International Journal of Theoretical Physics, Vol. 26, No. 8, 1987

Critical Behavior of Noninteracting Spin Waves

Lincoln Stoller[1]

Received March 6, 1987

Monte Carlo techniques are used to show that a free-spin-wave theory of the quantum Heisenberg spin-1/2 ferromagnet, formulated to accurately represent the state space of the underlying system, predicts the critical temperature of the interacting theory to within 1% and critical exponents to within a factor of two.

1. INTRODUCTION

In an earlier paper (Stoller, 1987a) (referred to as I), I derived a spin-wave theory that embodies the correct free-spin-wave statistics at all temperatures. It was asserted that this corrected form of free-spin-wave theory would provide a means of extending perturbation theory beyond the low-temperature region. In order to substantiate this claim, it is necessary to show that free-spin-wave theory, without any interactions, produces reasonably accurate predictions of long-wavelength phenomena at all temperatures. In a second paper (Stoller, 1987b) (referred to as II) I showed that this theory reproduces the Bloch $T^{3/2}$ law for the magnetization and the boson distribution for the density of states in k space at low temperatures. This paper presents the results of applying free-spin-wave theory at critical temperatures.

Previous work (Dyson, 1956a,b; Silberglitt and Harris, 1968; Wortis, 1965; Vaks *et al.*, 1968) using spin waves has used the boson distribution function as a free-particle propagator. While theoretical arguments and experimental data validate this approximation at low temperatures, the divergent behavior of the boson distribution precludes its use at higher temperatures. This means that even though interactions between long-wavelength spin waves remain weak at all temperatures, the boson distribution places an inherent limitation on the use of perturbation theory. This

[1]Department of Chemistry, University of Texas at Austin, Texas 78712.

0020-7748/87/0800-0805$05.00/0

is independent from the problem of strong spin-wave coupling at short wavelengths. Because of the boson distribution, we cannot use perturbation theory to approach problems dominated by long-wavelength modes if these problems occur away from $T \approx 0$. With an accurate distribution function it should be possible to extend thermal perturbation theory to higher temperatures. Perturbation theory will still be limited to the consideration of long-wavelength phenomena.

In Sections 2.1 and 2.2, respectively, I develop the analytic quantities used to describe critical phenomena, and then the form of the Monte Carlo simulation in Fock space. In Section 3 the problem is reformulated as a fully bosonic system with a nonlocal potential so as to be more amenable to the Monte Carlo technique. The results of the simulations, the values of T_c, and the critical exponents are given in Section 4, conclusions are in Section 5, and details of the implementation and optimization of the Monte Carlo algorithm are in the Appendix.

2. CRITICAL BEHAVIOR

2.1. The Calculation of Critical Quantities

I compute the bulk magnetic order and the bulk susceptibility of a three-dimensional Heisenberg ferromagnet in the free-spin-wave approximation given in I. The sum over a restricted set of states developed in I was amended in II to embody the reflection symmetry about the z axis of the state space. It was shown that for the computation of the magnetization and susceptibility the thermal average taken over m-spin-wave states for $m = 0, \ldots, N^d$ must be identical to that taken over $m = 0, \ldots, N^d/2$. If $\hat{\Theta}$ is a magnetization or susceptibility operator, the resulting expectation values for $d = 3$ are of the form

$$\langle \hat{\Theta} \rangle_\beta = \frac{\sum_{m=0}^{N^3/2} \sum_{\{\Phi(m)\}} \langle \hat{\Theta} \rangle \, e^{-\beta E(\Phi(m))}}{\sum_{m=0}^{N^3/2} \sum_{\{\Phi(m)\}} e^{-\beta E(\Phi(m))}}$$

The bulk magnetic order is given by the thermal expectation value of the absolute magnetization $|\mathbf{M}| \equiv (\langle \hat{\mathbf{M}}^2 \rangle)^{1/2}$. One must compute $|\mathbf{M}|$ rather than the z component of the magnetization, because $\langle \hat{\mathbf{M}}^z \rangle$ is unstable in finite systems in the absence of an external magnetic field. This instability causes $\langle \hat{\mathbf{M}}^z \rangle$ to fluctuate about zero and equilibrate to zero even in the spontaneously ordered regime. The expectation value of $\mathbf{M}^{x,y,z}$ is μ_B times

$\mathfrak{S}^{x,y,z}$, where $\mathfrak{S}^{x,y,z}$ is the x, y, or z component of the total spin operator $\mathfrak{S}$;

$$\mathfrak{S}^{x,y,z} = \sum_j \hat{S}_j^{x,y,z}$$

where $\hat{S}^{x,y,z} = (\hbar/2)\sigma^{x,y,z}$, and $\sigma^{x,y,z}$ are the Pauli matrices.

For any m-spin-wave state, $|\varphi\{n\}\rangle$, where $\{n\} = \{n_{k_1}, n_{k_2}, \ldots\}$ gives the occupation numbers at different wavenumbers and $\sum_j n(k_j) = m$, the expectation value of the total spin squared, is

$$\langle \varphi\{n\}|\mathfrak{S}^2|\varphi\{n\}\rangle = \mathfrak{S}^2_{\max} - \Delta(N^d + 1 - \Delta) \equiv \langle \mathfrak{S}^2\{n\}\rangle \tag{1}$$

Δ is the number of spin waves excited at nonzero wave vector, $\Delta = \sum_{\mathbf{k}\neq 0} n_{\mathbf{k}}$, and $\mathfrak{S}^2_{\max} = (N^d/2)[(N^d/2)+1]$ is the value of $\mathfrak{S}^2$ evaluated in the fully ordered ground state, $|0\rangle$, of the N^d-site system. $|\mathbf{M}|_\beta$ is given by

$$|\mathbf{M}|_\beta = Z^{-1} \sum_\varphi \langle \varphi | \left\{ \mathfrak{S}^2_{\max} - \left(\sum_{k\neq 0} n_{\mathbf{k}} \right) \left[N + 1 - \left(\sum_{k\neq 0} n_{\mathbf{k}} \right) \right] \right\}^{1/2} e^{-\beta \hat{H}} |\varphi\rangle \mu_B$$

Using the fluctuation-dissipation theorem, we define the susceptibility χ as

$$\chi = (N/k_B T)[\langle \hat{\mathbf{M}}^2 \rangle_\beta - |\mathbf{M}|^2_\beta] \tag{2}$$

Thermal averages are computed using two different models of the state space. The first, called the fixed-cutoff model, is a simplified version of the correct state space. In this model the only restriction on the system is that it contain no more than $N^d/2$ spin-wave excitations. The model does not introduce a density-dependent k-space cutoff as in the next model. The fixed-cutoff space is overcomplete, containing more high-energy degrees of freedom than in an exact formulation (over completeness is discussed in I).

The fixed-cutoff model is presented for two reasons. First, by comparing the critical properties of this and the second model with the known values for the interacting system we can see how improvements in the model affect the results. Second, exact results for the magnetization and susceptibility of the three-cube can be obtained using the fixed-cutoff model. This supplies a check of the Monte Carlo simulation.

