Coupled-cluster method for quantum lattices: Application to square $S=\frac{1}{2}$ Heisenberg antiferromagnets

Frank E. Harris
Department of Chemistry, University of Utah, Salt Lake City, Utah 84112
(Received 25 September 1992)

A coupled-cluster formalism similar to that used in electronic-structure theory is developed for quantum-mechanical systems localized to sites on a lattice. It is shown how diagrammatic methods may be used to represent the cluster contributions in a convenient and systematic manner. The formalism is applied to the square $S=\frac{1}{2}$ Heisenberg antiferromagnet, yielding with minor computational effort results of quality comparable to those from computationally intensive quantum Monte Carlo calculations.

I. INTRODUCTION

The widespread success of the coupled-cluster method[1] in electronic structure calculations makes it of interest to examine the possibility of its application to quantum systems on lattices. This possibility has been identified by Roger and Hetherington[2] and initial calculations have been reported by them, by Bishop[3], and by Bishop, Parkinson, and Xian.[4] However, these calculations were carried out without the explicit development of a formalism that would eludicate the identification of cancelling contributions and assist in systematizing and simplifying the algebra. Such a formalism must recognize that lattice problems differ from that of mobile fermions due to the antisymmetry requirements in the latter case. This paper derives coupled-cluster equations for lattice problems in a diagrammatic formulation which is similar in spirit, but different in detail from that used in electronic structure theory.[5] We illustrate the use of the equations by applying them to the $S=1/2$ Heisenberg antiferromagnet.[6]

For simplicity we discuss the coupled-cluster formalism for a system of $S=1/2$ spins on a square lattice with nearest-neighbor interaction. Extension to other systems introduces more complexity but is conceptually straightforward. At each lattice point μ we assign a reference wave function ϕ_μ (called a hole state), and note that in the two-dimensional spin space at that lattice point there will be one other state $\phi_{\mu'}$ (a particle state) orthogonal to ϕ_μ. It is convenient, but unnecessary to assume that ϕ_μ for different sites μ are the same or simply related. From the ϕ_μ we build a reference many-spin wave function

$$\Phi=\prod_\mu \phi_\mu \ , \tag{1}$$

where the product is over all lattice sites.

The Hamiltonian H for our system is assumed to be of the Heisenberg form

$$H=\frac{1}{2}\sum_{\mu\nu} h(\mu,\nu)=\frac{1}{2}J\sum_{\mu\nu} n_{\mu\nu}\mathbf{s}_\mu\cdot\mathbf{s}_\nu \ , \tag{2}$$

where $\mathbf{s}_\mu$ and $\mathbf{s}_\nu$ are spin operators, $n_{\mu\nu}$ is unity if μ and ν are nearest neighbors and zero otherwise, and J is a constant we shall assume to be positive, so Eq. (2) describes an antiferromagnet. If we form a matrix element $\langle\Psi|H|\Psi'\rangle$, where Ψ and Ψ' are products of single spin states (but not necessarily Φ), the use of Eq. (2) reduces it to a sum of two-body matrix elements of types $\langle\mu\nu|h|\mu\nu\rangle, \langle\mu\nu|h|\mu\nu'\rangle, \langle\mu\nu|h|\mu'\nu\rangle, \ldots, \langle\mu'\nu'|h|\mu'\nu'\rangle$. We wish to represent the possibilities diagrammatically, and do so by letting the presence (or absence) of a line labeled μ refer to the presence (or absence) of a particle-hole pair at site μ, i.e., the state μ' (or μ). If the two-body matrix elements are thought of as inducing particle-hole transitions, they then correspond to diagram fragments such as those shown in Fig. 1. As in fermion theory, "bubbles" are used in matrix elements where a hole state appears in both left and right half-brackets. There are three main differences between the present representation and that most often used in fermion theory: (i) since a particle and a hole of the same index (i.e., on the same site) must be created or destroyed together, one line suffices to represent both; (ii) the lack of antisymmetry makes it inappropriate to drop index restrictions (in Fig. 1, μ and ν must be distinct); however, we retain the usual convention that unlabeled closed lines (those connecting two operator fragments) are summed over all labels, but observing the restrictions; and (iii) there are no "energy denominators" in the present formulation.

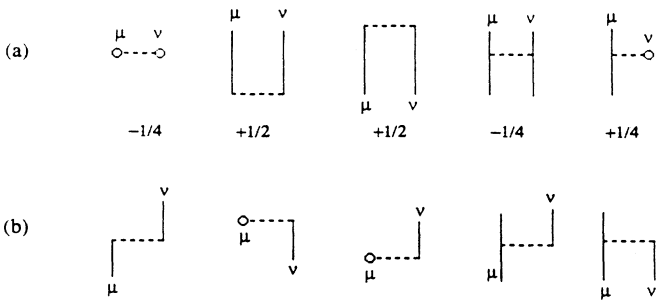

FIG. 1. Diagram fragments of H: (a) all those that conserve spin in the $S=1/2$ antiferromagnet with a Néel reference state, (b) all others. Below each diagram in (a) is its value in units of $\hbar^2 J$ for the H of Eq. (2), with μ and ν nearest neighbors, Néel reference state.

