Spin Waves and Long-Range Order in the Diluted Heisenberg Ferromagnet at Zero Temperature[*]

Deepak Kumar[†] and A. Brooks Harris[‡]

Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania 19104

(Received 16 October 1972)

The variational treatment of spin waves in the randomly diluted Heisenberg ferromagnet at zero temperature is studied. It is argued that a convenient variational principle for the excitation spectrum of an alloy is difficult to formulate rigorously. However, at extremely long wavelengths a simple variational calculatation probably has approximate validity providing localized excitations in isolated clusters of magnetic sites are excluded. Within the approximation which ignores this exclusion we obtain a second-order variational bound for the spin-wave energy which, unlike that found by Murray, is positive at all concentrations. Similar results for concentrations above the critical percolation concentration x_p are obtained when the localized excitations are excluded in which case the restricted configuration averages can only be evaluated approximately. We point out that the critical concentration x_c for the occurance of long-range order depends only on the properties of the infinite cluster. Thus the thermal stability of spin waves depends on the dimensionality of the infinite cluster. An argument is given to show that the infinite cluster is not one dimensional. The range of concentrations for which the infinite cluster is two dimensional is either nonexistent or small. We conclude, then, that x_c is close, if not exactly equal, to x_p. The condition for a discontinuity in $T_C(x)$ for $x \rightarrow x_c$ is discussed in terms of a simple periodic model for dilution.

I. INTRODUCTION

A long-standing and fundamental problem in statistical physics is to understand the nature of elementary excitations in an alloy. The basic difficulty is to find a way to incorporate local fluctuations in alloy configuration into the description of collective effects, particularly for long-wavelength excitations. In this paper we consider the case of a randomly diluted Heisenberg ferromagnet. Such a model is an attractive one to study because the properties of the pure system are so well understood. In contrast, the antiferromagnet presents additional theoretical difficulties due to the presence of zero-point motion. However, the greater prospect of comparison with experimental data[1] makes the extension of our ideas to antiferromagnetic systems desirable.

A promising approach to spin waves in the diluted Heisenberg ferromagnetic was made by Murray[2] in her variational formulation. She found that if one writes the spin-wave energy for the system with a concentration x of magnetic sites and $1 - x$ of nonmagnetic sites as $E_k(x) = D(x)a^2k^2$, then the variational upper bound for $D(x)$ vanishes for a value of x about 0.44 for a simple cubic lattice. This result and the general formulation raise a number of questions, several of which we have sought to answer. Our main conclusions and results were published previously,[3] and it is the purpose of the present paper to describe our work in more detail.

The most important difficulty with Murray's work was that the variational bound for $D(x)$ became zero

for some x. Since the fully aligned state of the Heisenberg ferromagnet is the ground state, it is impossible to have $D(x)$ negative, and thus such a result is likely to indicate an error in the calculation. Accordingly, we have repeated the variational calculation and find that the method does give a positive upper bound for $D(x)$.

A conceptual difficulty concerned the role in Murray's theory of the percolation concentration,[4] x_p, the minimum concentration at which an infinitely large connected cluster of magnetic sites is formed. Although it is clear that spin waves cannot exist for $x < x_p$, there seemed to be no restriction in the formalism to prevent application of the variational principle in this regime, where a finite bound for $D(x)$ was obtained. To avoid this problem we note that since spin waves only exist in the infinite cluster, appropriate variational wave functions should not contain local excitations on isolated clusters. (For brevity such excitations will henceforth sometimes be called simply "localized excitations." These excitations are not the usual local modes associated with impurities.) This restriction automatically limits spin waves to concentrations $x \geq x_p$. Furthermore, we have examined the basis for using the variational principle to bound $D(x)$. It appears to us that a rigorous formulation for such alloy problems is difficult. Arguments for the approximate validity of the variational principle for extremely-long-wavelength spin waves in the dilute ferromagnet are given. Even for $x > x_p$, the condition that admissible trial functions should contain excitations only in the infinite cluster is shown to be essential for obtaining valid bounds on

$D(x)$.

Finally, the problem of determining the critical concentration x_c for the occurrence of long-range order (LRO) in the zero-temperature limit is considered. This problem has been studied by several authors in recent years.[5-7] An important early work was a mathematical argument given by Elliott *et al.*[5] showing that x_c coincides with x_p for both the Ising and Heisenberg models. It was soon pointed out,[6] however, that the geometrical argument cannot be used to locate x_c in view of the fact that the two-dimensional Heisenberg model does not support LRO for any x. It is thus necessary to study the density of states at low energy, which depends both on $D(x)$ and on the dimensionality of the infinite cluster. Although we have obtained a positive upper bound for $D(x)$, the possibility that D might vanish must be considered, and recently Last[8] has given arguments which indicate that $D(x) \rightarrow 0$ as $x \rightarrow x_p$. Nevertheless, the determination of x_c depends on the dimensionality of the infinite cluster. If the infinite cluster were one dimensional then neither the Heisenberg nor Ising alloys would order.[9,10] If it were two dimensional then the Heisenberg alloy would not order,[11] while if it were three dimensional, both alloys would order. We give an argument that a one-dimensional cluster never exists. A two-dimensional cluster probably exists at most over a small range of concentrations. Thus, we believe that x_c is close, if not precisely equal, to x_p.

Briefly, this paper is organized as follows. In Sec. II we discuss qualitatively the possibility of applying the variational principle to obtain bounds on $D(x)$. Here we also give arguments which suggest that the infinite cluster is probably always three dimensional, so that $x_c \approx x_p$. In Sec. III we discuss the variational principle in more detail and carry out a second-order variational calculation of $D(x)$ in which localized excitations are not excluded. As noted above, in contrast to Murray's result,[2] our bound for $D(x)$ is positive. In Sec. IV we give an evaluation of the variational bound in which localized excitations are excluded. This calculation, were it done exactly, would probably represent a true variational result. Several of our qualitative conclusions are illustrated in Sec. V using a simple periodic model for dilution. Here the condition that $T_c(x)$ has a discontinuity for $x \rightarrow x_c$ is discussed. Our conclusions are summarized in Sec. VI. In Appendix A we show that the trial functions for wave vector k have an overlap with trial functions for wave vector k' which does not vanish in the $k = 0$ limit. In Appendix B we discuss the use of the variational principle in such a nonorthogonal basis set. In Appencix C it is explicitly shown that the second-order trial energy coincides with third-order perturbation theory.

Finally, a simple, but rigorous, argument that the average number of nearest neighbors occupied in the infinite cluster is at least $2/z$, where z is the number of nearest neighbors, is given in Appendix D.

II. QUALITATIVE DISCUSSION

This discussion will be organized as follows. First, in Sec. II A, we define the model used to treat the dilution problem. Next, in Sec. II B, we argue that there exist two qualitatively different regimes, namely, one for short wavelengths where localized effects and configurational fluctuations are important, and the other at long wavelengths where configurational fluctuations are unimportant and spin waves become good modes of the system. It is argued in Sec. II C that in the latter regime the variational principle is approximately valid providing localized excitations in isolated clusters of magnetic spins are properly projected out of the calculations. In Sec. II D we discuss briefly the theoretical predictions that $D(x)$ vanishes as $x \rightarrow x_p$. Finally, in Sec. II E the possibility of a thermal instability is considered and is shown to depend on the dimensionality of the infinite cluster. Our discussion indicates that the infinite cluster is probably three dimensional, in which case $x_c = x_p$.

A. Definition of Model

The model we consider is described by the Hamiltonian

$$\mathcal{H}(\theta) = -2J \sum_{\langle ij \rangle} p_i p_j \vec{S}_i \cdot \vec{S}_j \ , \tag{2.1}$$

where $\langle ij \rangle$ indicates a sum over pairs of nearest neighbors and $p_i \equiv p_i(\theta)$ is defined as

$$p_i(\theta) = 1 \ , \quad i \in \theta \tag{2.2a}$$

$$p_i(\theta) = 0 \ , \quad i = \theta \tag{2.2b}$$

where $i \in \theta$ means that the site i is occupied in the configuration θ. Except where specifically noted, we will treat only a simple cubic lattice. The quantities of interest will be averaged over all configurations, the configuration θ having a probability $x^{N(\theta)}(1 - x)^{N-N(\theta)}$, where $N(\theta) = \sum_i p_i(\theta)$, N is the total number of sites, and x is the average concentration. It has been shown[12] that this ensemble where only the average concentration and not the total number of magnetic spins is fixed gives thermodynamic results identical to the ensemble where $N(\theta)$ is fixed and all distinct configurations have equal weight. Thus, we have the configurational averages, denoted by $\langle\langle \ \rangle\rangle$:

$$\langle\langle p_i \rangle\rangle = x \ , \tag{2.3a}$$

$$\langle\langle p_i p_j \rangle\rangle = x^2 + x(1 - x)\Delta_{ij} \ , \tag{2.3b}$$

etc., where Δ is the Kronecker Δ.

With the exception of our dimensionality arguments, we shall treat the model in the harmonic approximation; in other words, we shall study only the single-spin-wave manifold and will neglect spin-wave interactions. This is most conveniently done by using the Holstein–Primakoff transformation[13] to boson operators $a_i^\dagger$ in terms of which Eq. (2.1) becomes

$$\mathcal{H} = -2JS^2 \sum_{\langle ij \rangle} p_i p_j + 2JS \sum_{i,\delta} a_i^\dagger (a_i - a_{i+\delta}) p_i p_{i+\delta} , \quad (2.4)$$

where δ is a nearest-neighbor vector. Since we are not interested in the thermodynamic properties of the system we will henceforth omit the first term in Eq. (2.4).

B. Regimes for Configurational Fluctuations

Before attempting to calculate the spin-wave energy of the diluted system, it is necessary to discuss the excitation spectrum of an alloy. Within the harmonic approximation, spin waves are perfect modes for the pure system and their energies are well defined. For the alloy, spin waves are no longer perfect modes. Then elementary excitations are defined as resonances in the configurationally averaged spectral weight function. For this purpose one introduces the Green's function[14]

$$G_\theta(r, t; r', t') \equiv -i\theta(t - t')\langle [S_r^+(t), S_{r'}^-(t')] \rangle \quad (2.5)$$

and its transform

$$G_\theta(r, r'; \omega) = \int_{-\infty}^{+\infty} G_\theta(r, t; r', t') e^{i\omega(t-t')} dt . \quad (2.6)$$

Both these functions depend on the configuration θ as is indicated by the subscript notation. After the configurational average is taken, one recovers translational invariance:

$$\langle\langle G_\theta(r, r'; \omega) \rangle\rangle \equiv g_x(r - r'; \omega) , \quad (2.7)$$

so that one can write

$$g_x(k, \omega) = \sum_r g_x(r; \omega) e^{-i\vec{k}\cdot\vec{r}} . \quad (2.8)$$

The configurationally averaged spectral weight function $\rho_x(k, \omega)$ is then given as

$$\rho_x(k, \omega) = i[g_x(k, \omega + i0^+) - g_x(k, \omega - i0^+)] . \quad (2.9)$$

For the pure system, $\rho_x(k, \omega)$ is a δ function at $\omega = E_k = 2z JS(1 - \gamma_k)$ in the usual notation.[15] For the diluted system, $\rho_x(k, \omega)$ will have a resonant behavior at $\omega = E_k(x)$, where $E_k(x)$ is the spin-wave energy for the concentration x. Thus even at zero temperature $E_k(x)$ is not well defined, because although k is a good quantum number on the average,

it is not conserved for each configuration and hence $\rho_x(k, \omega)$ is not sharp. Clearly, we cannot hope to obtain a rigorous variational bound on such an ill-defined quantity. One possible way to define the spin-wave energy would be to set it equal to the average energy of the excitation in the sense of

$$E_k(x) = \frac{\int \rho_x(k, \omega)\omega \, d\omega}{\int \rho_x(k, \omega) \, d\omega} . \quad (2.10)$$

Let us now consider how we might apply the variational principle to this problem. Of course, the variational principle can be applied to each configuration. However, in that case one cannot easily relate the spectrum to the momentum k. The difficult thing to prove, probably because, in fact, it is not rigorously true, is that the variational principle can be applied *after* the configurational average is taken. We therefore have formulated a weaker proposition as follows. We assume that for a fixed concentration $x > x_p$ one can define a length $l_0(x) \equiv k_c(x)^{-1}$ such that all significant configurational fluctuations occur over a scale of length less than or equal to $l_0(x)$. If the number of spins in the volume l_0^3 is denoted by n, it is clear that $\langle\langle n \rangle\rangle = xl_0^3$. Furthermore, it is also easy to show that $\langle\langle (n - \langle\langle n \rangle\rangle)^2 \rangle\rangle / \langle\langle n \rangle\rangle^2 = (1 - x)/(xl_0^3)$. Thus, by making l_0 large enough, we can clearly reduce the relative size of fluctuations in n. What is less obvious is that making l_0 large leads to volume elements which are uniform enough to support well-defined spin waves. Possibly l_0 in this context is defined so that $\Delta \equiv \langle\langle (\hat{n} - \langle\langle \hat{n} \rangle\rangle)^2 \rangle\rangle / \langle\langle \hat{n} \rangle\rangle^2$ is small, where $\hat{n}$ is the number of sites in the volume l_0^3 which are in the infinite cluster. Roughly speaking, the quantity Δ is of order $(1 - \hat{x})/(\hat{x}l_0^3)$, where $\hat{x}$ is the fraction of sites in the infinite cluster. Since[4] $\hat{x} \to 0$ as $x \to x_c$, we believe that l_0 becomes infinite as $x \to x_c$. This conclusion is appealing in view of the analogy with critical phenomena in pure systems. However, in contrast with pure systems where good modes occur for[16] $k\xi \gg 1$, where ξ is the order-parameter–order-parameter correlation length, here we have good modes for $kl_0 \ll 1$. The difference may be explained by the fact that in pure systems ξ measures the decay of correlations which tend to stabilize the modes involving the oscillations in long-range order, whereas in the dilute case l_0 measures the decay of fluctuations which destabilize the modes in question.