The second model is called the free-spin-wave model, or spin-wave model for short. It is a realization of the correct state space, as given in I and II, and is constructed to allow analysis using the Monte Carlo method. This model achieves a complete set of spin-wave states by forbidding spin waves from occupying a density-dependent number of short-wavelength modes. Specifically, the state space does not include states with spin-wave excitations at wavenumbers $|k^{x,y,z}| > Q(m)/2$, where

$$Q(m)/2 = (N^d - m)^{1/3}/2 \tag{3}$$

Equation (3) is strictly valid only in the thermodynamic limit. In applying

it to a finite system, one runs into the problem that wavenumbers take integer values, whereas the cutoff given by (3) is essentially noninteger. An interpolation scheme is used to obtain integer values of $Q(m)$. This generates errors in the size of the state space that vanish as $1/N$. This is discussed further in the Appendix. In both models the symmetry of the state space about the x-y plane makes it necessary to consider only the subset of m spin waves for m less than $N^d/2$, as discussed in II.

Both models impose global conditions on the state space. This is apparent in the range of $n_{\mathbf{k}}$, the number of spin waves at $\mathbf{k}$, which is a function of $n_{\mathbf{k}}$ at all values of k. As a consequence, the total state space in the second quantized picture cannot be constructed as a product of factors localized in k space. This is in contrast with a noninteracting Bose space, which contains all allowable excitation numbers $n_{\mathbf{k}}$ without regard to the occupation numbers at other wavelengths.

The global properties expressed in free-spin-wave theory imply a nonseparable wave function—a consequence of wavefunction symmetry particular to $SU(2)$. The zeroth-order particle approximation embodies an interaction that is kinematic in origin. This is analogous to the statistical repulsion that fermions experience as a consequence of the exclusion principle. In this sense, then, the spin-wave distribution given in equation (3) is not a *completely* independent particle theory, though its dynamics is governed by a noninteracting Hamiltonian. The name "free spin-wave theory" is retained with this *caveat* implicitly understood.

2.2. The Monte Carlo Method

Monte Carlo simulations of quantum systems usually involve the Trotter product formula (Suzuki *et al.*, 1977). This approach transforms a d-dimensional quantum system into a $(d+1)$-dimensional classical problem, which can be analyzed using standard Monte Carlo methods. The Trotter method is not used here. I have instead applied the Monte Carlo algorithm to states defined on a d-dimensional Fock space. Because of the novelty of this approach, it will be described in detail.

The object of the Monte Carlo method (Wood, 1968; Valleau and Torrie, 1977a; Valleau and Whittington, 1977b) is to construct a relatively small set of states such that the sum of contributions to the expectation value from these states approximately equals the thermal average. This set, which I call the Monte Carlo set or sequence, is constructed by an iterative process whereby one starts with a given state and from this generates another state at random. The resultant state is tested according to criteria that reflect the population of states given by the thermal average. It is either accepted or rejected as the next member of the set. If it is accepted, it becomes the

seed for subsequent modifications; if it is rejected, its parent state is duplicated as the next member of the set and is used again as the seed of the next modification. The way in which a subsequent state in the set is generated from a given state is determined by the general updating algorithm I now describe.

The first step in the general updating algorithm is to construct a new, untested state, given a state that is already a member of the Monte Carlo sequence. This is called updating the underlying Markov chain. If we define two labels for states—a superscript denoting the position of the state in the Monte Carlo sequence, and a subscript denoting the set of occupation numbers that fully characterize the state—then updating the Markov chain is described as follows. Given the αth state in the sequence $\Phi^\alpha = \Phi_i$, a new state $\Phi_j \neq \Phi_i$ is constructed from it. The probability function for generating state j from state i is defined as $T^*_{i\to j}$. I use the transition probability given by the Metropolis prescription where $T^*_{i\to j}$ is a constant for all states $j \neq i$ that lie within some "radius" about i. In Fock space this radius is the maximum number of particles that can be added or subtracted from the occupation number of the previous state. For states j outside this radius, or for $j = i$, the transition probability is zero. That is, for states Φ_j that differ from Φ_i by more than some predetermined number of particles, or for $\Phi_j = \Phi_i$, $T^*_{i\to j} = 0$.

The next step in the updating algorithm requires evaluating the acceptance probability $P_{i\to j}$ for the transition from state i to state j. The acceptance probability is given in terms of the weighting function W_j:

$$P_{i\to j} \equiv \min\left[1, \frac{T^*_{j\to i} W_j}{T^*_{i\to j} W_i}\right] \tag{4}$$

That is, $P_{i\to j}$ is given by whichever of the two quantities in the parentheses is less. In the Metropolis algorithm $T^*_{i\to j} = T^*_{j\to i}$, so the quotient of these factors is always one. It is standard to take W_j equal to the Boltzmann weighting $\exp[-\beta E(\Phi_j)]$.

A random number $\mathbf{R}$ is then chosen between 0 and 1 and is compared with $P_{i\to j}$. If $\mathbf{R}$ is less than or equal to $P_{i\to j}$, then Φ_j is accepted as the next state in the Monte Carlo sequence, that is, $\Phi^{\alpha+1} = \Phi_j$. If $\mathbf{R}$ is greater than $P_{i\to j}$, then the next state in the sequence is the same as the previous state, $\Phi^{\alpha+1} = \Phi_i$.

Notice that in equation (4) if the weighting function $W_j = 0$, then $P_{i\to j} = 0$. This means the transition probability $T_{i\to j}$ to state Φ_j is always zero from any state Φ_i for which $W_i \neq 0$. One can never reach states Φ_j even if they can be generated in the underlying Markov chain. Therefore, to exclude a portion of the parameter space, it is sufficient to associate a zero weighting for states that lie in that region.

3. NONLOCALITY OF THE ENERGY

According to the above prescription, a new state $\Phi'(n_{\mathbf{k}_1}, n_{\mathbf{k}_2}, \ldots, n'_{\mathbf{k}_j}, \ldots)$ is generated from a given state $\Phi(n_{\mathbf{k}_1}, n_{\mathbf{k}_2}, \ldots, n_{\mathbf{k}_j}, \ldots)$ by changing the occupation number at wavenumber $\mathbf{k}_j$. The change in the occupation number is given by

$$n'_{\mathbf{k}_j} = n_{\mathbf{k}_j} + \mathrm{Int}(D \cdot (\mathbf{R} + [\mathrm{sign}\,\mathbf{R}])) \tag{5}$$

The function $\mathrm{Int}(x)$ takes the integer component of a real number x. Here $\mathbf{R}$ is a random number between -1 and $+1$, and $[\mathrm{sign}\,\mathbf{R}]$ is the sign of $\mathbf{R}$. The quantity $(\mathbf{R}+[\mathrm{sign}\,\mathbf{R}])$ is uniformly distributed between -2, -1 and $+1$, $+2$. Here D is a free integer parameter that sets the range of $|n_{\mathbf{k}} - n'_{\mathbf{k}}|$, and plays the role of the radius mentioned earlier. Equation (5) tell how the new state with $n'_{\mathbf{k}_j} \neq n_{\mathbf{k}_j}$ is generated in the underlying Markov chain.

