The detector crystal responds to the difference in the incident power (or intensity) with and without the magnetic field. The crystal signal is then proportional to

$$
\begin{aligned}
&[I_+(\nu)+I_0(\nu)+I_-(\nu)]-I(\nu)\\
&\quad=\tfrac{1}{3}[I(\nu)-(\partial I/\partial\nu)\delta\nu+\tfrac{1}{2}(\partial^2 I/\partial\nu^2)\delta\nu^2+\cdots]+\tfrac{1}{3}I(\nu)\\
&\qquad\qquad+\tfrac{1}{3}[I(\nu)+(\partial I/\partial\nu)\delta\nu+\tfrac{1}{2}(\partial^2 I/\partial\nu^2)\delta\nu^2+\cdots]-I(\nu)\\
&\quad=\tfrac{1}{3}(\partial^2 I/\partial\nu^2)\delta\nu^2+\text{higher even degree terms.}
\end{aligned}
$$

Because relatively low modulating fields are used, $\delta\nu^2 \gg \delta\nu^4$. Hence, terms involving fourth and higher degrees may be neglected. Also, $\delta\nu$ is a constant for any one line. Thus, the detector responds to a signal that is proportional to $\partial^2 I/\partial\nu^2$.

The argument is good, even for the $(1+\cos\omega t)$ type modulation used, as contrasted with the ideal square pulse of the derivation. Since the actual applied field remains within 15 percent of its maximum and minimum values for half the period of modulation, it reasonably approximates the ideal square-pulse field. The more complex Zeeman splitting resulting from higher rotational states will yield more terms in this derivation, but the net effect will be the same as that developed for the simplest case.

PHYSICAL REVIEW VOLUME 87, NUMBER 4 AUGUST 15, 1952

The Spin-Wave Theory of Antiferromagnetics

RYOGO KUBO*
Institute for the Study of Metals, University of Chicago, Chicago, Illinois

(Received March 19, 1952)

The spin-wave theory of antiferromagnets, recently studied by Anderson for the absolute zero of temperature, is examined here for finite temperatures to derive the thermodynamic properties of antiferromagnets at low temperatures. Somewhat differently from Anderson's semiclassical treatment, the present theory has used the formulation devised by Holstein and Primakoff, upon which the thermodynamic quantities are derived quantum-statistically. The parallel susceptibility is shown to be proportional to T^2, while the perpendicular susceptibility is independent of the temperature in the first approximation but decreases with increase in temperature if calculated in the second approximation. A tentative discussion is given of the nature of the divergences which arise in the simple formulation of spin-wave treatments in the absence of any kind of anisotropy.

INTRODUCTION

RECENTLY Anderson[1] has presented an approximate quantum theory of antiferromagnets on the basis of the semiclassical spin-wave theory first introduced by Kramers and Heller.[2] He examined very carefully the zero-point energy terms and showed that the spin-wave theory can be used to give pretty good approximations to the ground states of antiferromagnets. His result lies between the limits predicted by his variation method,[3] and moreover, it is very close to the rigorous value in the case of a one-dimensional chain of spins equal to one-half, which was treated by Hulthén[4] and by Bethe.[5]

As a matter of fact, the spin-wave theory of antiferromagnets is far more complicated and probably poorer as an approximate method than that of ferromagnets. Unfortunately the spin-wave theories are still far from satisfactory in both cases, because we know very little about the limitations of the method, which necessarily introduces great simplifications to bring the problem into the tractable form of independent spin-wave oscillators.

In spite of this situation, the writer wishes here to present a statistical-thermodynamical theory of antiferromagnets based on the Hulthén-Anderson spin-wave theory. The reason is, in the first place, that the spin-wave theory is at present the only approach we can use to attack the problem at low temperatures, and secondly that Anderson's success suggests that the low-lying energy levels may be reasonably approximated by this method. Approximations for higher temperatures can be treated by several methods. Van Vleck's theory of antiferromagnetism[6] is one of them and should be regarded as the standard theory. Refinements of this theory have been tried by Li[7] using the Bethe-Peierls-Weiss method and also by the present author, Obata, and Ohno.[8]

The treatment to be presented here is of course to be regarded as a starting point from which we should proceed to more rigorous theories. Also it should be emphasized that such improvements will be achieved more easily by handling the density matrices directly, rather than the Hamiltonian, to yield the separate energy levels. We hope the present theory will be improved in the future by taking account of the higher terms in the Hamiltonian, which are omitted in the simple theory but are very important to establish any satisfactory theory. In the last section of this paper we shall consider, by first-order perturbation theory, some of the effects due to these higher terms together with

---

* On leave from the Department of Physics, University of Tokyo, Tokyo, Japan.

[1] P. W. Anderson, Phys. Rev. **86**, 694 (1952).

[2] G. Heller and H. A. Kramers, Proc. Roy. Acad. Sci. Amsterdam **37**, 378 (1934).

[3] P. W. Anderson, Phys. Rev. **83**, 1260 (1951).

[4] L. Hulthén, Arkiv. Mat. Astron. Fysik **26A**, 1 (1938).

[5] H. A. Bethe, Z. Physik **21**, 205 (1931).

[6] J. H. Van Vleck, J. Chem. Phys. **9**, 85 (1941).

[7] Yin-Yuan Li, Phys. Rev. **84**, 721 (1951).

[8] Kubo, Obata, and Ohno (unpublished).

some discussion as to the applicability of the customary perturbation methods.

We confine ourselves at present to some simple models of antiferromagnets, which are surely unsatisfactory for the understanding of any actual antiferromagnetic crystals. However, most of our theory may be extended to apply to the more complicated antiferromagnetic structures, which will be discussed on some other occasion.

1. THE HAMILTONIAN FOR THE SPIN WAVES

First we shall repeat some of Anderson's treatment in a somewhat different language, namely in the formalism invented by Holstein and Primakoff,[9] which seems useful to make clearer the nature of the approximation. Following Holstein and Primakoff, we express the spin operators in the forms

$$S_x+iS_y=(2S)^{\frac{1}{2}}[1-(a^*a/2S)]^{\frac{1}{2}}a, \tag{1.1a}$$

$$S_x-iS_y=(2S)^{\frac{1}{2}}a^*[1-(a^*a/2S)]^{\frac{1}{2}}, \tag{1.1b}$$

$$S_z=S-a^*a, \tag{1.1c}$$

where S is the magnitude of the spin in units of $\hbar$, and the operators a^* and a are defined by

$$(n+1|a^*|n)=(n+1)^{\frac{1}{2}}, \tag{1.2a}$$

$$(n-1|a|n)=n^{\frac{1}{2}}, \tag{1.2b}$$

in the representation diagonalizing S_z. They satisfy the commutation law,

$$aa^*-a^*a=1. \tag{1.3}$$

As one can see from Eq. (1.1c), the operator

$$a^*a=n \tag{1.4}$$

is the "spin deviation," so that a^* is to be regarded as the creation operator of the spin deviation and a is the annihilation operator.

In the ferromagnetic case, where the Hamiltonian is assumed to be

$$H_{\text{ex}}=-J\textstyle\sum_{\langle jk\rangle}\mathbf{S}_j\cdot\mathbf{S}_k, \tag{1.5}$$

we obtain

$$\begin{aligned}H_{\text{ex}}=-\tfrac{1}{2}NzJS^2+[zJS\textstyle\sum n_j-JS\sum_{\langle jk\rangle}\{a_j^*f_s(n_j)f_s(n_k)a_k\\+a_k^*f_s(n_k)f_s(n_j)a_j\}-J\textstyle\sum_{\langle jk\rangle}n_jn_k].\end{aligned} \tag{1.6}$$

In Eqs. (1.5) and (1.6) the exchange interactions are supposed to be present only between nearest neighbors. This assumption can easily be removed if necessary. In Eq. (1.6) we have introduced the abbreviation,

$$f_s(n)=[1-(n/2S)]^{\frac{1}{2}}. \tag{1.7}$$

In the antiferromagnetic case, where the lattice is assumed to be divided into two interpenetrating sublattices, we have to introduce two different definitions of the spin-deviation operators, that is,

$$S_{xj}+iS_{yj}=(2S)^{\frac{1}{2}}[1-(n_j/2S)]^{\frac{1}{2}}a_j, \tag{1.8a}$$

$$S_{xj}-iS_{yj}=(2S)^{\frac{1}{2}}a_j^*[1-(n_j/2S)]^{\frac{1}{2}}, \tag{1.8b}$$

$$S_{zj}=S-n_j, \tag{1.8c}$$

for a spin j on one of the sublattices, say the (+) lattice, and

$$S_{xk}+iS_{yk}=(2S)^{\frac{1}{2}}b_k^*[1-(n_k/2S)]^{\frac{1}{2}}, \tag{1.9a}$$

$$S_{xk}-iS_{yk}=(2S)^{\frac{1}{2}}[1-(n_k/2S)]b_k, \tag{1.9b}$$

$$S_{zk}=-S+n_k, \tag{1.9c}$$

for a spin k on the other lattice, say the (−) lattice. The operators b and b^* are naturally defined in the same way as a and a^*, and satisfy the equations

$$b_k^*b_k=n_k,\quad b_kb_k^*-b_k^*b_k=1. \tag{1.10}$$

The simplest form of the Hamiltonian of an antiferromagnet is usually assumed to be

$$H_{\text{ex}}=|J|\textstyle\sum_{\langle jk\rangle}\mathbf{S}_j\cdot\mathbf{S}_k, \tag{1.11}$$

with positive exchange interactions between the nearest neighbors. Inserting Eqs. (1.8) and (1.9) into Eq. (1.11), we obtain

$$\begin{aligned}H_{\text{ex}}=-\tfrac{1}{2}Nz|J|S^2+z|J|S(\textstyle\sum_j n_j+\sum_k n_k)\\+|J|S\textstyle\sum_{\langle jk\rangle}\{f_s(n_j)a_jf_s(n_k)b_k\\+a_j^*f_s(n_j)b_k^*f_s(n_k)\}-|J|\textstyle\sum_{\langle jk\rangle}n_jn_k.\end{aligned} \tag{1.12}$$

The semiclassical formulation is simply obtained by the approximation

$$f_s(n)=1, \tag{1.13}$$

and by omitting the last term of Eq. (1.12), which is square in the spin deviations. Smith and Klein[10] have used this kind of approach in their argument about the semiclassical theory of Kramers and Heller in the ferromagnetic case, and these approximations are completely equivalent to Anderson's treatment.