II. THEORY

The coupled-cluster wave function is obtained from the reference wave function by applying a cluster operator $\exp(T)$, with

$$T=\sum_n T_n \ , \tag{3}$$

where T_n is a cluster operator for n spins, and the sum over n is truncated at a point determined by practical considerations (in the example to be presented, our most complete calculations are cut off after $n=6$). The T_n are defined as follows:

$$\begin{aligned} T_1&=\sum_\mu t_\mu A_\mu^\dagger \ , \\ T_2&=\left[\frac{1}{2!}\right]\sum_{\mu\nu} t_{\mu\nu} A_\mu^\dagger A_\nu^\dagger \ , \\ T_3&=\left[\frac{1}{3!}\right]\sum_{\mu\nu\lambda} t_{\mu\nu\lambda} A_\mu^\dagger A_\nu^\dagger A_\lambda^\dagger \ , \end{aligned} \tag{4}$$

where $A_\mu^\dagger$ is a particle-hole creation operator for site μ. We note that all $A^\dagger$ for different sites commute, and $A_\mu^\dagger|\mu'\rangle=0$. The $t_\mu, t_{\mu\nu}, \ldots$ are coefficients that are determined as described below. The diagram fragments representing cluster operators are shown in Fig. 2.

In the coupled-cluster method, the energy and the t coefficients are determined by multiplying the Schrödinger equation, $H\exp(T)\Phi=E\exp(T)\Phi$, on the left by $\exp(-T)$ and requiring the satisfaction of its projections against $\Phi, \Phi_\mu=A_\mu^\dagger\Phi, \Phi_{\mu\nu}=A_\mu^\dagger A_\nu^\dagger\Phi, \ldots$. The projection against Φ yields a formula for the energy:

$$\langle\Phi|e^{-T}He^{T}|\Phi\rangle=E \ . \tag{5}$$

Each projection against a $\Phi_{\mu\nu\ldots}$ yields an equation for the t coefficients:

$$\begin{aligned} &\langle\Phi_\mu|e^{-T}He^{T}|\Phi\rangle=0 \quad (\text{each } \mu) \ , \\ &\langle\Phi_{\mu\nu}|e^{-T}He^{T}|\Phi\rangle=0 \quad (\text{each } \mu\nu) \ldots . \end{aligned} \tag{6}$$

In general, if the coefficients for clusters through T_n are to be determined, the equations needed for this purpose will involve projections on $\Phi_{\mu\nu\ldots}$ with up to n indices. Since E does not appear in Eqs. (6), a useful strategy will be to solve Eqs. (6) for the t coefficients and then to evaluate Eq. (5).

The quantity $\exp(-T)H\exp(T)\Phi$ may be written in terms of a series of nested commutators:

$$\begin{aligned} e^{-T}He^{T}\Phi=H\Phi+[H,T]\Phi+\left[\frac{1}{2!}\right][[H,T],T]\Phi \\ +\left[\frac{1}{3!}\right][[[H,T],T],T]\Phi+\cdots \ . \end{aligned} \tag{7}$$

Equation (7) is a starting point for the diagrammatic evaluation of Eqs. (5) and (6). To proceed, we note that successive operations on Φ by cluster operators and the Hamiltonian correspond to all diagrams that can be formed when the included operator fragments are placed in a vertical order (bottom to top) corresponding to the operator string (right to left), with no lines extending below the lowest operator. Moreover, when the lines are labeled, the index restrictions require that there be no label duplication at the same vertical position; in the language of fermion theory, there can be no "exclusion principle violating" (EPV) diagrams.

We discuss now the evaluation of the terms of Eq. (7), staring with $[H,T]\Phi=HT\Phi-TH\Phi$. We note (i) all diagrams of $TH\Phi$ are disconnected, while $HT\Phi$ is represented by both disconnected and connected diagrams; (ii) because there are no energy denominators, disconnected diagrams with the same structure and labeling will have the same value irrespective of the operator ordering; and (iii) some labeled diagrams of $TH\Phi$ would become EPV if the order of T and H were interchanged, but all the other diagrams of $TH\Phi$ are equal to corresponding disconnected diagrams of $HT\Phi$ and cancel all the disconnected $HT\Phi$ diagrams. We therefore conclude that $[H,T]\Phi$ is represented by all connected diagrams of $HT\Phi$ minus all disconnected diagrams of $HT\Phi$ which are EPV by virtue of index duplication between the disconnected parts. We need a diagrammatic notation to indicate the requirement that two disconnected diagram parts must have an index duplication. We show such a notation in Fig. 3. If we now generalize "connection" to include these "EPV connections," we retain the simple prescription from fermion theory that $[H,T]\Phi$ is represented by the sum of the "connected" diagrams of $HT\Phi$.

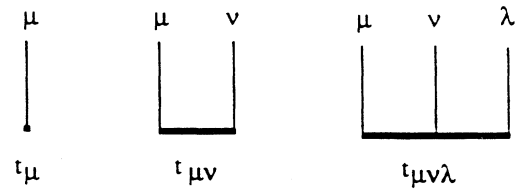

FIG. 2. Diagram fragments of cluster operators of sizes 1, 2, and 3. Below each is shown its value.

