agreement with the rigorous value $c = 0.375$ calculated for $S = \frac{1}{2}$.[15] This agreement is more or less accidental, because the complete Hamiltonian evaluated by the wave function in the simple spin-wave approximation diverges in one dimensions. The most favorable values of α are actually smaller than 1. It is, however, remarkable that our c's are rather close to c_∞ for $S \geqq 1$.

In three dimensions, A and C are small even for $\alpha = 1$, and our results are very close to the spin-wave approximation. This is because of a tendency for the higher terms to cancel. For example, the third term on the right-hand side of Eq. (3.4) contributes to c by 0.011 for $S = 1$, but at the same time the higher expansion terms in the fourth term nearly compensate this increase.

One point which seems unfavorable for the variational method is that the theory fails to show the ground states of one-dimensional chains to be disordered. For example, the minimum of the energy plotted against $\alpha = \sin\psi$ occurs around $\psi \sim 65°$ for $S = 1$, which gives $A \sim 0.23$. Thus the ground state of our variational method has a finite magnetization of the sublattice, which is generally given by the formula

$$M = \frac{1}{2}Ng\mu_0[S - AG(A)].$$

Most probably the one-dimensional antiferromagnets are disordered even at absolute zero.[2] The off-diagonal elements, neglected in the variational method, should then be very important in this case.

TABLE I. The functions A and C.

ψ	$A_1(\sin\psi)$	$C_1(\sin\psi)$	$A_2(\sin\psi)$	$C_2(\sin\psi)$	$A_3(\sin\psi)$	$C_3(\sin\psi)$
90°	∞	$-\infty$	0.1966	-0.2756	0.0593	-0.0958
80°	0.5038	-0.6831	0.1469	-0.2258	0.0534	-0.0898
70°	0.2973	-0.4696	0.1074	-0.1847	0.0433	-0.0795
60°	0.1865	-0.3475	0.0759	-0.1499	0.0329	-0.0679
50°	0.1161	-0.2618	0.0510	-0.1196		

[14] P. W. Anderson, Phys. Rev. **83**, 1260 (1951).
[15] H. A. Bethe, Z. Physik **71**, 205 (1931).

TABLE II. Numerical examples of c calculated by the variational method.

S	1/2	1	3/2	2	∞	1/z
Linear chain	0.198	0.323	0.345	0.330	0.363	0.5
Quadratic layer	0.147	0.197	0.212		0.1580	0.25
CsCl type	0.069	0.073	0.074	0.073	0.0730	0.125

5. ANTIFERROMAGNETS IN ANISOTROPIC AND EXTERNAL FIELDS

In the molecular field theory of antiferromagnets, which has been adopted by Nagamiya[16] and by Keffer and Kittel[17] for the calculation of resonance frequencies, each of the resultant moments of the sublattices, is pictured precessing in the resultant field due to the exchange forces, the anisotropic field, and the external field. The counterpart of this picture in the present analysis is as follows:

In Fig. 1, the crystallographic axes are given by (x_0, y_0, z_0). Let us assume the axes of the precessional motions of the spins of sublattices to be z' and z'', and take the bisector of these two axes as the y axis, the z axis being chosen orthogonal to y in the plane (z', z'', y), and the x axis orthogonal to y and z. Then the components of the spins on one of the sublattices can be represented by Eq. (2.6), choosing the axes (x, y', z') where y' is orthogonal to z' and x. Similarly, the spin-components of the other sublattices can be expressed with respect to the axes (x, y'', z''). Then the Hamiltonian of the spin-system is a function of the spin-deviation operators referred to the prescribed axes of precession, of the relative orientation of (x, y, z) with respect to the crystallographic axes and the applied magnetic field, and of the assumed distortion angle δ, the angle between z' and $-z''$, which can be assumed to be small. In this form, the odd power terms of the spin-deviation operators do not vanish, which means that torques remain making the precession axes deviate from the assumed directions. The variational method can be applied here in various ways. For instance, we can put δ equal to zero, and instead assume some finite shifts of equilibrium point, say ϵ, for the spin-wave oscillator of infinite wavelength, which should be determined in a variational way (in this case the spin-waves of finite wave numbers have the precessional axes in $z(-z)$ direction). Or, assuming ϵ to be zero, the most favorable value of δ can be determined. More generally we can seek the most favorable values of ϵ and δ, which means the spin-waves are precessing around z' and z'', except the longest spin-wave which has a little different axis of its precession.

In this way, we can first obtain the free energy of the system which depends on the other parameters left, that is, the orientation of (x, y, z) axes, the anisotropy constants, and the field strength. The most stable state of the antiferromagnet is that in which the free energy

[16] T. Nagamiya, Prog. Theoret. Phys. **6**, 342, 350 (1951).
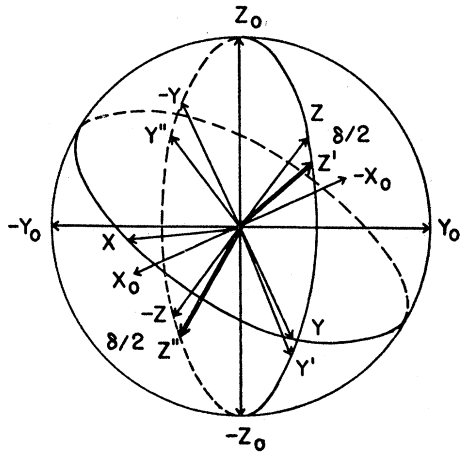[17] F. Keffer and C. Kittel, Phys. Rev. **85**, 329 (1952).