Thus our picture is that for energies corresponding to $k \ll k_c$, spin waves are well defined and most configurations have an energy spectrum which is essentially that of the configurationally averaged system. In this regime it is plausible to apply the variational principle to the configurationally averaged system. Even though $l_0 \to \infty$ as $x \to x_p + 0$, we still expect that for any $x > x_p$ there is a small regime of k's for which the variational reasoning is

valid. In that case the variational procedure can be used to bound the spin-wave energies in the $k \to 0$ limit for all $x > x_p$.

The above argument is probably not valid if the infinite cluster is one dimensional, because the correct single-spin-deviation eigenstates of a chainlike configuration cannot be associated with a three-dimensional wave vector. Thus, assuming the existence of a regime $k \ll k_c$, where the configurational fluctuations can be ignored, may be equivalent to asserting that the infinite cluster is three dimensional. In fact, as we have shown elsewhere,[3] arguments of the type given by Kikuchi[17] suggest that these two propositions are equivalent.

Finally, we comment on the earlier arguments based on calculations of the scattering from a single defect which lead to an energy width $\Delta \omega_k$ of order k^5. This result would lead one to predict the existence of well-defined spin waves when

$$g(x)k^5 \ll D(x)a^2k^2 \quad , \tag{2.11}$$

where $g(x) \propto (1 - x)$ for $x \to 1$. Clearly, as long as $D(x)$ is nonzero such a regime will exist. However, for x near x_p there is no reason to believe that such formulation is appropriate.

More generally, one would expect the energy width to be of the form

$$\Delta \omega_k = G(x)k^n \quad , \tag{2.12}$$

where $G(x)$ is an unknown function of x and sum-rule arguments of the type used by Lange[18] guarantee that $n \geq 2$. If $n > 2$, then spin waves are good modes in the long-wavelength regime defined by $k \ll k_c(x)$, where $k_c(x)$ now satisfies

$$G(x)k_c(x)^n = D(x)a^2k_c(x)^2 \quad . \tag{2.13}$$

As $x \to 1$, $G(x) \to g(x)$, but for $x \to x_p$, probably $G(x)$ diverges, so that $k_c(x) \to 0$ as $x \to x_p$. Since these assertions involve the introduction of a three-dimensional wave vector, they require justification of the type we have tried to give two paragraphs earlier, rather than one based on a single-defect calculation.

C. Exclusion of Excitations in Isolated Clusters

As we have seen, the variational principle is only valid at long wavelengths where $k \ll k_c(x)$. However, in applying the variational principle we must be sure that trial wave functions used actually describe spin waves. For instance, consider a cluster of magnetic spins surrounded by nonmagnetic sites. An exciation in such a cluster is a localized nonpropagating excitation confined to the finite volume of the cluster. Furthermore, the energy spectrum of such a cluster, being discrete, is quite unlike that of spin waves in the infinite cluster. Consequently, admissible trial functions contain

only excitations in the infinite cluster. Such functions exist only for concentrations larger than the critical percolation concentration x_p. Hence excluding localized excitations leads automatically to the conclusion that there are no spin waves for $x < x_p$. In contrast, Murray's treatment[2] paradoxically gives a finite upper bound for spin-wave energies for all nonzero x.

Another simple conclusion is reached by writing the configurationally averaged transverse susceptibility in zero field χ as

$$\chi(k, \omega) = \chi_{IC}(k, \omega) + \chi_{Inf}(k, \omega) \tag{2.14}$$

in terms of a contribution, χ_{IC} from isolated clusters, and one χ_{Inf} from the infinite cluster. For $k \ll k_c(x)$, we have approximately

$$N^{-1}\chi''_{Inf}(k, \omega) = R(k, \omega)\delta(\omega - E_k(x)) \quad , \tag{2.15}$$

where the double prime indicates the imaginary part. Further, since χ'' obeys the sum rule

$$\int \chi''(k, \omega)(d\omega/2\pi) = 2\langle S_z \rangle_{H \to 0^+} \quad , \tag{2.16}$$

one sees that in the thermodynamic limit, i.e., for $N \to \infty$, one has

$$R(k, \omega) = 4\pi S \hat{x} \quad , \tag{2.17}$$

where $N\hat{x}$ is the number of spins in the infinite cluster. (For the finite-sized isolated clusters, $\langle S_z \rangle$ goes smoothly to zero as $H \to 0$.) Here and below, use of carets will imply a restriction to the infinite cluster. Approximate calculations[4] show that $\hat{x} \to 0$ as $x \to x_p$. This may be seen in Fig. 1, where we have plotted $Q(x) \equiv \hat{x}/x$ vs x for a simple cubic lattice. From Eq. (2.17) it follows that the amplitude of the spin-wave pole in the transverse
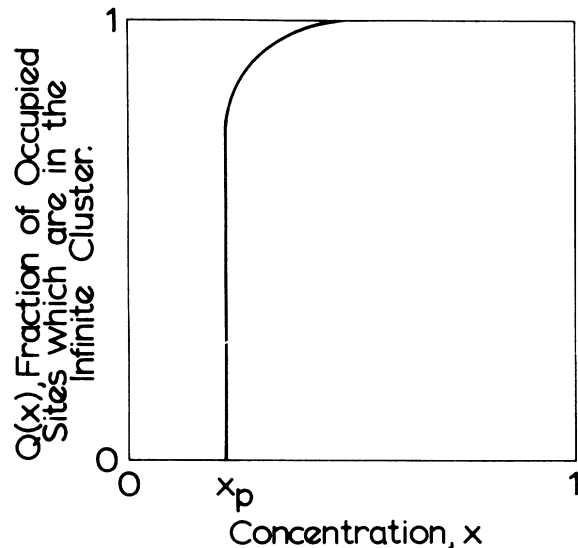

FIG. 1. $Q(x) \equiv \hat{x}/x$, the fraction of occupied sites which are in the infinite cluster, vs x.

susceptibility vanishes as $x \to x_p$. Thus, as $x \to x_p$, χ is dominated by the localized response. Nevertheless, the occurrence of LRO is determined by $\chi_{\mathrm{Inf}}(k, \omega)$.

D. Mechanical Stability

An interesting question is whether the diluted system is mechanically stable, i. e., has nonzero stiffness, as $x \to x_p$. The pure Heisenberg system has a dispersion relation at zero temperature of the form

$$E_k = 2JSa^2k^2 \equiv Da^2k^2 \qquad (2.18)$$

in the long-wavelength limit $(k \to 0)$, where a is the lattice constant, henceforth taken to be unity, and D is the stiffness. Assuming that for $k \ll k_c(x)$ spin waves are good modes in the alloy, one can write

$$E_k(x) = D(x)k^2 \equiv 2 JS\phi(x)k^2 . \qquad (2.19)$$

As we have mentioned, Murray[2] previously concluded that $\phi(x)$ vanished for $x \to x_p$. However, as we shall discuss in more detail below, that conclusion was based on an erroneous variational calculation. Recently, several authors have studied the behavior of $D(x)$ using scattering[19,20] treatments of the diluted alloy. These treatments are based on summing incoherent repeated scatterings, either using low-concentration theory or by an approach similar to the coherent-potential approximation.[21] They give the result that $D(x)$ vanishes for some critical value of the concentration. However, it must be remembered that these approximations represent approximations to the total susceptibility. As our discussion following Eq. (2.1) indicates, for $x \to x_p$, χ is dominated by isolated clusters. It is also implausible that these methods should give reliable information concerning $D(x)$ for x near x_p, since in that regime the global geometry of the infinite cluster determines the nature of $D(x)$. Thus, in our opinion, the only convincing study of $D(x)$ for x near x_p is that of Last,[8] who relates it to the electrical response of a dilute network of resistors. Then, from numerical experiments[22] for $x \approx x_p$ on such networks he is able to infer that $D(x) \to 0$ as $x \to x_p$. Previously,[3] and in the preprint of this work we gave an incorrect argument that $D(x)$ should not vanish for $x \to x_p$.

E. Thermal Stability

We now consider the possibility of a thermal instability. We recall that because of thermal instabilities neither the one-dimensional Ising model[9,10] nor the two-dimensional Heisenberg model can support LRO. Basically, the reason is that there are too many thermal spin waves excited for LRO to persist. This type of instability

could be relevant for the alloy. In principle one could have the situation shown in Fig. 2. Here we suppose that for $x_p \leq x \leq x^{(1)}$ the infinite cluster is one dimensional. Mathematically, one could define a cluster to be one dimensional if the density of spin-wave states $g(\epsilon)$ for the Heisenberg model varies as

$$g(\epsilon) \sim \epsilon^{-1/2}, \quad \epsilon \to 0. \qquad (2.20)$$

Similarly, for $x^{(1)} \leq x \leq {}^{(2)}$ the infinite cluster is assumed to be two dimensional, i. e.,

$$g(\epsilon) \sim \mathrm{const.}, \quad \epsilon \to 0, \qquad (2.21)$$

and for $x^{(2)} \leq x \leq 1$, the infinite cluster is three dimensional with

$$g(\epsilon) \sim \epsilon^{1/2}, \quad \epsilon \to 0. \qquad (2.22)$$

However, since the dimensionality of the infinite cluster is not a local property, cluster methods[23,24] or even repeated-scattering calculations[25,26] where configurations are treated from a local point of view are unsuitable tools for locating x_c.

It is also difficult to approach the critical regime by extrapolating high-temperature series expansions.[7] Physically, this is because the system behaves one dimensionally for $T \gg T_c(x)$ and three dimensionally only for T sufficiently close to $T_c(x)$. This phenomenon is difficult to elucidate using high-temperature expansions and hence their extrapolations which give predictions (see Fig. 3)
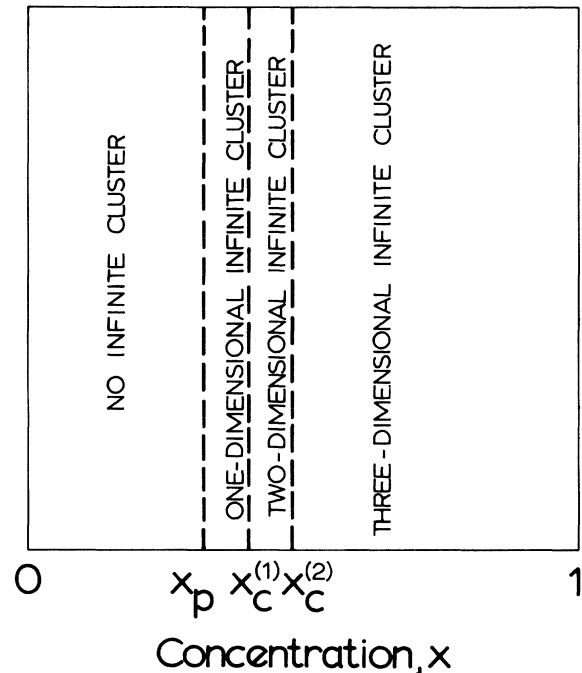
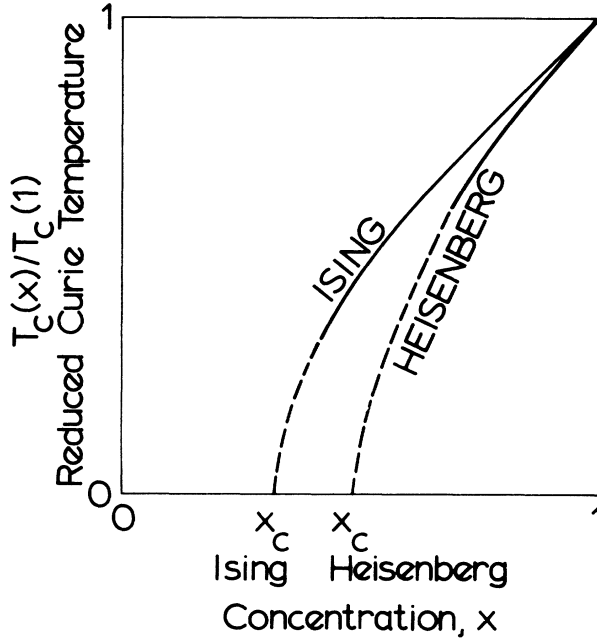FIG. 2. Concentration regimes for the dimensionality of the infinite cluster.

FIG. 3. Results of the high-temperature expansions for T_C for the diluted Heisenberg and Ising systems (see Ref. 7).

for $T_C(x)$ for $x \sim x_p$ are not to be trusted. These points will be illustrated in Sec. V where we consider a simple model for the dilute system. Mathematically, the use of high-temperature expansions is uncertain because, as Griffiths has shown, [27] one cannot expect them to converge for $T < T_C^{(1)}$ even though $T_C(x) < T_C^{(1)}$. This objection may be academic for $x \sim 1$, but is surely an essential problem as $x \to x_p$.

III. VARIATIONAL CALCULATION WITH LOCALIZED EXCITATIONS NOT EXCLUDED

In this section we discuss the use of the variational principle when localized excitations are not excluded. In Sec. III A we state the variational principle which we use and in Sec. III B we give the evaluation of the spin-wave energy using second-order trial wave functions.

A. Formulation

A variational principle for excited states is not completely trivial. For instance, if one contructs an orthonormal set ϕ_1, ϕ_2, $\ldots$, ϕ_n, the resulting trial energies $(\phi_j \mid \mathcal{H} \mid \phi_j)$ have no variational relation to the actual energies. Our approach is based on the following theorem. [28]

Consider the Hamiltonian $\mathcal{H}$ having N energy eigenvalues E_α:

$$\mathcal{H}\psi_\alpha = E_\alpha \psi_\alpha . \tag{3.1}$$

Consider a set of n-independent states ϕ_1, ϕ_2, $\ldots$,

ϕ_n. These states span a subspace S of the original set of states. We define the projection operator P_S to be unity in S and zero outside S. The eigenvalues λ_1, λ_2, $\ldots$, λ_n obtained from the solution of the eigenvalue problem within S, which may be written

$$P_S \mathcal{H} P_S \xi_i = \lambda_i \xi_i , \tag{3.2}$$

can be shown to have the following variational significance. Each member of the sequence $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ is an upper bound to the corresponding member of the sequence $E_1 \le E_2 \le \cdots \le E_n$ of the lowest eigenvalues of $\mathcal{H}$, i.e.,

$$E_1 \le \lambda_1 ; \quad E_2 \le \lambda_2 ; \quad \ldots ; \quad E_n \le \lambda_n . \tag{3.3}$$

We cannot apply this theorem directly to our case because we are unable to diagonalize even the truncated Hamiltonian of Eq. (3.2). We shall use a weaker form of the result Eq. (3.3), namely,

$$\mathrm{Tr} \, P_S \mathcal{H} P_S \ge \sum_{i=1}^n E_i , \tag{3.4}$$

which is obtained by adding all the inequalities of Eq. (3.3),

$$\sum_{i=1}^n \lambda_i \ge \sum_{i=1}^n E_i . \tag{3.5}$$

Equation (3.4) is sufficient for our purpose because the k dependence of the spin-wave-type eigenvalues is known. Then the sum on the right-hand side of Eq. (3.4) depends only on the single parameter $D(x)$ for which a bound is desired. The left-hand side of Eq. (3.4), being a trace, does not require the solution of an eigenvalue problem, but can be evaluated as a diagonal sum in any representation.