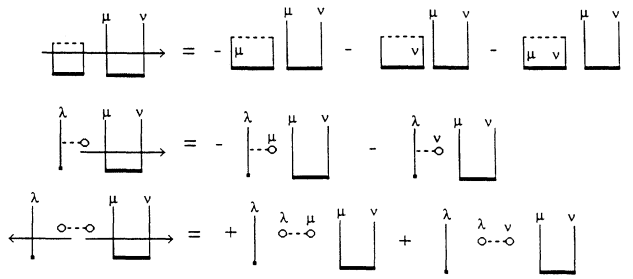

FIG. 3. Examples of the notation for "EPV connected" diagram and the labelings and signs thereby implied. The horizontal arrow indicates that one or more of the labels on the open line(s) it traverses must also be attached to traversed closed lines in all distinct ways, and each resultant diagram must have its sign changed.

Continuing now to the multiple commutators, we see that $[[H,T],T]\Phi$ is represented by the "connected" diagrams of $[H,T]T\Phi$ and therefore by the "connected" diagrams of $HTT\Phi$. The result generalizes to higher commutators; we note that because H has but two labels, there can at most be two EPV connection lines, and that EPV connections only result when a line from H is closed (terminates on a T fragment). Thus, an H fragment can at most tie together four T fragments, meaning that contributions from the series in Eq. (7) terminate with the fourfold commutator.

The coefficients in Eqs. (4) and (7) are such that if all T fragments are drawn at the same vertical position, the right side of Eq. (7) is represented by the sum of all distinct legitimate labelings of all "connected" diagrams which are distinct (distinctness implies nonsuperposability even after interchange of equivalent connections or horizontal distortions, including left-right reversal). Therefore, Eq. (5), which involves the Φ component of Eq. (7), gives E as the sum of all distinct closed "connected" diagrams containing one H fragment and below it zero or more T fragments. For each $\Phi_{\mu\nu\ldots}$, the corresponding Eq. (6) requires the vanishing of the sum of all distinct "connected" diagrams containing one H fragment and below it zero or more T fragments, with open lines at the top bearing labels $\mu\nu\ldots$.

III. APPLICATION

Specializing to the illustrative example, we take the reference state Φ to be the Néel state,[7] with individual spin states $s_z=\hbar/2$ on a sublattice consisting of half the lattice sites, and individual spin states $s_z=-\hbar/2$ on the sublattice consisting of the remaining sites, with all nearest neighbors on different sublattices. The Hamiltonian of Eq. (2) then causes the five diagram fragments in Fig. 1(a), with μ and ν nearest neighbors, to have the respective values shown there (all energies in this paper are stated in units of $\hbar^2J$). All other matrix elements of H vanish.

The requirement that S_z, the z component of the total spin, be conserved limits the cluster operators T_n to even n, with half the indices of each $t_{\mu\nu\ldots}$ on each sublattice. The diagrammatic equation for E is therefore that shown in Fig. 4(a). The first diagram gives the energy of the reference state, i.e., the Néel energy. For an N-site system, this diagram occurs for $2N$ distinct nearest-neighbor pairs, so $E_{\text{Néel}}=2N(-1/4)=-0.5N$, or (per site) $E_{\text{Néel}}/N=-0.5$. The other diagram of Fig. 4(a) gives the remainder of the energy; it also occurs for $2N$ distinct nearest-neighbor pairs. If we assume that all t coefficients have values that depend only on the relative geometry (but not on orientation) and denote the $t_{\mu\nu}$ for nearest neighbors $t_{2,1}$ [the indices refer to the cluster size and geometry (see Fig. 6)], this diagram contributes $2N(+1/2)t_{2,1}$, or simply $t_{2,1}$ per site. Summarizing,

$$E/N=-0.5+t_{2,1}\ . \tag{8}$$

Although the energy depends only on the cluster coefficient $t_{2,1}$, that coefficient ultimately depends upon contributions from clusters of all sizes and geometries. Actual computations will require a cluster-type truncation.

The diagrammatic equation for T_2 is shown in Fig. 4(b). Before evaluating this equation it is helpful to rearrange it by expanding the EPV diagrams and then replacing the resulting disconnected diagrams by connected diagrams to which they are numerically equal. The process is illustrated for the first EPV diagram in Fig. 5(a). A similar process for the other EPV diagram is presented in Fig. 5(b). Then, combining similar diagrams and rearranging, the T_2 equation can be brought to the form shown in Fig. 5(c). Continuing with the assumption that the t coefficients depend only on relative geometry, we now evaluate all the diagrams except those labeled $F_{\mu\nu}$ and $G_{\mu\nu}$, obtaining for each diagram the value written below it. The values for the third and fourth diagrams reflect the fact that the bubble can have four equivalent label assignments unless μ and ν are nearest neighbors, in which case only three label assignments are available.

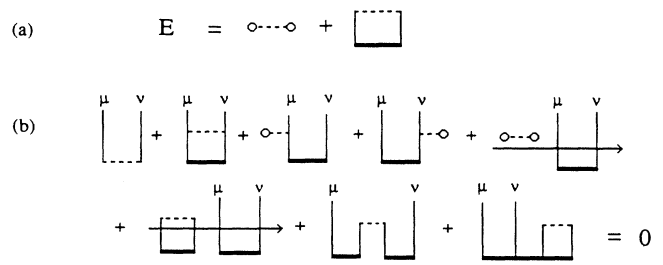

FIG. 4. Diagrammatic equations for (a) E and (b) T_2, for the $S=1/2$ Heisenberg antiferromagnet with Néel reference state (for details, see text).