As was noticed by Holstein and Primakoff, the operators a, a^*, b, and b^* are to be considered as matrices with infinite dimensions in order to satisfy the commutation rule (1.3), which is very useful in simplifying the mathematics and makes the spin-wave particles obey the Bose statistics. This mathematical trick is, of course, quite rigorous as far as we treat the whole expression of the Hamiltonian such as Eqs. (1.6) and (1.12). In the n-representation, that is, the representation where all the spin deviations are diagonalized, the eigen spaces are completely separated into three parts, namely, for the n_j, $n_k<0$, $0\leqq n_j$, $n_k\leqq 2S$, and for n_j, $n_k>2S$. Therefore the apparent infinity of dimensions does not introduce anything wrong.

---

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

[10] M. J. Klein and R. S. Smith, Phys. Rev. **80**, 1111 (1950).

Now the function $f_s(n)$, Eq. (1.7), can be developed as

$$\begin{aligned} f_s(n) &= [1-(n/2S)]^{\frac{1}{2}} \\ &= 1+f_1 n+f_2 n^2+\cdots+f_{2S} n^{2S}, \end{aligned} \tag{1.14}$$

which is rigorous for $0 \leqq n \leqq 2S$. Since the subspace which corresponds to this range of values of n is completely separated from other parts, Eq. (1.14) can be used for rigorous treatments of the Hamiltonian (1.12). The examples are

$$f_{\frac{1}{2}}(n)=1-n,$$

$$f_1(n)=1-(\tfrac{3}{2}-\sqrt{2})n-\tfrac{1}{2}(\sqrt{2}-1)n^2,$$

$$\begin{aligned} f_{\frac{3}{2}}(n)=1-\tfrac{1}{6}(11+3\sqrt{3}-6\sqrt{6})n+\tfrac{1}{6}(6+4\sqrt{3}-5\sqrt{6})n^2 \\ +\tfrac{1}{6}(-1-\sqrt{3}+\sqrt{6})n^3. \end{aligned}$$

For large values of S, the polynomial expansions (1.14) are naturally approximated by the binomial expansions.

As a matter of fact, the inclusion of higher terms of (1.14) and the square terms in n presents insurmountable difficulties for mathematical manipulation. Consequently, we usually omit these terms to get simple spin-wave Hamiltonians. However, with this simplification we cannot avoid introducing other difficulties, because it will certainly destroy the separability of the space and it causes the appearance of any large number of spin deviations. This is the chief reason for the fact that we have to have many divergences in the ordinary spin-wave theories. We shall return to this point in a later section.

One might think that the spin-wave theory may reasonably be applied to the ferromagnetic case, while it is not so good for the antiferromagnetic cases. In fact, the spin-wave theory is perfectly rigorous for the ground state of the ferromagnet, which is, however, trivial, and on the other hand, it remains only an approximation if applied to the antiferromagnetic case. However, the difficulties with regard to the excited states seem almost of the same nature in both cases. Though it may sound somewhat paradoxical, this situation seems to suggest that the spin-wave theory may be applied to antiferromagnetics as well as to ferromagnetics.

Postponing more discussion of this point to later sections, we proceed to establish the Hamiltonian of a spin-wave system following Anderson's treatment. In Eq. (1.12) the operators follow the commutation rules

$$\begin{aligned} a_j a_{j'}^* - a_{j'}^* a_j = \delta_{jj'}, \quad & b_k b_{k'}^* - b_{k'}^* b_k = \delta_{kk'}, \\ a_j b_k^* - b_k^* a_j = 0, \quad & a_j^* b_k - b_k a_j^* = 0, \\ a_j b_k - b_k a_j = 0, \quad & a_j^* b_k^* - b_k^* a_j^* = 0. \end{aligned} \tag{1.15}$$

By the canonical transformation defined by

$$\begin{aligned} a_j=\tfrac{1}{2}(Q_j+iP_j), \quad & a_j^*=\tfrac{1}{2}(Q_j-iP_j), \\ b_k=\tfrac{1}{2}(R_k+iS_k), \quad & b_k^*=\tfrac{1}{2}(R_k-iS_k), \end{aligned} \tag{1.16}$$

Eqs. (1.15) are transformed into

$$\begin{aligned} Q_j P_{j'} - P_{j'} Q_j = i\delta_{jj'}, \quad & R_k S_k - S_k R_k = i\delta_{kk'}, \\ Q_j R_k - R_k Q_j = 0, \quad & P_j R_k - R_k P_j = 0, \\ Q_j S_k - S_k Q_j = 0, \quad & P_j S_k - S_k P_j = 0. \end{aligned} \tag{1.17}$$

We introduce the Fourier transforms by

$$\begin{aligned} P_\lambda=(2/N)^{\frac{1}{2}}\textstyle\sum_j e^{i\lambda j}P_j, \quad & Q_\lambda=(2/N)^{\frac{1}{2}}\textstyle\sum_j e^{-i\lambda j}Q_j, \\ S_\lambda=(2/N)^{\frac{1}{2}}\textstyle\sum_k e^{-i\lambda k}S_k, \quad & R_\lambda=(2/N)^{\frac{1}{2}}\textstyle\sum_k e^{i\lambda k}R_k, \end{aligned} \tag{1.18}$$

which satisfy the commutation rules

$$\begin{aligned} Q_\lambda P_{\lambda'} - P_{\lambda'} Q_\lambda = i\delta_{\lambda\lambda'}, \quad & R_\lambda S_{\lambda'} - S_{\lambda'} R_\lambda = i\delta_{\lambda\lambda'}, \\ Q_\lambda R_{\lambda'} - R_\lambda' Q_\lambda = P_\lambda R_{\lambda'} & - R_{\lambda'} P_\lambda = 0, \\ Q_\lambda S_{\lambda'} - S_{\lambda'} Q_\lambda = P_\lambda S_{\lambda'} & - S_{\lambda'} P_\lambda = 0, \end{aligned} \tag{1.19}$$

The Hamiltonian of the spin waves is

$$\begin{aligned} H_{\mathrm{ex}}{}^0 = -\tfrac{1}{2}nz|J|S^2+(z)|J|S(\textstyle\sum_j n_j+\sum_k n_k) \\ +|J|S\textstyle\sum_{\langle jk\rangle}(a_j b_k+a_j^* b_k^*), \end{aligned} \tag{1.20}$$

which is obtained from Eq. (1.12) after the simplifications mentioned before. In terms of the Fourier transforms Eq. (1.20) is written as

$$\begin{aligned} H_{\mathrm{ex}}{}^0 = -\tfrac{1}{2}Nz|J|S(S+1)+(z/2)|J|S\textstyle\sum_\lambda\{P_\lambda{}^2+Q_\lambda{}^2 \\ +S_\lambda{}^2+R_\lambda{}^2+2\gamma_\lambda(Q_\lambda R_\lambda-P_\lambda S_\lambda)\}, \end{aligned} \tag{1.21}$$

where γ_λ is defined by

$$\gamma_\lambda=(1/z)\textstyle\sum_\rho e^{-i\lambda\rho}, \tag{1.22}$$

with ρ denoting the vectors to the nearest neighbors from a lattice point. The diagonalization of Eq. (1.21) is accomplished by the transformation,

$$\begin{aligned} P_\lambda=(p_{1\lambda}+p_{2\lambda})/\sqrt{2}, \quad & Q_\lambda=(q_{1\lambda}+q_{2\lambda})/\sqrt{2}, \\ S_\lambda=(p_{1\lambda}-p_{2\lambda})/\sqrt{2}, \quad & R_\lambda=(q_{1\lambda}-q_{2\lambda})/\sqrt{2}, \end{aligned} \tag{1.23}$$

which gives

$$\begin{aligned} H_{\mathrm{ex}}{}^0 &= -\tfrac{1}{2}Nz|J|S(S+1)+\tfrac{1}{2}z|J|S[\textstyle\sum_\lambda\{q_{1\lambda}{}^2(1+\gamma_\lambda) \\ &\quad +p_{1\lambda}{}^2(1-\gamma_\lambda)+q_{2\lambda}{}^2(1-\gamma_\lambda)+p_{2\lambda}{}^2(1+\gamma_\lambda)\}] \\ &= -\tfrac{1}{2}Nz|J|S(S+1)+\tfrac{1}{2}z|J|S[\textstyle\sum\{(2n_{1\lambda}+1) \\ &\quad +(2n_{2\lambda}+1)\}(1-\gamma_\lambda)^{\frac{1}{2}}]. \end{aligned} \tag{1.24}$$

As a typical form of antisotropic forces, we assume here that the crystalline fields cause anisotropic fields on the spins with an additional Hamiltonian,

$$\begin{aligned} H_{\mathrm{anis}} &= -K(\textstyle\sum_j S_{zj}{}^2+\sum_k S_{zk}{}^2) \\ &= -KS(S+1)+2KS(\textstyle\sum_j n_j+\sum_k n_k) \\ &\quad -K(\textstyle\sum_j n_j{}^2+\sum_k n_k{}^2), \end{aligned} \tag{1.25}$$

which is only reasonable for $S>\frac{1}{2}$ and for tetragonal structures. Anisotropic fields with symmetry higher than the form (1.25) can also be treated without any serious modification. Equation (1.25) is transformed as

$$\begin{aligned}H_{\mathrm{anis}}&=-NKS+2KS(\textstyle\sum_j n_j+\sum_k n_k)\\&\qquad-K(\textstyle\sum_j n_j^2+\sum_k n_k^2)\\&=-NKS(S+1)+KS\textstyle\sum_\lambda(P_\lambda^2+Q_\lambda^2+S_\lambda^2+R_\lambda^2)\\&\qquad-K(\textstyle\sum_j n_j^2+\sum_k n_k^2)\\&=-NKS(S+1)+KS\textstyle\sum(q_{1\lambda}^2+p_{1\lambda}^2+q_{2\lambda}^2+p_{2\lambda}^2)\\&\qquad-K(\textstyle\sum_j n_j^2+\sum_k n_k^2).\end{aligned}\tag{1.26}$$

Thus, including the effect of anisotropy, the Hamiltonian of the spin waves is assumed to be

$$\begin{aligned}H^0&=H_{\mathrm{ex}}^0+H_{\mathrm{anis}}^0\\&=-\tfrac{1}{2}Nz|J|(1+\alpha)S(S+1)\\&\quad+\tfrac{1}{2}(z|J|S)[\textstyle\sum_\lambda q_{1\lambda}^2(1+\alpha+\gamma_\lambda)+p_{1\lambda}^2(1+\alpha-\gamma_\lambda)\\&\qquad+q_{2\lambda}^2(1+\alpha-\gamma_\lambda)+p_{2\lambda}^2(1+\alpha+\gamma_\lambda)]\\&=-\tfrac{1}{2}Nz|J|(1+\alpha)S(S+1)+\tfrac{1}{2}(z|J|S)\textstyle\sum_\lambda\{(2n_{1\lambda}+1)\\&\qquad+(2n_{2\lambda}+1)\}[(1+\alpha)^2-\gamma_\lambda^2]^{\frac{1}{2}},\end{aligned}\tag{1.27}$$

where α means

$$\alpha=2K/z|J|.\tag{1.28}$$

Equations (1.24) and (1.27) were given by Anderson, who has examined the zero-point energies of the oscillators.

