Statistical Mechanical Theory of Ferromagnetism. High Density Behavior*

R. BROUT†
Laboratory for Atomic and Solid-State Physics and Department of Physics, Cornell University, Ithaca, New York
(Received November 30, 1959)

The partition function of the Ising model of ferromagnetism is examined in the limit of high density in the anticipation that in the limit of infinite density one recovers the Weiss molecular field. The formal parameter of expansion is $1/z$ where z is the number of spins in the range of the exchange potential (not restricted to nearest neighbor interactions). In the absence of long-range order, only ring diagrams in the cluster expansion contribute. These give a divergence in the specific heat at $kT_c=\Sigma_{j\neq i} v_{ij}$ where v_{ij} is the exchange potential. This is the molecular field value for the Curie point T_c. In the presence of a magnetic field the partition function is evaluated for fixed magnetic moment M in the same approximation, M being determined by minimization. This results in a susceptibility differing from the molecular field theory and hence an inconsistency in the theory.

The inconsistency is traced back to the observation that the acceptance of ring diagrams is equivalent to the gaussian model of Kac and Berlin which violates the sum rule $\Sigma_{i=1}^{N}\mu_i^2=N$. Here μ_i is the "spin" per particle and N is the total number of particles. This condition is reinstated by insuring the sum rule. The result leads to the spherical model. Thus, a consistent high density approximation to the Ising model is the spherical model. Below the Curie point or for fixed magnetic field, M is again held fixed and only the Fourier components of the spin density with nonvanishing wave vector are "sphericalized." The result leads to a physically acceptable model which becomes the molecular field theory at low temperatures or high fields and deviates in $O(1/z)$ in general. Formally, the results are simply expressed in terms of a temperature dependent Weiss field. These results differ from the ordinary spherical model which is physically unacceptable below the Curie point. However, a molecular field modification of the spherical model due to M. Lax yields the same result when properly interpreted.

It is shown that the above results are also valid (to the same approximation) in the quantum mechanical Heisenberg model, for temperatures above the Curie point.

I. INTRODUCTION

THE molecular field theory of ferromagnetism founded by Weiss has all in all been remarkably successful in the interpretation of ferromagnetism.[1] The origin of the molecular field is, as is well known, due to the exchange interaction among spins. However, in order to get the molecular field theory out of an exchange interaction model, seemingly drastic approximations must be made in the statistical mechanics of the model. Various refinements introducing short-range order have been introduced in the past by Bethe and others,[2] and exact moment expansion have also been used to attack the statistical mechanical problem.[3] In all of these theories, it is always remarked that in the limit that the exchange potential becomes very long range ($z\rightarrow\infty$ where z is the number of spins in the range of the exchange interaction), the molecular field theory is recovered as a limit. This indicates that the evaluation of the partition function could be studied with profit by adopting $1/z$ as an expansion parameter —or alternatively to develop the free energy in a power series in the inverse density.

It is the purpose of this paper to present the leading (first two) terms in such an expansion through the use of a systematic cluster development of the free energy developed by the author in the study of random ferromagnets.[4] In the present paper, the Ising model is studied and some quantum mechanical results are given. The full quantum theory is reserved for a subsequent study where the present work will be related to spin waves.

Straightforward evaluation of terms of $O(1)$ and $O(1/z)$ amounts to a summation over ring diagrams in the cluster expansion. This is akin to the random phase approximation in the theory of correlation due to Coulomb forces.[5] In the present case one finds a singularity in the specific heat at that temperature given by the Curie point calculated from molecular field theory. However, the susceptibility remains finite at this point having a singularity at some lower temperature. Moreover, the theory in this lower temperature range is undefined. It is therefore seen that the simple random phase approximation fails.

It is a simple matter to discover the difficulty. We define μ_i as the z component of the ith spin, (taking on the value ±1), together with the Fourier transform

$$\mu(\mathbf{q})=(1/\sqrt{N})\sum_{\mathbf{i}}\mu_i\exp(i\mathbf{q}\cdot\mathbf{i}), \qquad (1.1)$$

where $\mathbf{q}$ is a wave vector in the first Brillouin zone, $\mathbf{i}$ the position of the ith spin. The lattice spacing is unity. From (1.1) follows the sum rule

$$\sum_{\mathbf{q}}|\mu_{\mathbf{q}}|^2=N. \qquad (1.2)$$

Now, it is the essence of the random phase approximation to treat each Fourier component $\mu_{\mathbf{q}}$ independ-

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

† Part of this work was done by the author as a consultant for the Bell Telephone Laboratories.

[1] See, for example, C. Kittel, *Introduction to Solid-State Physics* (John Wiley & Sons, New York, 1956), 2nd ed., Chap. XV. Also W. Bragg and E. Williams, Proc. Roy. Soc. (London) **A145**, 699 (1934); W. Heisenberg, Z. Physik **49**, 619 (1928).

[2] H. A. Bethe, Proc. Roy. Soc. (London) **A166**, 552 (1935).

[3] J. G. Kirkwood, J. Chem. Phys. **6**, 70 (1938), W. Opechowski, Physica **4**, 181 (1937).

[4] R. Brout, Phys. Rev. **115**, 824 (1959), hereafter referred to as I.

[5] J. E. Mayer, J. Chem. Phys. **18**, 1426 (1950); M. Gell-Mann and K. Brueckner, Phys. Rev. **106**, 364 (1957); N. Hugenholz, Physica **23**, 533 (1957); K. Sawada, K. A. Brueckner, N. Fukuda, and R. Brout, Phys. Rev. **108**, 507 (1957).

ently which, therefore, violates Eq. (1.2). In order to repair this damage, we still maintain the essence of the random phase approximation but impose (1.2) as a restraint. This amounts to the introduction of a LaGrange multiplier or a new parameter in the theory which is determined by (1.2). It then turns out that above the Curie point one recovers the spherical model[6] which theory is consistent in that the susceptibility and specific heat anomalies occur at a common temperature. This theory is presented in a different notation from reference 6, in such manner that the analogy to molecular field theory is made clear. It also becomes obvious that as z becomes large one approaches molecular field theory.

Below the Curie point, we develop the theory for a fixed long-range order which is fixed by minimization of the free energy. It then follows that the molecular field theory is correct at low temperatures (for the Ising model), deviations setting in as one approaches the Curie point. The specific heat turns out to be continuous at the Curie point having a cusp there which is approached by a function having an infinite derivative.

Above the Curie temperature these results are quantum mechanically valid in the same approximation as the classical theory. Below the Curie point, the situation is not clear and further work remains to establish the correspondence with spin-wave theory.

II. GENERAL CLUSTER DEVELOPMENT

In this section, we review for completeness the cluster development of the Ising model presented in I, specialized in the present case to a fixed array of spins on a lattice. For variety, we present a slightly different version of the derivation which may further one's understanding of the existence of the "dotted line" diagrams in I.

We first work out the complete partition function with no restriction to long-range order. We must evaluate

$$\ln Z=\ln \sum_{\{\mu_i\}} \exp[\beta \sum_{1\leq i<j\leq N} v_{ij}\mu_i\mu_j], \tag{2.1}$$

$\mu_i=\pm 1$ and the summation over $\{\mu_i\}$ means that one sums over all configurations of μ_i's. v_{ij} is an arbitrary function of $\mathbf{i}-\mathbf{j}$ whose Fourier transform is assumed to exist. As it is convenient to use the notation of probability theory, Eq. (2.1) is rewritten

$$\ln Z=\ln\langle \exp\beta \sum v_{ij}\mu_i\mu_j\rangle+N\ln 2, \tag{2.2}$$

where the symbol $\langle O(\{\mu_i\})\rangle$ means

$$\langle 0(\{\mu_i\})\rangle\equiv(1/2^N)\sum_{\{\mu_i\}} O(\{\mu_i\}). \tag{2.3}$$

Thus we regard $\sum v_{ij}\mu_i\mu_k$ as a random variable distributed according to the law that the μ_i's are a set of N independent random variables of value ± 1 chosen with equal probability. Equation (2.2) is of the form $\ln\langle\exp\beta X\rangle$ where X is a random variable. The expansion of this function in a power series in β is called the semi-invariant expansion. The defining equation of the M_n's is

$$\ln\langle\exp\beta X\rangle=\sum[(\beta)^n/n!]M_n. \tag{2.4}$$

The M_n's have the following simple, but extremely important property. If X, Y are two *independent* random variables and $M_n^{(X)}$, $M_n^{(Y)}$, and $M_n^{(X+Y)}$ are the semi-invariants generated by X, Y, and $X+Y$, respectively, then

$$M_n^{(X+Y)}=M_n^{(X)}+M_n^{(Y)}, \tag{2.5}$$

We now consider the first few M_n's

$$M_1=\sum_{i<j}\langle\mu_i\mu_j\rangle=0. \tag{2.6}$$

This result obtains because any odd power of μ_i has an average $=0$ and the μ_i's are independent.

$$M_2=\sum_{i<j}\sum_{k<l} v_{ij}v_{kl}[\langle\mu_i\mu_j\mu_k\mu_l\rangle-\langle\mu_i\mu_j\rangle\langle\mu_k\mu_l\rangle]. \tag{2.7}$$

The second term on the right-hand side of Eq. (2.7) is zero for the same reason that $M_1=0$. The first term is nonvanishing only if two indices are paired. Hence

$$M_2=\sum_{i<j}(v_{ij})^2\langle\mu_i\mu_j\mu_i\mu_j\rangle=\sum_{i<j}(v_{ij})^2. \tag{2.8}$$

We diagram this term in Fig. 1(a). The notation is the same as in I. M_3 follows the same pattern. The only nonvanishing term is the cycle term given in Fig. 1(b).