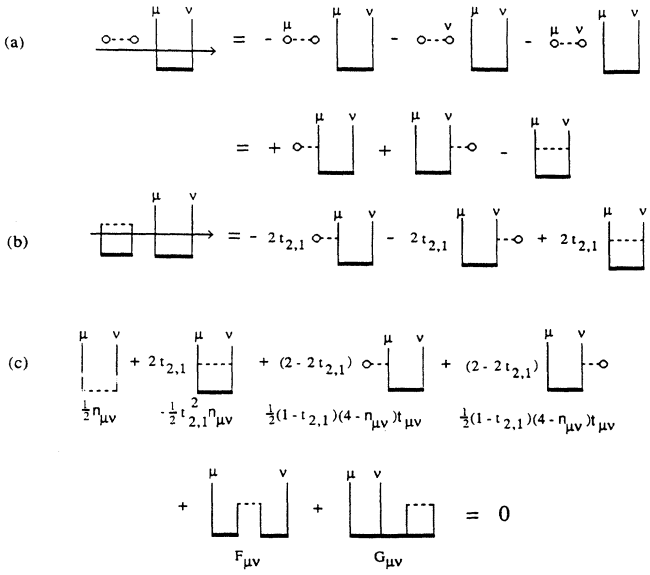

FIG. 5. (a) and (b) Reduction of the EPV diagrams of Fig. 4(b); (c) result of substituting from (a) and (b) into the equation of Fig. 4(b), with values shown under the diagrams.

The equation represented by Fig. 5(c) can be rearranged to our final computational form:

$$t_{\mu\nu}=-\frac{(1+t_{2,1})n_{\mu\nu}}{6}-\frac{(F_{\mu\nu}+G_{\mu\nu})}{(4-n_{\mu\nu})(1-t_{2,1})}\ . \tag{9}$$

The symbol $n_{\mu\nu}$ was defined in Eq. (2).

Each choice of labels μ,ν in Eq. (9) yields a separate equation; there will be one independent equation for each relative geometry. It will be found that $t_{\mu\nu}$ rapidly becomes smaller with increasing distance between sites μ and ν, so that only a few relative geometries actually have non-negligible $t_{\mu\nu}$. The geometries and the identification numbers we have assigned to them are shown in Fig. 6. The diagram $F_{\mu\nu}$ can now be evaluated as the sum over index numbers i,j of $\frac{1}{2}t_{2,i}t_{2,j}$ times the number of ways the displacement $2,i$ from μ and the displacement $2,j$ from ν (in any direction) can be on adjacent unoccupied sites. (The "$\frac{1}{2}$" is from the matrix element of h). We illustrate by exhibiting $F_{2,1}$, $F_{2,2}$, and $F_{2,3}$ in an approximation retaining $t_{2,j}$ for $j\leq 3$:

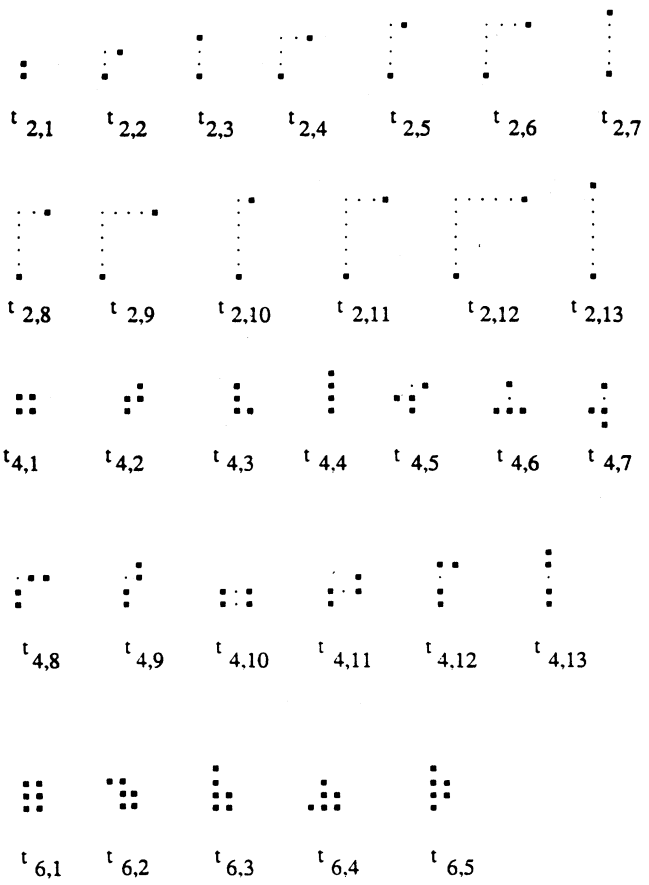

FIG. 6. Clusters of sizes 2, 4, and 6 and the identification numbers assigned to them. The assumptions used in calculations cause a single identification number to apply to all clusters related to each other by rotation or reflection.

$$\begin{aligned}
F_{2,1}&=t_{2,1}(t_{2,1}+6t_{2,2}+t_{2,3})+t_{2,2}(7t_{2,2}+4t_{2,3})+2t_{2,3}^2\ ,\\
F_{2,2}&=t_{2,1}(\tfrac{3}{2}t_{2,1}+3t_{2,2}+2t_{2,3})+t_{2,2}(3t_{2,2}+4t_{2,3})\ ,\\
F_{3,3}&=t_{2,1}(\tfrac{1}{2}t_{2,1}+4t_{2,2})+4t_{2,2}^2\ .
\end{aligned} \tag{10}$$