In the following we shall examine some of the thermodynamic properties of the spin-wave system expressed by Eqs. (1.24) and (1.27).

2. MAGNETIZATION OF THE SUBLATTICES, AND THE FREE ENERGY

The magnetizations of the sublattices are defined by the operators

$$M_z^+=g\mu_0\textstyle\sum_j S_{zj},\quad M_z^-=g\mu_0\sum_k S_{zk},\tag{2.1}$$

where g is the Landé g-factor, μ_0 the Bohr magneton, and the superscripts, + and −, refer to each of the sublattices. With the use of Eqs. (1.8), (1.16), (1.18), and (1.23), Eq. (2.1) is transformed to

$$\begin{aligned}M_z^+&=\tfrac{1}{2}Ng\mu_0S-g\mu_0\textstyle\sum n_j\\&=\tfrac{1}{2}Ng\mu_0[S+\tfrac{1}{2}-(2N)^{-1}\textstyle\sum(p_{1\lambda}^2+p_{2\lambda}^2+q_{1\lambda}^2\\&\qquad+q_{2\lambda}^2+2p_{1\lambda}p_{2\lambda}+2q_{1\lambda}q_{2\lambda})]\\M_z^-&=\tfrac{1}{2}Ng\mu_0S-g\mu_0\textstyle\sum n_k\\&=\tfrac{1}{2}Ng\mu_0[S+\tfrac{1}{2}-(2N)^{-1}\textstyle\sum(p_{1\lambda}^2+p_{2\lambda}^2+q_{1\lambda}^2\\&\qquad+q_{2\lambda}^2-2p_{1\lambda}p_{2\lambda}-2q_{1\lambda}q_{2\lambda})].\end{aligned}\tag{2.2}$$

The magnetizations of the sublattice, M_z^+ and M_z^-, are not, of course, constants of motion for the Hamiltonian (1.27). However, it is important to recognize that this situation does not prevent us from asking for the probability distributions of the observables M_z^+ and M_z^- (long-range order parameters), which are phase functions constructed from a great number of dynamical variables (S_{zi} and S_{zk}). Usually such phase functions are expected to have normal distributions (Gaussian distributions) with extremely sharp maxima. If such distributions are really proved for the expectation values of M_z^+ and M_z^-, we should be justified in conceiving the magnetizations of the sublattices to be thermodynamic quantities. And if these magnetizations are proved to be finite, then our theory may be considered as self-consistent.

Generally speaking, the probability distribution of dynamical quantities $A_1, A_2, \cdots A_k$, in a system with a Hamiltonian H and the canonical distribution $\exp(-H/kT)$, should be defined by its moment-generating function

$$\begin{aligned}&\exp\Phi(\xi_1,\xi_2,\cdots)\\&\quad=\mathrm{trace}[e^{-H/kT}\cdot e^{\xi_1A_1+\xi_2A_2+\cdots}]/\mathrm{trace}\ e^{-H/kT},\end{aligned}\tag{2.3}$$

which has a natural correspondence in classical statistical mechanics. A complexity in the quantum statistics is the product of operators $A_1, \cdots A_k$, but these are reasonably interpreted in Eq. (2.3) as the symmetrized products. All expectations of dynamical quantities are defined by Eq. (2.3) in agreement with the simple definition

$$\langle A\rangle=\mathrm{trace}(e^{-H/kT}A)/\mathrm{trace}\ e^{-H/kT}.$$

Equation (2.3) can be transformed inversely to give the probability distribution of "thermodynamical variables" $\bar{A}_1, \bar{A}_2, \cdots\bar{A}_k$, by

$$\begin{aligned}P(\bar{A}_1,\bar{A}_2,\cdots\bar{A}_k)&=\frac{1}{(2\pi i)^k}\int\cdots\int\exp[\Phi(\xi_1,\xi_2\cdots)\\&\qquad-\xi_1\bar{A}_1-\xi_2\bar{A}_2\cdots]d\xi_1d\xi_2\cdots,\end{aligned}\tag{2.4}$$

where the $\bar{A}_i$'s should be distinguished from their averages $\langle A_i\rangle$'s.

Although we lack any exact proof, it seems very reasonable to expect a certain form of the central limit theorem to hold for the distribution (2.4) in quantum statistics, if the dynamical variables $A_1, \cdots A_k$ have really their counterparts as thermodynamical variables. But of course we need some conditions for this statement. Unfortunately we know nothing about such conditions, so that we have to check the situation in each case.

If one assumes a Gaussian distribution for (2.4), we need to know the expansion of $\Phi(\xi_1, \xi_2\cdots)$ only to second powers, which can be conveniently written as

$$\Phi=\textstyle\sum\xi_l\langle A_l\rangle+\tfrac{1}{2}\sum\sum\Delta_{jl}\xi_j\xi_l+\cdots.\tag{2.5}$$

Then Eq. (2.4) is explicitly written as

$$\begin{aligned}P(\bar{A}_1,\bar{A}_2,\cdots\bar{A}_k)&=(2\pi)^{-k/2}[\det\Delta_{jl}]^{-\frac{1}{2}}\\&\quad\times\exp\{-\tfrac{1}{2}\textstyle\sum_{jl}(\Delta^{-1})_{jl}(\bar{A}_j-\langle A_j\rangle)(\bar{A}_l-\langle A_l\rangle)\},\end{aligned}\tag{2.6}$$

where

$$\begin{aligned} \langle A_l\rangle &= \mathrm{trace}(e^{-H/kT}\cdot A_l)/\mathrm{trace}(e^{-H/kT}), \\ \Delta_{jj} &= \langle A_j^2\rangle - \langle A_j\rangle^2, \qquad (2.7) \\ \Delta_{jl} &= \langle \tfrac{1}{2}(A_jA_l + A_lA_j)\rangle - \langle A_j\rangle\langle A_l\rangle. \end{aligned}$$

Returning to our problem, we take ξ_+ and ξ_- as the generating parameters. Then we have

$$\begin{aligned} \xi_+M_z^+ + \xi_-M_z^- = g\mu_0[\tfrac{1}{2}N(S+\tfrac{1}{2})(\xi_+-\xi_-) \\ -\tfrac{1}{4}(\xi_+-\xi_-)\textstyle\sum_\lambda(q_{1\lambda}^2+p_{1\lambda}^2+q_{2\lambda}^2+p_{2\lambda}^2 \\ -\tfrac{1}{2}(\xi_++\xi_-)\textstyle\sum_\lambda(q_{1\lambda}q_{2\lambda}+p_{1\lambda}p_{2\lambda})], \end{aligned} \tag{2.8}$$

which indicates that we had better ask for the distribution of $(M_z^+ + M_z^-)/2$ and $(M_z^+ - M_z^-)/2$, taking

$$\xi' = \xi_+ + \xi_-, \quad \xi'' = \xi_+ - \xi_- \tag{2.9}$$

because we have

$$\begin{aligned} \xi_+M_z^+ + \xi_-M_z^- &= \tfrac{1}{2}(M_z^+ + M_z^-)\xi' + \tfrac{1}{2}(M_z^+ - M_z^-)\xi'' \\ &= g\mu_0[-\tfrac{1}{2}\xi'\textstyle\sum(q_{1\lambda}q_{2\lambda}+p_{1\lambda}p_{2\lambda}) \\ &\quad +\xi''\{\tfrac{1}{2}NS-\tfrac{1}{4}\textstyle\sum_\lambda(q_{1\lambda}^2+p_{1\lambda}^2 \\ &\quad +q_{2\lambda}^2+p_{2\lambda}^2-2)\}. \end{aligned} \tag{2.10}$$

In our problem the whole expression for the generating function can be calculated without difficulty. But in the present paper we do not make use of it, so we give here only the first and the second moments, assuming a canonical distribution with the Hamiltonian H^0, (1.27). For convenience, we introduce the following notation:

$$\begin{aligned} (Z|J|S/2kT)(1+\alpha+\gamma_\lambda) &= \beta_\lambda', \\ (Z|J|S/2kT)(1+\alpha-\gamma_\lambda) &= \beta_\lambda'', \qquad (2.11) \\ 2(\beta_\lambda'\beta_\lambda'')^{\frac{1}{2}} = (Z|J|S/kT)\{(1+\alpha)^2-\gamma_\lambda^2\}^{\frac{1}{2}} &= \beta_\lambda. \end{aligned}$$

The well-known formulas for harmonic oscillators give us

$$\langle\beta_\lambda' q_{1\lambda}^2+\beta_\lambda'' p_{1\lambda}^2\rangle = \langle\beta_\lambda'' q_{2\lambda}^2+\beta_\lambda' p_{2\lambda}^2\rangle = \beta_\lambda\left(\frac{1}{2}+\frac{1}{e^{\beta_\lambda}-1}\right), \tag{2.12}$$

which shows that

$$\begin{aligned} \langle q_{1\lambda}^2\rangle = \langle p_{2\lambda}^2\rangle &= \tfrac{1}{2}(\beta_\lambda''/\beta_\lambda')^{\frac{1}{2}}\coth\tfrac{1}{2}\beta_\lambda \\ \langle p_{1\lambda}^2\rangle = \langle q_{2\lambda}^2\rangle &= \tfrac{1}{2}(\beta_\lambda'/\beta_\lambda'')^{\frac{1}{2}}\coth\tfrac{1}{2}\beta_\lambda. \end{aligned} \tag{2.13}$$

We also have evidently

$$\langle q_{1\lambda}q_{2\lambda}\rangle = \langle p_{1\lambda}p_{2\lambda}\rangle = 0. \tag{2.14}$$

From Eqs. (2.13) and (2.14), we obtain

$$\langle M_z^+\rangle + \langle M_z^-\rangle = 0,$$

and

$$\begin{aligned} M_s &= \langle M_z^+\rangle = -\langle M_z^-\rangle \\ &= \tfrac{1}{2}Ng\mu_0[S-(2/N)\tfrac{1}{4}\textstyle\sum_\lambda\langle(q_{1\lambda}^2+p_{1\lambda}^2+q_{2\lambda}^2+p_{2\lambda}^2-2)\rangle] \\ &= \frac{Ng\mu_0}{2}\left[S+\frac{1}{2}-\frac{2}{N}\sum\frac{(1+\alpha)}{\{(1+\alpha)^2-\gamma_\lambda^2\}^{\frac{1}{2}}}\right. \\ &\qquad\left.\times\left\{\frac{1}{2}+\frac{1}{e^{\beta_\lambda}-1}\right\}\right]. \end{aligned} \tag{2.15}$$