This theorem can be applied to any configuration θ. Let us restrict ourselves to spin-wave excitations for which $k \ll k_c(x)$. Then, since the right-hand side of Eq. (3.4) does not depend sensitively on configuration, we evaluate it for the configurationally averaged system. We carry the sum over a very small sphere of radius k_0 in momentum space and write

$$\sum_{i=1}^n E_i = \frac{V}{8\pi^3} \int d\vec{k} \; E_k(x) \tag{3.6a}$$

$$= \frac{V}{8\pi^3} \frac{4}{5}\pi k_0^5 D(x) , \tag{3.6b}$$

where $n = (V/8\pi^3)(\tfrac{4}{3}\pi)k_0^3$. The left-hand side of Eq. (3.4) is evaluated using an orthonormal set $\{\Phi_k\}$. Then we have

$$D(x) \le \lim_{k_0 \to 0} \left(\frac{5}{4\pi}\right) k_0^{-5} \int_0^{k_0} d\vec{k} \, (\Phi_k \mid \mathcal{H} \mid \Phi_k) . \tag{3.7}$$

For any reasonable choice of Φ_k's the trial energy is proportional to k^2 and consequently Eq. (3.7) becomes

$$D(x) \lesssim \lim_{k \to 0} (\Phi_k \, | \, \mathcal{H} \, | \, \Phi_k)/k^2 \qquad (3.8)$$

A final complication is caused by the fact that it proves inconvenient to use an orthonormal set to evaluate Eq. (3.8). In Appendix A we show that, even for the best trial wave functions we have used, the mean-square overlap $\langle\langle \, | \, \langle \Phi_k \, | \, \Phi_{k'} \rangle \, |^2 \rangle\rangle$, does not vanish as $k \to 0$ for fixed $k' \neq 0$. This result seems to contradict that of Ref. 2, Eq. (2.13), where the overlap is stated to be of order $k^2 k'^2$ at low energy. Accordingly, in Appendix B we give a phase-space argument, which shows that Eq. (3.8) is still valid for the types of nonorthogonal wave functions used here.

B. Evaluation of Second-Order Trial Energy

Murray's perturbative series for the spin-wave energy may be written, in the long-wavelength limit, as[2]

$$E_k(x) = \sum_n E_k^n (x), \qquad (3.9)$$

where

$$E_k^{(1)}(x) = x E_k , \qquad (3.10a)$$

$$E_k^{(2)}(x) = -\frac{1}{x} \sum_f \frac{H_{kf} H_{fk}}{E_f} , \qquad (3.10b)$$

$$E_k^{(3)}(x) = \frac{1}{x^2} \sum_{f,f'} \frac{H_{kf} H_{ff'} H_{f'k}}{E_f E_{f'}} , \qquad (3.10c)$$

and so forth. Here H_{kf} is the spin-wave interaction relative to the configurationally averaged system:

$$H_{kf} = \langle k \, | \, \mathcal{H} \, | \, f \rangle - x E_k \Delta_{kf} , \qquad (3.11)$$

where Δ is the Kronecker Δ and the matrix elements are taken between the states,

$$| \, k \, \rangle = (Nx)^{-1/2} \sum_r p_r e^{i \vec{k} \cdot \vec{r}} \, | \, r \, \rangle , \qquad (3.12)$$

where $| \, r \, \rangle$ is a state with a single-spin deviation at r. It is important to note that truncation of the infinite series in Eq. (3.9) does not in general yield a result with any variational significance. However, a sequence of variational bounds can be obtained by using as trial functions the wave functions obtained for $k \to 0$ in the nth-order perturbation theory:

$$\psi_k^{(1)} = | \, k \, \rangle , \qquad (3.13a)$$

$$\psi_k^{(2)} = | \, k \, \rangle - \sum_f (H_{kf}/x E_f) \, | \, f \, \rangle , \qquad (3.13b)$$

etc. To leading order in k one finds that

$$(\psi_k^{(1)} \, | \mathcal{H} | \, \psi_k^{(1)})/(\psi_k^{(1)} \, | \, \psi_k^{(1)}) = E_k^{(1)}(x) = x E_k , \qquad (3.14a)$$

$$(\psi_k^{(2)} \, | \mathcal{H} | \, \psi_k^{(2)})/(\psi_k^{(2)} \, | \, \psi_k^{(2)}) = \sum_{i=1}^{3} E_k^{(i)}(x) . \qquad (3.14b)$$

Equation (3.14b) is derived in Appendix C.
It is not surprising that the approximation for the

energy $E_k^{(1)}(x) + E_k^{(2)}(x)$ becomes negative for some x, since this expression has no variational significance. Addition of the third term in Eq. (3.14b) must lead to a positive definite result, and it is here that Murray's error has occurred. Previously,[3] we gave the result of including the third term in Eq. (3.14b) without any details of the calculation. Subsequently, almost the same result was obtained by Last.[8]

We now evaluate the right-hand side of Eq. (3.14b). The first two terms were given by Murray, $E_k^{(1)}(x)$ as in Eq. (3.14a) and

$$E_k^{(2)}(x) = 2 J S k^2 [-I(1-x) - \tfrac{1}{3}(1-x)^2/x] , \qquad (3.15)$$

where[2]

$$I = (4/N) \sum_k \sin^2 k_x \, \hat{\epsilon}_k^{-1} \approx 0.42, \qquad (3.16)$$

where $\hat{\epsilon}_k = 6 - 2\cos k_x - 2\cos k_y - 2\cos k_z$ is the dimensionless spin-wave energy. To evaluate the third term we need the expression for matrix element of H_{kf}. According to the definitions of Eqs. (3.11) and (3.12), we have

$$H_{kf} = \frac{JS}{Nx} \sum_{r,\delta} e^{i(\vec{k}-\vec{f})\cdot\vec{r}} (p_r p_{r+\delta} - x^2)$$
$$\times (1 - e^{-i\vec{f}\cdot\vec{\delta}})(1 - e^{i\vec{k}\cdot\vec{\delta}}) , \qquad (3.17)$$

where δ is a nearest-neighbor vector and $\vec{r}$ ranges over the N sites of a simple cubic lattice. Then we have

$$E_k^{(3)}(x) = 2 J S (8 N^3 x^5)^{-1} \sum_{f,f'} \sum_{r,\delta} \sum_{r',\delta'} \sum_{r'',\delta''} \hat{\epsilon}_f^{-1} \hat{\epsilon}_{f'}^{-1}$$
$$\times e^{i(\vec{k}-\vec{f})\cdot\vec{r}} e^{i(\vec{f}-\vec{f'})\cdot\vec{r'}} e^{i(\vec{f'}-\vec{k})\cdot\vec{r''}}$$
$$\times (1 - e^{-i\vec{f}\cdot\vec{\delta}})(1 - e^{i\vec{k}\cdot\vec{\delta}})(1 - e^{-i\vec{f'}\cdot\vec{\delta'}})(1 - e^{i\vec{f}\cdot\vec{\delta'}})$$
$$\times (1 - e^{-i\vec{k}\cdot\vec{\delta''}})(1 - e^{i\vec{f'}\cdot\vec{\delta''}})$$
$$\times \langle\langle (p_r p_{r+\delta} - x^2)(p_{r'} p_{r'+\delta'} - x^2)$$
$$\times (p_{r''} p_{r''+\delta''} - x^2) \rangle\rangle . \qquad (3.18)$$

Here and below, the limit $k \to 0$ is implied. The most convenient way of keeping track of all the terms obtained from the configurational average is to use a diagrammatic prescription such as that proposed by Langer[29] or by Edwards and Beeby.[30] In this formulation each vertex of the type shown in Fig. 4 represents a factor H_{kf} when the factors
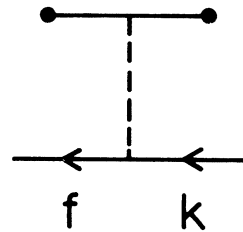

FIG. 4. Diagrammatic representation of impurity scattering. The lower directed lines represent initial and final spin-wave states. The upper line connects dots representing the two lattice sites r and $r+\delta$ of Eq. (3.17).

$(pp - x^2)$ are elided. The p's are then taken into account by the upper vertices in the diagrams of Fig. 5, where each vertex with n solid lines represents the cumulant average of x^n, denoted $f_n(x)$. For instance,

$$f_1(x) = x ,\qquad (3.19a)$$

$$f_2(x) = x(1 - x),\qquad (3.19b)$$

$$f_3(x) = x(1 - x)(1 - 2x),\qquad (3.19c)$$

etc. It is concluded that (i) only "connected" diagrams (i.e., those where the upper vertices are connected) contribute and (ii) many terms give identical contributions in view of the various symmetries, e.g., interchanging r and $r + \delta$, interchanging the pairs (r, δ) and (r', δ'), etc. To take account of these symmetries each diagram is given an associated multiplicity factor as indicated in Fig. 5. Consequently, the configurational average in Eq. (3.18) may be replaced by the factor

$$8x^4(1-x)(1-2x)\Delta_{r,r'}\,\Delta_{r,r''}+8x^4(1-x)^2\Delta_{r+\delta,r'}\,\Delta_{r'+\delta',r''}+8x^4(1-x)^2\Delta_{r+\delta,r''}\,\Delta_{r''+\delta'',r'}$$

$$+\,8x^4(1-x)^2\Delta_{r''+\delta'',r}\,\Delta_{r+\delta,r'}+8x^3(1-x)^2(1-2x)\,\Delta_{r,r'}\,\Delta_{r,r''}\,\Delta_{\delta',\delta''}+8x^3(1-x)^2(1-2x)\,\Delta_{r,r'}\,\Delta_{r,r''}\,\Delta_{\delta,\delta'}$$

$$+\,8x^3(1-x)^2(1-2x)\,\Delta_{r,r'}\,\Delta_{r,r''}\,\Delta_{\delta,\delta''}+4x^2(1-x)^2(1-2x)^2\Delta_{r,r'}\,\Delta_{r,r''}\,\Delta_{\delta,\delta'}\,\Delta_{\delta,\delta''}\;.\qquad (3.20)$$

Substituting this eight-term expression into Eq. (3.17) leads to an eight-term expression for $E_k^{(3)}$, which we write as $E_k^{(3)} = \Sigma_i T_i$, where T_i corresponds to the ith diagram of Fig. 5. We now evaluate the T_i's. We have

$$T_1 = (2JS/N^3x^5)x^4(1-x)(1-2x)\sum_{f,f'}\sum_r\sum_{\delta,\delta',\delta''}\hat{\epsilon}_f^{-1}\hat{\epsilon}_{f'}^{-1}\,(1-e^{-i\vec{f}\cdot\vec{\delta}})(1-e^{i\vec{k}\cdot\vec{\delta}})(1-e^{-i\vec{f}\cdot\vec{\delta}'})(1-e^{i\vec{f}\cdot\vec{\delta}'})(1-e^{-i\vec{k}\cdot\vec{\delta}''})$$

$$\times(1-e^{i\vec{f}\cdot\vec{\delta}''}).\qquad (3.21)$$

Taking the $k \to 0$ limit and using the fact that then $T_1 \propto k^2$, we find that

$$T_1 = \tfrac{2}{3}JSk^2[(1-x)/x](1-2x)N^{-2}\sum_{f,f'}\sum_{\delta,\delta',\delta''}\hat{\epsilon}_f^{-1}\hat{\epsilon}_{f'}^{-1}\,\vec{\delta}\cdot\vec{\delta}''\,(1-e^{-i\vec{f}\cdot\vec{\delta}})(1-e^{-i\vec{f}\cdot\vec{\delta}'})(1-e^{i\vec{f}\cdot\vec{\delta}'})(1-e^{i\vec{f}\cdot\delta''})\;.\qquad (3.22)$$

Only terms even in all variables need be kept, whence

$$T_1 = \tfrac{2}{3}JSk^2[(1-x)/x](1-2x)N^{-2}\sum_{f,f'}\sum_{\delta,\delta',\delta''}\hat{\epsilon}_f^{-1}\hat{\epsilon}_{f'}^{-1}\,\vec{\delta}\cdot\vec{\delta}''\,\sin\vec{f}\cdot\vec{\delta}\,\sin\vec{f}\cdot\vec{\delta}'\,\sin\vec{f}'\cdot\vec{\delta}'\,\sin\vec{f}'\cdot\vec{\delta}''\;.\qquad (3.23)$$