FIG. 1. Axes of quantization in anisotropic and external fields.

(or the energy at 0°K) is minimized with respect to the orientation. The calculation will not be carried out here, but we can show that the free energy is given in the form

$$F = F_{ex} + F_{anis} - \tfrac{1}{2}\chi_{\perp}(H_x^2 + H_y^2) - \tfrac{1}{2}\chi_{\parallel}H_z^2. \quad (5.1)$$

Here F_{ex} means the free energy of the system in the absence of the anisotropic and the external fields, and F_{anis} is the additional free energy due to the anisotropic field. The last two terms are the magnetic energy due to the external field, and $\chi_{\parallel}$ and $\chi_{\perp}$ are the parallel and perpendicular susceptibilities. This expression is proved on the assumption that both the anisotropic field and the external field are far weaker than the exchange force, which allows us to use a perturbation method. If, for instance, the anisotropy is strong enough, F cannot be separated into such terms as in Eq. (5.1).

In Eq. (5.1) the first term F_{ex} is isotropic with respect to the orientation of the precessional axes, while the others are anisotropic. The stable orientation is dependent on the relative magnitudes of the anisotropic field and the applied field, so that the effective susceptibility is generally dependent on both the direction and the magnitude of the applied field.

These are the same conclusions one gets from the molecular field theory[16,18] except that the quantities in Eq. (5.1) have different functional forms as regards temperature dependence, and that the anisotropic part of the free energy is calculated in a dynamical way. Some of such thermodynamic quantities have been studied by the author[3] in the frame of the spin-wave theory. The variational method does not change the results significantly for three-dimensional lattices, where the corrections due to the higher terms are small except in the case of fluctuations. For instance,

$$\chi_{\perp} \sim \frac{N}{2} \frac{g^2\mu_0^2}{zJ}$$

$$\chi_{\parallel} \sim \frac{4}{3} \frac{Ng^2\mu_0^2}{zJS}\left(\frac{kT}{zJS}\right)^2$$

[18] L. Néel, Ann. phys. 113 Serie 5, 232 (1936).

at low temperatures. More detailed discussion is out of place here and will be given elsewhere.

APPENDIX

Any function of the operator $n = a^*a$ can be represented by

$$f(n) = \frac{1}{2\pi i}\oint f(x)x^n\frac{dx}{x}, \quad (A.1)$$

with the generating function f defined by

$$f(x) = \sum_{n=0}^{\infty} f(n)x^{-n}. \quad (A.2)$$

The operator x^n is conveniently expanded as

$$x^n = \sum_{p=0}^{\infty} \frac{(x-1)^p}{p!}a^{*p}a^p. \quad (A.3)$$

Thus we have

$$\sum_j f(n_j)a_j{}^r f_2(n_{j+\rho})b_{j+\rho}{}^r$$

$$= (2\pi i)^{-2}\oint\oint f_1(x)x^{-1}dx f_2(y)y^{-1}dy$$

$$\times \sum_p \sum_q \frac{(x-1)^p}{p!}\frac{(y-1)^q}{q!}$$

$$\times \sum_j a_j{}^{*p}a_j{}^{p+r}b_{j+\rho}{}^{*q}b_{j+\rho}{}^{q+r}. \quad (A.4)$$

Now we define the notation

$$\text{diag}\sum_j a_j{}^{*p}a_j{}^{p+r}b_{j+\rho}{}^{*q}b_{j+\rho}{}^{q+r}$$

$$= \frac{N}{2}(p,\ p+r,\ q,\ q+r)_\rho, \quad (A.5)$$

where "diag" means the diagonal elements of the operator in the representations satisfying the conditions (3.9) and (3.10). Then we can show that

$$(p,\ p+r,\ q,\ q+r)_\rho = (p+r)!\ q!$$

$$\times \frac{1}{2\pi i}\oint (A + C_\rho{}^*\zeta^{-1})^p(B + C_\rho\zeta)^{q+r}\zeta^{-r-1}d\zeta. \quad (A.6)$$

For instance,

$$\frac{N}{2}(p,\ p,\ q,\ q)_\rho$$

$$= (2/N)^{p+q}\sum_j\sum_\lambda\cdots\sum_\mu\cdots\sum_\nu\cdots\sum_\kappa\cdots$$

$$\text{diag } a_{\lambda 1}{}^*\cdots a_{\lambda p}{}^*a_{\mu 1}\cdots a_{\mu p}b_{\nu 1}{}^*\cdots b_{\nu q}{}^*b_{\kappa 1}\cdots b_{\kappa q}$$

$$\times\exp[-i(\sum\lambda - \sum\mu - \sum\nu + \sum\kappa)j + i(\sum\kappa - \sum\nu)\rho]$$

$$= \frac{N}{2}\sum_t p!q!\binom{p}{t}\binom{q}{q-t}A^t C_\rho{}^{*p-t}C_\rho{}^{p-t}B^{q-p-t}$$

$$= \frac{N}{2}p!q!\frac{1}{2\pi i}\oint (A + C_\rho{}^*\zeta^{-1})^p(B + C_\rho\zeta)^q\zeta^{-1}d\zeta,$$