The fluctuations are calculated as follows: One can easily show that

$$\langle(\beta_\lambda' q_{1\lambda}^2+\beta_\lambda'' p_{1\lambda}^2)^2\rangle = \langle(\beta_\lambda'' q_{2\lambda}^2+\beta_\lambda' p_{2\lambda}^2)^2\rangle = \tfrac{1}{4}\beta_\lambda^2(2\coth^2\tfrac{1}{2}\beta_\lambda-1). \tag{2.16}$$

On the other hand, we know that the equalities

$$\langle p_1^4\rangle = 3\langle p_1^2\rangle^2, \quad \langle q_1^4\rangle = 3\langle q_1^2\rangle^2, \quad \text{etc.,} \tag{2.17}$$

hold because of the Gaussian properties of the canonical distribution for harmonic oscillators. Equations (2.16) and (2.17) give at once

$$\langle p_{1\lambda}^2q_{1\lambda}^2+q_{1\lambda}^2p_{1\lambda}^2\rangle = \langle p_{2\lambda}^2q_{2\lambda}^2+q_{2\lambda}^2p_{2\lambda}^2\rangle = \tfrac{1}{2}\coth^2\tfrac{1}{2}\beta_\lambda - 1. \tag{2.18}$$

Further, we evidently have

$$-\langle p_{1\lambda}q_{1\lambda}\rangle = \langle q_{1\lambda}p_{1\lambda}\rangle = -\langle p_{2\lambda}q_{2\lambda}\rangle = \langle q_{2\lambda}p_{2\lambda}\rangle = i/2. \tag{2.19}$$

Equations (2.17), (2.18), and (2.19) are enough to yield the formulas for the fluctuations. For brevity we put

$$\langle p_{1\lambda}^2+q_{1\lambda}^2\rangle = \langle p_{2\lambda}^2+q_{2\lambda}^2\rangle = m_\lambda = \frac{(1+\alpha)}{[(1+\alpha)^2-\gamma_\lambda^2]^{\frac{1}{2}}}\coth\tfrac{1}{2}\beta_\lambda, \tag{2.20}$$

and then we have

$$\begin{aligned} \langle(p_{1\lambda}^2+q_{1\lambda}^2)^2\rangle = \langle(p_{2\lambda}^2+q_{2\lambda}^2)^2\rangle &= 3m_\lambda^2-\delta_\lambda, \\ \langle(p_{1\lambda}^2+q_{1\lambda}^2)(p_{2\lambda}^2+q_{2\lambda}^2)\rangle &= m_\lambda^2, \end{aligned} \tag{2.21}$$

where

$$\delta_\lambda = 1+\coth^2\tfrac{1}{2}\beta_\lambda.$$

Now one can easily see from Eq. (2.8) that

$$\begin{aligned} &\langle\exp\{\tfrac{1}{2}(M_z^+ + M_z^-)\xi' + \tfrac{1}{2}(M_z^+ - M_z^-)\xi''\}\rangle \\ &\quad = \exp\{\tfrac{1}{2}Ng\mu_0(S+1)\xi''\}\times[1-\tfrac{1}{4}g\mu_0\xi''\textstyle\sum_\lambda 2m_\lambda \\ &\qquad +\tfrac{1}{2}(\tfrac{1}{4}g\mu_0\xi'')^2\{\displaystyle\sum_{\lambda\neq\lambda'}\sum 4m_\lambda m_\lambda' + \textstyle\sum_\lambda[2(3m^2-\delta_\lambda)+2m^2]\} \\ &\qquad +\tfrac{1}{2}(\tfrac{1}{2}g\mu_0\xi')^2\tfrac{1}{2}\textstyle\sum_\lambda(\coth^2\tfrac{1}{2}\beta_\lambda-1)+\cdots] \\ &\quad = \exp(M_s\xi''+\tfrac{1}{2}\xi''^2\Delta''+\tfrac{1}{2}\xi'^2\Delta'+), \end{aligned}$$

with Δ' and Δ'' defined by

$$\Delta' = \tfrac{1}{8}(g\mu_0)^2\textstyle\sum_\lambda(\coth^2\tfrac{1}{2}\beta_\lambda-1), \tag{2.22}$$

$$\Delta''=\tfrac{1}{8}(g\mu_0)^2\sum(2m_\lambda{}^2-\delta_\lambda)$$
$$=\tfrac{1}{8}(g\mu_0)^2\sum\nolimits_\lambda\left[\frac{2(1+\alpha)^2}{(1+\alpha)^2-\gamma_\lambda{}^2}-2+\left\{\frac{2(1+\alpha)^2}{(1+\alpha)^2-\gamma_\lambda{}^2}-1\right\}(\coth^2\tfrac{1}{2}\beta_\lambda-1)\right]. \quad (2.23)$$

The free energy $F(\bar{M}_z{}^+,\bar{M}_z{}^-)$ is defined by

$$e^{-F/T}=(\text{trace } e^{-H/kT})P(M_z{}^+,M_z{}^-), \quad (2.24)$$

so that we have

$$F(\bar{M}_z{}^+,\bar{M}_z{}^-,T)=-\tfrac{1}{2}Nz|J|(1+\alpha)S(S+1)$$
$$+z|J|S\sum\nolimits_\lambda\{(1+\alpha)^2-\gamma_\lambda{}^2\}^{\frac{1}{2}}$$
$$+2kT\sum\nolimits_\lambda\log(1-e^{-\beta\lambda})$$
$$+\frac{kT}{2\Delta''}\{\tfrac{1}{2}(\bar{M}_z{}^+-\bar{M}_z{}^-)-M_s\}^2$$
$$+\frac{kT}{2\Delta'}\left(\frac{\bar{M}_z{}^++\bar{M}_z{}^-}{2}\right)^2+\cdots, \quad (2.25)$$

which is an expansion in the neighborhood of the equilibrium values of the magnetizations

$$\bar{M}_z{}^+=-\bar{M}_z{}^-=M_s.$$

In Eq. (2.25) the second term is the zero-point energy of spin waves, which has been discussed by Anderson.

The equilibrium values of the energy, the entropy, and the specific heat are derived from Eq. (2.25) by the standard methods. They are given by

$$E(T)=-\tfrac{1}{2}Nz|J|(1+\alpha)S(S+1)$$
$$+z|J|S\sum\nolimits_\lambda\{(1+\alpha)^2-\gamma_\lambda{}^2\}^{\frac{1}{2}}$$
$$+2z|J|S\sum\nolimits_\lambda\{(1+\alpha)^2-\gamma_\lambda{}^2\}^{\frac{1}{2}}(e^{\beta\lambda}-1)^{-1}, \quad (2.26)$$

$$S(T)=-2k\sum\nolimits_\lambda\log(1-e^{-\beta\lambda})+(2z|J|S/T)$$
$$\times\sum\nolimits_\lambda\{(1+\alpha)^2-\gamma_\lambda{}^2\}^{\frac{1}{2}}(e^{\beta\lambda}-1)^{-1}, \quad (2.27)$$

$$C(T)=2[(z|J|S)^2/kT^2]$$
$$\times\sum\nolimits_\lambda\{(1+\alpha)^2-\gamma_\lambda{}^2\}e^{\beta\lambda}(e^{\beta\lambda}-1)^{-2}. \quad (2.28)$$

3. MAGNETIC SUSCEPTIBILITY

The antiferromagnet we are now considering is magnetically anisotropic, the anisotropy being caused not only by the antisotropic energy H_{anis}, but also by the approximate form of the Hamiltonian (1.21), though in fact the direction of the magnetization can be settled only by the existence of anisotropic energy. Thus we have two different susceptibilities $\chi_\perp$ and $\chi_\parallel$, the former being observed in the external field parallel to the spontaneous magnetization, the latter in a perpendicular field. The parallel susceptibility $\chi_\parallel$ can be calculated in two ways. In the presence of an external field H_z, the potential energy term,

$$H_{\text{field}}=H_z(M_z{}^++M_z{}^-)$$
$$=-H_zg\mu_0\sum\nolimits_\lambda(q_{1\lambda}q_{2\lambda}+p_{1\lambda}p_{2\lambda}) \quad (3.1)$$

is added to the Hamiltonian (1.27). The new Hamiltonian can still be brought into a diagonal form, so that we can calculate the free energy as a function of H_z, $\bar{M}_z{}^+$, $\bar{M}_z{}^-$, and T. The second procedure is a thermodynamical one, namely with the use of the equilibrium conditions

$$\partial F/\partial(\bar{M}_z{}^++\bar{M}_z{}^-)=H_z,\quad \partial F/\partial(\bar{M}_z{}^+-\bar{M}_z{}^-)=0. \quad (3.2)$$

Of course, this thermodynamical method is equivalent to the first one, as one can easily see from the fact that H_{field}, (3.1), commutes with the Hamiltonian (1.27).

Inserting Eq. (2.25) into (3.2), we obtain

$$\bar{M}_z{}^++\bar{M}_z{}^-=(4\Delta'/kT)H_z,\quad \bar{M}_z{}^+-\bar{M}_z{}^-=2M_s.$$

Thus we arrive at

$$\chi_\parallel=4\Delta'/kT=[(g\mu_0)^2/2kT]\sum\nolimits_\lambda(\coth^2\tfrac{1}{2}\beta_\lambda-1), \quad (3.3)$$

which clearly vanishes at the absolute zero.

The total magnetization in the x-direction is given by the operator

$$M_x=g\mu_0(\sum\nolimits_j S_{xj}+\sum\nolimits_k S_{xk}), \quad (3.4)$$

which is expressed in the Fourier component as

$$M_x=g\mu_0(NS)^{\frac{1}{2}}q_{10}, \quad (3.5)$$

using the simplification (1.13). The perpendicular susceptibility $\chi_\perp$ is obtained from the equation

$$\chi_\perp=\frac{M_x}{H_x}$$
$$=\frac{kT}{H_x}\frac{\partial}{\partial H_x}\log\text{trace}\left[\exp\left\{-\frac{1}{kT}(H^0-H_xM_x)\right\}\right]. \quad (3.6)$$

Inserting Eq. (1.27), (3.6) is simplified to

$$\chi_\perp=\frac{kT}{H_x}\frac{\partial}{\partial H_x}\log\text{trace}\left[\exp\left\{-\frac{z|J|S}{2kT}[q_{10}{}^2(1+\alpha+\gamma_0)+p_{10}{}^2(1+\alpha-\gamma_0)]+\frac{g\mu_0}{kT}(NS)^{\frac{1}{2}}H_xq_{10}\right\}\right]. \quad (3.7)$$

The magnetic energy $-H_xM_x$ gives simply a shift of

$$\epsilon=g\mu_0(NS)^{\frac{1}{2}}H_x/[z|J|S(1+\alpha+\gamma_0)] \quad (3.8)$$

to the equilibrium point of the oscillator (10), so that the exponent of Eq. (3.7) can be brought into the standard form by a simple canonical transformation $\exp(i\epsilon p_{10})$. Hence we get

$$\chi_\perp=\frac{kT}{H_x}\frac{\partial}{\partial H_x}\frac{g^2\mu_0{}^2NSH_x{}^2}{2z|J|SkT(1+\alpha+\gamma_0)}$$
$$=\frac{Ng^2\mu_0{}^2}{(1+\alpha+\gamma_0)z|J|}, \quad (3.9)$$

which can be written as

$$\chi_\perp = Ng^2\mu_0^2/2zJ \tag{3.10}$$

if we neglect α relative to 1 and put $\gamma_0=1$. Equation (3.10) is exactly the same as the susceptibility χ_c at the Curie point derived by Van Vleck.[6] (Note that J used in the present paper is just twice the ordinary definition of the exchange integral.)