Now consider M_4. Two obvious diagrams are the cycle given in Fig. 1(c) and the ladder in Fig. 1(d), corresponding to which are the semi-invariants

$$M_4^{\text{cycle}}=\sum v_{ij}v_{jk}v_{kl}v_{li},$$

$$M_4^{\text{ladder}}=\sum v_{ij}{}^4[\langle\mu_i{}^4\mu_j{}^4\rangle-3\langle\mu_i{}^2\mu_j{}^2\rangle^2]=-2\sum v_{ij}{}^4.$$

We must also consider the reducible cluster given by

$$\sum_{i,k}(v_{ij})^2(v_{jk})^2[\langle\mu_i{}^2\mu_j{}^2\mu_j{}^2\mu_k{}^2\rangle-\langle\mu_i{}^2\mu_j{}^2\rangle\langle\mu_j{}^2\mu_k{}^2\rangle]=0. \tag{2.9}$$

Thus reducible as well as unlinked clusters vanish when a complete trace is taken. This is in contrast to the case where long-range order is present as discussed in I.

Let us now go back to the diagram of Fig. 1(c). Here there is the combination

$$\sum v_{ij}v_{jk}v_{kl}v_{li}, \tag{2.10}$$

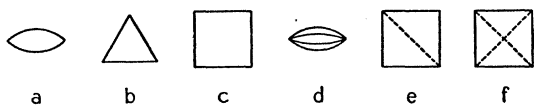


FIG. 1. Cluster diagrams contributing to $\ln Z$ in the absence of long-range order.

[6] T. Berlin and M. Kac, Phys. Rev. 86, 821 (1952). See also the general review article on the Ising model by G. F. Newell and E. Montroll, Revs. Modern Phys. 25, 353 (1953).

where it is specified that none of the indices i, j, k, l may be equal. As this is an inconvenience on summation, add and subtract the terms with $k=i$ or $j=l$, or both. Now sums are free and the new terms are given by the diagrams Fig. 1(e) and Fig. 1(f). Here we have inserted the dotted line convention of I. A dotted line symbolizes the short-range *spatial* function $(-\delta_{ij})$. All sums over indices in a diagram are now free (except that the indices attached by a bond are unequal).

The complete diagrammatic description is then the following. An even number of solid bonds emanate from each vertex. All irreducible linked graphs are allowed as such [Figs. 1(a), (b), (c), (d)], as well as all graphs generated from them by dotted line insertions [Figs. 1(e), (f)]. No reducible linked graphs or unlinked graphs occur. To each graph there corresponds an M_n which in general is not trivial and must be worked out in each case.

Before studying this cluster development in detail, we first turn to the corresponding expansion in the presence of long-range order. For this case we are interested in the expression

$$\ln Z_R=\ln \sum_{\{\mu_i\}}{}' \exp[\beta \sum_{i<j} v_{ij}\mu_i\mu_j], \tag{2.11}$$

where the prime over $\sum$ means a sum over all configurations $\{\mu_i\}$ such that

$$(1/N)\sum \mu_i=R. \tag{2.12}$$

R is taken to be a number of order unity. Note also that $-1\leq R\leq +1$. Equation (2.11) is recast into convenient form

$$\ln Z_R=\ln\langle \exp \sum_{i<j} v_{ij}\mu_i\mu_j\rangle_R+\ln W(R), \tag{2.13}$$

where

$$W(R)=\binom{N}{\frac{1}{2}N(1+R)}, \tag{2.14}$$

$$\langle O(\{\mu_i\})\rangle_R=\frac{1}{W(R)}\sum_{\{\mu_i\}}{}' O(\{\mu_i\}). \tag{2.15}$$

In the form (2.13) we may now use the semi-invariant expansion. As the steps leading to the diagrams are the same as before we shall not present the details here. For further exposition the reader is referred to I. In Fig. 2, all nonvanishing diagrams are included up to M_3 inclusive. These diagrams are more numerous than for the case of vanishing long-range order, as now one may have an odd number of bonds emanating from a vertex and reducible graphs of solid bonds do occur if they are irreducibly connected by dotted lines. In addition each diagram represents a more complicated expression than for the $R=0$ case since the entire semi-invariant now enters. For some detailed examples, see the appendix of I.

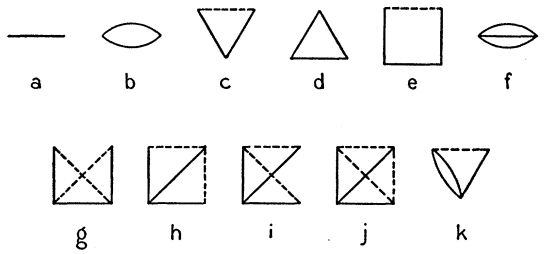


FIG. 2. Cluster diagrams contributing to $\ln Z$ in the presence of long-range order.

III. HIGH DENSITY LIMIT

We first give the argument why the molecular field theory is a high density limit. The energy is

$$\tfrac{1}{2}\sum_{i\neq j} v_{ij}\mu_i\mu_j. \tag{3.1}$$

In the case where v_{ij} is very long range we may consider the following limiting process. Let the number of spins z in the range of the potential grow without limit and let the strength of the potential v_{ij} decrease like $1/z$. (This latter in order to keep the Curie point fixed.) Also, for convenience we take the potential to be a constant (J/z) up to the end of its range and zero thereafter. Then the limit of (3.1) as $z\rightarrow N$ is

$$\frac{1}{2}\frac{J}{N}\sum_{i\neq j}\mu_i\mu_j=\frac{1}{2}\frac{J}{N}(\sum\mu_i)(\sum\mu_j)+O(1)$$
$$=\frac{N}{2}JR^2=\frac{N}{2}R\cdot H_{\text{mol}}, \tag{3.2}$$

where JR is the molecular field, H_{mol}; and we have used (2.12).

The qualitative argument given above is the motivation for choosing an expansion parameter, the quantity $(1/z)$. It is then anticipated that the rigorous leading term is the molecular field theory and that corrections should give a reasonable qualitative description of the expected deviations, viz., short-range order. Furthermore, in real ferromagnets z is indeed large at least of $O(10)$ so that an expansion parameter of ~ 0.1 is quite appropriate. Indeed, it would be rather surprising if the qualitative as well as semiquantitative features of this work did not correspond to reality when the correct choice of v_{ij} is made and the full exchange coupling is used.

We now classify graphs according to their order of magnitude in $1/z$. We first consider $R=0$ and estimate the order of magnitude of diagrams in Fig. 1. As the temperature scale is entirely determined by the strength of the potential we have in the region of the Curie point

$$\beta zJ=O(1);\quad \beta J=O(1/z), \tag{3.3}$$

where J characterizes the strength of the potential. Thus every solid bond in a graph is $O(1/z)$. We then see that there are no terms of $O(1)$ in Fig. 1. The

leading terms are all of $O(1/z)$. In particular, Fig. 1(a) has two solid bonds and one summation and hence is $O[(\beta J)^2 z]=O(1/z)$. Figures 1(b) and (c) are also $O(1/z)$ whereas Figs. 1(d) and (f) are $O(1/z^3)$ and Fig. 1(e) of $O(1/z^2)$.