It is now clear that since we treat a simple cubic lattice, only terms with $(\vec{\delta}\cdot\vec{\delta}')^2 = (\vec{\delta}'\cdot\vec{\delta}'')^2 = 1$ need be kept. Thus, we write

$$T_1 = \tfrac{16}{3}JSk^2[(1-x)/x]N^{-2}\sum_{f,f'}\sum_{\alpha}\hat{\epsilon}_f^{-1}\hat{\epsilon}_{f'}^{-1}\sin^2 f_\alpha\sin^2 f_\alpha'\;,\qquad (3.24)$$

where α takes on the values x, y, or z. Thus,

$$T_1 = \tfrac{16}{3}JSk^2[(1-x)/x](1-2x)\sum_\alpha\left(N^{-1}\sum_f\sin^2 f_\alpha\,\hat{\epsilon}_f^{-1}\right)^2\qquad (3.25a)$$

$$= I^2JSk^2(1-x)(1-2x)/x\;.\qquad (3.25b)$$

Similarly, we have

$$T_2 = (2JS/N^3x^5)x^4(1-x)^2\sum_{f,f'}\sum_r\sum_{\delta,\delta',\delta''}\hat{\epsilon}_f^{-1}\hat{\epsilon}_{f'}^{-1}$$

$$\times e^{i(\vec{f}-\vec{D}\cdot\vec{\delta}}e^{i(\vec{f}'-\vec{D}\cdot\vec{\delta}'}(1-e^{-i\vec{f}\cdot\vec{\delta}})(1-e^{i\vec{k}\cdot\vec{\delta}})\,(1-e^{-i\vec{f}'\cdot\vec{\delta}'})(1-e^{i\vec{f}\cdot\vec{\delta}'})(1-e^{-i\vec{k}\cdot\vec{\delta}''})(1-e^{i\vec{f}'\cdot\vec{\delta}''})\qquad (3.26a)$$

$$= \tfrac{2}{3}JSk^2(1-x)^2x^{-1}N^{-2}\sum_{f,f'}\sum_{\delta,\delta',\delta''}\hat{\epsilon}_f^{-1}\hat{\epsilon}_{f'}^{-1}\vec{\delta}\cdot\vec{\delta}''\,(e^{i\vec{f}\cdot\vec{\delta}}-1)(e^{i\vec{f}'\cdot\vec{\delta}'}-1)(1-e^{i\vec{f}\cdot\vec{\delta}'})(1-e^{i\vec{f}'\cdot\vec{\delta}''})\;.\qquad (3.26b)$$

If the signs of δ, δ'', and f' are reversed, this expression is seen to be proportional to Eq. (3.22) for T_1, so that

$$T_2 = I^2JSk^2(1-x)^2/x\;.\qquad (3.27)$$

For T_3 and T_4, we write

$$T_4 = T_3 = (2JS/N^3x^5)x^4(1-x)^2\sum_{f,f'}\sum_r\sum_{\delta,\delta',\delta''}\hat{\epsilon}_f^{-1}\hat{\epsilon}_{f'}^{-1}$$

$$\times e^{i(\vec{l}-\vec{b})\cdot\vec{\delta}}e^{i(\vec{l}-\vec{l}')\cdot\vec{\delta}''}(1-e^{-i\vec{l}\cdot\vec{\delta}})(1-e^{i\vec{k}\cdot\vec{\delta}})(1-e^{-i\vec{l}'\cdot\vec{\delta}'})(1-e^{i\vec{l}\cdot\vec{\delta}'})(1-e^{-i\vec{k}\cdot\vec{\delta}''})(1-e^{i\vec{l}'\cdot\vec{\delta}''}) \quad (3.28a)$$

$$=\tfrac{2}{3}JSk^2(1-x)^2x^{-1}N^{-2}\sum_{f,f'}\sum_{\delta,\delta',\delta''}\hat{\varepsilon}_f^{-1}\hat{\varepsilon}_{f'}^{-1}\vec{\delta}\cdot\vec{\delta}''(e^{i\vec{l}\cdot\vec{\delta}}-1)(1-e^{-i\vec{l}'\cdot\vec{\delta}'})(1-e^{i\vec{l}\cdot\vec{\delta}'})(e^{-i\vec{l}'\cdot\vec{\delta}''}-1)e^{i\vec{l}\cdot\vec{\delta}''} \; . \quad (3.28b)$$

Taking only the contributions even in the δ's and replacing the sums over δ's by sums over α's, we obtain

$$T_4=T_3=\tfrac{16}{3}JSk^2(1-x)^2x^{-1}N^{-2}\sum_{f,f'}\sum_{\alpha,\alpha',\alpha''}\hat{\varepsilon}_f^{-1}\hat{\varepsilon}_{f'}^{-1}\Delta_{\alpha,\alpha''}\sin f_\alpha$$

$$\times[(1-\cos f'_{\alpha'})(1-\cos f_{\alpha'})+\sin f'_{\alpha'}\sin f_{\alpha'}][\sin(f'_{\alpha''}-f_{\alpha''})+\sin f_{\alpha''}] \quad (3.29a)$$

$$=\tfrac{16}{3}JSk^2(1-x)^2x^{-1}N^{-2}\sum_{f,f'}\sum_{\alpha,\alpha'}\hat{\varepsilon}_f^{-1}\hat{\varepsilon}_{f'}^{-1}$$

$$\times[(1-\cos f'_{\alpha'})(1-\cos f_{\alpha'})\sin^2 f_\alpha(1-\cos f'_\alpha)+\sin f'_{\alpha'}\sin f_{\alpha'}\sin f'_\alpha\sin f_\alpha\cos f_\alpha] \quad (3.29b)$$

$$=\tfrac{16}{3}JSk^2(1-x)^2x^{-1}\left[\sum_{\alpha,\alpha'}Q_{\alpha,\alpha'}R_{\alpha,\alpha'}+\sum_\alpha\frac{I}{4}\left(\frac{I}{4}-Q_{\alpha\alpha}\right)\right]\;, \quad (3.29c)$$

where

$$Q_{\alpha,\alpha'}=N^{-1}\sum_f(1-\cos f_\alpha)\sin^2 f_{\alpha'}\hat{\varepsilon}_f^{-1}\;, \quad (3.30a)$$

$$R_{\alpha,\alpha'}=N^{-1}\sum_f(1-\cos f_\alpha)(1-\cos f_{\alpha'})\hat{\varepsilon}_f^{-1}\;. \quad (3.30b)$$

Numerically, we have[31-33]

$$Q_{\alpha,\alpha'}\approx 0.08\;,\quad \alpha\neq\alpha'\;; \quad (3.31a)$$

$$Q_{\alpha,\alpha}\approx 0.09\;; \quad (3.31b)$$

$$R_{\alpha,\alpha'}\approx 0.135\;,\quad \alpha\neq\alpha'\;; \quad (3.31c)$$

$$R_{\alpha,\alpha}\approx 0.23\;. \quad (3.31d)$$

Next, for T_5 and T_6, we write

$$T_6=T_5=(2JS/N^3x^5)x^3(1-x)^2(1-2x)\sum_{f,f'}\sum_r\sum_{\delta,\delta'}\hat{\varepsilon}_f^{-1}\hat{\varepsilon}_{f'}^{-1}(1-e^{-i\vec{l}\cdot\vec{\delta}})(1-e^{i\vec{k}\cdot\vec{\delta}})(1-e^{-i\vec{l}'\cdot\vec{\delta}'})(1-e^{i\vec{l}\cdot\vec{\delta}'})$$

$$\times(1-e^{-i\vec{k}\cdot\vec{\delta}'})(1-e^{i\vec{l}'\cdot\vec{\delta}'}) \quad (3.32a)$$

$$=\tfrac{2}{3}JSk^2\left(\frac{1-x}{x}\right)^2(1-2x)N^{-2}\sum_{f,f'}\sum_{\delta,\delta'}\hat{\varepsilon}_f^{-1}\hat{\varepsilon}_{f'}^{-1}\vec{\delta}\cdot\vec{\delta}'(1-e^{-i\vec{l}\cdot\vec{\delta}})(1-e^{i\vec{l}\cdot\vec{\delta}'})(1-e^{-i\vec{l}'\cdot\vec{\delta}'})(1-e^{i\vec{l}'\cdot\vec{\delta}'})\;. \quad (3.32b)$$

Taking only terms even in all variables, we have

$$T_6=T_5=\tfrac{4}{3}JSk^2\left(\frac{1-x}{x}\right)^2(1-2x)N^{-2}\sum_{f,f'}\sum_{\delta,\delta'}\hat{\varepsilon}_f^{-1}\hat{\varepsilon}_{f'}^{-1}\vec{\delta}\cdot\vec{\delta}'(1-\cos\vec{f}'\cdot\vec{\delta}')\sin\vec{f}\cdot\vec{\delta}\sin\vec{f}\cdot\vec{\delta}' \quad (3.33a)$$

$$=\tfrac{8}{9}JSk^2\left(\frac{1-x}{x}\right)^2(1-2x)N^{-1}\sum_{f,\alpha}\hat{\varepsilon}_f^{-1}\sin^2 f_\alpha \quad (3.33b)$$

$$=\tfrac{2}{3}JSk^2\left(\frac{1-x}{x}\right)^2(1-2x)I\;. \quad (3.33c)$$

For T_7, we have

$$T_7=\frac{2JS}{N^3x^5}x^3(1-x)^2(1-2x)\sum_{f,f'}\sum_r\sum_{\delta,\delta'}\hat{\varepsilon}_f^{-1}\hat{\varepsilon}_{f'}^{-1}$$

$$\times(1-e^{-i\vec{l}\cdot\vec{\delta}})(1-e^{i\vec{k}\cdot\vec{\delta}})(1-e^{-i\vec{l}'\cdot\vec{\delta}'})(1-e^{i\vec{l}\cdot\vec{\delta}'})(1-e^{-i\vec{k}\cdot\vec{\delta}})(1-e^{i\vec{l}'\cdot\vec{\delta}}) \quad (3.34a)$$

$$=\tfrac{2}{3}JSk^2\left(\frac{1-x}{x}\right)^2(1-2x)N^{-2}\sum_{f,f'}\sum_{\delta,\delta'}\hat{\varepsilon}_f^{-1}\hat{\varepsilon}_{f'}^{-1}(1-e^{-i\vec{l}\cdot\vec{\delta}})(1-e^{i\vec{l}\cdot\vec{\delta}'})(1-e^{-i\vec{l}'\cdot\vec{\delta}'})(1-e^{i\vec{l}'\cdot\vec{\delta}'})\;. \quad (3.34b)$$

Again we keep only that part of the integrand which is even in all variables:

$$T_7=\tfrac{2}{3}JSk^2\left(\frac{1-x}{x}\right)^2(1-2x)\sum_{\delta,\delta'}\left[\left(N^{-1}\sum_f\hat{\varepsilon}_f^{-1}(1-\cos\vec{f}\cdot\vec{\delta})(1-\cos\vec{f}\cdot\vec{\delta}')\right)\right.$$

$$\times\left(N^{-1}\sum_{f'}\hat{\epsilon}_{f'}^{-1}(1-\cos\vec{f}'\cdot\vec{\delta})(1-\cos\vec{f}'\cdot\vec{\delta}')\right)+\left(N^{-1}\sum_f\hat{\epsilon}_f^{-1}\sin\vec{f}\cdot\vec{\delta}\sin\vec{f}\cdot\vec{\delta}'\right)$$

$$\times\left(N^{-1}\sum_{f'}\hat{\epsilon}_{f'}^{-1}\sin\vec{f}'\cdot\vec{\delta}\sin\vec{f}'\cdot\vec{\delta}'\right)\Big]\qquad(3.35a)$$

$$=\tfrac{8}{3}JSk^2\left(\frac{1-x}{x}\right)^2(1-2x)\left(\sum_{\alpha,\alpha'}R^2_{\alpha,\alpha'}+\sum_\alpha\frac{I^2}{16}\right)\ .\qquad(3.35b)$$

Finally, for T_8, we write

$$T_8=\frac{JS}{N^3x^5}\,x^2(1-x)^2(1-2x)^2\sum_{f,f'}\sum_r\sum_\delta\hat{\epsilon}_f^{-1}\hat{\epsilon}_{f'}^{-1}(1-e^{-i\vec{f}\cdot\vec{\delta}})(1-e^{i\vec{k}\cdot\vec{\delta}})(1-e^{-i\vec{f}\cdot\vec{\delta}})(1-e^{i\vec{f}\cdot\vec{\delta}})(1-e^{-i\vec{k}\cdot\vec{\delta}})(1-e^{i\vec{f}'\cdot\vec{\delta}})$$

$$(3.36a)$$

$$=\tfrac{1}{3}JSk^2(1-x)^2(1-2x)^2x^{-3}N^{-2}\sum_{f,f'}\sum_\delta\hat{\epsilon}_f^{-1}\hat{\epsilon}_{f'}^{-1}(1-e^{-i\vec{f}\cdot\vec{\delta}})(1-e^{i\vec{f}\cdot\vec{\delta}})(1-e^{-i\vec{f}'\cdot\vec{\delta}})(1-e^{i\vec{f}'\cdot\vec{\delta}})\qquad(3.36b)$$

$$=\tfrac{8}{3}JSk^2(1-x)^2(1-2x)^2x^{-3}\sum_\alpha\left(N^{-1}\sum_f\hat{\epsilon}_f^{-1}(1-\cos f_\alpha)\right)^2\qquad(3.36c)$$

$$=\tfrac{2}{9}JSk^2(1-x)^2(1-2x)^2x^{-3}\ .\qquad(3.36d)$$

Collecting all the above results we have

$$E_k^{(3)}(x)=2JSk^2\left[A\frac{(1-x)(1-2x)}{x}+B\frac{(1-x)^2}{x}+C\left(\frac{1-x}{x}\right)^2(1-2x)+D\frac{(1-x)^2(1-2x)^2}{x^3}\right]\ ,\qquad(3.37)$$

with

$$A=\tfrac{1}{2}I^2\approx0.088\ ,\qquad(3.38a)$$

$$B=\tfrac{3}{2}I^2+16Q_{xx}R_{xx}+32Q_{xy}\,R_{xy}-4IQ_{xx}\approx0.77\ ,\qquad(3.38b)$$

$$C=\tfrac{2}{3}I+4R^2_{xx}+8R^2_{xy}+\tfrac{1}{4}I^2\approx0.68\ ,\qquad(3.38c)$$

$$D=\tfrac{1}{9}\ .\qquad(3.38d)$$

These results for the coefficients A, B, and D agree with those since found by Last,[8] but our evaluation of C is different from his. However, numerically this difference is unimportant. The corresponding variational results are shown in Fig. 6. Note that the result of Eq. (3.14b) never becomes negative. Also note the upturn in the energy near $x=\tfrac{1}{2}$. This is an indication that the trial function is becoming unreliable in this range of concentrations. As we shall see in Sec. IV, better results can be obtained by taking the optimal linear combination of the two wave functions given in Eqs. (3.13a) and (3.13b). Of course, all these conclusions will have to be suitably modified when localized excitations are excluded from the wave function, as we shall do in Sec. IV.