The diagram $G_{\mu\nu}$ is a sum over index numbers i of $\frac{1}{2}t_{4,i}$ times the number of ways the cluster $4,i$ can be positioned to cover sites μ,ν, and two additional sites adjacent to each other, but not necessarily to μ and ν. (Again the "$\frac{1}{2}$" is from h). We illustrate $G_{2,1}$, $G_{2,2}$, and $G_{2,3}$ in an approximation retaining the 13 most important four-clusters (see Fig. 6).

$$\begin{aligned}
G_{2,1}&=\tfrac{1}{2}(2t_{4,1}+4t_{4,2}+8t_{4,3}+2t_{4,4}+4t_{4,8}+4t_{4,9}\\
&\quad +2t_{4,10}+4t_{4,11}+8t_{4,12}+2t_{4,13})\ ,\\
G_{2,2}&=\tfrac{1}{2}(t_{4,2}+2t_{4,3}+2t_{4,5}+2t_{4,6}+2t_{4,7})\ ,\\
G_{3,3}&=\tfrac{1}{2}(t_{4,4}+4t_{4,7})\ .
\end{aligned} \tag{11}$$

Our initial calculations will be with truncation after T_2, corresponding to $G_{\mu\nu}$ tentatively set to zero. Equations (9) and (10), though nonlinear, are easily solved simultaneously by a successive approximation scheme in which the $(n+1)$th iterations of the $t_{2,i}$ are obtained by evaluating Eqs. (10) with the nth iterations on the right-hand sides. We start with $t_{2,1}=-1/6$ and all other $t_{2,i}$ zero. We report calculations keeping $t_{2,j}$ through $j_{\max}=3,5,8$, and 13. Table I shows the values of $t_{2,1}$ thereby obtained, also all fully converged $t_{2,j}$ through $j=8$ (i.e., those obtained in the limit of large $j_{\max}$). Already at $j_{\max}=3$ we have over 99.9% of the T_2 energy; it is practical to reach a $j_{\max}$ value for which convergence is essentially complete. The T_2 energy leads to a total energy (per site) -0.6508. Bishop *et al.*[4] report results for this problem only at the T_2 approximation; their T_2 result is in agreement with ours.

Our next calculations include T_4. The diagrammatic equation for T_4 is shown in Fig. 7. From Fig. 7(b) we reach the final computational form

$$t_{\mu\nu\lambda\sigma}=\frac{-F_{\mu\nu\lambda\sigma}-G_{\mu\nu\lambda\sigma}+K_{\mu\nu\lambda\sigma}-L_{\mu\nu\lambda\sigma}}{8-n_{\mu\nu\lambda\sigma}(1+t_{2,1})}\ , \tag{12}$$

where $n_{\mu\nu\lambda\sigma}$ is the number of adjacent pairs in the set $\mu\nu\lambda\sigma$. Assuming μ and λ to be on one sublattice and ν and σ on the other,

TABLE I. Values of cluster coefficient $t_{2,1}$ for square Heisenberg antiferromagnet in T_2 approximation with cutoff after $t_{2,j_{\max}}$; also converged values of $t_{2,j}$. Energy unit: $\hbar^2J$.

			Values for $j_{\max}=\infty$		
$j_{\max}$	$t_{2,1}$	j	$t_{2,j}$	j	$t_{2,j}$
3	−0.150721	1	−0.150834	5	−0.000774
5	−0.150824	2	−0.008791	6	−0.000239
8	−0.150833	3	−0.003915	7	−0.000316
13	−0.150834	4	−0.001242	8	−0.000171

$$F_{\mu\nu\lambda\sigma}=4(-\tfrac{1}{4})[t_{\mu\nu}t_{\lambda\sigma}(n_{\nu\lambda}+n_{\mu\sigma}) + t_{\mu\sigma}t_{\lambda\nu}(n_{\sigma\lambda}+n_{\mu\nu})] , \quad (13)$$

$$K_{\mu\nu\lambda\sigma}=\frac{1}{2}\sum_{\rho\tau} n_{\rho\tau}(t_{\mu\nu\lambda\rho}t_{\tau\sigma}+t_{\mu\nu\rho\sigma}t_{\tau\lambda} + t_{\mu\rho\lambda\sigma}t_{\tau\nu}+t_{\rho\nu\lambda\sigma}t_{\tau\mu}) , \quad (14)$$

$$L_{\mu\nu\lambda\sigma}=2\left[\frac{1}{2}\right]\sum_{\tau}(n_{\mu\tau}t_{\lambda\tau}t_{\mu\nu}t_{\mu\sigma}+n_{\lambda\tau}t_{\mu\tau}t_{\lambda\nu}t_{\lambda\sigma} + n_{\nu\tau}t_{\sigma\tau}t_{\mu\nu}t_{\lambda\nu}+n_{\sigma\tau}t_{\nu\tau}t_{\mu\sigma}t_{\lambda\sigma}) . \quad (15)$$

The summation in Eq. (15) is limited to sites τ neighboring the cluster points; that in Eq. (14) is not limited but sites τ not neighboring the cluster points give extremely small contributions.