4. TEMPERATURE DEPENDENCE

Now we turn to the discussion of the temperature dependence of the thermodynamic quantities derived in the previous sections. Equation (1.22) gives the following expressions for γ_λ:

1. one-dimensional chain,

$$\gamma_\lambda = \cos(\lambda/2); \tag{4.1}$$

2. two-dimensional square lattice,

$$\gamma_\lambda = \cos\tfrac{1}{2}\lambda_1 \cos\tfrac{1}{2}\lambda_2; \tag{4.2}$$

3. NaCl-type,

$$\gamma_\lambda = \tfrac{1}{3}[\cos\tfrac{1}{2}(\lambda_1-\lambda_2-\lambda_3)+\cos\tfrac{1}{2}(\lambda_2-\lambda_3-\lambda_1) + \cos\tfrac{1}{2}(\lambda_3-\lambda_1-\lambda_2); \tag{4.3}$$

4. CsCl-type,

$$\gamma_\lambda = \cos\tfrac{1}{2}\lambda_1 \cos\tfrac{1}{2}\lambda_2 \cos\tfrac{1}{2}\lambda_3; \tag{4.4}$$

where the wave numbers are referred to the reciprocal lattice of the sublattice, so that each component varies independently from $-\pi$ to π. Thus we can replace all the sums over wave numbers such as (2.15), (2.22), (2.23), (2.26), (2.27), and (2.28), by integrations over λ's, namely,

$$\Sigma_\lambda[\quad] = \frac{N}{2}\frac{1}{(2\pi)^D}\int\cdots\int_{-\pi}^{\pi}[\quad]d\lambda_1\cdots d\lambda_D, \tag{4.5}$$

where $D=1$, 2, 3 denotes the dimensionality of the lattice. In Eq. (4.3) the axes are not orthogonal, so that it is sometimes more convenient to change them into orthogonal coordinates defined by $\lambda_1'=\tfrac{1}{2}(\lambda_2+\lambda_3-\lambda_1)$, etc., the integration domain being changed at the same time into the first Brillouin zone of the face-centered cubic system whose volume is equal to $4\pi^3$. At first sight Eqs. (4.1)–(4.4) look different from the simple formula,

$$\gamma_\lambda = \frac{1}{D}\sum_{i=1}^{D}\cos\lambda_i, \tag{4.6}$$

used by Anderson, but it turns out that except for Eq. (4.4) they are equivalent to (4.6), because of the relations,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\cos\frac{\lambda}{2}\right)^{2n}d\lambda = \frac{1}{2\pi}\int_{-\pi}^{\pi}(\cos\lambda)^{2n}d\lambda = 2^{-2n}\binom{2n}{n}, \tag{4.7}$$

$$\frac{1}{(2\pi)^2}\int\int_{-\pi}^{\pi}\left(\cos\frac{\lambda_1}{2}\cos\frac{\lambda_2}{2}\right)^{2n}d\lambda_1 d\lambda_2 = \frac{1}{(2\pi)^2}\int\int_{-\pi}^{\pi}[\tfrac{1}{2}(\cos\lambda_1+\cos\lambda_2)]^{2n}d\lambda_1 d\lambda_2 = \left[2^{-2n}\binom{2n}{n}\right]^2, \tag{4.8}$$

$$\begin{aligned}\frac{1}{(2\pi)^3}\int\int\int_{-\pi}^{\pi}&[\tfrac{1}{3}\{\cos\tfrac{1}{2}(\lambda_1-\lambda_2-\lambda_3)+\cos\tfrac{1}{2}(\lambda_2-\lambda_3-\lambda_1) + \cos\tfrac{1}{2}(\lambda_3-\lambda_1-\lambda_2)\}]^{2n}d\lambda_1 d\lambda_2 d\lambda_3 \\ &= \frac{1}{(2\pi)^3}\int\int\int_{-\pi}^{\pi}[\tfrac{1}{3}(\cos\lambda_1+\cos\lambda_2+\cos\lambda_3)]^{2n}d\lambda_1 d\lambda_2 d\lambda_3 \\ &= 3^{-2n}2^{-2n}\sum_{p+q+r=n}\frac{(2n)!}{(p!q!r!)^2} \\ &= 3^{-2n}2^{-2n}\binom{2n}{n}\sum_{m=0}^{n}\binom{2m}{m}^2\binom{n}{m}^2.\end{aligned} \tag{4.9}$$

However,

$$\frac{1}{(2\pi)^3}\int\int\int_{-\pi}^{\pi}\left(\cos\frac{\lambda_1}{2}\cos\frac{\lambda_2}{2}\cos\frac{\lambda_3}{2}\right)^{2n}d\lambda_1 d\lambda_2 d\lambda_3 = \left[2^{-2n}\binom{2n}{n}\right]^3 \tag{4.10}$$

is not equal to (4.9). Equations (4.7)–(4.10) are useful for the numerical evaluation of integrals.

At absolute zero, $\chi_{\shortparallel}$, $S(T)$, and $C(T)$ are zero, as they should be, while E, M_s, Δ'', and $\chi_\perp$ are finite. But in the absence of the anisotropy K, M_s diverges for the one-dimensional case ($D=1$), and Δ'' diverges both for $D=1$, and $D=2$. As will be seen later, this divergent character is emphasized more at finite temperatures, where M_s and $\chi_{\shortparallel}(\Delta')$ diverge for $D=1$, 2, and Δ'' for all cases. These divergent quantities remain finite if we assume a finite anisotropy K, which may be much smaller in magnitude than the exchange constant J. Therefore, in the following we may neglect K for convergent integrals, while the divergent integrals are evaluated with finite $K \ll z|J|$; but in some cases the convergence thus attained is to be considered as superficial, as will be discussed in the next section.

For brevity we use the following notation:

$$\theta = kT/z|J|S, \tag{4.11}$$

$$M_\infty = Ng\mu_0 S/2, \tag{4.12}$$

and for the averages over wave number space we use the symbol $[\ \ ]_{\mathrm{Av}}$; for example,

$$[(1-\gamma_\lambda^2)^{\frac{1}{2}}]_{\mathrm{Av}}=\frac{2}{N}\sum(1-\gamma_\lambda^2)^{\frac{1}{2}}$$

$$=\frac{1}{(2\pi)^D}\int_{-\pi}^{\pi}(1-\gamma_\lambda^2)^{\frac{1}{2}}(d\lambda)^D,$$

or

$$[(1-\gamma_\lambda^2)^{\frac{1}{2}}(e^{\beta\lambda}-1)^{-1}]_{\mathrm{Av}}$$

$$=\frac{1}{(2\pi)^D}\int_{-\pi}^{\pi}(1-\gamma_\lambda^2)^{\frac{1}{2}}(e^{\beta\lambda}-1)^{-1}(d\lambda)^D. \tag{4.13}$$

For convenience we rewrite here the thermodynamic quantities derived in the preceding sections in the following forms:

$$E=E_0+E_T,$$

$$E_0=-\tfrac{1}{2}Nz|J|S\{S+1-[(1-\gamma_\lambda^2)^{\frac{1}{2}}]_{\mathrm{Av}}\}, \tag{4.14}$$

$$E_T=Nz|J|S[(1-\gamma_\lambda^2)^{\frac{1}{2}}(e^{\beta\lambda}-1)^{-1}]_{\mathrm{Av}},$$

$$S_T=Nk\{-[\log(1-e^{-\beta\lambda})]_{\mathrm{Av}}$$
$$+\theta^{-1}[(1-\gamma_\lambda^2)^{\frac{1}{2}}(e^{\beta\lambda}-1)]_{\mathrm{Av}}, \tag{4.15}$$

$$C_T=Nk\theta^{-2}[(1-\gamma_\lambda^2)e^{\beta\lambda}(e^{\beta\lambda}-1)^{-2}]_{\mathrm{Av}}, \tag{4.16}$$

$$M_S=M_{S0}+M_{ST},$$

$$M_{S0}=M_\infty[1+\tfrac{1}{2}S^{-1}\{1-(1+\alpha)[\{(1+\alpha)^2$$
$$-\gamma_\lambda^2\}^{-\frac{1}{2}}]_{\mathrm{Av}}\}], \tag{4.17}$$

$$M_{ST}=-(M_\infty/S)[(1+\alpha)\{(1+\alpha)^2-\gamma_\lambda^2\}^{-\frac{1}{2}}(e^{\beta\lambda}-1)^{-1}]_{\mathrm{Av}},$$

$$\chi_{\parallel}=(Ng^2\mu_0^2/4kT)[\coth^2\tfrac{1}{2}\beta_\lambda-1]_{\mathrm{Av}}, \tag{4.18}$$

$$\Delta''=\Delta_0''+\Delta_T'',$$

$$\Delta_0''=\tfrac{1}{8}Ng^2\mu_0^2\left[\frac{(1+\alpha)^2}{(1+\alpha)^2-\gamma_\lambda^2}-1\right]_{\mathrm{Av}}, \tag{4.19}$$

$$\Delta_T''=\tfrac{1}{16}Ng^2\mu_0^2\left[\left\{\frac{2(1+\alpha)^2}{(1+\alpha)^2-\gamma_\lambda^2}-1\right\}\right.$$
$$\left.\times(\coth^2\tfrac{1}{2}\beta_\lambda-1)\right]_{\mathrm{Av}}.$$