IV. VARIATIONAL CALCULATION WITH LOCALIZED EXCITATIONS EXCLUDED

In this section we remedy the defect mentioned earlier of Murray's unrestricted variational principle.[2] Now we shall use wave functions which are confined to the infinite cluster and hence are appropriate wave functions for spin waves. The resulting modifications in the formalism are discussed in Sec. IV A. The first-order variational result for the spin-wave energy is given in Sec. IV B. In Sec. IV C an optimal second-order variational result is given. In all these calculations the restricted averages required can only be evaluated approximately.

A. Formalism

The excitations are restricted to the infinite cluster by everywhere replacing p_i by $\hat{p}_i$, where

$$\hat{p}_i=1,\quad i\in\theta,\quad i\in\infty\qquad(4.1a)$$

$$\hat{p}_i=0,\quad\text{otherwise}\qquad(4.1b)$$

where $i\in\infty$ means that the ith site is in the infinite cluster. The wave functions where the p's are replaced by the $\hat{p}$'s will be denoted $|\hat{k}\rangle$. Also, matrix elements with $\hat{p}$'s replacing p's will be denoted $\hat{H}_{kf}$, etc. Thus, the configurational averages required to evaluate $E_k^{(n)}(x)$ are somewhat different from those used above, especially as x approaches x_p. For the simplest average, $\langle\langle\hat{p}_i\rangle\rangle\equiv\hat{x}$, this difference can be seen from Fig. 1.

Finally, as mentioned earlier, we shall optimize the second-order result by replacing $\hat{\psi}_k^{(2)}$ by $(1-\lambda)\hat{\psi}_k^{(1)}+\lambda\hat{\psi}_k^{(2)}$, where λ is a variational parameter.[34] The resulting trial energy $E_k^{(2)}(x,\lambda)$ is then

$$E_k^{(2)}(x,\lambda)=\left(\langle\hat{k}|\mathcal{K}|\hat{k}\rangle-2\lambda\sum_f\frac{\hat{H}_{kf}\hat{H}_{fk}}{xE_f}\right.$$

$$+ \lambda^2 \sum_{f,f'} \frac{\hat{H}_{kf}\langle \hat{f}|\mathcal{H}| \hat{f}'\rangle \hat{H}_{f'k}}{x^2 E_f E_{f'}} \Bigg)$$

$$\times \Bigg(1 - \lambda \sum_f \frac{\hat{H}_{kf}\langle \hat{f}|\hat{k}\rangle + \langle \hat{k}|\hat{f}\rangle \hat{H}_{fk}}{x E_f}$$

$$+ \lambda^2 \sum_{f,f'} \frac{\hat{H}_{kf}\langle \hat{f}|\hat{f}'\rangle \hat{H}_{f'k}}{x^2 E_f E_{f'}} \Bigg)^{-1}. \qquad (4.2)$$

The deviation of the denominator from unity will be neglected, since it leads to corrections higher order in k, whence

$$E_k^{(2)}(x, \lambda) = \langle \hat{k}|\mathcal{H}|\hat{k}\rangle + (\lambda^2 - 2\lambda)\sum_f \frac{\hat{H}_{kf}\hat{H}_{fk}}{x E_f}$$
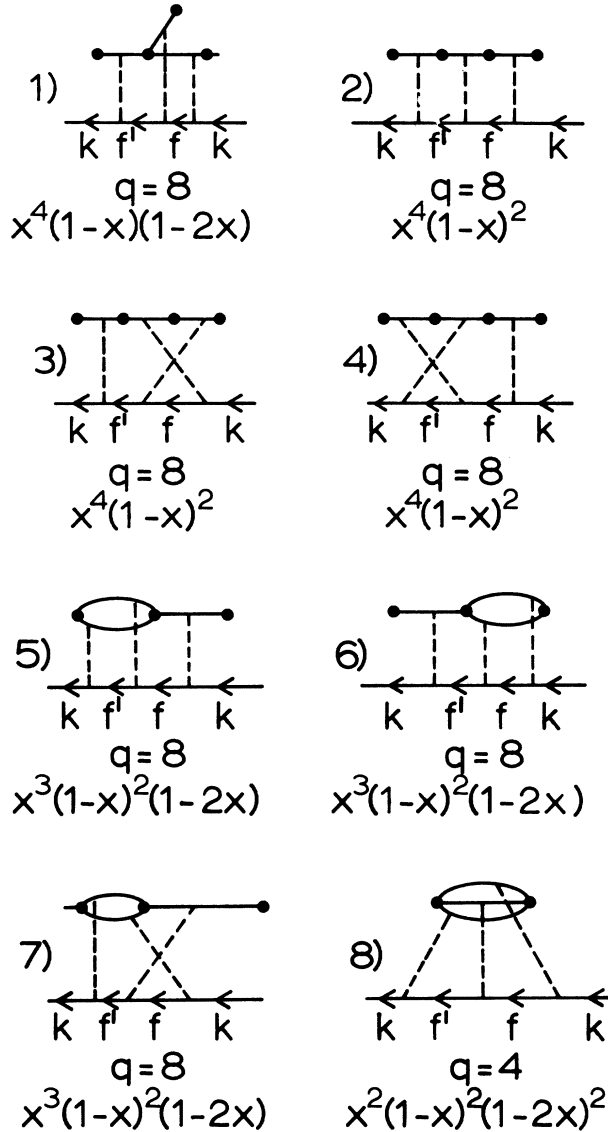

FIG. 5. Diagrammatic representation of contributions to $E_k^{(3)}$. The multiplicity and concentration factors associated with each diagram are given below it.
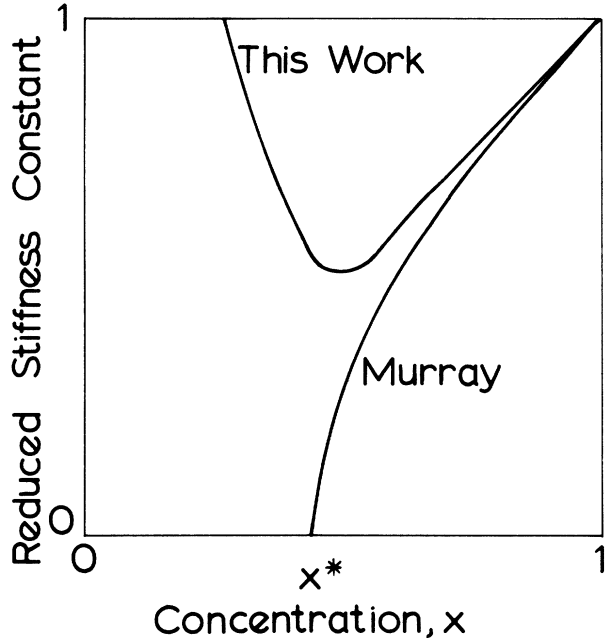

FIG. 6. Reduced stiffness constant $E_k(x)/E_k(1)$ from Eq. (3.14b) vs x, as evaluated by Murray (Ref. 2) and by us. The value of x^* is 0.44.

$$+ \lambda^2 \sum_{f,f'} \frac{\hat{H}_{kf}\hat{H}_{ff'}\hat{H}_{f'k}}{x^2 E_f E_{f'}}. \qquad (4.3)$$

When this expression is minimized with respect to λ, one obtains the optimal second-order energy $E_k^{(2)}(x)_{opt}$ as

$$E_k^{(2)}(x)_{opt} = \Big(\langle \hat{k}|\mathcal{H}|\hat{k}\rangle - \Big(\sum_f \frac{\hat{H}_{kf}\hat{H}_{fk}}{x E_f}\Big)^2$$

$$\times \Big(\sum_f \frac{\hat{H}_{kf}\hat{H}_{fk}}{x E_f} + \sum_{f,f'} \frac{\hat{H}_{kf}\hat{H}_{ff'}\hat{H}_{f'k}}{x^2 E_f E_{f'}}\Big)^{-1}. \qquad (4.4)$$

In deriving this result it is essential that the coefficient of λ^2 in Eq. (4.3) be positive. A negative value would indicate a mechanical instability, which is not to be expected.

B. First-Order Variational Result

The evaluation of the first-order variational energy when localized excitations are excluded yields the following result:

$$E_k^{(1)}(x) = \langle \hat{k}|\mathcal{H}|\hat{k}\rangle \qquad (4.5a)$$

$$= (2JS/N\hat{x})\sum_{\langle ij\rangle} \langle\langle \hat{p}_i \hat{p}_j \rangle\rangle (1 - e^{i\vec{k}\cdot(\vec{r}_i - \vec{r}_j)}). \qquad (4.5b)$$

To evaluate the average, we need the correlated average $\langle\langle \hat{p}_i \hat{p}_j \rangle\rangle$ for which we adopt the following approximation:

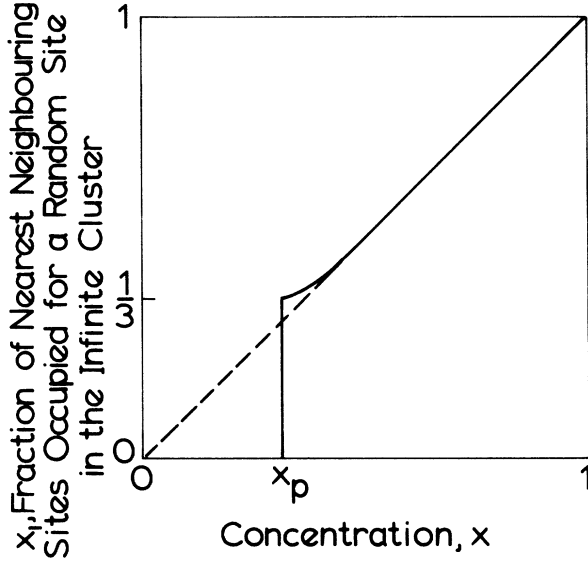$$\langle\langle \hat{p}_i \hat{p}_j \rangle\rangle = \hat{x}, \quad i = j \qquad (4.6a)$$

FIG. 7. Schematic plot of x_1 vs x. Here x_1 is the probability that a neighboring site of a site in the infinite cluster is occupied. Note that x_1 is expected to be different from x only for x near x_p. Also $x_1 \geq 2/z$, as is shown in Appendix B. This is also a plot of the reduced spin-wave energy $E_k^{(1)}(x)/E_k(1)$ obtained by the first-order variational calculation of Eq. (4.9).

$$\langle\langle \hat{p}_i \hat{p}_j \rangle\rangle = \hat{x} x_1, \quad i = j + \delta \tag{4.6b}$$

$$\langle\langle \hat{p}_i \hat{p}_j \rangle\rangle = \hat{x}^2, \quad \text{otherwise} \tag{4.6c}$$

where x_1 is the probability that if a spin belongs to the infinite cluster, a given neighboring site is occupied. Thus,

$$\langle\langle \hat{p}_i \hat{p}_j \rangle\rangle = \hat{x}^2 + \hat{x}(1-\hat{x})\,\Delta_{ij} + \sum_\delta \hat{x}(x_1 - \hat{x})\,\Delta_{i,j+\delta}. \tag{4.7}$$

The reason for adopting this approximation is that the nearest-neighboring correlations are the most important ones in the infinite cluster. Clearly, spins in the infinite cluster have more neighbors than a spin chosen randomly. In fact, in Appendix D it is shown that $x_1 \gtrless 2/z$, since, crudely speaking, each spin has at least two neighbors as required for connectivity. So, it follows that

$$x_1 > x > \hat{x}, \quad x_1 \geq 2/z. \tag{4.8}$$

In Fig. 7, we have given a schematic plot of x_1. We expect x_1 to be appreciably different from x only for x near x_p. Substituting Eq. (4.7) into Eq. (4.5b), one obtains

$$E_k^{(1)}(x) = 2JzS(1-\gamma_k)x_1. \tag{4.9}$$

This result is physically very reasonable, because zx_1 is the effective number of nearest neighbors seen by a given spin in the infinite cluster. Since $E_k^{(1)}(x)/E_k^{(1)}(1) = x_1$, Fig. 7 can also be regarded as

the plot for the first-order energy. Exclusion of localized excitations makes this energy larger than that obtained without exclusions.

C. Second-Order Variational Result

To evaluate the optimal second-order variational energy given in Eq. (4.4), it is clear that we must evaluate the quantities

$$\mathcal{E}_1 \equiv \sum_f \frac{\hat{H}_{kf}\hat{H}_{fk}}{xE_f}, \tag{4.10a}$$

$$\mathcal{E}_2 \equiv \sum_{f,f'} \frac{\hat{H}_{kf}\hat{H}_{ff'}\hat{H}_{f'k}}{x^2 E_f E_{f'}}. \tag{4.10b}$$

The expression on the right-hand side of Eq. (4.4) really can only be applied configuration by configuration. To configurationally average this expression, we have to introduce a large number of approximations. First, we shall average the numerator and denominator separately. This procedure is reasonable as long as these quantitites do not fluctuate too wildly. This condition is certainly fulfilled when x is near unity. However, for x near x_p this procedure does represent an uncontrolled approximation. Even with this simplification, the averages involve a large number of $\hat{p}$ factors, and thus cannot be evaluated exactly. We evaluate $\langle\langle \hat{p}_i \hat{p}_j \ldots \hat{p}_m \rangle\rangle$ approximately as follows. We draw the diagrams for this average as if we were evaluating the corresponding average of p's. Sites which are restricted by exchange interactions to be nearest neighbors of one another are connected by solid lines. The first approximation we make is to assume disconnected clusters to be statistically independent. This, of course, is not quite correct, even for the first diagram of Fig. 8 and Table I. Even if i, $i+\delta$, j, and $j+\delta'$ are all distinct, we cannot write

$$\langle\langle \hat{p}_i \hat{p}_{i+\delta} \hat{p}_j \hat{p}_{j+\delta'} \rangle\rangle = \langle\langle \hat{p}_i \hat{p}_{i+\delta} \rangle\rangle \langle\langle \hat{p}_j \hat{p}_{j+\delta'} \rangle\rangle \tag{4.11}$$

as we would for the corresponding variables p. The approximation of Eq. (4.11) is especially bad if one of the pair i, $i+\delta$ is a nearest neighbor of one of the pair j, $j+\delta'$. However, these correlations will be neglected.