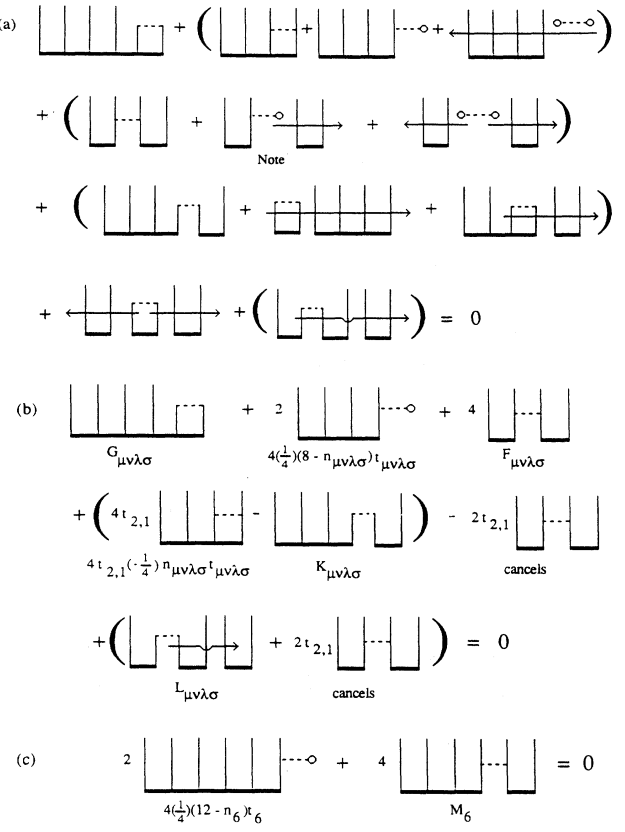

FIG. 7. Diagrammatic equations for T_4 and T_6: (a) For T_4, before simplifying; (b) for T_4, after simplifying; each parenthesized group in (a) has been reduced to a single diagram or a simpler parenthesized group; (c) for T_6 (leading terms only). The indication "Note" marks a diagram in which the two inequivalent pair clusters will be reduced to a form in which the pairs are equivalent, so the resulting diagram must be multiplied by 2. The open lines are unlabeled; labels must be attached to them in all distinguishable ways. Values are shown under the diagrams in (b) and (c). Note that $F_{\mu\nu\lambda\sigma}$ includes the coefficient "4".

The only T_6 contributions we shall include in $G_{\mu\nu\lambda\sigma}$ are from the five six-clusters shown in Fig. 6, and we only use them for the 13 most important T_4 clusters. The formulas for the first four $G_{4,j}$ are

$$\begin{aligned} G_{4,1}&=-\tfrac{1}{2}(4t_{6,1}+8t_{6,2}+8t_{6,3}) , \\ G_{4,2}&=-\tfrac{1}{2}(t_{6,2}+2t_{6,5}) , \\ G_{4,3}&=-\tfrac{1}{2}(t_{6,1}+t_{6,2}+t_{6,3}+2t_{6,4}) , \\ G_{4,4}&=-\tfrac{1}{2}(t_{6,3}+t_{6,5}) . \end{aligned} \quad (16)$$

Tentatively setting the $G_{4,j}$ to zero, we use Eq. (12), along with Eq. (9), to determine simultaneous solutions for the $t_{2,j}$ and the $t_{4,j}$. We use the same iterative scheme as for the $t_{2,j}$ alone. We wrote a computer program to generate about 600 candidate four-clusters, finding that energy convergence to 10^{-5} was reached with eight T_2 clusters and 43 T_4 clusters, and that only slightly poorer results were still obtained when the number of T_4 clusters was reduced to 13. Table II shows $t_{2,j}$ ($j=1,8$) and $t_{4,j}$ ($j=1,13$) for fully converged (large $j_{\max}$) calculations and also for a calculation truncated to the clusters shown. We note that the inclusion of T_4 lowers the total energy per site from the T_2 value of -0.6508 to -0.6649. The only previously reported calculation for this problem that includes a T_4 contribution is that of Roger and Hetherington.[2] Their calculation is not directly comparable with ours because it includes only a portion of T_2 and only one T_4 cluster, but it did indicate that T_4 makes a significant energy contribution.

Our final calculation included the five T_6 clusters shown in Fig. 6. We shall be satisfied with a 10% error in these T_6 coefficients, and therefore keep only the leading terms in the diagrammatic equation, as shown in Fig. 7(c),

$$t_6=-M_6/(12-n_6) . \quad (17)$$

TABLE II. Values of cluster coefficients $t_{2,j}$ and $t_{4,j}$ for square Heisenberg antiferromagnet in T_2+T_4 approximation, for fully converged and for truncated sets of $t_{2,j}$ and $t_{4,j}$. Energy unit: $\hbar^2 J$.

	$t_{2,j}$			$t_{4,j}$	
j	$j_{\max}=8$	$j_{\max}=\infty$	j	$j_{\max}=13$	$j_{\max}=\infty$
1	−0.16490	−0.16491	1	+0.02221	+0.02226
2	−0.01165	−0.01155	2	+0.00372	+0.00373
3	−0.00535	−0.00533	3	+0.00414	+0.00414
4	−0.00185	−0.00198	4	+0.00482	+0.00480
5	−0.00118	−0.00127	5	−0.00111	−0.00109
6	−0.00039	−0.00045	6	−0.00045	−0.00045
7	−0.00049	−0.00055	7	−0.00052	−0.00052
8	−0.00027	−0.00033	8	−0.00026	−0.00024
			9	−0.00027	−0.00025
			10	−0.00039	−0.00036
			11	−0.00016	−0.00015
			12	−0.00011	−0.00010
			13	−0.00004	−0.00003

TABLE III. Values of cluster coefficients $t_{2,j}, t_{4,j}$, and $t_{6,j}$ for square Heisenberg antiferromagnets in $T_2+T_4+T_6$ approximation, with $t_{6,j}$ cut off after $j=5$, for fully converged and for truncated sets of $t_{2,j}$ and $t_{4,j}$. Energy unit: $\hbar^2 J$.