It is then seen that the leading terms are of $O(1/z)$ and that these are the cycle graphs. We now proceed to their detailed evaluation. We remark that the same diagrams correspond to the Debye Hückel theory in the case of Coulomb forces so that the same techniques apply.[5]

In particular the contribution to M_n from closed cycles is

$$g_n \sum_{i_1 \cdots i_n} v_{i_1 i_2} v_{i_2 i_3} \cdots v_{i_n i_1},$$

where g_n is a combinatorial factor and the spin factor is obviously unity. To evaluate g_n we fix one index in a cycle, say i, and ask how many distinct ways there are to arrange the remaining $(n-1)$ indices which give distinct graphs. (By a distinct graph is meant one which contains different factors of v_{ij} but which is still a cycle, e.g., $v_{12}v_{23}v_{34}v_{41}$ is distinct from $v_{12}v_{24}v_{43}v_{31}$ but is not distinct from $v_{23}v_{34}v_{41}v_{12}$ or $v_{14}v_{43}v_{32}v_{21}$.) The answer to this question is $(n-1)!/2$ since $(n-1)!$ is the number of permutations of the $(n-1)$ particles, but half of these permutations do not produce distinct graphs since they correspond to the symmetric permutation about the bisecting diagonal of a graph (this diagonal starting at i_1). The net result is

$$(M_n)_{\text{from cycles}}=N\frac{(n-1)!}{2}\sum_{i_2 \cdots i_n} v_{i_1 i_2} v_{i_2 i_3} \cdots v_{i_n i_1}. \quad (3.4)$$

The intermediate indices in (3.4) may be allowed to overlap previous indices since this contingency has been allowed for in the subtraction procedure that led to dotted line graphs. This is a very important observation and it is the central reason for the introduction of the dotted line graphs. It goes without saying that the indices on a single v_{ij} are always unequal, i.e., $i_\nu \neq i_\nu+1$ in Eq. (3.4).

Using Eq. (3.4) we then have from Eqs. (2.2) and (2.4)

$$\frac{1}{N}\ln Z=\tfrac{1}{2}\sum_{n=2}^{\infty}\left\{\frac{\beta^n}{n}\sum_{i_2 \cdots i_n}[v_{i_1 i_2} \cdots v_{i_n i_1}]\right\} +\ln 2+O\left(\frac{1}{z^2}\right). \quad (3.5)$$

We now turn to the evaluation of the nth term in (3.5). The central remark is that if v_{ij} is considered as a matrix then the cyclic combination that appears in M_n is a matrix product. If the eigenvalues of the matrix v_{ij} are known then the evaluation is trivial. However, by translational symmetry of the lattice, v_{ij} is a function of $(\mathbf{i}-\mathbf{j})$ alone. Therefore its eigenvectors are $\exp(i\mathbf{q}\cdot\mathbf{j})$, with eigenvalue $v(\mathbf{q})$ where $v(\mathbf{q})$ is the Fourier transform of v_{ij}. Here $\mathbf{q}$ is a vector in the reciprocal lattice space and may be taken to be restricted to the first Brillouin zone. For simplicity we have assumed one spin per unit cell. The net result is that Eq. (3.5) may be written

$$\frac{1}{N}\ln Z=\frac{1}{2N}\sum_{n=2}^{\infty}\sum_{\mathbf{q}}\frac{1}{n}[\beta v(\mathbf{q})]^n+\ln 2$$

$$=-\frac{1}{2N}\sum_{\mathbf{q}}\ln[1-\beta v(\mathbf{q})]+\ln 2+O\left(\frac{1}{z^2}\right), \quad (3.6)$$

where

$$v(\mathbf{q})=\sum_{j\neq i} v_{ij}\exp[i\mathbf{q}\cdot(\mathbf{i}-\mathbf{j})]. \quad (3.7)$$

$\mathbf{q}$ is in the first zone and we have used the result $\sum_{\mathbf{q}} v(\mathbf{q})=0$. Further, we have assumed that the sum on n and $\mathbf{q}$ are interchangeable. A detailed investigation of (3.6) is given in Sec. IV.

We now turn to the case where R is fixed. R is considered to be a constant of $O(1)$. Then there exists a single diagram of $O(1)$ which is Fig. 2(a). No other diagram exists of $O(1)$. Hence, the total contribution to $\ln Z$ of $O(1)$ is

$$\sum_{i<j} v_{ij}\langle \mu_i \mu_j\rangle=\frac{N}{2}v(0)R^2, \quad (3.8)$$

where $v(0)$ is defined by (3.7) for $\mathbf{q}=0$ and R^2 is assumed to be $O(1)$. Equation (3.8) is the same as (3.2) as expected.

Terms of $O(1/z)$ are given by Figs. 2(b), (c), (d), and (e). Generalizing to arbitrary order, one sees that included in $O(1/z)$ are all closed cycles of solid bonds as well as all open chains of solid bonds whose ends are attached by a dotted line. In nth order there are obviously n more of the latter than of the former corresponding to the n places where a dotted line may be inserted into a closed cycle of n bonds. The contribution to $M_n/n!$ from these diagrams is then

$$N\frac{(n-1)!}{2}\Big\{\sum_{i_2 \cdots i_n}[v_{i_1 i_2} \cdots v_{i_n i_1}]S_{n;\mathrm{cl}} -n\sum_{i_2 \cdots i_{n+1}}[v_{i_1 i_2} \cdots v_{i_n i_{n+1}}\delta_{i_{n+1} i_1}]S_{n;\mathrm{op}}\Big\}, \quad (3.9)$$

where $S_{n;\mathrm{cl}}$ and $S_{n;\mathrm{op}}$ are the spin contributions to the nth semi-invariant from the spins arranged in the closed and open chains, respectively. The spatial part is the same as in (3.5) and hence we have

$$N\frac{(n-1)!}{2}\sum_{\mathbf{q}}[v(\mathbf{q})]^n[S_{n;\mathrm{cl}}-nS_{n;\mathrm{op}}]. \quad (3.10)$$

For orientation we calculate the S_n's for the first few n's

$$S_{2;\mathrm{cl}}=\langle(\mu_1\mu_2)^2\rangle-\langle\mu_1\mu_2\rangle^2=1-R^4,$$

$$S_{2;\mathrm{op}}=\langle\mu_1\mu_2\mu_2\mu_3\rangle-\langle\mu_1\mu_2\rangle\langle\mu_2\mu_3\rangle=R^2(1-R^2),$$

$$S_{2;\mathrm{cl}}-2S_{2;\mathrm{op}}=(1-R^2)^2,$$

$$S_{3;\mathrm{cl}}=\langle\mu_1\mu_2\mu_2\mu_3\mu_3\mu_1\rangle-3\langle\mu_1\mu_2\rangle\langle\mu_2\mu_3\mu_3\mu_1\rangle+2\langle\mu_1\mu_2\rangle\langle\mu_2\mu_3\rangle\langle\mu_3\mu_1\rangle$$
$$=1-3R^4+2R^6,$$

$$S_{3;\mathrm{op}}=\langle\mu_1\mu_2\mu_2\mu_3\mu_3\mu_4\rangle-2\langle\mu_1\mu_2\rangle\langle\mu_2\mu_3\mu_3\mu_4\rangle-\langle\mu_2\mu_3\rangle\langle\mu_1\mu_2\mu_3\mu_4\rangle+2\langle\mu_1\mu_2\rangle\langle\mu_2\mu_3\rangle\langle\mu_3\mu_1\rangle$$
$$=R^2-2R^4-R^6+2R^6=R^2(1-R^2)^2,$$

$$S_{3;\mathrm{cl}}-3S_{3;\mathrm{op}}=(1-R^2)^3.$$

We shall now prove that $S_{n;\mathrm{cl}}-nS_{n;\mathrm{op}}=(1-R^2)^n$. This may be done in two ways: by recursion relations and induction or by trickery. We give the latter since it will turn out to be essential to further understanding We assert that the expression

$$\ln\langle\exp\beta v(\mathbf{q})\mu_{\mathbf{q}}{}^*\mu_{\mathbf{q}}\rangle, \tag{3.13}$$

where

$$\mu_{\mathbf{q}}=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mu_i\exp(i\mathbf{q}\cdot\mathbf{i})$$

is the same as $\sum[(\beta)^n/n!][M_n]_{\mathrm{cyc}}{}^{(\mathbf{q})}$ where $[M_n]_{\mathrm{cyc}}{}^{(\mathbf{q})}$ is given by the $\mathbf{q}$th term in Eq. (3.10). To see this, first expand (3.13) in β

$$\ln\langle\exp\beta v(\mathbf{q})\mu_{\mathbf{q}}{}^*\mu_{\mathbf{q}}\rangle=\textstyle\sum_n[\beta v(\mathbf{q})]^n/n!M_n{}^{(\mu_{\mathbf{q}}{}^*\mu_{\mathbf{q}})}. \tag{3.14}$$