It remains to assign the appropriate factors for connected diagrams. This we do recursively, as follows. Assume factors for all diagrams of n connected points have been determined, and let us discuss the determination of the factor for a diagram of $n+1$ connected points. We first assign to this diagram a "primary" factor $\hat{x} x_1^n$. This factor is a rough approximation to the probability of finding a site in the infinite cluster with a sequence of n neighbors. Now we must build up the analog of the cumulant expansion. A simple prescription for doing this is to open the diagram, i.e., to break it wherever two or more sites
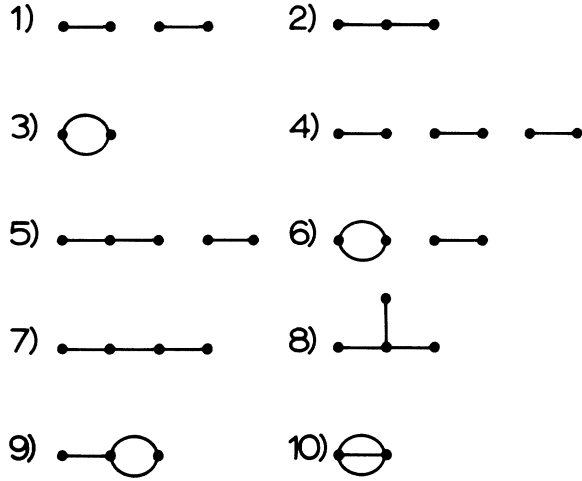
FIG. 8. Diagrams used in Table I to calculate cumulant averages for sites in the infinite cluster. Each line represents a factor of the type p_r p_{r+6}.

coalesce. For instance, in the case of the third diagram of Fig. 8 and Table I, we would obtain the broken diagrams shown in Fig. 9. Now we subtract the contribution of these broken diagrams, assumed to be determined previously, from the "primary" factor. The results of this procedure are shown in Table I where they are compared to the analogous results for the averages of p's, where the "primary" factor is x^{m+1}. Using the computations in Table I, we obtain

$$\mathcal{E}_1 = \sum_f \left\langle\!\!\left\langle \frac{\hat{H}_{kf}\hat{H}_{fk}}{x E_f} \right\rangle\!\!\right\rangle = 2JSk^2\hat{x}$$

$$\times [0.42 x_1(1-\hat{x})\hat{x}^{-2} + 3(1 - 2x_1 + x_1\hat{x})\hat{x}^{-2}],$$

(4.12)

$$\mathcal{E}_2 = \sum_{f,f'} \left\langle\!\!\left\langle \frac{\hat{H}_{kf}\hat{H}_{ff'}\hat{H}_{f'k}}{x^2 E_f E_{f'}} \right\rangle\!\!\right\rangle = 2JSk^2\hat{x}$$

$$\times [0.09 x_1(1-\hat{x})(1-2\hat{x})\hat{x}^{-3}$$

$$+ 0.68(1 - 2x_1 + x_1\hat{x})\hat{x}^{-2}$$

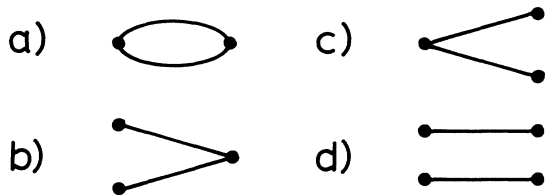$$+ 0.77(1 - 4x_1 + 5x_1\hat{x} - 2x_1\hat{x}^2)\hat{x}^{-3}$$

FIG. 9. Diagrammatic prescription for obtaining cumulant averages. The original bubble diagram (a) can be broken open either at the left vertex as in (b), at the right vertex as in (c), or at both vertices as in (d).
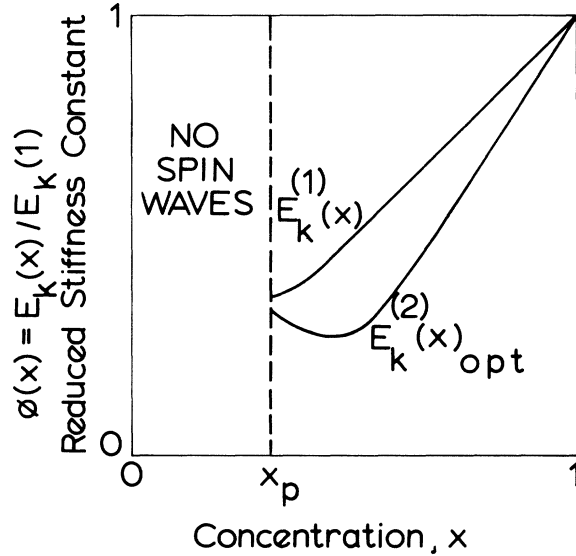

Concentration, x

FIG. 10. Variational bounds for the spin-wave energy with localized excitations excluded. The first-order result, taken from Eq. (4.9), is the same as in Fig. 7. The optimal second-order result was evaluated by inserting Eqs. (4.12) and (4.13) into Eq. (4.4).

$$+ \tfrac{1}{9}(1 - 6x_1 - 3x_1\hat{x} + 16x_1^2$$

$$- 12\hat{x}x_1^2 + 4\hat{x}^2x_1^2)x_1^{-1}\hat{x}^{-2}]. \qquad (4.13)$$

Unfortunately, $\hat{x}$ and x_1 are not known functions of x. For $\hat{x}$ we have used the result of Frisch *et al.*[4] shown in Fig. 1. For x_1 we simply used Fig. 7 which was drawn by fixing x_1 to be $2/z$ for $x = x_p$, and to join smoothly to the curve $x_1 = x$. The answers were not sensitive to the details of this procedure.

Using the results given in Eqs. (4.12) and (4.13), we have evaluated $E_k^{(2)}(x)_{opt}$ as given in Eq. (4.4) and the result is plotted in Fig. 10. As it must, $E_k^{(2)}(x)_{opt}$ lies below $E_k^{(1)}(x)$. The upturn in $E_k^{(2)}(x)_{opt}$ for x near x_p is probably an artifact of our calculation, which becomes less reliable as $x \to x_p$.

V. SIMPLE MODEL FOR DILUTION

We may illustrate some of the qualitative conclusions we have reached by a consideration of the following model. The crystal consists of N_{uc} unit cells of a simple cubic lattice. Each simple-cubic-lattice site is connected to each of its nearest neighbors by a chain of $(L + 1)$ spins (see Fig. 11). Adjacent spins on a chain are separated by a distance a and have an exchange coupling J. The unit cell thus contains $3L + 1$ spins and for $L = 0$ we recover the usual simple cubic lattice. Henceforth, we shall consider L to be large and will work to lowest order in L^{-1}.

The spin-wave stiffness D in a cubic system of

the type we consider is given by[35]

$$D = \sum_{i,j} J_{ij} S_i S_j r_{ij}^2 \left(3 \sum_i S_i\right)^{-1}, \qquad (5.1)$$

and, for the present model, this gives

$$D = \tfrac{2}{3} JSa^2 (1 + 1/L)/(1 + 1/3L) \qquad (5.2a)$$

$$\approx \tfrac{2}{3} JSa^2. \qquad (5.2b)$$

Correspondingly, the density of states is

$$g(E) = N_{uc} L^3 (2\pi)^{-2} E^{1/2} (\tfrac{2}{3} JS)^{-3/2}. \qquad (5.3)$$

Clearly, this description in terms of acoustic spin waves will only be correct at low energy. Moreover, the acoustic spin waves will only be described by the stiffness of Eq. (5.2) as long as $kLa \ll 1$. Thus, Eq. (5.3) only holds for energies less than $E_c \equiv 2JSL^{-2}$.

Outside the low-energy regime the system must look one dimensional. Using the usual dispersion relation for a linear Heisenberg ferromagnet, we find that for a system of $3N_{uc}$ chains of length La the density of states is

$$g(E) = 3N_{uc} (L/2\pi)[(8JS - E)E]^{-1/2}. \qquad (5.4)$$

This formula will be appropriate for energies larger than E_c. Note that Eqs. (5.3) and (5.4) are of the same order of magnitude at E_c where the density of states per spin [i.e., $g(E_c)/3LN_{uc}$] assumes its maximum value $\sim L/JS$.

Since the low-energy maximum in the density of states sharpens and moves to lower energy as L increases, we expect the transition temperature to decrease with increasing L. To see this more explicitly, we calculate the number of excitations needed to destroy long-range order and thereby locate T_c:

$$3LN_{uc}S = \int_0^\infty g(E)(e^{E/kT_c} - 1)^{-1} dE. \qquad (5.5)$$

Assuming $kT_c \gg E_c$, as is justified *a posteriori*, we find the contribution to the right-hand side of Eq. (5.5) from $E < E_c$ to be

$$\int_0^{E_c} dE \, (e^{E/kT_c} - 1)^{-1} N_{uc} L^3 (2\pi)^{-2} E^{1/2} (\tfrac{2}{3} JS)^{-3/2}$$

$$\sim L^3 E_c^{1/2} N_{uc} T_c S^{-3/2}. \qquad (5.6)$$

The contribution for $E > E_c$ is found, using Eq. (5.4), to be of order $LT_c^{1/2} S^{-1/2}$. Inserting these evaluations into Eq. (5.5) we see that T_c is of order S^2/L. This calculation shows that although the peak

TABLE I. Cumulant averages for sites randomly occupied and for sites within the infinite cluster.

| Diagram No. in Fig. 8	Sites randomly occupied			Sites within the infinite cluster		
	Primary factor	Contributions from broken diagrams	Total result	Primary factor	Contributions from broken diagrams	Total result
1	x^4		x^4	$\hat{x}^2 x_1^2$		$\hat{x}^2 x_1^2$
2	x^3	x^4	$x^3(1-x)$	$\hat{x} x_1^2$	$\hat{x}^2 x_1^2$	$\hat{x} x_1^2(1-\hat{x})$
3	x^2	$2x^3(1-x)$ x^4	$x^2(1-x)^2$	$\hat{x} x_1$	$2\hat{x} x_1^2(1-\hat{x})$ $x_1^2 \hat{x}^2$	$\hat{x} x_1(1-2x_1+\hat{x}x_1)$
4	x^6		x^6	$\hat{x}^3 x_1^3$		$\hat{x} x_1^3$
5	x^5	x^6	$x^5(1-x)$	$\hat{x}^2 x_1^3$	$\hat{x}^3 x_1^3$	$\hat{x}^2 x_1^3(1-\hat{x})$
6	x^4	$2x^5(1-x)$ x^6	$x^4(1-x)^2$	$\hat{x}^2 x_1^3$	$2\hat{x}^2 x_1^3(1-\hat{x})$ $\hat{x}^3 x_1^3$	$\hat{x}^2 x_1^3(1-2x_1+\hat{x}x_1)$
7	x^4	$2x^5(1-x)$ x^6	$x^4(1-x)^2$	$\hat{x} x_1^3$	$2\hat{x}^2 x_1^3(1-\hat{x})$ $\hat{x}^3 x_1^3$	$\hat{x} x_1^3(1-2\hat{x}+\hat{x}^2)$
8	x^4	$3x^5(1-x)$ x^6	$x^4(1-x)(1-2x)$	$\hat{x} x_1^3$	$3\hat{x}^2 x_1^3(1-\hat{x})$ $x_1^3 \hat{x}^2$	$\hat{x} x_1^3(1-\hat{x})(1-2\hat{x})$
9	x^3	$x^4(1-x)(1-2x)$ $3x^4(1-x)^2$ $4x^5(1-x)$ x^6	$x^3(1-x)^2(1-2x)$	$\hat{x} x_1^2$	$\hat{x} x_1^3(1-\hat{x})(1-2\hat{x})$ $3\hat{x} x_1^3(1-\hat{x})^2$ $4\hat{x}^2 x_1^3(1-\hat{x})$ $\hat{x}^3 x_1^3$	$\hat{x} x_1^2(1-4x_1 +5x_1\hat{x}-2x_1\hat{x}^2)$
10	x^2	$6x^3(1-x)^2(1-2x)$ $2x^4(1-x)(1-2x)$ $3x^4(1-x)^2$ $6x^4(1-x)^2$ $6x^5(1-x)$ x^6	$x^2(1-x)^2(1-2x)^2$	$\hat{x} x_1$	$6\hat{x} x_1^2(1-4x_1 +5x_1\hat{x}-2x_1\hat{x}^2)$ $2\hat{x} x_1^3(1-\hat{x})(1-2\hat{x})$ $3\hat{x}^2 x_1^2(1-2x_1+\hat{x}x_1)$ $6\hat{x} x_1^3(1-\hat{x})^2$ $6\hat{x}^2 x_1^3(1-\hat{x})$ $\hat{x}^3 x_1^3$	$\hat{x} x_1(1-6x_1 -3\hat{x}x_1+16x_1^2 -12x_1^2\hat{x}+4x_1^2\hat{x}^2)$
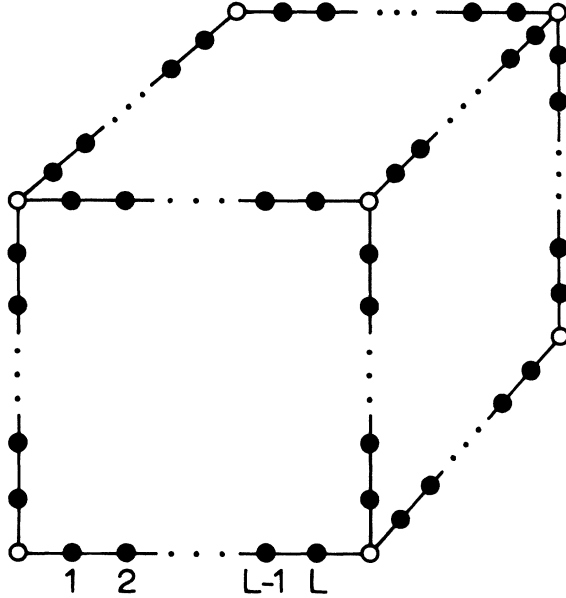
FIG. 11. Periodic model used to study the effects of dilution. The open circles represent spins on sites of the simple-cubic lattice. The filled circles represent spins on a chain joining simple-cubic-lattice sites.

in the density of states occurs at an energy of order L^{-2}, T_c varies as L^{-1}. In any event, the main conclusion is that T_c vanishes as L tends to infinity.