	$t_{2,j}$			$t_{4,j}$			$t_{6,j}$
j	$j_{\max}=8$	$j_{\max}=\infty$	j	$j_{\max}=13$	$j_{\max}=\infty$	j	$j_{\max}=5$
1	−0.16816	−0.16817	1	+0.02696	+0.02701	1	−0.00049
2	−0.01245	−0.01236	2	+0.00406	+0.00406	2	−0.00011
3	−0.00572	−0.00571	3	+0.00507	+0.00507	3	−0.00012
4	−0.00203	−0.00218	4	+0.00517	+0.00515	4	−0.00008
5	−0.00130	−0.00140	5	−0.00103	−0.00102	5	−0.00008
6	−0.00044	−0.00052	6	−0.00034	−0.00034		
7	−0.00055	−0.00061	7	−0.00049	−0.00048		
8	−0.00031	−0.00038	8	−0.00011	−0.00009		
			9	−0.00010	−0.00009		
			10	−0.00004	−0.00001		
			11	−0.00000	−0.00002		
			12	−0.00003	−0.00002		
			13	−0.00004	−0.00003		

Here n_6 is the number of adjacent pairs in the index set for t_6 and M_6 is the diagram so labeled in Fig. 7(c).

We illustrate the evaluation of Eq. (17) for $t_{6,j}$ $(j=1,5)$ in an approximation dropping all terms containing $t_{4,j}$ with $j>4$:

$$t_{6,1}=\tfrac{1}{5}[t_{2,1}(4t_{4,1}+12t_{4,3})+8t_{2,2}t_{4,2}]\ ,$$

$$t_{6,2}=\tfrac{1}{6}[t_{2,1}(t_{4,1}+2t_{4,2}+2t_{4,3})+3t_{2,2}(t_{4,2}+t_{4,3})]\ ,$$

$$t_{6,3}=\tfrac{1}{6}[t_{2,1}(t_{4,1}+2t_{4,3}+2t_{4,4})+3t_{2,2}t_{4,3}+3t_{2,3}t_{4,2}]\ , \tag{18}$$

$$t_{6,4}=\tfrac{1}{6}[t_{2,1}(2t_{4,2}+4t_{4,3})+t_{2,2}(2t_{4,1}+3t_{4,2}+3t_{4,3})]\ ,$$

$$t_{6,5}=\tfrac{1}{6}[t_{2,1}(4t_{4,2}+2t_{4,4})+6t_{2,2}t_{4,3}+2t_{2,3}t_{4,1}]\ .$$

Iteratively solving simultaneously Eqs. (9), (12), and (17), we obtained the coefficient values shown in Table III. The total energy per site has now been lowered to −0.6682.

IV. STAGGERED MAGNETIZATION

In a coupled-cluster formalism, the most consistent way to calculate properties is by including an appropriate perturbation in the Hamiltonian and then determining its first-order (or, if relevant, higher-order) contributions to the energy. For the staggered magnetization (denoted $M^\dagger$) we add to H the operator $H'=g\sum_\mu \varepsilon_\mu s_{z\mu}$, where $\varepsilon_\mu=+1$ on the sublattice with $s_z=\hbar/2$ in the reference state, and $\varepsilon_\mu=-1$ on the other sublattice. Then, because H' is diagonal and has only the diagram fragments shown

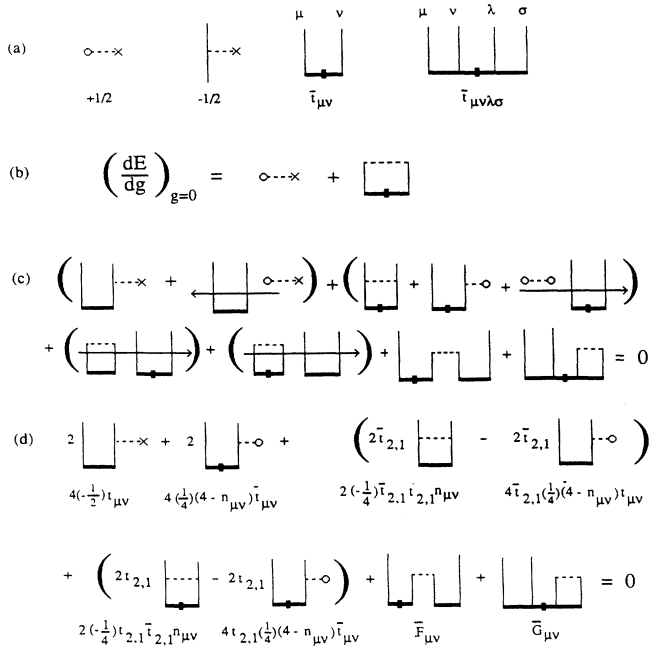

FIG. 8. Diagrams for calculation of staggered magnetization: (a) fragments of the staggered magnetization operator and of first-order cluster contributions, in units of $\hbar$, with values under the diagrams; (b) first-order perturbation of the energy; (c) unsimplified diagrammatic equation for the first-order contributions $\bar{t}_{\mu\nu}$; (d) simplified equation for $\bar{t}_{\mu\nu}$, with values under the diagrams.