The acceptance probability is computed to determine whether or not the new state is accepted,

$$P_{i\to j} = \min[1, e^{-\beta(E(\Phi_j)-E(\Phi_i))}] \tag{6}$$

The implementation of the Monte Carlo scheme used here requires $E(\Phi_j)$ to be different from the energy given by free-spin-wave theory. Beginning with the simplest form for the energy, I will show that the restricted sum, which includes only valid spin-wave states, can be replaced with an unrestricted sum if modifications are made to the form of $E(\Phi_j)$. The idea is to remove the constraints on the underlying Markov chain and reexpress them as an altered form of the weighting function. Once the underlying Markov chain is updated without constraints, this nonstandard system can be simulated using standard Monte Carlo methods.

Begin by considering the case where $E(\Phi_i)$ is defined to be the sum of the noninteracting spin-wave energies $\varepsilon_{\mathbf{k}}$:

$$E(\Phi_i) = \sum_{k} n_{\mathbf{k}} \varepsilon_{\mathbf{k}} \tag{7}$$

Now consider the question of how to handle the boundaries of the state space. Suppose, in the fixed-cutoff model, one is given an αth state in the Monte Carlo sequence such that Φ^α contains exactly $N^d/2$ excitations. The updating scheme of equation (5) cannot be used if one wants to realize the condition that states containing more than $N^d/2$ excitations are excluded from the Monte Carlo average. Equation (5) pays no attention to the total number of spin waves in the system and, if repeated enough times, will generate any value of $n_{\mathbf{k}}$, positive or negative. This can be corrected without altering equation (5). Instead, one can allow unphysical states to be constructed in the underlying chain, and then give them an infinite energy. Infinite-energy states will never be accepted and so will never be included in the

average. Accordingly, an additional term is amended to the free-spin-wave energy to enforce the maximum particle number constraint. In the fixed-cutoff model this additional energy $V_{\mathrm{nonlocal}}(\Phi_j)$ is a function of the total excitation number m_{tot}. In the spin-wave model, $V_{\mathrm{nonlocal}}(\Phi_j)$ is a function of the number of excitations in the new state Φ_j at wavenumbers $\mathbf{k}$ any of whose components $|k^a| > Q(m_{\mathrm{tot}})/2$, for $a = x, y, z$. This number of states will be given by $m_{\mathrm{tot}}(Q(m))$. Previously no states were allowed with nonzero occupation numbers at $|k^a| > Q(m_{\mathrm{tot}})/2$. They are now allowed in the underlying chain, but are given energies that make them thermodynamically inaccessible. For both models we have

$$E(\Phi_i) = \sum_{\mathbf{k}} n(\mathbf{k})\varepsilon(\mathbf{k}) + V_{\mathrm{nonlocal}}(x) \tag{8}$$

where

$$V_{\mathrm{nonlocal}}(x) = \begin{cases} 0 & \text{if } x = 0 \\ \infty & \text{if } x > 0 \end{cases}$$

where x equals either $m_{\mathrm{tot}} - N/2$ or $m_{\mathrm{tot}}(Q(m))$, depending on the model.

$V_{\mathrm{nonlocal}}(\Phi)$ is a nonlocal potential energy that can cause the energy to change an infinite amount in response to the smallest variation in any of the parameters that characterize a state. Despite their initial appearance as being noninteracting, both of these models embody strong spin-wave correlations that can be traced back to the statistical ansatz resolving the overcompleteness problem.

Remove next the constraint that all states in the underlying Markov chain must have nonnegative occupation numbers. Since such states have no physical meaning, they are excluded by being accorded an additional energy $\sum_{\mathbf{k}} V_{\mathrm{stable}}(n_{\mathbf{k}})$ that is zero for all $n_{\mathbf{k}} \geq 0$, and infinite for negative values. This fictitious potential does not represent any physical interaction between spin waves and is a necessity for the simulation of any system in Fock space. The final form of the energy is

$$E(\Phi_i) = \sum_{\mathbf{k}} n_{\mathbf{k}}\varepsilon_{\mathbf{k}} + V_{\mathrm{nonlocal}}(x) + \sum_{\mathbf{k}} V_{\mathrm{stable}}(n_{\mathbf{k}}) \tag{9}$$

where

$$V_{\mathrm{stable}}(n_{\mathbf{k}}) = \begin{cases} 0 & \text{if } n_{\mathbf{k}} \geq 0 \\ \infty & \text{if } n_{\mathbf{k}} < 0 \end{cases}$$

According to (6) and (9), any newly constructed state that does conform to the restrictions of the complete spin-wave space is given an infinite energy and is rejected as the next state in the Monte Carlo sequence. Any state

with a finite energy lies within the restricted state space.[2] If such a state has an energy greater than $E(\Phi_i)$, it is accepted or rejected depending on the comparison of $P_{i\to j}$ with a random number. If its energy is less than $E(\Phi_i)$, it is automatically accepted and becomes the next state in the Monte Carlo sequence.

A full-system update is obtained after each occupation number in the original state has been updated. After γ full system updates have been performed, $\gamma \geq 1$ is some integral, the state is used to compute the next contribution to the Monte Carlo sum.

The Fock space feature of this Monte Carlo approach allows me to treat the quantum system much like any classical system—in terms of a c-number "configuration." The analog of the classical configuration is the quantum state described in terms of (real) occupation numbers. In the case of classical systems, discretization of the space introduces computational errors. In the quantum case, the discretization is an exact expression of the finite lattice structure. The Fock space description is practicable because the Hamiltonian is approximately diagonal in this space.

4. CRITICAL TEMPERATURE, AND THE MAGNETIZATION AND SUSCEPTIBILITY EXPONENTS

4.1. Finite-Size Data

The magnetization and susceptibility of systems of 3-, 5-, 7-, 9-, and 13- or 15-cubed sites have been computed in the fixed-cutoff and spin-wave models. The results are given in Figures 1 and 2. The lengths of the Monte Carlo sequences were chosen so as to achieve 1–2% expected accuracy in the susceptibility at its peak. Each sequence typically contained between 10,000 and 50,000 full-system updates at each temperature.

The value of T where the susceptibility reaches its peak is defined as the pseudo-critical temperature $T_c(N)$ for a system of N^3 sites. The increasing height of the maxima indicates a phase transition in the thermodynamic limit.

The results do not exhibit critical slowing down (Binder, 1979) as occurs in position-space formulations of Monte Carlo. To understand this, recall from scaling theory that critical phenomena are strongly influenced by long-wavelength behavior. A single position-space update generates only a localized change in configuration. Large-scale collective behavior, in which distant parts of the system interact, requires that these coupling effects propagate across distance of order N. In order to get position-space Monte Carlo to reflect long-range effects, the length of the Monte Carlo chains

[2]The actual implementation of the cutoff is more complicated than appears in equation (9) and is discussed further in the Appendix.