It is clear that we should not identify the concentration of the random alloy with that of the model introduced here, since the present system remains three dimensional in the limit as L^{-1}, and hence the concentration goes to zero. Instead it seems plausible to identify L^3 as a characteristic volume over which the system appears to be three dimensional. With this interpretation we may discuss the behavior of $T_c(x)$ for a random alloy as $x \to x_c$. For $x > x_c$, the infinite cluster must be three dimensional. If, in the limit $x \to x_c$, it is "strongly three dimensional" in the sense that L^3 remains finite, then $T_c(x)$ will drop discontinuously to zero at $x = x_c$. If, however, L^3 diverges as $x \to x_c$, then $T_c(x)$ will go continuously to zero as $x \to x_c$. In the latter case one might even hope that the critical indices for L and T_c are related in the same way as for our simple model. Some possible behaviors of $T_c(x)$ are shown in Fig. 12. Our guess is that $L \to \infty$ as $x \to x_c$, so that a "second-order" transition occurs. A more rigorous version of the above argument would be desirable, and the approach used by Elliott[36] based on random walks may be useful.

This model also demonstrates our remarks about the use of high-temperature expansions to determine $T_c(x)$. For $L \gg 1$, it is clear that the terms in the high-temperature expansion will be essentially one dimensional up to order $(J/kT)^m$, where

m is of order L. Physically, this is because the system behaves one dimensionally for $kT \gg E_c$.

This model is unphysical in certain of its properties because of the artificiality of the periodicity present in the model. For example, this model gives $\varphi(x) \geq \frac{1}{3}$ at all dilutions, in contrast to the predictions of Ref. 8. As Thouless has pointed out,[37] one defect of the model is that all circuits "lead somewhere." In the language of the conductivity models it is that the nonvanishing mobility is an artifact of the periodic model we are using. To illustrate the argument of Thouless one can study the simple cubic lattice where each site is decorated by addition of a loop of N_0 spins each of magnitude S connected by isotropic exchange interactions. That is, we add to the Hamiltonian for a simple cubic lattice the term

$$\mathcal{K}' = - 2J \sum_r \sum_{i=0}^{N_0} \vec{S}_{r,i} \cdot \vec{S}_{r,i+1} , \tag{5.7}$$

where $\vec{S}_{r,N_0+1} \equiv \vec{S}_{r,0}$ is the spin on the simple-cubic-lattice site at r and $S_{r,1}$, $S_{r,2}$, $\cdots S_{r,N_0}$ are the spins on the loop at r.

For this model the unit cell contains $N_0 + 1$ spins and the dynamical matrix describing these degrees of freedom is quite simple. From this matrix the energy of the acoustic mode in the long-wavelength limit is found to be

$$E_k = 2JSk^2/(1 + N_0) . \tag{5.8}$$

This result shows that loading the system by circuits that "do not lead anywhere" is at least partially responsible for $\Phi(x)$ vanishing for $x \to x_p$.

VI. CONCLUSIONS

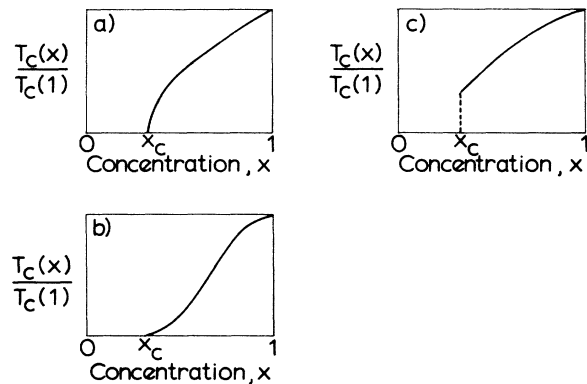We have studied spin waves in the diluted Heisenberg ferromagnet at zero temperature and have

FIG. 12. Some simple possible behaviors for $T_c(x)$. In (a) and (b) we show a second-order "transition" with $T_c(x)$ having either infinite slope [as in (a)] or finite slope [as in (b)] as $x \to x_c$. In (c) we show a first-order "transition." Elliott's arguments (Ref. 36) indicate that $T_c(x)$ for the Ising model resembles diagram (a).

reached the following conclusions:

(a_1) It is important to take account of the fact that spin waves can occur only in that part of the spin system which is connected (via interactions between magnetically occupied sites) into an infinite cluster. An automatic consequence of this formulation is that spin waves cannot exist for concentrations x less than the critical percolation concentration x_p.

(a_2) A rigorous variational principle has yet to be formulated for spin waves in the diluted system. If fluctuations in the spin-wave energy due to configurational disorder are small, then a variational calculation will have approximate validity, providing localized excitations in isolated clusters are excluded.

(a_3) Even when localized excitations are not excluded, the variational expression is positive for all concentrations, in contrast to Murray's result which becomes zero at $x = 0.44$ for a simple-cubic lattice.

(a_4) Excluding localized excitations in an approximate calculation yields results for $x > x_p$ qualitatively similar to those obtained using the unrestricted variational wave function.

We have also considered the question of determining the critical concentration x_c for the occurrence of long-range order and have reached the following conclusions:

(b_1) Although the susceptibility for $x \approx x_p$ is dominated by localized excitations, x_c is determined exclusively by the response of the infinite cluster.

(b_2) Although the infinite cluster is mechanically stable for $x = x_p$, its thermal stability determines whether x_c is in fact equal to x_p. Thus, if the in-

finite cluster is one dimensional, neither the Heisenberg nor the Ising systems will order. If it is two-dimensional, only the Ising system will order, and if it is three-dimensional, both systems will order.

(b_3) The same type of arguments which rule out long-range order in one dimension suggest that the infinite cluster is not one dimensional. It is conceivable, but unlikely, that the cluster is two dimensional over a small range of concentration above x_p. We thus conclude that x_c is close, if not exactly equal, to x_p.

(b_4) Whether or not $T_c(x)$ is discontinuous at $x = x_c$ is shown to depend on the degree to which the infinite cluster is three dimensional. If the characteristic volume over which the infinite cluster is three dimensional becomes infinite as $x \to x_c$, then $T_c(x)$ goes continuously to zero as $x \to x_c$.

ACKNOWLEDGMENTS

APPENDIX A: OVERLAP OF TRIAL FUNCTIONS AT LONG WAVELENGTHS

We use Murray's expressions, slightly amended,[38] for the overlap between second-order trial functions:

$$\langle \psi_{k'}^{(2)} \mid \psi_k^{(2)} \rangle = \Delta_{k,k'} + \sum_f{}' \frac{E_k P_{fk} - x^{-1} H_{fk}}{E_f - E_k} P_{k'f} + \sum_f{}' P_{fk} \frac{E_{k'} P_{k'f} - x^{-1} H_{k'f}}{E_f - E_{k'}}$$

$$+ \sum_f{}' \left(\frac{E_k P_{fk} - x^{-1} H_{fk}}{E_f - E_k} \right) \left(\frac{E_{k'} P_{k'f} - x^{-1} H_{k'f}}{E_f - E_{k'}} \right) + \sum_{f,f'}{}' \left(\frac{E_k P_{kf} - x^{-1} H_{fk}}{E_f - E_k} \right) P_{f'f} \left(\frac{E_{k'} P_{k'f} - x^{-1} H_{k'f}}{E_f - E_{k'}} \right) , \quad \text{(A1)}$$

where $P_{kf} = \langle k \mid f \rangle - \Delta_{kf}$ and the primes on the summations indicate that terms with zero denominators are to be omitted. We now consider the $k = 0$ limit and write for finite k',

$$\Lambda_{k'} = \lim_{k \to 0} \mid \langle \psi_{k'}^{(2)} \mid \psi_k^{(2)} \rangle \mid^2 \quad \text{(A2a)}$$

$$= \lim_{k \to 0} \mid \sum_f{}' P_{fk} \frac{E_{k'} p_{k'f} - x^{-1} H_{k'f}}{E_f - E_k} \mid^2 \quad \text{(A2b)}$$

$$= \lim_{k \to 0} \sum_{f,f'} \sum_{r,s,t,u} (E_f - E_k)^{-1} (E_{f'} - E_{k'})^{-1} (Nx)^{-4} \, e^{i(\vec{f} - \vec{k}) \cdot \vec{r}} \, e^{i(\vec{k} - \vec{f}) \cdot \vec{s}} \, e^{i(\vec{k}' - \vec{f}) \cdot \vec{t}} \, e^{i(\vec{f}' - \vec{k}) \cdot \vec{u}} (p_r - x)(p_s - x)$$

$$\times \left(E_{k'} (p_t - x) - \frac{JS}{x} \sum_{\delta'} (1 - e^{-i\vec{f}\cdot\vec{\delta}})(1 - e^{i\vec{k}\cdot\vec{\delta}'})(p_t\, p_{t+\delta} - x^2) \right)$$

$$\times \left(E_{k'} (p_u - x) - \frac{JS}{x} \sum_{\delta'} (1 - e^{i\vec{f}'\cdot\vec{\delta}'})(1 - e^{-i\vec{k}'\cdot\vec{\delta}'})(p_u\, p_{u+\delta'} - x^2) \right). \tag{A3}$$

One can evaluate the expression in Eq. (A3) by a diagrammatic formulation similar to that introduced in connection with Fig. 5. However, our more limited aim here is to show that $\langle\langle \Lambda_{k'} \rangle\rangle \neq 0$. Accordingly, we will work to lowest order in $(1-x)$. This can be done easily by selecting the terms in the configurational average where Kronecker Δ functions occur only at a single site. This is, we set

$$\langle\langle(p_r - x)(p_s - x)(p_t - x)(p_u - x)\rangle\rangle = f_4(x)\Delta_{r,s}\Delta_{s,t}\Delta_{t,u} \tag{A4a}$$

$$\langle\langle(p_r - x)(p_s - x)(p_t - x)(p_u p_{u+\delta'} - x^2)\rangle\rangle = 2xf_4(x)\Delta_{r,s}\Delta_{s,t}\Delta_{t,u}, \tag{A4b}$$

$$\langle\langle(p_r - x)(p_s - x)(p_t p_{t+\delta} - x^2)(p_u p_{u+\delta'} - x^2)\rangle\rangle = 4x^2 f_4(x)\Delta_{r,s}\Delta_{s,t}\Delta_{t,u}, \tag{A4c}$$

where $f_4(x)$ is the cumulant average of x^4. Here the factors 2 and 4 in Eqs. (A4b) and (A4c) take account of similar terms where t and $u + \delta'$, for example, are equal. The terms we have omitted in Eq. (A4) are of order $(1-x)^2$, whereas those retained are of order $f_4(x) \approx (1-x)$, for $x \approx 1$. Using Eq. (A4) we write Eq. (A3) to lowest order in $(1-x)$ as

$$\langle\langle \Lambda_{k'} \rangle\rangle = f_4(x)N^{-3}x^{-4} \sum_{f,f'} E_f^{-1}(E_f - E_{k'})^{-1} \left[E_{k'}^2 - E_{k'}(2JS) \sum_\delta (1 - e^{-i\vec{f}\cdot\vec{\delta}})(1 - e^{i\vec{k}\cdot\vec{\delta}}) \right.$$

$$\left. - E_{k'}(2JS) \sum_\delta (1 - e^{i\vec{f}'\cdot\vec{\delta}})(1 - e^{-i\vec{k}'\cdot\vec{\delta}}) + (2JS)^2 \sum_{\delta,\delta'} (1 - e^{-i\vec{f}\cdot\vec{\delta}})(1 - e^{i\vec{k}\cdot\vec{\delta}})(1 - e^{i\vec{f}'\cdot\vec{\delta}'})(1 - e^{-i\vec{k}'\cdot\vec{\delta}'}) \right], \tag{A5a}$$

$$= \frac{E_{k'}^2 f_4(x)}{Nx^4} \left(N^{-1} \sum_f \gamma_f E_f^{-1} \right)\left(N^{-1} \sum_{f'} \gamma_{f'}(E_{f'} - E_{k'})^{-1} \right). \tag{A5b}$$

Thus $\langle\langle \Lambda_{k'} \rangle\rangle \neq 0$.

APPENDIX B: NONORTHOGONALITY EFFECTS

The purpose of this Appendix is to show that it is possible to apply the variational principle in the form of Eq. (3.8) even when the set of wave functions used is not perfectly orthonormalized. The gist of the argument is that for small values of k_0 the set of trial wave functions for $k < k_0$ is a small one and these states are nearly orthogonal. There is no need even to consider the overlap between these states and any states with $k > k_0$. We shall carry out the argument for the set of states of Eq. (3.12),

$$|k\rangle = (Nx)^{-1/2} \sum_r p_r e^{i\vec{k}\cdot\vec{r}}|r\rangle. \tag{B1}$$

The necessary generalizations for the states $|\hat{k}\rangle$ or to second-order wave functions are straightforward.

We define the overlap matrix S by

$$\langle k'|k\rangle = \Delta_{k,k'} + S_{k',k}. \tag{B2}$$

Comparison with Eq. (B1) shows that

$$S_{k',k} = (Nx)^{-1} \sum_r (p_r - x)e^{i(\vec{k}'-\vec{k})\cdot\vec{r}}. \tag{B3}$$

An orthonormal set of states can formally be gen-

erated by setting

$$\Phi_k = \sum_p T_{kp}|p\rangle, \tag{B4}$$

where the matrices T and S are related by

$$T = (1 + S)^{-1/2} \tag{B5a}$$

$$= 1 - \tfrac{1}{2}S + \tfrac{3}{8}S^2 + \cdots. \tag{B5b}$$

It is natural to consider the sequence of wave functions of the type of Eq. (B4) when Eq. (B5b) is taken to nth order in S. However, this approach is not satisfactory, because unless $(1 - x)$ is small, the overlap between these states does not become smaller as n increases. Mathematically, this can be verified by evaluating either $\langle\langle(\Phi_k^{(n)}|\Phi_{k'}^{(n)})\rangle\rangle$ or $\langle\langle|(\Phi_k^{(n)}|\Phi_{k'}^{(n)})|^2\rangle\rangle$ for small values of n. Physically, the poor convergence of Eq. (B5b) is because the expansion is in powers of $(p_r - x)/x$, a fluctuating quantity which is only small for $x \approx 1$.