Using the definition of $\mu_{\mathbf{q}}$ and M_n we have by direct substitution

$$M_n{}^{(\mu_{\mathbf{q}}{}^*\mu_{\mathbf{q}})}=1/N^n\sum_{\substack{i_1\cdots i_n\\ j_1\cdots j_n}}\{\langle\mu_{i_1}\mu_{j_1}\cdots\mu_{i_n}\mu_{j_n}\rangle-\langle\mu_{i_1}\mu_{j_1}\rangle\langle\mu_{i_2}\mu_{j_2}\cdots\mu_{i_n}\mu_{j_n}\rangle-\cdots+(-1)^{n-1}n\langle\mu_{i_1}\mu_{j_1}\rangle\cdots\langle\mu_{i_n}\mu_{j_n}\rangle\}\times\exp[i\mathbf{q}\cdot(\mathbf{i}_1-\mathbf{j}_1+\cdots+\mathbf{i}_n-\mathbf{j}_n)]. \tag{3.15}$$

The exact structure of sums and differences in Eq. (3.15) is unimportant. What is important is that the unlinked terms that come into (3.15) are cancelled out by the semi-invariant structure, by virtue of (2.5). We must then examine the linked terms only. Of these, the closed cycles of which there are $N^n\times(n-1)!/2$ give $S_{n;\mathrm{cl}}$. Open chains give

$$\frac{n!}{2}S_{n;\mathrm{op}}\frac{1}{N^n}\sum_{\substack{i_1\cdots i_nj_n\\ j_n\neq i_1}}\exp[i\mathbf{q}\cdot(\mathbf{i}_1-\mathbf{j}_n)]$$
$$=\frac{n!}{2}S_{n;\mathrm{op}}\sum_{j\neq i}\exp[i\mathbf{q}\cdot(\mathbf{i}-\mathbf{j})]. \tag{3.16}$$

The inequality $j\neq i$ arises from the fact that we are taking open chains. Now $\sum_j\exp[i\mathbf{q}\cdot(\mathbf{i}-\mathbf{j})]=0$ for $q\neq0$ and hence $\sum_{j\neq i}\exp[i\mathbf{q}\cdot(\mathbf{i}-\mathbf{j})]=-1$. Therefore, (3.16) contributes $-(n!/2)S_{n;\mathrm{op}}$. Thus, (3.10) is contained in (3.14). Finally, we point out that there are no other linked contributions in (3.15) since this would involve contraction over more than n indices and hence would give terms of $O(1/N)$. Recalling the definition (2.15) we thus have shown that[7]

$$\ln Z_R=\frac{N}{2}v(0)R^2+\tfrac{1}{2}\sum_{\mathbf{q}\neq0}\ln\langle\exp\beta v(\mathbf{q})\mu_{\mathbf{q}}{}^*\mu_{\mathbf{q}}\rangle_R+\ln W(R)+O(1/z^2). \tag{3.17}$$

Equation (3.6) is included in (3.17) by putting $R=0$ and including $q=0$ in the second term.

Equation (3.17) is remarkable when it is noticed that the original problem is the evaluation of

$$\ln\langle\exp\tfrac{1}{2}\textstyle\sum_{\mathbf{q}}\beta v(\mathbf{q})\mu_{\mathbf{q}}{}^*\mu_{\mathbf{q}}\rangle.$$

In fact we have

$$\ln\left\langle\exp\frac{\beta}{2}\sum_{\mathbf{q}}v(\mathbf{q})\mu_{\mathbf{q}}{}^*\mu_{\mathbf{q}}\right\rangle=\tfrac{1}{2}\sum_{\mathbf{q}}\ln\langle\exp\beta v(\mathbf{q})\mu_{\mathbf{q}}{}^*\mu_{\mathbf{q}}\rangle+O\left(\frac{1}{z^2}\right). \tag{3.18}$$

Equation (3.18) is the analog of the random phase approximation in the Coulomb problem[5] and it arises for the same reasons. The high density limit (restriction to ring diagrams) effectively decouples the various Fourier components of the (spin) density fluctuation.

The evaluation of $\langle\exp\beta v(\mathbf{q})\mu_{\mathbf{q}}{}^*\mu_{\mathbf{q}}\rangle$ is effected by the calculation of the moments $\langle(\mu_{\mathbf{q}}{}^*\mu_{\mathbf{q}})^n\rangle$. This is carried out in the Appendix where it is shown that

$$\langle(\mu_{\mathbf{q}}{}^*\mu_{\mathbf{q}})^n\rangle=n!\langle\mu_{\mathbf{q}}{}^*\mu_{\mathbf{q}}\rangle^n, \tag{3.19}$$

$$\langle\mu_{\mathbf{q}}{}^*\mu_{\mathbf{q}}\rangle=1-R^2\equiv\epsilon. \tag{3.20}$$

Equation (3.20) is a consequence of the sum rule $\sum_{\mathbf{q}}\langle\mu_{\mathbf{q}}{}^*\mu_{\mathbf{q}}\rangle=N$ when it is recognized that for $q\neq0$, $\langle\mu_{\mathbf{q}}{}^*\mu_{\mathbf{q}}\rangle$ is independent of $\mathbf{q}$ and $|\mu_0|^2=NR^2$, Hence we get

$$\langle\exp\beta v(q)\mu_{\mathbf{q}}{}^*\mu_{\mathbf{q}}\rangle=\frac{1}{1-\beta\epsilon v(\mathbf{q})}. \tag{3.21}$$

Substituting into Eq. (3.18), gives

$$\ln Z_R=\frac{N}{2}v(0)R^2-\tfrac{1}{2}\sum_{\mathbf{q}\neq0}\ln[1-\beta\epsilon v(\mathbf{q})]+\ln W(R), \tag{3.22}$$

where we have again used $\sum v(\mathbf{q})=0$.

Equation (3.22) is the central result of this section. However, we should also call attention to Eq. (3.17) as an interesting by-product of the analysis leading to (3.22). We point out parenthetically that for the problem of Coulomb correlation in the Debye-Hückel

[7] Equation (3.17) may be rewritten more concisely by using $\mu_0=(N)^{\frac{1}{2}}R$

$$\ln Z_R=\tfrac{1}{2}\textstyle\sum_{\mathbf{q}}\ln\langle\exp\beta v(\mathbf{q})\mu_{\mathbf{q}}{}^*\mu_{\mathbf{q}}\rangle+\ln W(R).$$

limit one makes the approximation

$$\ln Z \equiv \ln(1/\Omega)^N \int \exp(-\beta \sum v_{ij}) d\mathbf{r}_i \cdots d\mathbf{r}_n$$

$$\rightarrow \tfrac{1}{2} \sum_{\mathbf{q}} \ln\langle \exp\beta v(\mathbf{q}) \rho_{\mathbf{q}} {}^* \rho_{\mathbf{q}} \rangle,$$

where

$$\rho_{\mathbf{q}} = \sum_i e^{i\mathbf{q}\cdot\mathbf{r}_i} \quad \text{and} \quad v(q) = \int v(r_{ij}) e^{i\mathbf{q}\cdot\mathbf{r}_{ij}} d\mathbf{r}_{ij}$$

$$\langle (\rho_q {}^* \rho_q)^n \rangle = n!. \tag{3.23}$$

IV. THERMODYNAMIC CONSEQUENCES AND DISCUSSION

We first investigate our results in the absence of a magnetic field with the assumption of no long-range order. Using Eq. (3.6) we get

$$\frac{E}{N}\beta = -\frac{\beta}{N}\frac{\partial \ln Z}{\partial \beta} = -\frac{1}{2N}\sum_{\mathbf{q}} \frac{\beta v(\mathbf{q})}{1-\beta v(\mathbf{q})}, \tag{4.1}$$

and the specific heat is

$$\frac{C_v}{Nk} = \frac{1}{Nk}\frac{\partial E}{\partial T} = -\frac{\beta^2}{N}\frac{\partial E}{\partial \beta} = \frac{1}{2N}\sum_q \frac{[\beta v(\mathbf{q})]^2}{(1-\beta v(\mathbf{q}))^2}. \tag{4.2}$$

The functions arising in (4.1) and (4.2) have been studied by Lax[8] and Kac and Berlin[6] for nearest neighbor interactions. One introduces the function $F(x)$ defined by

$$F(x) = \sum_{\mathbf{q}} \frac{1}{x - v(\mathbf{q})/v(0)}. \tag{4.3}$$

In this notation we find

$$E/Nv(0) = -[x^2 F(x) - x], \tag{4.4}$$

$$C_v/Nk = -[x^2 F'(x) - 2xF(x) + 1], \tag{4.5}$$

where $x = 1/\beta v(0)$. x is the temperature measured in units of $v(0)$. $F(x)$ is plotted in reference 8, Fig. 2.