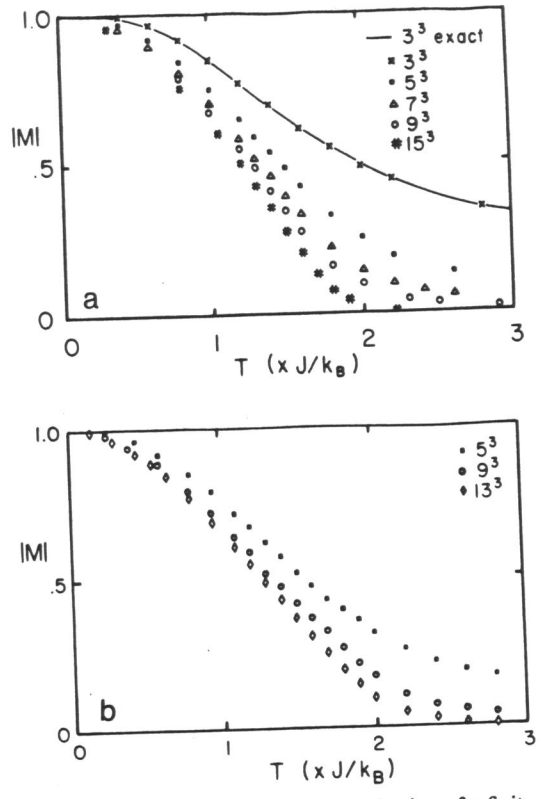

Fig. 1. Monte Carlo results for the lowest order magnetization of a finite, three-dimensional Heisenberg ferromagnet: (a) Fixedcutoff model, (b) spin-wave model.

must grow in proportion to the size of the system. This growth in the number of updates to obtain equilibration over large scales generates the critical slowing-down phenomena.

In wavenumber space, on the other hand, each state is a superposition of nonlocalized excitations. Spin-wave states already embody correlations that extend over the size of the system, so it is not necessary to wait for local effects to diffuse throughout the lattice. As a result, the simulation does not suffer this sort of equilibration problem in the critical region.

4.2. Extrapolation to Infinite Size

According to finite-size scaling theory (Barber, 1983; Landau, 1976), the size dependence of the pseudo-critical temperature as a function of N is given by the shift exponent λ as

$$T_c(N) \approx (bN^{-\lambda} + 1)T_c$$

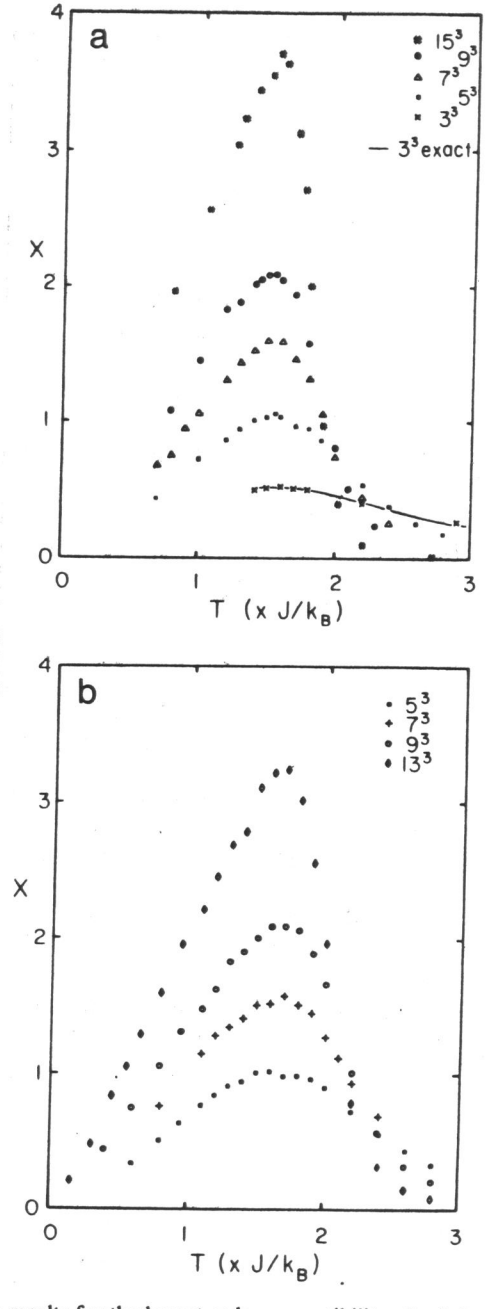

Fig. 2. Monte Carlo results for the lowest order susceptibility of a finite, three-dimensional Heisenberg ferromagnetic: (a) fixed cutoff model, (b) spin-wave model.

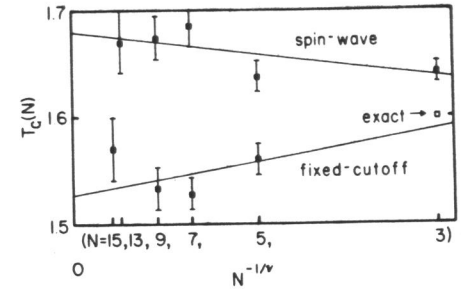

Fig. 3. Pseudo-critical temperatures extrapolated to the infinite-size limit.

where $\lambda = 1/\nu$, and $\nu = 0.735 \pm 0.015$ from high-temperature series expansions. Therefore, the extrapolation to infinite size is done by fitting $T_c(N)$ to a linear function of $N^{-1.36}$. The results are shown in Figure 3, where an estimation of standard deviation is given by the error bars. The extrapolated values for the fixed-cutoff and the spin-wave models are $T_c(\infty, \text{FIXED}) = 1.53 \pm 0.02$ and $T_c(\infty, \text{SPIN}) = 1.68 \pm 0.02$, in units of J/k_B.

Finite-size scaling theory gives the N dependence of $\mathbf{M}$ and χ as

$$\log(\mathbf{M}N^{\beta\lambda_1}) = \beta \log(tN^{\lambda_2}) + \text{const.}$$
$$\log(\chi TN^{-\gamma_3}) = -\gamma \log(tN^{\lambda_4}) + \text{const.} \tag{10}$$

with $t \equiv |(1-T)/T_c|$. The λ_i are determined by adjusting the magnetization and susceptibility data for different N so that they fall along the same magnetization and susceptibility log curves—their values do not affect the critical exponents. The data for $\mathbf{M}_{\text{SPIN}}(N, T)$ for T less than $T_C(\text{SPIN})$ are plotted in Figure 4. In Figure 5 $\chi_{\text{SPIN}}(N, T)$ is plotted for T above the estimated critical temperature.