At temperatures lower than the order of magnitude $z|J|S/k$, the low frequency spin waves are most important, so that we may approximate in the customary way the frequency distribution of the spin waves by retaining the first-square term in the wave number λ in the expansion of γ_λ. Thus, for all types of crystals we can put

$$\gamma_\lambda^2=1-\lambda^2, \tag{4.20}$$

with normalization factors properly chosen for the integrals. The approximation formula is

$$[f(\gamma_\lambda^2)]_{\mathrm{Av}}\simeq A_l\int_0^\infty f(1-\lambda^2)\lambda^{D-1}d\lambda, \tag{4.21}$$

with

$$\begin{aligned}A_1&=2/\pi && \text{for the linear chain,}\\ A_2&=2/\pi && \text{for the quadratic layer,}\\ A_3&=3^{\frac{3}{2}}/\pi^2 && \text{for the NaCl-type,}\\ A_4&=4/\pi^2 && \text{for the CsCl-type.}\end{aligned} \tag{4.22}$$

For example, E_T is calculated as follows:

$$E_T/Nz|J|S=[(1-\gamma_\lambda^2)^{\frac{1}{2}}(e^{\beta\lambda}-1)^{-1}]_{\mathrm{Av}}$$

$$=[\lambda(e^{\lambda/\theta}-1)^{-1}]_{\mathrm{Av}}$$

$$=A_l\int_0^\infty\lambda\sum_{n=1}^{\infty}e^{-n\lambda/\theta}\lambda^{D-1}d\lambda$$

$$=A_l\Gamma(D+1)\zeta(D+1)\theta^{D+1}, \tag{4.23}$$

and M_{ST}:

$$\frac{M_{ST}}{M_\infty/S}=-[(1-\gamma_\lambda^2)^{-\frac{1}{2}}(e^{\beta\lambda}-1)^{-1}]_{\mathrm{Av}}$$

$$=A_l\sum_{n=1}^{\infty}\int e^{-n\lambda/\theta}\lambda^{D-2}d\lambda$$

$$=A_l\Gamma(D-1)\zeta(D-1)\theta^{D-1}. \tag{4.24}$$

In Eqs. (4.23) and (4.24), ζ means Riemann's zeta-function, whose values are

$$\zeta(2)=\pi^2/6,\quad \zeta(3)=1.202,\quad \zeta(4)=\pi^4/90.$$

The results of the integrations are summarized in Table I, where the integrals involved in E_0, M_{S0}, and Δ_0'' are calculated numerically with use of formulas (4.7)–(4.10). Some of the numerical results have been borrowed from Anderson's work. The integrals which are divergent for $\alpha=0$ are evaluated crudely only to the leading terms.

Leaving to the next section the discussions on the nature of the divergences in the limit of vanishing anisotropy, we shall now make some comments on the results. The most interesting point is the temperature dependence of $\chi_{\parallel}$, which is given by

$$\chi_{\parallel}=\frac{4Ng^2\mu_0^2}{3z|J|S}\left(\frac{kT}{z|J|S}\right)^2. \tag{4.25}$$

The perpendicular susceptibility, $\chi_\perp$, has been predicted to be independent of temperature by Eq. (3.9). Therefore, the decrease of $\chi_\perp-\chi_{\parallel}$ with temperature is expected to be proportional to T^2, a fact which has been observed by Griffel and Stout[11] for the case of MnF_2 (see Fig. 1).

It is interesting to note that our theoretical formulas are consistent with experimental data on this crystal. From Eqs. (3.10) and (4.25) we obtain

$$\frac{1}{\chi_\perp}\frac{d(\chi_{\parallel}-\chi_\perp)}{dT^2}=\frac{8}{3}\left(\frac{k}{z|J|}\right)^2\frac{1}{S^3}, \tag{4.26}$$

---

[11] M. Griffel and J. W. Stout, J. Chem. Phys. **18**, 1455 (1950).

TABLE I. Thermodynamical quantities of antiferromagnets calculated by the spin-wave theory.

	Linear chain	Quadratic layer	NaCl-type	CsCl-type
$\frac{E_0}{(N/2)z\|J\|S}$	$S+0.363^{a}$	$S+0.158^{a}$	$S+0.097^{a}$	$S+0.073^{b}$
$\frac{E_T}{Nz\|J\|S}$	$\frac{\pi}{3}\theta^2$	$\frac{4.808}{\pi}\theta^3$	$\frac{3^{3/2}}{15}\pi^2\theta^4$	$\frac{4\pi^2}{15}\theta^4$
$\frac{S_T}{Nk}$	$\frac{2}{3}\pi\theta$	$\frac{7.212}{\pi}\theta^2$	$\frac{3^{3/2}4}{45}\pi^2\theta^3$	$\frac{16}{45}\pi^2\theta^3$
$\frac{C_T}{Nk}$	$\frac{2}{3}\pi\theta$	$\frac{14.424}{\pi}\theta^2$	$\frac{3^{3/2}4}{15}\pi^2\theta^3$	$\frac{16}{15}\pi^2\theta^3$
$\frac{M_{s0}}{M_\infty/S}$	$S+\frac{1}{2}+(1/2\pi)$ $\times\log 2\alpha$	$S-0.197^{a}$	$S-0.078^{a}$	$S-0.075^{b}$
$\frac{M_{sT}}{M_\infty/S}$	$(2\alpha)^{-\frac{1}{2}}\theta$	$\left(\frac{1}{\pi}\log\frac{1}{2\alpha}\right)\theta$	$\frac{3^{\frac{1}{2}}}{2}\theta^2$	$\frac{2}{3}\theta^2$
$\chi_{\shortparallel}\Big/\frac{Ng^2\mu_0^2}{kT}$	$\left(\frac{1}{2\alpha}\right)^{\frac{1}{2}}\theta^2$	$\left(\frac{1}{\pi}\log\frac{1}{2\alpha}\right)\theta^2$	$3^{\frac{1}{2}}\theta^3$	$\frac{4}{3}\theta^3$
$\frac{8\Delta_0''}{Ng^2\mu_0^2}$	$\left(\frac{1}{2\alpha}\right)^{\frac{1}{2}}$	$\frac{1}{\pi}\log\frac{1}{2\alpha}$	finite	0.396^{b}
$\frac{8\Delta_T''}{Ng^2\mu_0^2}$	$2\left(\frac{1}{2\alpha}\right)^{3/2}\theta^2$	$\frac{2}{\pi\alpha}\theta^2$	$\frac{3^{3/2}}{\pi}\frac{1}{(2\alpha)^{\frac{1}{2}}}\theta^2$	$\frac{4}{\pi}\frac{1}{(2\alpha)^{\frac{1}{2}}}\theta^2$

[a] Calculated by Anderson.
[b] Calculated by the present author.

so that the Curie point T_c is correlated with the slope of the plot $\chi_\perp-\chi_{\shortparallel}$ *versus* T^2, by the equation

$$T_c^2=\frac{8}{27}\frac{(S+1)^2}{S}\Big/\frac{1}{\chi_\perp}\frac{d(\chi_{\shortparallel}-\chi_\perp)}{dT^2}, \qquad (4.27)$$

if the Van Vleck formula[6]

$$T_c=\tfrac{1}{3}(z|J|/k)S(S+1)$$

is assumed. The data of Fig. 1 gives $z|J|/k=21.2$°K and $T_c=62$°K, while the specific heat anomaly has been observed at 66.5°K. (If the corrected formula for $\chi_\perp$, Eq. (6.17), is used instead of (3.10), one obtains $z|J|/k=24.2$°K and $T_c=71$°K.) According to Stout,[12] the experimental data on the FeF_2 crystal gives $T_c=113$°K calculated by Eq. (4.27). Experimentally the Curie point of this crystal is not known accurately, but from the magnetic data it is guessed to be around 80°K. These quantitative comparisons, however, are not to be emphasized too much, because the crystal structures and probably the interactions of the magnetic ions in these crystals are different from the simple models we have assumed in this paper.

Equation (4.25) is valid for the temperature range where $kT\ll z|J|S$. At higher temperatures, our simple treatment will be unsatisfactory because the higher terms of the original Hamiltonian will become important. However, if we neglect such complications and assume that the frequency spectrum of the spin waves is approximated by something like the Debye model of lattice vibrations, we expect that $\chi_{\shortparallel}$ will be proportional to T in the higher temperature regions. In this approximation Eq. (4.21) is replaced by

$$[f(\gamma_\lambda^2)]_{\text{Av}}=A_l\int_0^{\lambda_m} f(1-\lambda^2)\lambda^{D-1}d\lambda, \qquad (4.26)$$

with a cut-off wave number λ_m. Equation (4.26) predicts the behavior of $\chi_{\shortparallel}$ for $kT\gtrsim z|J|S$. Of course, this kind of treatment cannot be trusted too much, but it indicates that the increase of $\chi_{\shortparallel}$ will be slower than proportionality to T^2 as the temperature goes higher. Such a trend has also been observed for MnF_2.[11]

The temperature dependence of the quantities listed in Table I is derived on the assumption that $2(z|J|K)^{\frac{1}{2}}S \gg kT$. At low temperatures, where $kT\lesssim 2(z|J|K)^{\frac{1}{2}}S$, all quantities, except Δ'', are expressed in terms of functions $G_{\mu\nu}$ defined by

$$G_{\mu\nu}(\theta')=\frac{\Gamma(\nu+\frac{1}{2})}{\Gamma(\frac{1}{2})}\sum_{n=1}^{\infty}\frac{(2\theta')^\nu}{n^{\mu+\nu}}K_\nu(n/\theta'), \qquad (4.27)$$

where θ' means

$$\theta'=kT/[2(z|J|K)^{\frac{1}{2}}S], \qquad (4.28)$$

and K_ν means the Bessel functions of the second kind with imaginary arguments. For example, we have

$$E=E_0+Nz|J|SA_l2^{D+1}(K/z|J|)^{\frac{1}{2}(D+1)}$$
$$\times\{G_{D,\frac{1}{2}(D-1)}(\theta')+G_{0,\frac{1}{2}(D+1)}(\theta')\},$$
$$F=E_0+Nz|J|SA_lD^{-1}2^{D+1}$$
$$\times(K/z|J|)^{\frac{1}{2}(D+1)}G_{0,\frac{1}{2}(D+1)}(\theta')\},$$
$$M_{ST}=-\tfrac{1}{2}Ng\mu_0(1+\alpha)A_l2^{D-1}$$
$$\times(K/z|J|)^{\frac{1}{2}(D-1)}G_{0,\frac{1}{2}(D-1)}(\theta'),$$
$$\Delta'=\tfrac{1}{4}Ng^2\mu_0^2A_lD^{-1}2^D(K/z|J|)^{D/2}G_{-2,\frac{1}{2}(D+1)}(\theta')/\theta'.$$

At very low temperatures, Eq. (4.27) is approximated by

$$G_{\mu\nu}(\theta')\simeq\tfrac{1}{4}\Gamma(\nu+\tfrac{1}{2})(2\theta')^{\nu+\frac{1}{2}}\exp(-1/\theta'),$$

which gives exponential decreases for all the quantities. Δ_T'' cannot be expressed by the function (4.27), but it behaves in the same way at very low temperatures.