From the forms (4.1) and (4.2) we see immediately as $\beta \rightarrow 0$ $(T \rightarrow \infty)$ that $E \rightarrow 0$ since $\sum v(q) \rightarrow 0$. Similarly $C_v \rightarrow 0$ like β^2. This corresponds to the breaking up of short-range order as discussed below. From (4.2) we see that C_v becomes singular at $\beta_c v(q_m) = 1$ where $\mathbf{q}_m$ is the value of $\mathbf{q}$ for which $v(\mathbf{q})$ is a maximum. In a ferromagnet $\mathbf{q}_m = 0$. In an antiferromagnet $\mathbf{q}_m$ is an extreme point at the end of the Brillouin zone. [E.g., in a cubic lattice $\mathbf{q}_m$ is at the corner (π,π,π) if one of the cube corners is chosen at 0.] Intermediate values of $\mathbf{q}_m$ are also possible in some models (see Anderson and Suhl).[9] In this paper we shall only discuss the usual ferromagnet with $\mathbf{q}_m = 0$. For this case C_v blows up at the temperature T_c where

$$kT_c = v(0). \tag{4.6}$$

It is thus seen that retention of all terms in $O(1/z)$ yields a singularity in the specific heat at the Curie point calculated by molecular field theory.

To characterize the nature of the singularity we use Lax's work.[8] He shows that in the region $x \gtrsim 1$ that

$$F(x) = F(1) - C(x-1)^{\frac{1}{2}}.$$

From our analysis it is seen that $F(1) - 1 = O(1/z)$. C is a constant $\cong 0.8$ for Lax's three cases. Thus near the Curie point we have

$$\frac{1}{N}\frac{C_v}{k} \cong \frac{C/2}{(x-1)^{\frac{1}{2}}} = \frac{1}{2}\frac{C}{(T/T_c - 1)^{\frac{1}{2}}}; \tag{4.7}$$

We now interpret this result in terms of short-range order. This is obtained by noting that

$$E = -\tfrac{1}{2}\sum_{\mathbf{q}} v(\mathbf{q}) \langle \mu_{\mathbf{q}} {}^* \mu_{\mathbf{q}} \rangle_{\text{ensemble av}}, \tag{4.8}$$

where $\langle O \rangle_{\text{ensemble av}}$ means the ensemble average of O as distinct from $\langle O \rangle$. Then identification with (4.1) leads to[10]

$$\langle \mu_{\mathbf{q}} {}^* \mu_{\mathbf{q}} \rangle_{\text{ensemble av}} = \frac{1}{1-\beta v(\mathbf{q})}. \tag{4.9}$$

Using the result that

$$\mu_{\mathbf{q}} {}^* \mu_{\mathbf{q}} = 1 + \frac{1}{N}\sum_{i \neq j} \langle \mu_i \mu_j \rangle_{\text{ensemble av}} \exp[i\mathbf{q}\cdot(\mathbf{i}-\mathbf{j})], \tag{4.10}$$

and inverting the Fourier transform, we have

$$\langle \mu_i \mu_j \rangle_{\text{ensemble av}} = \frac{1}{N}\sum_{\mathbf{q}} \frac{\beta v(\mathbf{q})}{1-\beta v(\mathbf{q})} \exp[i\mathbf{q}\cdot(\mathbf{i}-\mathbf{j})]; \quad \mathbf{i} \neq \mathbf{j}. \tag{4.11}$$

For $|\mathbf{i}-\mathbf{j}|$ large, one can estimate (4.11) very simply by writing $v(\mathbf{q}) \cong v(0)[1-\alpha q^2]$ where $\alpha = O(z^{\frac{2}{3}})$. Inserting into (4.11) and carrying out the integration gives

$$\langle \mu_i \mu_j \rangle_{\text{ensemble av}} = \frac{1}{2}\frac{\exp(-\kappa|\mathbf{i}-\mathbf{j}|)}{|\mathbf{i}-\mathbf{j}|} \text{ for } |\mathbf{i}-\mathbf{j}| \gg 1, \tag{4.12}$$

where $\kappa = \alpha^{-\frac{1}{2}}[T/T_c - 1]^{\frac{1}{2}}$. As $T \rightarrow T_c$, $\kappa \rightarrow 0$ and the correlation length becomes infinite at the Curie point, i.e., short-range order becomes long-range order at $T = T_c$.

We now examine the susceptibility by assuming the existence of long-range order R. To the free energy one must add $N\gamma RH$ where γ is the magnetic moment per spin and H is the magnetic field. R is determined by minimizing the free energy with respect to it. [Equivalently, one may sum on R and carry out the integration by steepest descents. This adds a negligible

[8] M. Lax, Phys. Rev. **97**, 629 (1955).
[9] P. W. Anderson and H. Suhl (to be published).
[10] Alternatively $\langle \mu_{\mathbf{q}} {}^* \mu_{\mathbf{q}} \rangle_{\text{ensemble av}} = \partial \ln Z / \partial \beta v(\mathbf{q})$.

term of $O(\ln N)$ to F.] We then have from (3.22)

$$\ln Z(H)=\frac{N}{2}v(0)R^2-\tfrac{1}{2}\sum_{\mathbf{q}}\ln[1-\beta\epsilon v(\mathbf{q})]+\ln W(R)-N\gamma RH, \quad (4.13)$$

with R determined by

$$\tfrac{1}{2}\ln\frac{1+R}{1-R}=+\left[\beta v_0-\frac{1}{N}\sum_{\mathbf{q}}\frac{\beta v(\mathbf{q})}{1-\beta\epsilon v(\mathbf{q})}\right]R=\beta\gamma H. \quad (4.14)$$

Equation (4.14) is suggestively written

$$R=\tanh[\beta\gamma H_{\text{eff}}], \quad (4.15a)$$

$$H_{\text{eff}}=H+\frac{1}{\gamma}\left[v(0)-\frac{S(\beta,\epsilon)}{\beta}\right]R, \quad (4.15b)$$

where

$$S(\beta,\epsilon)\equiv\frac{1}{N}\sum_{q}\frac{\beta v(q)}{1-\beta\epsilon v(q)}.$$

The deviation from the usual molecular field is then $(1/\gamma)S(\beta,\epsilon)$. This is a temperature dependent effect which vanishes at high field ($\epsilon\to 0$) or high temperature ($\beta\to 0$).

For small fields Eq. (4.15a) can be expanded out and solved to give the susceptibility (defined as magnetic moment per spin)

$$\chi\equiv\lim_{H\to 0}\frac{\gamma R}{H}=\frac{\beta\gamma^2}{1-\beta v_0+S(\beta;1)} \quad (4.16)$$

$$=\frac{\gamma^2}{v_0}\frac{1}{x^2F(x)-1},\quad x=\frac{1}{\beta v_0}, \quad (4.17)$$

where we have used (4.3). From (4.16) it is then seen that the Curie point, defined as the place where $\chi\to\infty$, no longer occurs at $\beta v(0)=1$.

This, then, is the dilemma of the present theory. Whereas C_v blows up at $T_c=v(0)/k$, χ remains finite. The difference in the two expressions is that C_v contains terms in $O(1/z)$ only whereas $1/\chi$ contains terms in $O(1)$ and $O(1/z)$ both. If only terms of $O(1)$ are retained in χ, then the divergencies in C_v and χ coincide. This, however, is clearly improper on physical grounds, for the terms in $O(1/z)$ represent short-range order and should appear in the presence of a magnetic field as well. Furthermore, it is very easy to show that the infinity in χ does not occur at $T>T_c(\beta v_0<1)$. However, $S(\beta;1)$ is undefined in this region and the theory is left ambiguous. The resolution of these difficulties is given in the next section.