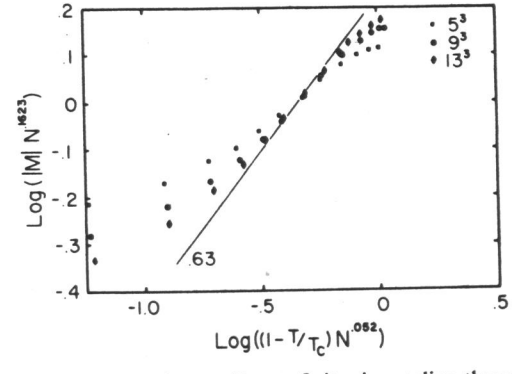

Fig. 4. Magnetization data plotted according to finite-size scaling theory, spin-wave model.

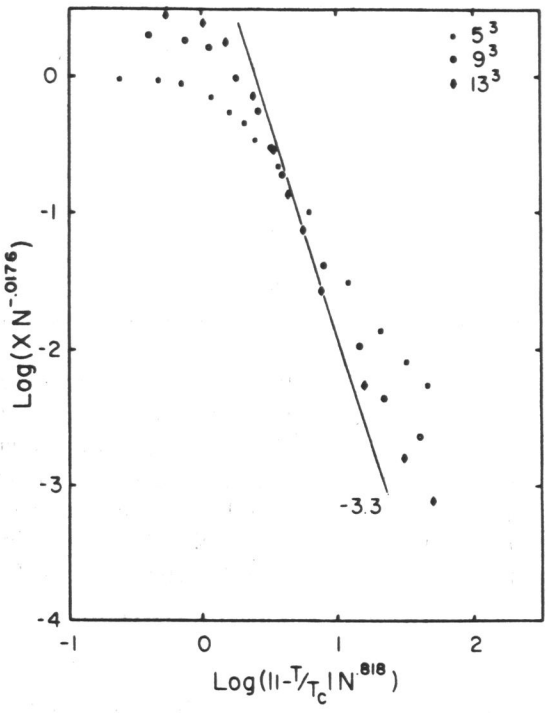

Fig. 5. Susceptibility data plotted according to finite-size scaling theory spin-wave model.

The critical exponents for the infinite system are found by extrapolation. The exponent as a function of N is the slope of the straight line that best fits the log **M** or log χ points in the region where the scaling relations are expected to hold. The log t ranges over which the straight-line fits are made and the quality of the straight-line fits over this region both increase with N.

The graphs for β and γ are shown in Figures 6a and 6b.

The critical behavior of the Heisenberg ferromagnet, according to the theory I have developed and as calculated by the Monte Carlo method, is summarized in Table I. For comparison, I have also listed the values obtained from classical theory ($S\to\infty$), mean-field theory, and high-temperature quantum mechanical series expansions (Rushbrooke *et al.*, 1974). The high-temperature expansions are considered to be the most accurate predictions.

5. CONCLUSIONS

The effects of spin statistics that cause the correct spin-wave distribution to differ from the Bose distribution are called kinematical interactions. The

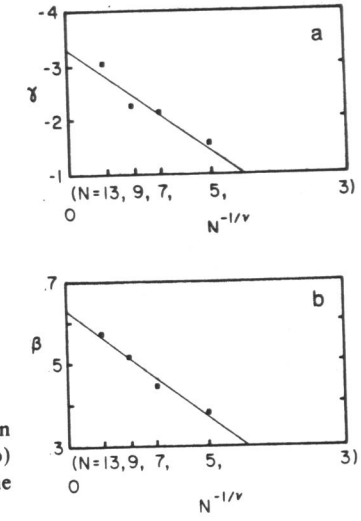

Fig. 6. (a) Extrapolation of the magnetization exponent to the infinite-size limit, spin-wave model. (b) Extrapolation of the susceptibility exponent to the infinite-size limit, spin-wave model.

spin-wave state space presented in I resolves spin-wave overcompleteness effects in the form of a new distribution function and thereby provides a partial solution[3] to the kinematical problem.

In this paper I have shown that a theory based on this improvement alone generates the correct value of the critical temperature and predicts the critical exponents for the magnetization and susceptibility to within a

Table I. Comparison of Different Predicted Values of Critical Temperature and Critical Indices

				Noninteracting models	
	Classical	Mean field	High-temperature series	Flat cutoff	Spin-wave cutoff
γ	1.405	1.00	1.43 ± 0.01	4.1 ± 0.2	3.3 ± 0.2
β	0.365 ± 0.035	0.50	0.385 ± 0.025	0.53 ± 0.02	0.63 ± 0.02
T_c	2.89 ± 0.02	3	1.68 ± 0.01	1.53 ± 0.02	1.68 ± 0.02

[3]Spin-wave nonorthogonality is one of two parts of the kinematical problem. Its effects are separate from that of establishing a complete state space and computing the effects of particle interaction. It is not included in the present spin distribution, but can be included within the formalism of thermal perturbation theory (Stoller, 1985).

factor of 2. The differences between the exponents calculated in the noninteracting model and those expected in the full theory are due either to the effects of spin-wave interaction,[4] or spin-wave nonorthogonality.

Although one must conclude that free-spin-wave theory by itself does not describe the critical phenomena of the Heisenberg ferromagnet as well as existing theories, these predictions may rapidly improve with the addition of the particle interactions of first-order perturbation theory.[5]

APPENDIX. IMPLEMENTATION OF THE MONTE CARLO ALGORITHM

Two features of the Monte Carlo work that have not yet been discussed include, first, a technique that exploits the symmetry of the system to reduce the state space and speed up program execution, and second, the implementation of the nonlocal cutoff. The use of symmetry properties to reduce the size of the state space can be applied generally and is discussed in detail.

In such cases where the weighting function does not depend on the signs of the wavenumbers of the component spin waves it is possible to reduce the state space to wavenumbers in the first octant of k space, i.e., to states composed of spin waves at wavenumbers $\mathbf{k}$ for which k^x, k^y, and k^z are each either positive or zero. As a result of reducing the number of parameters that must be sampled, the simulation converges more rapidly.

Consider the symmetry properties of the state space. The energy is linear in its components: $E(\Phi\{\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_n\}) = \sum_{\mathbf{k}} n_{\mathbf{k}}(\Phi)\varepsilon_{\mathbf{k}}$, where $\varepsilon_{\mathbf{k}} = \varepsilon(k^x) + \varepsilon(k^y) + \varepsilon(k^z)$, and $\varepsilon(k^i)$ is symmetric about $k^i = 0$. Given three positive wavenumbers, k^x, k^y and k^z, there are eight ways they can be assigned plus and minus signs. Each set of signs specifies a three-dimensional wavenumber in one of eight octants in parameter space. Because $\varepsilon(k^i)$ is symmetric, the points $(\pm k^x, \pm k^y, \pm k^z)$ in different octants are degenerate. The magnetization and susceptibility are also symmetric about $k_i = 0$ and so are also degenerate in these octants. This symmetry is exploited by restricting the Monte Carlo simulation to the approximately $(N/2)^3$ wavenumbers in the first octant where $k^x \geq 0$, $k^y \geq 0$, $k^z \geq 0$. This reduction must be accompanied by changes in the updating algorithm to account for this degeneracy in order that the simulation retain information about full system.