5. DISCUSSIONS ON THE NATURE OF THE DIVERGENCE

Many of the quantities listed in Table I will diverge in the limit of vanishing anisotropy. These divergences are very important, because some of them are usually interpreted as proof that particular types of lattices cannot be ordered. In Table I one can see at once that the divergent character (as $\alpha\rightarrow 0$) increases as one goes downward in a column and leftward in a line.

The situation is almost the same also in the case of the spin-wave theory of ferromagnets, where M_s diverges for $D=1$ and 2, and its fluctuation diverges for all cases. It seems that there is no essential difference between the two, and the spin-wave theory can be

---

[12] Private communication.

applied to antiferromagnetic cases with almost the same rigor as to ferromagnetic cases, though the approximation is certainly somewhat worse for the former.

The divergence of the fluctuations needs to be more carefully examined. As a matter of fact, they are of the order of $N^{4/D}$ ($D=1, 2, 3$), N being the total number of spins under consideration.[13] This estimation is obtained from Eq. (4.19) by the approximation

$$\Delta_T''\simeq\left(\frac{kT}{z|J|S}\right)^2 g^2\mu_0^2\frac{1}{(2\pi)^4}\sum{}'\frac{L^4}{(l_1^2+\cdots+l_D^2)^2}, \tag{5.1}$$

where L is the length of the crystal (which is supposed for simplicity to be a cube) in a proper unit, and the sum is over integral values of $l_1, \cdots l_D$, except $l_1=\cdots=l_D=0$. The same result is obtained also for ferromagnetics, where the fluctuation vanishes at $T=0$.

The extremely large fluctuations for one- and two-dimensional cases may be regarded as strong indications that any ordering, either ferromagnetic or antiferromagnetic, is impossible in these cases, in agreement with the customary interpretation of the divergence of M_s.

For three-dimensional lattices, the spin-wave theory gave finite magnetizations. Their fluctuations are of the order of $N^{4/3}$, which is sufficiently small to define sharp distributions. However, these fluctuations are abnormal in the sense that they are larger than $O(N)$, which should be expected if the magnetizations can really be regarded as thermodynamical variables. Of course, any small anisotropy can change them to normal. But, this seems an important point of the spin-wave theory, so that we shall discuss it a little further.

With regard to this point, one may take one of the following points of view:

(a) The abnormal fluctuations are physically significant, and they are probably related to the free rotation of magnetization vectors. Therefore it is necessary to take account of anisotropy in order to obtain results which are consistent with the thermodynamic point of view. If we do this, the customary spin-wave theory will give good approximations to reality.

(b) The fluctuations should be normal, even in the absence of any anisotropy, so that rigorous calculations will remove the difficulty of abnormal fluctuations if we could take into consideration the higher terms in the Hamiltonian which are omitted in the customary formulation of the spin-wave theory.

If one takes the viewpoint (a), he can derive all the thermodynamical quantities of ferromagnets and antiferromagnets without any divergence difficulty if an anisotropy is assumed, although the results apply only to very low temperatures. The calculation of Holstein and Primakoff[9] is an example, which gave the susceptibility (which is proportional to the fluctuation of the magnetization) of ferromagnets taking into account the dipolar interactions.

However, the anisotropy energy is usually supposed to be so small that it can be treated by a sort of perturbation method. In other words, the spins are aligned by the exchange forces, and the resultant magnetization vector is rotating in the field of the anisotropic force and of the external forces. As a matter of fact, the anisotropy field is a free energy field rather than a simple potential-energy field. All the same, this picture assumes that the magnetization vector is a well-defined quantity with a normal fluctuation. Therefore it seems to be inconsistent with the viewpoint (a) and must be justified, if possible, from the viewpoint (b).

At present we are not ready to decide between these two viewpoints, because, in order to do this, we have to find some mathematical device to deal with the troublesome nonquadratic terms which represent the interaction between the spin waves.† Unfortunately, the usual perturbation method does not work well for this purpose, although it can give useful results in some cases.

It might be worth while to note here the following remarks. If no anisotropy is present, the spin wave of infinite wavelength has a quantum of energy equal to zero, a fact which corresponds to the free rotation of all spins in phase. Consequently the number of quanta belonging to this oscillator is primarily indefinite. This indeterminism, however, has been removed in the usual spin-wave theory, in which the sums over a wave number of the oscillators are calculated assuming

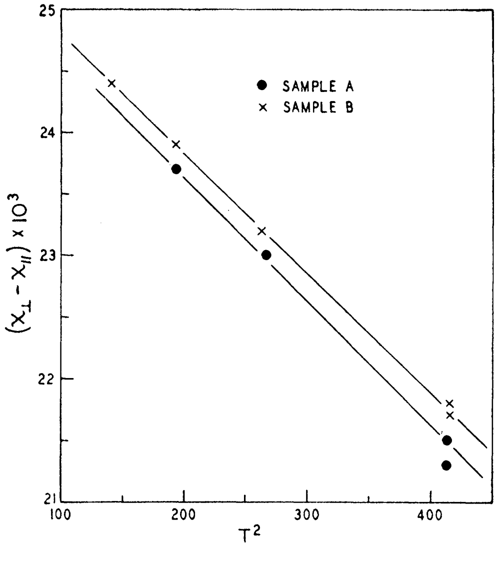

FIG. 1. Molal magnetic anisotropy $\chi_\perp-\chi_\parallel$ of MnF$_2$. Experimental data by Griffel and Stout.

---

[13] This fact was pointed out to the author by P. W. Anderson.

† This point will be discussed in detail in a following paper from the viewpoint of variational principle.

the quantum number of the spin wave with infinite wavelength always to be zero, so that the sums are replaced by the integrals. (Otherwise we should obtain, for instance, any value of the magnetization ranging from M_s to $-\infty$.) Therefore the divergence in the spin-wave theory has nothing to do with the free rotation of the magnetization vectors. Thus it seems probable that the abnormal fluctuations of three-dimensional spin lattices should be attributed to the approximate nature of the spin-wave method.

Evidently the approximation (1.13) makes it possible for any spin wave to assume any large number of quanta, so that the distribution probability of M_s will have too large a tail towards $-\infty$, which may result in an overestimation of the fluctuations. This is also responsible for the failure of the perturbation method, as will be seen in the next section.

The assumption, that $n_\lambda=0$ for $\lambda=0$, may possibly play an important role in any extension of the spin-wave theory to include the nonquadratic terms in the Hamiltonian. In such a theory, the spin waves are no longer independent, and it may happen that the inhibition introduced by this assumption will change the situation essentially and give convergent results which are quite normal thermodynamically, even if no anisotropy is assumed. Thus, refined calculations on this assumption may give magnetizations equal to zero for $D=1$ and presumably for $D=2$, finite magnetizations for $D=3$, and normal fluctuations for all cases.

These considerations seem to throw some doubt on the validity of the results obtained by the spin-wave theory assuming a finite anisotropy. The fluctuations of the magnetizations can be very different from their rigorous values, at least when the actual anisotropy is small.

Our simple results show that even the one-dimensional lattice is antiferromagnetic if the anisotropy is sufficiently large. The same is true also for the ferromagnetic case. This is, of course, very improbable, because a linear chain can be neither ferromagnetic nor antiferromagnetic if the spins are replaced by a semi-classical Ising model, which corresponds to the quantum-mechanical model where the anisotropy is extremely large. The apparent failure of the spin-wave theory in this case can easily be traced. This is also caused by the omission of higher order terms, the most important term in this case being the last term of Eq. (1.26).

According to the spin-wave theory, a two-dimensional lattice is ordered at very low temperatures if a finite anisotropy is present. Although this is not unreasonable, we cannot have much confidence in it. Similarly for the one-dimensional case, the last term of Eq. (1.26) will diminish the tendency of ordering. On the other hand, other higher terms may result in rapid convergence of the magnetizations and the fluctuation. Thus the problem of two-dimensional ordering seems still to be an open question.

As we have seen in the above, we have many questions left for the future which should be solved before we can find how far the spin-wave theory can give trustworthy results. At present we only assume that the convergent results of the spin-wave theory may be physically significant and will be good approximations if they are continuous in the anisotropy parameter.

6. SECOND APPROXIMATIONS

In the preceding section, we have emphasized the importance of higher terms, in the Hamiltonian, which are omitted in the customary formulation of spin-wave theories. They should be included in any satisfactory improvement of the theories. However, it turns out that this cannot be done usually with the perturbation method. Leaving detailed investigations about this point to some other occasion, we shall examine here some of the results obtained by the application of perturbation theory to our problem. In such applications we suffer very often from divergence difficulties, which limit the usefulness of the method. This comes from the fact that our unperturbed Hamiltonian (1.12) is not good enough to give unperturbed wave functions sufficiently concentrated in a region which corresponds to small numbers of spin deviations. Therefore we feel that we need a more refined approach than the ordinary perturbation methods. This is also true even for ferromagnetic cases if one wants to improve the spin-wave theory to apply a little higher temperature.

However, the ordinary perturbation method can give convergent results in some cases, which may be supposed physically significant. So it seems worth while to try higher approximations in this frame. In the following we shall confine ourselves to discussions of the energy and the perpendicular susceptibility $\chi_\perp$.

The classical treatment of spins in the limit of $S\rightarrow\infty$ will be called the zeroth approximation. The first approximation is the approach by the spin-wave theory on the basis of the simplified Hamiltonian (1.20). Taking account of some of the higher terms omitted in the first approximation and applying the first-order perturbation theory, we may obtain the second approximation, provided that the results are convergent. This successive approximation is an expansion in powers of $1/S$, but it cannot be applied always, because it often happens that the results are divergent. Therefore it has only limited applications depending on the dimensionality and also on the nature of the physical quantity in question.

Our perturbation Hamiltonian is obtained from Eq. (1.12), the fourth-order terms of which are

$$H_{\mathrm{ex}}'=f_1|J|S\sum_{\langle jk\rangle}(n_ja_jb_k+a_jn_kb_k+a_j^*n_jb_k+a_jb_k^*n_k) \\ -|J|\sum_{\langle jk\rangle}a_j^*a_jb_k^*b_k. \quad (6.1)$$

Now, let us introduce the Fourier components of the

TABLE II. Ground-state energies of antiferromagnets.