V. INTRODUCTION OF THE SPHERICAL CONSTRAINT

After Eq. (3.20) we pointed out the existence of the sum rule

$$\sum|\mu_{\mathbf{q}}|^2=N, \quad (5.1a)$$

or alternatively

$$\sum_i\mu_i^2=N. \quad (5.1b)$$

It is immediately seen that the ring diagram approximation, of which Eq. (4.9) is a consequence, leads to a violation of Eq. (5.1). This is not surprising since it was shown at the end of Sec. III that restriction to ring diagrams is equivalent to assuming that the *all* the $\mu_{\mathbf{q}}$'s are independent, a clear violation of the condition (5.1). It thus appears that a straightforward expansion in powers of $1/z$ fails because of the violation of (5.1). We therefore abandon the original plan and modify the term in $1/z$ in order to assure that (5.1) is satisfied. At this point we can no longer claim that the answer obtained is a rigorous high density limit, though the author feels that this is the case. This feeling is based on the fact that the sphericalization introduced below still keeps the Fourier components of the spin density fluctuations as independent as possible. Hence, one stays as close to the random phase approximation (sum on ring diagrams) as possible. In the Coulomb problem this approximation gives a valid limit since the density fluctuations remain bounded [$|\rho_q|^2\sim 1/(q^2+\kappa^2)$]. In our problem, the free random phase approximation leads to unbounded fluctuations $\{|\mu_q|^2=1/[1-\beta v(q)]\}$ and hence a restraint is necessary. It is possible that a formulation of the problem in the grand ensemble may lead directly to the results of this section, but we have not carried through such a proof.

To show how to introduce the restraint, we review the steps of Sec. III first in the absence of long-range order. The original problem was the computation of

$$\langle\prod_{\mathbf{q}}\exp\tfrac{1}{2}\beta v(\mathbf{q})|\mu_{\mathbf{q}}|^2\rangle, \quad (5.2)$$

where we have written the Hamiltonian in Fourier transform space. Making use of $\mu_{\mathbf{q}}^*=\mu_{-\mathbf{q}}$ and $v(\mathbf{q})=v(-\mathbf{q})$ (this latter assumes inversion symmetry), (5.2) may be written

$$\langle\prod_{\mathbf{q}\geq 0}\exp\beta v(\mathbf{q})|\mu_{\mathbf{q}}|^2\rangle. \quad (5.3)$$

The essence of the approximation of Sec. 3 is to evaluate (5.3) as

$$\prod_{\mathbf{q}\geq 0}\langle\exp\beta v(\mathbf{q})|\mu_{\mathbf{q}}|^2\rangle, \quad (5.4)$$

which together with

$$\langle(|\mu_{\mathbf{q}}|^2)^n\rangle=n!\langle|\mu_{\mathbf{q}}|^2\rangle^n \quad (5.5)$$

yields the results of Sec. 3. It is readily seen that (5.5) may be satisfied by adopting a continuous probability distribution given by

$$P(\{|\mu_{\mathbf{q}}|^2\})=\frac{1}{\langle|\mu_{\mathbf{q}}|^2\rangle^N}\prod_{\mathbf{q}\geq 0}\exp[-|\mu_{\mathbf{q}}|^2/\langle|\mu_{\mathbf{q}}|^2\rangle]$$

$$0\leq|\mu_{\mathbf{q}}|^2\leq\infty. \quad (5.6)$$

Evaluating (5.3) with (5.6) obviously yields (3.6) with $\langle|\mu_q|^2\rangle=1$. This is the Gaussian model of reference 6 and yields the same results. (Recall μ_q is complex so that a Gaussian distribution on μ_i leads to a Poisson distribution on $|\mu_q|^2$.)

To include the restraint (5.1) we retain (5.6) but with (5.1) imposed. For $\langle|\mu_q|^2\rangle=1$, this is

$$P(\{|\mu_q|^2\})=\prod_q \exp[-|\mu_q|^2]\delta(\sum \mu_q{}^2-N)$$
$$=e^{-N}\delta(\sum \mu_q{}^2-N) \quad 0\leq|\mu_q{}^2|\leq\infty. \quad (5.6a)$$

For large N, (5.6a) is normalized as it stands since

$$\int_0^\infty dx_1\cdots\int_0^\infty dx_N\,\delta(\sum x_i-N)=e^N$$

in the limit $N\to\infty$. It should be noted that the new distribution (5.6a) is the analog of the microcanonical ensemble on μ_q whereas (5.6) is the canonical ensemble. In both cases the subdistribution functions in the limit as $N\to\infty$, are the same.

Inserting (5.6) into (5.3) and using the integral representation of the δ function gives

$$\prod_{q\geq 0}\langle\exp\beta v(q)\,|\mu_q|^2\rangle_{\text{spherical}}$$
$$=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}d(\beta\delta)\Big[\prod_{q\geq 0}\int_0^\infty|\mu_q|^2\exp\{-|\mu_q|^2$$
$$+\beta v(\mathbf{q})\,|\mu_q|^2-\beta\delta\,|\mu_q|^2\}\Big]e^{\beta\delta N/2}. \quad (5.7)$$

Each of the $|\mu_q|^2$ integrals is trivial and the last integral is done by steepest descents[6] resulting in

$$(1/N)\ln Z=-\tfrac{1}{2}\textstyle\sum_q\ln[1-\beta w(\mathbf{q})]+\tfrac{1}{2}\beta\delta+\ln 2, \quad (5.8)$$

where

$$w(\mathbf{q})=v(\mathbf{q})-\delta, \quad (5.9)$$

and the saddle parameter δ is determined by

$$\sum 1/[1-\beta w(\mathbf{q})]=N, \quad (5.10)$$

which is the same as the restraint (5.1) when it is recognized that $\langle|\mu_q|^2\rangle_{\text{ensemble av}}=1/[1-\beta w(\mathbf{q})]$. Equations (5.8), (5.9), and (5.10) are the equations of the spherical model as given in references 6 and 8 in slightly different notation (adopted in order to bring out the analogy with the work of the previous sections). The analogy with Eq. (3.6) is marked, the whole effect being the substitution of $w(\mathbf{q})$ for $v(\mathbf{q})$ and the addition of the δ term in (5.8).

The thermodynamic consequences of (5.8) are easy to evaluate. The energy is given by

$$\frac{E}{N}=-\frac{1}{N}\frac{\partial\ln Z}{\partial\beta}=-\frac{\delta}{2}, \quad (5.11)$$

where we have used (5.10) and $\sum_q 1=N$. The parameter δ is determined by (5.10) which is cast into the form [using $F(x)$ defined by (4.3)]

$$\beta v(0)=F(x), \quad (5.12)$$

where

$$x=(1+\beta\delta)/\beta v(0). \quad (5.13)$$

As discussed before, $x=1$ is a singularity in $F(x)$ according to $F(x)\cong F(1)-C(1-x)^{\frac{1}{2}}$ for small $(x-1)$. From (5.11) the specific heats is $-\frac{1}{2}\partial\delta/\partial T$ or in terms of the parameter x

$$\frac{1}{k}\frac{C_v}{N}=-\frac{1}{2}\frac{d\delta}{dT}=\frac{1}{2}\Big[1+\beta v(0)\frac{\partial x}{\partial\beta}\Big]. \quad (5.14)$$

At the singular point $x=1$ it is evident that $\partial x/\partial\beta=0$. We therefore have $C_v/N=(1/2)k$ and $dC_v/dT\to-\infty$, corresponding to singular behavior in the slope of C_v just above the Curie point rather than C_v itself as in Sec. 4. The Curie point is given at $x=1$ which may be written

$$\beta_c w(0)=1, \quad (5.15)$$

where we have used the definition (5.9). The analogy with the molecular field result $\beta_c v(0)=1$ is evident. We shall discuss this point at the end of this section

Again one may interpret the onset of ferromagnetism in terms of increasing the short-range correlation distance κ^{-1} defined in Eq. (4.12). In fact, the entire analysis (4.9)–(4.12) may be taken over with $v(\mathbf{q})$ replaced by $w(\mathbf{q})$ since $\langle|\mu_q|^2\rangle_{\text{ensemble av}}=[1-\beta w(\mathbf{q})]^{-1}$. Remembering that $w(\mathbf{q})-w(0)=v(\mathbf{q})-v(0)\cong-v(0)\alpha q^2$, we now have for $T\gtrsim T_c$ $\kappa=\alpha^{-\frac{1}{2}}[1+\beta_c\delta]^{-\frac{1}{2}}[T/T_c-1]^{\frac{1}{2}}$ where $\beta_c\delta=O(1/z)$.