[4]Within the context of renormalization group we have the assertion that all types of interactions within the same universality class generate the same critical exponents. The current statement is not in contradiction with this, because free and interacting spin-wave theories are in different universality classes.

[5]Because the propagator is the thermal expectation value of the number operator, we are assured that it remains finite throughout the critical region and can thus be used for perturbation theory.

The difference between working in the full, eight-octant space and the one-octant space is the number of ways signs can be assigned wavenumber components. Consider a single excitation state. There is only one state at $k^x = k^y = k^z = 0$; it is the state that lies at the corners of the eight octants. For $k^x = k^y = 0$, $k^z \neq 0$, there are two states, depending on the sign of k^z; these states lie along the edges of the octants. For $k^x = 0$, $k^y \neq 0$, $k^z \neq 0$ there are four ways to choose the signs, and if none of the k_i equals zero, then there are eight states. The number of components of $\mathbf{k}$ that are zero will be called the *class* of $\mathbf{k}$. The class of any point in k space is therefore either 0, 1, 2, or 3.

The situation for a two-excitation state is more complicated. In this case each of the excitations can be assigned in either of the four possible classes. If both of the excitations are of class zero, then there are 28 ways the two excitations can be arranged in the eight octants. The number of ways that excitations of fixed wavenumbers can be distributed in the eight octants is called the *wavenumber degeneracy* of the state, or simply the degeneracy. Given a general, m-excitation state with $n(\mathbf{k})$ excitations at each $\mathbf{k}$, the degeneracy of the whole is given by $\mathcal{D}(\tilde{n})$, where

$$\mathcal{D}(\tilde{n}) = \prod_{k} D(n(\mathbf{k})) \tag{A.1}$$

$D(n(\mathbf{k}))$ is the degeneracy associated with each occupation number:

$$D(n(\mathbf{k})) = \begin{cases} 1 & \text{if } \mathbf{k} \text{ is class 0,} \\ n(\mathbf{k}) & \text{if } \mathbf{k} \text{ is class 1} \\ \binom{n(\mathbf{k})+3}{3} & \text{if } \mathbf{k} \text{ is class 2} \\ \binom{n(\mathbf{k})+7}{7} & \text{if } \mathbf{k} \text{ is class 3} \end{cases} \tag{A.2}$$

Large systems mostly contain points of class 0, which means they have a vanishing number of degrees of freedom at zero wavenumber. In the thermodynamic limit the distinction between the multiplicity of various classes is unnecessary. However, because I will extrapolate to this limit using finite-size scaling, I have chosen to preserve the correct finite-size behavior.

By limiting the trace over spin-wave states to the first octant we obtain a new form for the thermal expectation value $\{\hat{A}\rangle_\beta$:

$$\langle \hat{A} \rangle_B = \frac{\sum_{(\Phi)}' \langle A(\tilde{n}) \rangle \mathcal{D}(\tilde{n}(\Phi))\, e^{-\beta E(\Phi)}}{\sum_{(\Phi)}' \mathcal{D}(n(\Phi))\, e^{-\beta E(\Phi)}} \tag{A.3}$$

where the primes limit the summations to states in the first octant. In transcribing equation (A.3) into the Monte Carlo scheme, the multiplicity $\mathcal{D}(\tilde{n})$ is taken as part of the weighting function $W(\tilde{n}(\Phi))$. From equation (4) the new expression for the acceptance probability is

$$P_{i\to j} = \min\left[1, \frac{\mathcal{D}(\tilde{n}(\Phi_j))\exp[-\beta E(\tilde{n}_j)]}{\mathcal{D}(\tilde{n}(\Phi_i))\exp[-\beta E(\tilde{n}_i)]}\right] \tag{A.4}$$

$\mathcal{D}(\tilde{n}(\Phi_j))$ requires lengthy computation. However, $\mathcal{D}(\tilde{n}(\Phi_j))$ itself is not needed, only the ratio $\mathcal{D}(\tilde{n}(\Phi_j))/\mathcal{D}(\tilde{n}(\Phi_j))$ is used. Since the occupation numbers at each $\mathbf{k}$ are updated independently, any given state Φ_i and the next state Φ_j in the underlying Markov chain differ only by the occupation numbers at a single wavenumber, say $\mathbf{k}_0$. According to (A.1), the degeneracy ratio is then given by

$$\frac{\mathcal{D}(\tilde{n}(\Phi_j))}{\mathcal{D}(\tilde{n}(\Phi_i))} = \frac{D(n(\mathbf{k}_0)_j)}{D(n(\mathbf{k}_0)_i)} \tag{A.5}$$

where $n(\mathbf{k}_0)_i$ and $n(\mathbf{k}_0)_j$ refer to the old and new occupation numbers at $\mathbf{k}_0$. The end result is that I can compute the Monte Carlo average for an N^3 system by computing the sum over a sequence of states generated in an $(N/2)^3$ system. Unfortunately, this saving of a factor of eight in the size of the system does not translate into an eightfold increase in speed—rather, the saving is about a factor of three.

The reason for this has to do with sampling efficiency. The problem is that by altering the weighting according to (A.4), the acceptance probability for going from one configuration to another is no longer determined by energy differences alone. There is now a higher probability of making the transition from a state of low energy to one of higher energy just because the higher energy state has many more ways it can be obtained. Similarly, where the original implementation of equation (4) might have moved into a region of representatively low energy in some number of steps, the present realization will take longer to move to lower energies because lower energy states effectively occupy a smaller region of phase space and, in a sense, are harder to find.

This means that part of the movement of the simulation through parameter space is now being determined by the geometry of the space without regard to energy. In the limit that sampling is determined by geometry alone the energy difference between states is not taken into account. In such a case state-space sums are obtained by a random choice of states. Such random sampling is horrendously inefficient. Historically, it was only through the inclusion of a "drift" toward more important regions of phase space, achieved through the introduction of the energy-dependent weighting, that the Monte Carlo method became useful.

Since the modification of equation (A.4) is a step toward geometrically determined sampling, its effect is to erode the efficiency of the original algorithm. As a result, some of the speed gained by going to a smaller system is lost in less efficient sampling.

The situation can also be described by saying that the convergence of the Monte Carlo simulation is strongly dependent on the form of the weighting function. A smooth, steep gradient in phase space speeds up convergence. The weight function of (A.4) has a more gradual gradient than that of equation (4) and the simulation converges more slowly as a result.

Next consider the implementation of the nonlocal expression of the energy given by (9). Nonlocality in the energy means that the energy associated with one location in k space is a function of the occupation numbers at all other regions of k space. There are two choices in how to handle this situation numerically. One either programs the computer to review the whole configuration every time the occupation number at one site is changed, which takes order N^d operations per site, or one writes code that records the relevant global variables each time the state is modified, which takes a constant number of operations. This latter choice keeps continuous track of global properties and is clearly the method of choice when the relevant global parameters are sufficiently few in number.