Lattice	z	$-E_{\text{zero}}/(Nz\|J\|S/2)$
Linear chain	2	$S+0.363+0.033S^{-1}$
Quadratic layer	4	$S+0.158+0.0062S^{-1}$
NaCl-type	6	$S+0.097+0.0024S^{-1}$
CsCl-type	8	$S+0.073+0.0013S^{-1}$

creation and annihilation operators by

$$\begin{aligned} a_\lambda &= (2/N)^{\frac{1}{2}}\sum a_j e^{i\lambda j} = \tfrac{1}{2}(X_{1\lambda}+X_{2\lambda}), \\ a_\lambda^* &= (2/N)^{\frac{1}{2}}\sum a_j^* e^{-i\lambda j} = \tfrac{1}{2}(X_{1\lambda}^*+X_{2\lambda}^*), \\ b_\lambda &= (2/N)^{\frac{1}{2}}\sum b_k e^{-i\lambda k} = \tfrac{1}{2}(X_{1\lambda}-X_{2\lambda}), \\ b^* &= (2/N)^{\frac{1}{2}}\sum b_k^* e^{i\lambda k} = \tfrac{1}{2}(X_{1\lambda}^*-X_{2\lambda}^*), \end{aligned} \tag{6.2}$$

where $X_{1\lambda}$, etc., mean

$$X_{1\lambda} = q_{1\lambda}+ip_{1\lambda}, \quad X_{1\lambda}^* = q_{1\lambda}-ip_{1\lambda}, \quad \text{etc.}, \tag{6.3}$$

with q's and p's defined by Eq. (1.23). In terms of these Fourier components Eq. (6.1) is written as

$$\begin{aligned} H_{\text{ex}}' = f_1|J|SN\sum\sum\sum_{\lambda+\kappa=\mu+\nu} \sum_{\rho}[&a_\lambda^* a_\mu a_\nu b_\kappa + a_\kappa b_\lambda^* b_\mu b_\nu \\ &+ a_\mu^* a_\nu^* a_\lambda b_\kappa^* + a_\kappa^* b_\mu^* b_\nu^* b_\lambda]e^{i\kappa\rho} \\ &-(1/N)|J|\sum\sum\sum_{\lambda+\nu=\mu+\kappa}\sum_{\rho} a_\lambda^* a_\mu b_\nu^* b_\kappa e^{i(\kappa-\nu)\rho}. \end{aligned} \tag{6.4}$$

The unperturbed Hamiltonian H^0 is that of Eq. (1.20), with which the averages of operators are defined by

$$\langle A\rangle = \text{trace}[\exp(-H^0/kT)\cdot A]/\text{trace}[\exp(-H^0/kT)]. \tag{6.5}$$

For our calculations the following averages are useful:

$$\langle a_\lambda^* a_\lambda\rangle = \langle b_\lambda^* b_\lambda\rangle = \tfrac{1}{4}\langle q_{1\lambda}^2+p_{1\lambda}^2+q_{2\lambda}^2+p_{2\lambda}^2-2\rangle \tag{6.6a}$$

$$\langle a_\lambda^* b_\lambda\rangle = \langle a_\lambda b_\lambda^*\rangle = 0, \tag{6.6b}$$

$$\langle a_\lambda^2\rangle = \langle b_\lambda^2\rangle = \langle a_\lambda^{*2}\rangle = \langle b_\lambda^{*2}\rangle = 0, \tag{6.6c}$$

$$\langle a_\lambda b_\lambda\rangle = \langle a_\lambda^* b_\lambda^*\rangle = \tfrac{1}{4}\langle q_{1\lambda}^2-p_{1\lambda}^2-q_{2\lambda}^2+p_{2\lambda}^2\rangle. \tag{6.6d}$$

Now, to the first order of H_{ex}', the partition function can be written as

$$\begin{aligned} e^{-F/kT} &= \text{trace}[\exp(-H^0/kT)(1-H_{\text{ex}}'/kT)] \\ &= [\text{trace}(\exp(-H^0/kT))](1-\langle H_{\text{ex}}'\rangle/kT), \end{aligned} \tag{6.7}$$

and the free energy as

$$-F/kT = \log[\text{trace}(\exp(-H^0/kT))] - \langle H_{\text{ex}}'\rangle/kT. \tag{6.8}$$

The derivation of (6.8) from (6.7) is far from satisfactory from the mathematical point of view, but nevertheless Eq. (6.8) is rigorous in the first order of H_{ex}'. Noticing that all products of two operators, such as $a_\lambda a_\mu$, $a_\lambda b_\mu$, $b_\lambda b_\mu$, $a_\lambda^* a_\mu$, and so on, have averages equal to zero if the wave number λ and μ are different, we find easily that

$$\langle H_{\text{ex}}'\rangle = -\tfrac{1}{2}NZ|J|(A+C)^2 + N|J|(1+4f_1S)AC, \tag{6.9}$$

which is correct to $O(N)$. In Eq. (6.9), A and C are defined by

$$\begin{aligned} A &= (2/N)\textstyle\sum_\lambda\langle a_\lambda^* a_\lambda\rangle \\ &= (2/N)\textstyle\sum_\lambda\langle b_\lambda^* b_\lambda\rangle \\ &= S/M_\infty - M_s)/M_\infty, \end{aligned} \tag{6.10}$$

and

$$\begin{aligned} C &= (2/N)\textstyle\sum\gamma_\lambda\langle a_\lambda b_\lambda\rangle \\ &= (2/N)\textstyle\sum\gamma_\lambda\langle a_\lambda^* b_\lambda^*\rangle, \end{aligned} \tag{6.11}$$

for which we have the relation

$$A+C = \frac{\langle H^0{}_{\text{ex}}\rangle + \frac{1}{2}Nz|J|S^2}{2Nz|J|S}. \tag{6.12}$$

If we assume

$$f_1 = -1/4S \tag{6.13}$$

in Eq. (6.9) and insert the expressions of E_0 and E_T given in Table I, we obtain the energy in the second approximation as

$$\begin{aligned} E = -\tfrac{1}{2}Nz|J|S(S+c_0+\tfrac{1}{4}c_0^2S^{-1}) \\ + Nz|J|S(1+\tfrac{1}{2}Dc_0S^{-1})c_1\theta^{D+1}, \end{aligned} \tag{6.14}$$

where c_0 and c_1 are easily found from Table I. The corrected formulas for the ground-state energies are shown in Table II. One can see that the first approximation obtained from the Anderson's spin-wave theory is very good as far as the zero-point energies are concerned. It should be remarked that if we abandon the assumption (6.13) using instead, for instance, the polynomial expression of $f_S(n)$, Eq. (1.14), the last term of Eq. (6.9) diverges always for the one-dimensional case and for the two-dimensional case except at absolute zero, unless a certain anisotropy is assumed.

Finally we shall examine the perpendicular susceptibility $\chi_\perp$ a little further than the first approximation. In this case Eq. (3.4) should be replaced by

$$M_x = M_x^0 + M_x',$$

where

$$M_x^0 = g\mu_0(NS)^{\frac{1}{2}}q_{10}$$

$$\begin{aligned} M_x' = g\mu_0(2S)^{\frac{1}{2}}f_1[&\textstyle\sum_j(n_j a_j + a_j^* n_j) \\ &+ \textstyle\sum_k(n_k b_k + b_k^* n_k)] \end{aligned} \tag{6.15}$$

so that the partition function in the presence of external field H_x is written as

$$\begin{aligned} \text{trace}[\exp\{-(H^0 - H_xM_x^0 + H' - H_xM_x)/kT\}] \\ = \text{trace}[\exp\{-(H^0 - H_xM_x^0)/kT\}] \\ \cdot\{1 - \langle H'\rangle/kT + \langle H_xM_x'\rangle/kT\} \end{aligned}$$

to the first order of H' and M_x'. In terms of the Fourier components (6.2), Eq. (6.15) is written as

$$M_x' = \tfrac{1}{2} g\mu_0 (2S)^{\frac{1}{2}} f_1 (2/N)^{\frac{1}{2}} \sum_{\lambda=\mu+\nu}\sum\sum [a_\lambda^* a_\mu a_\nu + a_\mu^* a_\nu^* a_\lambda b_\lambda^* b_\mu b_\nu + b_\mu^* b_\nu^* b_\lambda]. \quad (6.16)$$

Now, the external field H_x causes the shift of equilibrium position of the oscillator 10 by the amount ϵ, Eq. (3.8). Considering this, we easily see that

$$\langle M_x'\rangle = \tfrac{1}{2} f_1 g\mu_0 (4S/N)^{\frac{1}{2}} 2\{\langle a_0 + a_0^*\rangle \sum \langle a_\lambda^* a_\lambda\rangle + \langle b_0 + b_0^*\rangle \sum \langle b_\lambda^* b_\lambda\rangle\},$$

$$\begin{aligned} H_x \langle M_x'\rangle &= 2 f_1 g\mu_0 H_x (SN)^{\frac{1}{2}} A\epsilon \\ &= 4 f_1 z|J| A\epsilon^2, \end{aligned}$$

and that the part of H_{ex}' Eq. (6.4), which depends on the external field H_x, is given by

$$\langle H_{\text{ex}}'\rangle = -\tfrac{1}{2} z|J|(1-4f_1S)(A+C)\epsilon^2 + \cdots.$$

Thus the correction F' to the free energy is obtained as

$$F' = -\tfrac{1}{2} z|J|\{(1+4f_1S)A + (1-4f_1S)C\}\epsilon^2,$$

which is simplified to

$$F' = -z|J|C\epsilon^2$$

if the assumption (6.13) is made. Hence, with use of Eqs. (6.10), (6.11), (6.12), and Table I, we arrive at the corrected expression of

$$\chi_\perp = \chi_\perp^0 (1 - c_0' S^{-1} - c_1' S^{-1}\theta^2 + c_1 S^{-1}\theta^4), \quad (6.17)$$

where the constants are

$$c_0' = 0.127, \quad c_1' = 3^{\frac{1}{2}}/2, \quad c_1 = 3^{\frac{3}{2}}\pi^2/15 \quad \text{for NaCl-type,}$$

and

$$c_0' = 0.112, \quad c_1' = 2/3, \quad c_1 = 4\pi^2/15 \quad \text{for CsCl-type.}$$

The result (6.17) is of some interest. Experimentally, Griffel and Stout[11] have shown for MnF_2 that $\chi_\perp$ decreases slowly at low temperatures with increasing temperature, while theoretical investigations[8] seem to suggest that the susceptibility at the Curie point will be somewhat lower than the value predicted by Van Vleck theory. Equation (6.17) shows, in agreement with this, that $\chi_\perp$ is very near to the value given by the Van Vleck theory at absolute zero, and that it decreases with increasing temperature.

The author wishes to thank Dr. P. W. Anderson for sending him his unpublished manuscript which has inspired the present work. He is also very grateful to Professor G. Wentzel and Professor J. W. Stout for their stimulating discussions.