When a magnetic field is present, we still follow the development of Sec. 3, considering the partition function for field R, later determined by minimization. The restraint condition may now be written

$$\mu_0{}^2=NR^2, \quad (5.16a)$$
$$\sum_{q\neq 0}\mu_q{}^2=N(1-R^2)=N\epsilon. \quad (5.16b)$$

Corresponding to (5.6), we now adopt the normalized probability distribution for $q\neq 0$

$$P(\{|\mu_q|^2\})=(1/\epsilon^N)\prod_{q\neq 0}\exp[-|\mu_q|^2/\epsilon]\delta(\sum_{q\neq 0}|\mu_q|^2-N\epsilon). \quad (5.17)$$

We have for fixed R

$$(1/N)\ln Z=\tfrac{1}{2}\beta v(0)R^2+\ln W(R)+\beta\gamma RH$$
$$+\ln\langle\exp\tfrac{1}{2}\sum_{q\neq 0}\beta v(\mathbf{q})\,|\mu_q|^2\rangle. \quad (5.18)$$

The last term is now to be evaluated with the distribution function (5.17). Going through the same steps as

those that led to (5.8), we obtain

$$(1/N)\ln Z=\tfrac{1}{2}\beta v(0)R^2+\ln W(R)+\beta\gamma RH-\tfrac{1}{2}\sum_{\mathbf{q}\neq 0}\ln[1-\beta\epsilon w(\mathbf{q})]+\tfrac{1}{2}\beta\epsilon\delta$$

$$=\tfrac{1}{2}\beta w(0)R^2+\ln W(R)+\beta\gamma RH-\tfrac{1}{2}\sum_{\mathbf{q}\neq 0}\ln[1-\beta\epsilon w(\mathbf{q})]+\tfrac{1}{2}\beta\delta, \quad (5.19)$$

where we have used the definition (5.9). The saddle parameter δ is determined by

$$\sum_{\mathbf{q}\neq 0}\frac{1}{1-\beta\epsilon w(\mathbf{q})}=N, \quad (5.20)$$

which is the restraint condition (5.16b) when it is recognized that $\langle|\mu_{\mathbf{q}}|^2\rangle_{\text{ensemble av}}=\epsilon/[1-\beta\epsilon w(\mathbf{q})]$. Finally R is determined by minimization which gives[11]

$$R=\tanh\beta\gamma H_{\text{eff}}R, \quad (5.21)$$

where

$$\beta\gamma H_{\text{eff}}=\beta\gamma H+\beta w(0). \quad (5.22)$$

To obtain (5.22) we have used the saddle condition (5.20) repeatedly, together with $\sum_{\mathbf{q}}1=N$ so that

$$\sum_{q}\frac{\beta\epsilon w(q)}{[1-\beta\epsilon w(\mathbf{q})]}=0.$$

To obtain the susceptibility χ, the linear approximation to (5.21) is used and gives

$$\chi=\frac{\gamma^2}{kT}\frac{1}{1-\beta w(0)}, \quad (5.23)$$

in marked analogy to molecular field theory which contains $v(0)$ in place of $w(0)$. The theory is now consistent in that the anomalies in C_v and χ occur at the same temperature, viz., $\beta_c w(0)=1$.

Equation (5.23) is the standard spherical model Eq. (8) for small H. For large H Eq. (5.21) is not included in the original spherical model[6] but is included by Lax[8] when the correction in footnote 11 is taken into account.

For $T<T_c$; Eqs. (5.19), (5.20), and (5.21) apply. In particular the spontaneous magnetization curve is obtained by setting $H=0$ for $T<T_c$. At very low temperatures, it follows from (5.21) that $\epsilon\to 0$ exponentially. From (5.20) one may then get the behavior of ϵ by expansion

$$\sum_{\mathbf{q}\neq 0}\frac{1}{1-\beta\epsilon w(\mathbf{q})}=\sum_{q\neq 0}\{1+\beta\epsilon w(\mathbf{q})+[\beta\epsilon w(\mathbf{q})]^2+\cdots\}=N.$$

Using $\sum v(\mathbf{q})=0$, we then obtain

$$\lim_{T\to 0}\beta\delta=\left(\beta^2\frac{1}{N}\sum[v(\mathbf{q})]^2\right)\epsilon, \quad (5.24)$$

so that $\beta\delta\to 0$ exponentially. Therefore in the region $T\gtrsim 0$ $w(0)\to v(0)$ and all equations become the molecular field equations. This must be since a given spin "sees" an average field at low T, spin fluctuations being exponentially unlikely.

The final remaining point of interest is the specific heat for $T<T_c$. We take $H=0$. From (5.19) follows

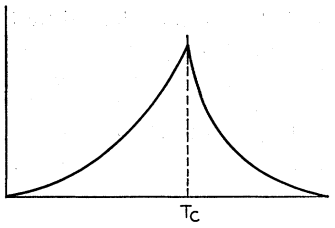
$$E/N=-\tfrac{1}{2}w(0)R^2-\tfrac{1}{2}\delta. \quad (5.25)$$

C_v is then calculated by differentiation with respect to T. At low T, $\delta\to 0$ and one recovers molecular field theory. The only interesting region is the Curie point itself. A straightforward but tedious differentiation yields

$$\lim_{T\to T_c}C_v=1/2k$$

so that in the present theory $\Delta C_v=0$ at the Curie point. The qualitative character of C_v is sketched in Fig. 3.

FIG. 3. Schematic plot of C_v vs T.



The net result is that below the Curie point our modified spherical model gives the molecular field theory at low T. As the Curie point is approached, deviations of $O(1/z)$ from the molecular field theory set in. A measure of the deviation is the value of the Curie point itself. In the molecular field theory this is given by

$$kT_c=v(0). \quad (5.25)$$

In the present theory we obtain

$$(kT_c)_{\text{sph}}=v(0)-\delta,$$

where $\delta>0$ so that $(kT_c)_{\text{sph}}<(kT_c)_{\text{mol. field}}$. Quantitatively we have from (5.12)

$$kT_c=\frac{v(0)}{F(1)}=v(0)\left(1-\frac{[F(1)-1]}{F(1)}\right), \quad (5.26)$$

where

$$F(1)=\frac{1}{N}\sum_{\mathbf{q}}\frac{1}{1-v(\mathbf{q})/v(0)}. \quad (5.27)$$

[11] Lax has examined the spherical model under the same circumstances as the present work.[8] His Eq. (2.21) corresponds to Eq. (5.20) with the identification of the saddle parameter t_s as $t_s=1/\epsilon+\beta\delta$. His formulas (3.17) (with $H_1=M'=0$ in the ferromagnetic case) and (3.18), and (5.1) (specialized to spin $\frac{1}{2}$) are the same as our (5.22) and (5.21), respectively. However the discussion after (3.19) in Lax's paper is in error. In fact, t_s should not be replaced by λ_m, but is given by the above equation with δ determined by (5.20). In other words Lax's equation, Eq. (2.21), must be taken literally even below the Curie point. I am grateful to Dr. Lax for a fruitful discussion of this point.

$F(1)$ has been evaluated by Lax for face centered cubic ($z=12$) body-centered cubic ($z=8$) and simple cubic lattices ($z=6$) for nearest neighbor interactions. The results are $F(1)=1.34(z=12)$; $F(1)=1.39(z=8)$; $F(1)=1.52(z=6)$. The correction $[F(1)-1]$ is roughly proportional to $1/z$ as expected. In the limit of large z we note that the form,

$$F(1)-1=\sum_q \frac{v(\mathbf{q})/v(0)}{1-v(\mathbf{q})/v(0)}, \tag{5.28}$$

shows that $F(1)-1=O(1/z)$ since only those $\mathbf{q}$'s lying within a volume of $O(1/z)$ of the origin contribute.

VI. QUANTUM THEORY

The quantum theory of the Heisenberg ferromagnet $[H=\frac{1}{2}\sum v_{ij}\mathbf{S}_i\cdot\mathbf{S}_j]$ is intrinsically more complicated than the Ising model. In this paper we shall not give a complete solution of the high density limit, since for $T<T_c$, one has spin waves which complicate the problem.

However, in the high density limit for $T>T_c$, the theory is identical to the Ising model. In fact restriction to ring diagrams leads to two simplifications. The first is that there is no commutator problem. It turns out that in the quantum case each ring diagram of the Ising model is replaced by $(1/n!)\sum[n!$ orderings of the n bonds in the ring$]$. Since each $\mathbf{S}_i$ appears bilinearly in a ring, the commutator that results from a change in the order of the factors is linear in $\mathbf{S}_i$ the trace of which vanishes. Thus all $n!$ orderings have the same trace. This simplification does not exist below the Curie point where the subtrace for fixed S_z is taken. It is this complication that gives rise to spin waves in the present theory.