In the fixed-cutoff model it is a simple matter to keep track of the total occupation number m through the updating process. Given the occupation number at $\mathbf{k}$ in the present state, $n_i(\mathbf{k})$, a new occupation number $n_j(\mathbf{k})$ is generated in the underlying Markov chain according to the scheme of equation (5). The program tests to see if $n_j(\mathbf{k})$ is greater than zero and less than $(N^3/2)$—if not, the state is immediately rejected. If the state meets this condition, the quantity $\Delta_{ji} \equiv n_j(k) - n_i(k)$ is computed to give the new total occupation number $M_j = M_i + \Delta_{ji}$, which also must be greater than zero and less than $N^3/2$. If this condition is satisfied, the energy of the state is finite. The program proceeds to accept or reject the configuration according to the standard algorithm [with the weights given by (A.4)].

Implementation of the m-dependent cutoff is more complicated (m is the number of excitations). Here one has the additional constraint that for each excitation added to the system the number of accessible k-space points must decrease by one (cf. I). Phase-space points are removed from the edges of the first Brillouin zone in accordance with the ansatz of equation (3). In the simulation restricted to the first octant of k space, each point $\mathbf{k}$ satisfying $k^x, k^y, k^z \neq 0$ represents eight points in the original k-space. To implement the m-dependent reduction, I require that one point in the first octant of k space become inaccessible (or completely vacant) for every eight excitations added to the system. This is done by first defining a fixed

set of k-space points $\{\mathbf{k}\} = \mathbf{k}_1, \mathbf{k}_2, \ldots$ (the actual choice of this set will be described below). For each point in $\{\mathbf{k}\}$ there is a number $m_{\max}(\mathbf{k}_i)$ which indicates the maximum total number of spin waves beyond which the occupation number at $\mathbf{k}_i$ must go to zero. As one k-space point becomes inaccessible for every eight excitations, the values of $m_{\max}(\mathbf{k}_i)$ for $\mathbf{k}_i = \mathbf{k}_1$, $\mathbf{k}_2$, $\mathbf{k}_3, \ldots$ are 8, 16, 24, ... Since the total number of spin waves lies between 0 and $N^3/2$, the set $\{\mathbf{k}\}$ contains $N^3/16$ elements.

To realize the state-space conditions in the updating scheme, I include in the weight $W(\Phi)$ for a state Φ, given in equation (A.4), an exponential damping factor F,

$$W(\Phi) \to W(\Phi)F$$

where

$$F = \prod_i \exp\left\{-\left[\frac{m(\Phi)}{m_{\max}(\mathbf{k}_i)}\right]^r n(\mathbf{k}_i)\right\}$$

r is a positive integer, $m(\Phi)$ is the total number of excitations, and $n(\mathbf{k}_i)$ is the occupation number at wavenumber $\mathbf{k}_i$ in state Φ. When $m(\Phi)^r n(\mathbf{k}_i)$ is much less than $m_{\max}(\mathbf{k}_i)^r$ the damping factor $\exp\{-[m(\Phi)/m_{\max}(\mathbf{k}_i)]^r n(\mathbf{k}_i)\}$ is close to 1 and the factor has no effect. When $m(\Phi)^r n(\mathbf{k}_i)$ is greater than $m_{\max}(\mathbf{k}_i)^r$ the damping factor inhibits such a state from being accepted. Thus, the factor F works to eliminate from the Monte Carlo average states with $m(\Phi)/m_{\max}(\mathbf{k}_i) > 1$ and $n(\mathbf{k}_i) > 0$.

In the limit that $r \to \infty$, F provides a sharp, density-dependent cutoff that vanishes whenever $m(\Phi)/m_{\max}(\mathbf{k}_i) > 1$ and $n(\mathbf{k}_i) \geq 1$, and goes to one for $m(\Phi)/m_{\max}(\mathbf{k}_i) < 1$ and $n(\mathbf{k}_i)$ finite. In this limit F reproduces the cutoff scheme of equation (3).

The thermodynamic behavior must not depend on the cutoff scheme. By varying r between 1 and 16, I have tested the sensitivity of the results to different cutoff schemes and found no significant effects. This implies that at wavenumbers $\mathbf{k}_i \in \{\mathbf{k}\}$, for temperatures in the critical region, $n(\mathbf{k}_i)$ generally takes the value 0 or 1. One might expect that the best results would be obtained by setting r to some very large value, but this leads to ergodic problems.

An ergodic problem in a Monte Carlo simulation occurs when the simulation tends to get stuck in certain regions of the phase space. These problem regions of quasistability prevent the simulation from taking a path through phase space that is representative of the thermal average. The general result is that convergence is slow and erratic, and that the answers are often untrustworthy.

The reason that discretely changing potentials, such as those that one obtains by taking $r \to \infty$, give rise to ergodic problems is that the simulation

only senses a sharp boundary when the boundary is contacted. Without a relatively long-range potential gradient, the simulation can only move away from the boundary by random diffusion. When modeling a system with a multidimensional potential surface, such as in the present case, a sharp cutoff, as suggested by equation (3), makes it easy for the simulation to wander into a multidimensional corner. Without a potential gradient to lead the simulation out of these corners, the simulation must find its way out by simple random sampling. In this regard, a finite value of r provides a "soft damping" in the sense that it allows some modes violating the restricted state-space condition to be accepted, and forces other modes that do not violate the condition to be rejected. This smooths the sharp boundaries of the original potential.

Returning to the implementation of the state-space restrictions, recall that the updating scheme I use changes $n(\mathbf{k})$ for one value of $\mathbf{k}$ at a time. The ratio of the damping factor of a new state Φ' to that of an old state Φ is then given by the ratio

$$\frac{F'}{F} = \frac{\exp(\{[m(\Phi)+\Delta]/m_{\max}(\mathbf{k}_i)\}^r [n(\mathbf{k}_i)+\Delta])}{\exp\{[m(\Phi)/m_{\max}(\mathbf{k}_i)]^r n(\mathbf{k}_i)\}} \tag{A.6}$$

where $\Delta = n'(\mathbf{k}_i) - n(\mathbf{k}_i)$.

Next consider the choice of the wavenumbers in the set $\{\mathbf{k}\}$ for which the values of $m_{\max}(\mathbf{k}_i)$ are finite. In order to simplify the combinatorics, I choose $\{\mathbf{k}\}$ such that wave numbers $\mathbf{k} \in \{\mathbf{k}\}$ have all nonzero components. According to the results obtained in I, there is a large amount of freedom in how the state space may be reduced. I choose $\{\mathbf{k}\}$ to consist of the cube of k-space points at the outer corner of the first octant, as shown in Figure 7. This is an approximation to the cutoff scheme given by equation (3).

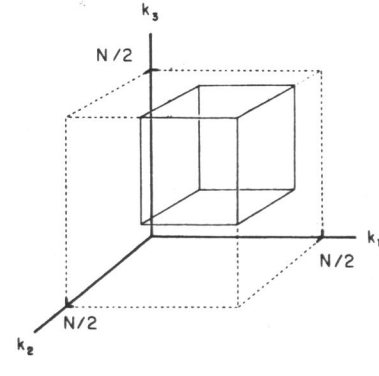

Fig. 7. The cube of points in the first octant of k space that becomes depleted of excitations as a function of increasing number of spin waves.