Instead, we only attempt to orthogonalize the states for which $k < k_0$. This we do using the Schmidt procedure in the following way. First, we order the wave functions with increasing k. That is, we set

$$\psi_n = |k_n\rangle, \tag{B6}$$

with $k_n \leq k_m$, if $n < m$. Now the orthonormal set is constructed as follows. Set $\Phi_1 = \alpha \psi_1$, where α is a normalization constant. Set $\Phi_2 = \alpha' (\psi_2 - \beta \Phi_1)$ and determine α' and β so that $(\Phi_2 | \Phi_1) = 0$ and $(\Phi_2 | \Phi_2) = 1$. Here β will be of order N^{-1}. In general we will have

$$\Phi_j = \alpha_j (\psi_j - \beta_1 \Phi_1 - \beta_2 \Phi_2 \cdots - \beta_{j-1} \Phi_{j-1}) , \qquad \text{(B7)}$$

where each β_k will be of order N^{-1}. But typically j is of order ξN, where ξ is of order the volume of the sphere in phase space over which the orthogonalization is carried, i.e., $\xi \sim k^3$. Thus, it is clear that the Φ's are different from the ψ's by terms of order ξ. More precisely, we set

$$\Phi_k = |k\rangle - \sum_{k' \leq k} V_{k'k} |k'\rangle . \qquad \text{(B8)}$$

Note in particular that the sum is carried only over $k' \leq k$. The orthonormality condition requires V to satisfy

$$V_{fk} = S_{fk} - \sum_{k' \leq k} V_{k'k} S_{fk'} - \sum_{f' \leq f} V^*_{f'f} S_{f'f}$$
$$+ \sum_{f' \leq f} V^*_{f'k} V_{f'k} + \sum_{f' \leq f} \sum_{k' \leq k} V^*_{f'f} V_{k'k} S_{f'k'}$$
$$- \Delta_{fk} V^*_{kk} , \qquad \text{(B9)}$$

where $f \leq k$. When this equation is solved by iteration the result is that V is given as a power series in ξ: terms involving n phase-space integrals each confined to a volume of order k^3 are nth order in ξ.

To illustrate the argument we consider the first iteration of Eq. (B9), which yields

$$\Phi_k^{(1)} = |k\rangle - \sum_{k' \leq k} S_{k'k} |k'\rangle . \qquad \text{(B10)}$$

Since we now have the restriction $k' \leq k$, the $\frac{1}{2}$ factor present in Eq. (B5b) does not appear. We wish to show that

$$(\Phi_k^{(1)} | \mathcal{K} | \Phi_k^{(1)}) \equiv F_k \qquad \text{(B11)}$$

is the same, to lowest order in k, as

$$\langle k | \mathcal{K} | k \rangle = E_k^{(1)} . \qquad \text{(B12)}$$

Using Eq. (B10) for $\Phi_k^{(1)}$, we have

$$F_k = E_k^{(1)} - \sum_{f<k} \langle k | \mathcal{K} | f \rangle S_{fk} - \sum_{f<k} S_{kf} \langle f | \mathcal{K} | k \rangle$$
$$+ \sum_{f'<k} \sum_{f<k} S_{kf'} \langle f' | \mathcal{K} | f \rangle S_{fk} . \qquad \text{(B13)}$$

Since the analysis of the other terms is quite similar, we shall consider here only the second term on the right-hand side of Eq. (B13). Its configurational average, denoted T_2, is

$$T_2 = - (JS/Nx)(Nx)^{-1} \sum_{f<k} \sum_{r,\delta} \sum_{r'} e^{i(\vec{k}-\vec{f})\cdot\vec{r}} (1 - e^{-i\vec{f}\cdot\vec{\delta}})$$
$$\times (1 - e^{i\vec{k}\cdot\vec{\delta}}) e^{i(\vec{f}-\vec{k})\cdot\vec{r}} \langle\langle (p_r p_{r+\delta} - x^2)(p_{r'} - x) \rangle\rangle .$$
$$\text{(B14)}$$

The configurational average vanishes unless $r = r'$ or $r = r' + \delta$, whence

$$T_2 = - JS(Nx)^{-2} \sum_{f<k} \sum_{r,\delta} (1 + e^{i(\vec{f}-\vec{k})\cdot\vec{\delta}}) x^2 (1 - x)$$
$$\times (1 - e^{-i\vec{f}\cdot\vec{\delta}})(1 - e^{i\vec{k}\cdot\vec{\delta}}) \quad \text{(B15a)}$$

$$= - 2JS(1-x)N^{-1} \sum_{f<k} \sum_{\delta} [(1 - \cos\vec{f}\cdot\vec{\delta})$$
$$\times (1 - \cos\vec{k}\cdot\vec{\delta}) + \sin\vec{f}\cdot\vec{\delta} \sin\vec{k}\cdot\vec{\delta}] \quad \text{(B15b)}$$

$$= - 2JS(1-x)N^{-1} \sum_{f<k} \sum_{\delta} (1 - \cos\vec{f}\cdot\vec{\delta})$$
$$\times (1 - \cos\vec{k}\cdot\vec{\delta}) , \quad \text{(B15c)}$$

which is of order

$$T_2 \sim \sum_{f<k} f^2 k^2 \qquad \text{(B16a)}$$

$$\sim k^7 , \qquad \text{(B16b)}$$

which is negligible. Hence $E_k^{(1)}$ is a sufficiently good approximation to F_k.

APPENDIX C: DERIVATION OF EQ. (3.14b)

In this Appendix we derive Eq. (3.14b) of the text. To evaluate the left-hand side of Eq. (3.14b) we use the analysis of Ref. 2 and also of Appendix A to write

$$\lim_{k \to 0} (\psi_k^{(2)} | \psi_k^{(2)}) = 1 . \qquad \text{(C1)}$$

Using Eq. (3.13b), we have that

$$\frac{(\psi_k^{(2)} | \mathcal{K} | \psi_k^{(2)})}{(\psi_k^{(2)} | \psi_k^{(2)})} = \langle k | \mathcal{K} | k \rangle$$
$$- x^{-1} \sum_f{}' H_{fk} \langle k | \mathcal{K} | f \rangle E_f^{-1}$$
$$- x^{-1} \sum_f{}' H_{kf} \langle f | \mathcal{K} | k \rangle E_f^{-1}$$
$$+ x^{-2} \sum_{f,f'}{}' H_{kf'} \langle f' | \mathcal{K} | f \rangle H_{fk} E_f^{-1} E_{f'}^{-1} . \quad \text{(C2)}$$

Here the prime on the summations indicates that terms with f or f' equal to k are to be omitted. We next use Eq. (3.11) to express $\langle k | \mathcal{K} | k \rangle$, $\langle k | \mathcal{K} | f \rangle$, $\langle f | \mathcal{K} | k \rangle$, and $\langle f' | \mathcal{K} | f \rangle$ in terms of the $H_{k,k'}$'s. We then may write Eq. (C2) as

$$(\psi_k^{(2)} | \mathcal{K} | \psi_k^{(2)}) / (\psi_k^{(2)} | \psi_k^{(2)})$$
$$= xE_k + H_{kk} - x^{-1} \sum_f{}' H_{kf} H_{fk} E_f^{-1}$$
$$+ x^{-2} \sum_{f,f'}{}' H_{kf'} H_{f'f} H_{fk} E_f^{-1} E_{f'}^{-1} . \qquad \text{(C3)}$$

Since the configurational average of $H_{k,k}$ vanishes, we may omit it, in which case Eq. (C3) becomes equivalent to Eq. (3.14b).

APPENDIX D: PROOF THAT $x_1 > 2/z$

For any connected cluster, consider the following quantities: N, the number of sites in the cluster, and Q, the number of nearest neighbors in the cluster:

$$Q = \sum_i q_i , \qquad (D1)$$

where q_i is the number of nearest neighbors of the ith site. For a single isolated site, $N = 1$ and $Q = 0$. For a two-site cluster, $N = 2$ and $Q = 2$. Now add sites to the cluster in any way. Each time a site is added one has for the changes in N and Q,

$$\Delta N = 1 , \qquad (D2a)$$

$$\Delta Q \geq 2 . \qquad (D2b)$$

Hence for an infinitely large cluster, $x_1 \equiv Q/N \geq 2$.

---

*Work supported in part by the National Science Foundation.

†Present address: Department of Physics, Roorkee University, Roorkee, U.P., India.

‡John Simon Guggenhein Fellow. Present address: Department of Theoretical Physics, Oxford University, Oxford, England (through 1973).

[1] E. C. Svensson, W. J. L. Buyers, T. M. Holden, R. A. Cowley, and R. W. H. Stevenson, AIP Conf. Proc. **5**, 1315 (1972); D. J. Breed, K. Gilijamse, J. W. E. Sterkenbur, and A. R. Miedema, J. Appl. Phys. **41**, 1267 (1970).

[2] G. A. Murray, Proc. Phys. Soc. Lond. **89**, 87 (1966); Proc. Phys. Soc. Lond. **89**, 111 (1966).

[3] D. Kumar and A. B. Harris, AIP Conf. Proc. **5**, 1345 (1972). See also D. Kumar, Ph.D. thesis (University of Pennsylvania, 1972) (unpublished).

[4] H. L. Frish, J. M. Hammersley, and D. J. A. Welsh, Phys. Rev. **126**, 949 (1962).

[5] R. J. Elliott, R. B. Heap, D. J. Morgan, and G. S. Rushbrooke, Phys. Rev. Lett. **5**, 366 (1960).

[6] D. J. Morgan and G. S. Rushbrooke, Mol. Phys. **4**, 291 (1961).

[7] D. J. Morgan and G. S. Rushbrooke, Mol. Phys. **6**, 477 (1963).

[8] B. J. Last, J. Phys. C **5**, 2805 (1972); B. J. Last, M.Sc. thesis (University of Birmingham, 1972) (unpublished); S. Kirkpatrick (unpublished).

[9] R. B. Griffiths, Phys. Rev. **136**, A437 (1964).

[10] L. Landau and E. Lifshitz, *Statistical Physics*, translated by D. Shoenberg (Oxford U. P., Oxford, England, 1938).

[11] N. D. Mermin and H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966).

[12] R. B. Griffiths and J. Lebowitz, J. Math. Phys. **9**, 1284 (1968).

[13] T. Holstein and H. Primakoff, Phys. Rev. **58**, 1098 (1940).

[14] D. N. Zubarev, Usp. Fiz. Nauk **71**, 71 (1960) [Sov. Phys.-Usp. **3**, 320 (1960)].

[15] T. Oguchi, Phys. Rev. **117**, 117 (1960).

[16] B. I. Halperin and P. C. Hohenberg, Phys. Rev. Lett. **19**, 700 (1967); J. L. Beeby and J. Hubbard, Phys. Lett. A **26**, 376 (1968).

[17] It is likely that two infinite clusters do not coexist. See R. Kikuchi, J. Chem. Phys. **53**, 2173 (1970).

[18] R. V. Lange, Phys. Rev. **146**, 301 (1966).

[19] R. A. Tahir-Kheli, Phys. Rev. B **6**, 2808 (1972).

[20] D. Pepper, Ph.D. thesis (Oxford University, 1972) (unpublished).

[21] P. Soven, Phys. Rev. **156**, 809 (1967); D. W. Taylor, Phys. Rev. **156**, 1017 (1967).

[22] B. J. Last and D. Thouless, Phys. Rev. Lett. **25**, 1719 (1971); S. Kirkpatrick, Phys. Rev. Lett. **25**, 1722 (1971).

[23] H. Sato, A. Arrott, and R. Kikuchi, J. Phys. Chem. Solids **10**, 19 (1958).

[24] T. Oguchi and T. Obakata, J. Phys. Soc. Jap. **27**, 1111 (1969).

[25] T. Kaneyoshi, Prog. Theor. Phys. **42**, 477 (1969).

[26] S. F. Edwards and R. C. Jones, J. Phys. C **4**, 2109 (1971).

[27] R. B. Griffiths, Phys. Rev. Lett. **23**, 17 (1969).

[28] J. K. L. MacDonald, Phys. Rev. **43**, 830 (1933).

[29] J. S. Langer, J. Math. Phys. **2**, 584 (1961).

[30] S. F. Edwards and J. Beeby, Proc. R. Soc. A **274**, 395 (1962).

[31] These values were obtained by N. Nafari, Ph.D. thesis (University of Pennsylvania, 1972) (unpublished).

[32] The Q's and the R's can be expressed in terms of I and Y, where I was defined in Eq. (3.16) and its numerical value was given in Ref. 2 and Y is defined by $Y = 6N^{-1} \sum_k \hat{e}_k^{-1} \sin^2 k_x (1 - \cos k_x)$. Gehring (Ref. 33) gives $Y = 0.48$. We have $Q_{xx} = Y/6$, $Q_{xy} = (3 - 2Y)/24$, $R_{xx} = (4 - 3I)/12$, and $R_{xy} = (2 - 3I)/24$.

[33] G. M. Gehring (unpublished) as quoted by Last in Ref. 8.

[34] Note that the wave functions given in Eq. (3.13) were not determined variationally; they are simply the result of perturbation theory.

[35] A. W. Saenz, Phys. Rev. **125**, 1940 (1962); A. B. Harris and H. Meyer, Phys. Rev. **127**, 101 (1962).

[36] R. J. Elliott (unpublished).

[37] D. Thouless (private communication).

[38] In Eq. (2.6) of Ref. 2, a_{fk} and a^*_{fk} should be interchanged. In Eq. (2.11) of Ref. 2, H_{kf} and P_{kf} should be replaced by H_{fk} and P_{fk}.