Having established that above the Curie point the commutator problem can be ignored for ring diagrams we then have

$$\ln\langle\exp\sum\beta v_{ij}\mathbf{S}_i\cdot\mathbf{S}_j\rangle\rightarrow\tfrac{1}{2}\sum_q\sum_{n=1}^{\infty}\frac{1}{n}\times[\beta v(q)]^n\frac{\mathrm{tr}[(\mathbf{S}_1\cdot\mathbf{S}_2)\cdots(\mathbf{S}_n\cdot\mathbf{S}_1)]}{\mathrm{tr}1}, \tag{6.1}$$

where

$$\langle O(\{S_i\})\rangle\equiv\mathrm{tr}O/\mathrm{tr}1.$$

We now turn to the second simplifying feature. If each of the scalar products is expanded out in (6.1) using $S_1\cdot S_2=S_{1z}S_{2z}+S_{1+}S_{2-}+S_{1-}S_{2+}$ where $S_\pm$ are the usual $(1/\sqrt{2})(S_x\pm iS_y)$, then it is immediately seen that mixed terms do not occur. If one begins with $S_{1z}S_{2z}$ then one is forced to keep only z components throughout the product for otherwise the trace vanishes. For simplicity we restrict ourselves to spin $\frac{1}{2}$ so that the S_i's are Pauli matrices. In this case we have $\sigma^2=1$ and $\sigma_+\sigma_-=2$ or 0 depending on whether the initial spin is up or down. For this simple case we then have

$$\frac{\mathrm{tr}[\sigma_1\cdot\sigma_2\cdots\sigma_n\cdot\sigma_1]}{\mathrm{tr}1}=\frac{\mathrm{tr}[\sigma_{1z}{}^2\cdots\sigma_{nz}{}^2]}{\mathrm{tr}1}+\frac{2\,\mathrm{tr}[\sigma_{1+}\sigma_{1-}\cdots\sigma_{n+}\sigma_{n-}]}{\mathrm{tr}1}=3. \tag{6.2}$$

To arrive at (6.2) we have used the following. The first term on the right-hand side of (6.2) is 1. The second term gives a nonvanishing result if the n spins in question are all up and zero otherwise. There are 2^{N-n} such states. But the 2^{-n} factor is just cancelled by the value of the nonvanishing matrix element which is 2^n. The result is that for spin $\frac{1}{2}$ the sum on ring diagrams gives

$$\ln\langle\exp\beta\sum v_{ij}\sigma_i\cdot\sigma_j\rangle=\tfrac{3}{2}\sum_q\ln[1-\beta v(\mathbf{q})]. \tag{6.3}$$

The first singularity is at $\beta_c v(0)=1$ or the same as the Ising model.

One may then improve on the model by spherical-ization noting that restriction to ring diagrams is analogous to assuming the probability distribution (5.6). The spherical condition would restrict the spins to the sphere $\delta(\sum\sigma_i{}^2-3N)$ which introduces the parameter δ in complete analogy to Eq. (5.8).

We reserve a full discussion of the quantum calculation for a future study where the relation of the present work to spin wave theory will be given. We remark however, that it appears as if the approximations of this paper neglect the dynamical interaction of Dyson.[12]

VII. CONCLUSION

We present a brief summary of the methods and results of this paper. A straightforward expansion in $1/z$ is given by a selection of ring diagrams alone. This corresponds to the Gaussian model of Kac and Berlin in which each of the Fourier components of the spin fluctuation density are independent. The analogy with the Debye-Hückel theory is complete except that in the present case $v(q)$ has an attractive domain in q space. This latter circumstance gives rise to very large fluctuations in the spin density leading to flagrant violation of the sum rule (5.1). It thus appears that no expansion in $1/z$ is possible.

Abandoning this approach, but retaining the fundamental spirit of the random phase approximation, we insure the sum rule is obeyed by introducing it as a δ function. This leads to the conventional spherical model above the Curie point. Below the Curie point, the magnetization is held fixed and later determined by minimization of the free energy. The Fourier components of spin density fluctuation corresponding to $\mathbf{q}\neq0$ are now separately sphericalized. This procedure leads to a magnetization curve describable in terms of

[12] F. J. Dyson, Phys. Rev. **102**, 1217 (1956).

a temperature dependent Weiss field. At low temperature, one recovers the usual Weiss field and near the Curie point deviations occur of $O(1/z)$. Above the Curie point the specific heat decreases corresponding to break up of short-range order. The short-range order has a range $\sim(T-T_c)^{-\frac{1}{2}}$ which shows how ferromagnetism sets in at T_c. Finally, it is shown that to the same approximation above the Curie point these results are valid for the Heisenberg model as well.

The difficulty introduced by sphericalization is that the original program of establishing a high density limit has not succeeded. Though the approximations seem reasonable, one has lost control of the original parameter of smallness. This is the big question mark of the method given in this paper. It also remains to find an approximate quantum mechanical expression below the Curie point which will enable one to connect the low-temperature or spin-wave region to the high-temperature or Ising model region. At the time of the writing of this paper an interpolation formula for $\langle\sigma_q^+\sigma_q^-\rangle$ has been found, but there are still difficulties in how to use it to find a consistent approximation for the free energy.

APPENDIX

We prove Eq. (3.19) by expanding the moment (we take $\mathbf{q}\neq 0$)

$$\langle(|\mu_{\mathbf{q}}|^2)^n\rangle=1/N^n\sum_{\substack{i_1\cdots i_n\\ j_1\cdots j_n}}\langle\mu_{i_1}\mu_{j_1}\cdots\mu_{i_n}\mu_{j_n}\rangle \times\exp[i\mathbf{q}\cdot(\mathbf{i}_1-\mathbf{j}_1+\cdots+\mathbf{i}_n-\mathbf{j}_n)]. \quad (A.1)$$

In particular

$$\langle|\mu_{\mathbf{q}}|^2\rangle=(1/N)\sum_{i,j}\langle\mu_i\mu_j\rangle\exp[i\mathbf{q}\cdot(\mathbf{i}-\mathbf{j})]$$
$$=1+(1/N)\sum_{i\neq j}\langle\mu_i\mu_j\rangle\exp[i\mathbf{q}\cdot(\mathbf{1}-\mathbf{j})]$$
$$=1+R^2\sum_{j\neq 0}\exp(i\mathbf{q}\cdot\mathbf{j})=1-R\equiv\epsilon, \quad (A.2)$$

where we have used $\sum_j\exp(i\mathbf{q}\cdot\mathbf{j})=0$. Notice that if $\langle\mu_i\mu_j\rangle$ had a value completely independent of its indices (i.e., μ_i did not fluctuate) we should find $\langle|\mu_q|^2\rangle=0$ and in general $\langle(|\mu_q|^2)^n\rangle=0$. This is in fact the case at $T=0$. It is the special value of the moments $\langle\mu_{i_1}\cdots\mu_{j_n}\rangle$ which arises when one has identical indices that gives nonvanishing values and in fact it must be overlapping of an i index with a j index from the structure of (A.1). In particular, in (A.1) a special value of the summand occurs if each index j_k is paired to a particular index i. This may be done in $n!$ ways and hence such terms give a total contribution to $\langle(|\mu_q|^2)^n\rangle$ of $n!$. We now study the effect of allowing two indices which were formerly paired in one of the $n!$ arrangements to become unequal. This may be done in $\binom{n}{1}$ ways and hence gives a total contribution to $\langle(|\mu_q|^2)^n\rangle$ of

$$n!\binom{n}{1}1/N^n\sum_{i_1\cdots i_{n-1}}1\sum_{j_n\neq i_n}\langle\mu_{i_n}\mu_{j_n}\rangle\exp[i\mathbf{q}(\mathbf{i}_n-\mathbf{j}_n)]$$
$$=n!\binom{n}{1}R^2\sum_{j\neq 0}\exp(i\mathbf{q}\cdot\mathbf{j})=-n!\binom{n}{1}R^2.$$

Similarly uncoupling two pairs gives

$$+n!\binom{n}{2}R^4,$$

etc. In this fashion by continuing to uncouple paired i and j indices it is seen that

$$\langle(\mu_q^*\mu_q)^n\rangle=n!\sum_{m=0}^{n}(-)^m\binom{n}{m}R^{2m}$$
$$=n!(1-R^2)^n=n!\langle|\mu_q|^2\rangle^n,$$

which was to be shown. It has been suspected by the author and later confirmed that the above constitutes a special case of the central limit theorem.[13]

[13] I am grateful to Professor Mark Kac for his confirmation of this statement.