The prescription for assigning $m_{max}(\mathbf{k}_i)$ values is as follows. Beginning at the outermost point in the Brillouin zone, the point at $\mathbf{k} = (N/2, N/2, N/2)$, and moving inward each wavenumber is assigned one of the increasing values $m_{max} = 8, 16, 24, \ldots$. Assignments continue to be made until $m_{max}(\mathbf{k}_i) > N^3/2$ for some $\mathbf{k}_i$. The remaining wavenumbers for which no assignment of $m_{max}(\mathbf{k}_i)$ has been made are distinguished in the computer program by a flag or marker. When the computer updates a value of $\mathbf{k}_i$ for which there is no value of $m_{max}(\mathbf{k}_i)$, it acts as if $m_{max}(\mathbf{k}_i)$ were infinite by automatically setting $F'/F = 1$. At those wavenumbers for which $m_{max}(\mathbf{k}_i)$ is defined and finite, the weighting as given by equations (A.4) and (A.6) is computed to determine the transition probability.

The use of a cutoff scheme other than that given in equation (3) is justified because there is no unique way of reducing an overcomplete space. The advantages of equation (3) are its simple form, its easily visualized effect, and its apparent property of eliminating the correct number of states from each of the subspaces of fixed total wavenumber, as discussed in I. It is replaced by the scheme illustrated in Figure 7 in the interests of computational efficiency and on the assumption that errors would be insignificant in the calculation of bulk behavior. An indirect test of this assumption was done using the critical theory of Bose condensation. In II it was shown that at low temperatures and in the thermodynamic limit the present spin-wave distribution function is identical to a Bose distribution. Using a Bose distribution function with a chemical potential (i.e., for a system of fixed boson density), it is easy to impose k-space cutoffs and observe their effect on the critical behavior of the boson system. It is found that the cutoffs specified by equation (3) and Figure 7 result in the same critical temperatures and β exponents, 1.95 and 1.01, respectively. This provides only indirect support for the equivalence of the cutoff schemes of equation (3) and Figure 7 because the spin-wave distribution is not bosonic outside the range of low temperatures.

The last topic to be considered is the actual choice of parameters used to tune the Monte Carlo algorithm, as touched on briefly in Section 2.2. There are three parameters that are adjusted to optimize the computer simulation. The first is the maximum change in the occupation number at any k value, referred to as the step width D. By making D small, one ensures that, as long as one is not too close to the state-space boundary, the energy of an updated state is close to the energy of the original state. This means that even if there is an energy increase, it will be small and the acceptance probability will be high. Making the step length small allows the simulation to climb over potential barriers and test the detailed structure of the state space, but it also makes it difficult for the simulation to cover large state-space distances. By making D large, one ensures that newly

constructed states in the underlying Markov chain are only weakly correlated with the original state. However, in this case it is likely that higher energy states will be much higher in energy and will have a low acceptance probability. As a result, one may find that the simulation quickly moves into regions of low energy and then cannot move away. The value chosen for D is such that about half of the new states generated in the underlying Markov chain are accepted.

The second parameter at one's disposal is the number of updates between each state that is included in the sum. That is, by choosing some integer $\gamma > 1$, one can calculate the Monte Carlo sum as

$$S \approx \frac{1}{M} \sum_{i=1}^{M} F(\Phi^{\gamma i}) \tag{A.7}$$

where $\Phi^{\gamma i}$ is the (γi)th state in the Monte Carlo sequence.

The third parameter is ρ, the number of initial states in the Monte Carlo sequence that are excluded because they are correlated with the arbitrarily chosen initial state. It is clear that if one includes all states generated from an initial seed, and one only generates a small number of states for the whole simulation, then the answer one obtains will have more to do with the state one started with than the actual state-space average. The system is said to have equilibrated when the states in the sequence are no longer correlated with the initial state. The final form of the Monte Carlo sum is

$$S \approx \frac{1}{M} \sum_{i=1}^{M} F(\Phi^{\gamma i + \rho})$$

The results presented here generally used simulations in which every fourth state in the Monte Carlo sequence ($\gamma = 4$) was counted. Simulations were allowed to equilibrate for 4000–12,000 cycles before counting states ($\rho = 1000$–4000), and generally contained 10,000 configurations ($M =$ 10,000), except at temperatures within 15% of $T_c(N)$, where they might contain up to 70,000 configurations. This was done both because the error bars were larger in this region and because high accuracy was needed to identify the susceptibility peaks that define $T_c(N)$. The susceptibility near $T_c(N)$ of the largest systems of 13^3 and 15^3 sites was found to an estimated 2.5% accuracy, while that of systems with fewer than 9^3 sites was found to within 1%. Computation times ran between 5×10^{-6} and 3×10^{-5} sec per $\mathbf{k}$-site update in both the fixed-cutoff and spin-wave models.

REFERENCES

Barber, M. N. (1983). In *Phase Transitions and Critical Phenomena*, Vol. 8, C. Domb and J. L. Lebowitz, eds., Academic Press, New York, pp. 145–266.

Binder, K. (1979). In *Monte Carlo Methods in Statistical Physics*, K. Binder, ed., Springer-Verlag, Berlin, pp. 1–46.
Dyson, F. J. (1956a). *Physical Review*, **102**, 1217.
Dyson, F. J. (1956b). *Physical Review*, **102**, 1230.
Landau, D. P. (1976). *Physical Review B*, **14**, 255.
Rushbrooke, G. S., Baker, Jr., G. A., and Wood, P. J. (1974). In *Phase Transitions and Critical Phenomena*, Vol. 3, C. Domb and M. S. Green, eds., Academic Press, New York, pp. 245–356.
Silberglitt, R., and Harris, A. B. (1968). *Physical Review*, **174**, 640.
Stoller, L. (1985). Extending spin-wave theory to critical temperatures, Ph.D. Dissertation, University of Texas at Austin.
Stoller, L. (1987a). Kinematical problem in spin-wave theory, *International Journal of Theoretical Physics*, this issue.
Stoller, L. (1987b). Low-temperature results of the statistical theory of noninteracting spin waves, *International Journal of Theoretical Physics*, this issue.
Suzuki, M., *et al.* (1977). *Progress in Theoretical Physics*, **58**, 1377.
Vaks, V. G., *et al.* (1968). *Soviet Physics-JETP*, **26**, 188.
Valleau, J. P., and Torrie, G. M. (1977a). In *Modern Theoretical Chemistry*, Vol. 5, B. J. Berne, ed., Plenum Press, New York, Chapter 5.
Valleau, J. P., and Whittington, S. G. (1977b). In *Modern Theoretical Chemistry*, Vol. 5, B. J. Berne, ed., Plenum Press, New York, Chapter 4.
Wood, W. W. (1968). In *Physics of Simple Liquids*, H. Temperley *et al.*, eds., North-Holland, Amsterdam, pp. 117–230.
Wortis, M. (1965). *Physical Review B*, **8**, A1126.