in Fig. 8(a), $M^\dagger=(dE/dg)_{g=0}$ can be represented diagrammatically as in Fig. 8(b). The $\bar{t}$ [represented diagrammatically in Fig. 8(a)] are the coefficients of g in the expansions of the corresponding t. The first diagram in Fig. 8(b) has (in units of $\hbar$) the value $+0.5N$, and is the staggered magnetization of the Néel state. The second diagram has the value $2N(+1/2)\bar{t}_{2,1}$. Thus,

$$M^\dagger/N=+0.5+\bar{t}_{2,1}\ . \tag{19}$$

Writing all cluster operators as a sum of zero-order (unperturbed) and first-order terms, the equation for $\bar{t}_{\mu\nu}$ consists of the first-order terms in the projection equation, Eq. (6), involving $\Phi_{\mu\nu}$. This equation, in diagrammatic form, is shown in Figs. 8(c) and 8(d).

The $\bar{t}_{\mu\nu}$ equation can be rearranged to the form

$$\bar{t}_{\mu\nu}=\frac{t_{\mu\nu}(2+4\bar{t}_{2,1})-\bar{F}_{\mu\nu}-\bar{G}_{\mu\nu}}{(4-n_{\mu\nu})(1-t_{2,1})}\ . \tag{20}$$

The diagrams $\bar{F}_{\mu\nu}$ and $\bar{G}_{\mu\nu}$ (first order in $\bar{t}$) can be evaluated by methods similar to those presented in Sec. III. The $\bar{G}_{\mu\nu}$ involve four-clusters $\bar{t}_{\mu\nu\lambda\sigma}$ which can be shown to satisfy

$$\bar{t}_{\mu\nu\lambda\sigma}=\frac{t_{\mu\nu\lambda\sigma}(4+n_{\mu\nu\lambda\sigma}\bar{t}_{2,1})-\bar{F}_{\mu\nu\lambda\sigma}-\bar{G}_{\mu\nu\lambda\sigma}+\bar{K}_{\mu\nu\lambda\sigma}-\bar{L}_{\mu\nu\lambda\sigma}}{8-n_{\mu\nu\lambda\sigma}(1+t_{2,1})}\ . \tag{21}$$

The $\bar{F}_{\mu\nu\lambda\sigma}, \bar{G}_{\mu\nu\lambda\sigma}, \bar{K}_{\mu\nu\lambda\sigma}$, and $\bar{L}_{\mu\nu\lambda\sigma}$ are first order in $\bar{t}$ and are evaluated by the methods of Sec. III. The $\bar{G}_{\mu\nu\lambda\sigma}$ involve clusters $\bar{t}_6$, given by

$$\bar{t}_6=(6t_6-\bar{M}_6)/(12-n_6) . \tag{22}$$

Here $\bar{M}_6$ is first order in $\bar{t}$.

Truncating $\bar{t}_{2,j}$ at $j_{\max}=5$, $\bar{t}_{4,j}$ at $j_{\max}=13$ and $\bar{t}_{6,j}$ at $j_{\max}=5$, we obtain in the T_2, T_2+T_4, and $T_2+T_4+T_6$ approximations the respective $\bar{t}_{2,1}$ values -0.097, -0.127, and -0.139, corresponding to staggered magnetizations (per site) 0.40, 0.37, and 0.36.

V. DISCUSSION

The results of the present calculations and those of two other groups of workers are summarized in Table IV. (For a more complete survey of earlier work, see Ref. 6.) The Green-function Monte Carlo (GFMC) study[6] is nonvariational, but becomes exact in the limit of zero statistical error. When run on finite systems with periodic boundary conditions, it shows a finite-size effect which can be extrapolated to give infinite-system results. The resonating valence bond (RVB) study[8] is variational, and was run on 32×32 lattices (the GFMC calculations indicate a negligible finite-size effect at this system size). There is no logical extrapolation procedure applicable to the RVB data.

Our results are also nonvariational, but become exact in the limit of including clusters of all sizes. The extrapolated values shown in the table were obtained by fitting a geometric series to the results of actual calculations. This procedure is reasonable and *a posteriori* justified by the results, but is not supported by a definitive theoretical argument. In any event, the present work supports a claim that both it and the GFMC calculations are free from unnoticed systematic errors.

One of the most significant differences between the three sets of calculations is the computational labor required. Both the GFMC and RVB studies involve numerical Monte Carlo procedures that can become computationally extremely intensive. The present study could have been carried out by hand; actually a PC was used to organize the work and reduce the possibility of error (typical running time: a few seconds). We note that the GFMC and present studies actually reached energies approximately equally close to the extrapolated value.

The main difference between the results of the three studies lies in the estimates of staggered magnetization. The RVB study used an *ad hoc* long-range wave function and could not be claimed to be converged (with the lack of convergence potentially affecting the staggered magnetization more than it would affect the energy). The GFMC and present studies do not completely agree, but the largest actual calculations of both types are quite close. We do not totally disclaim the possibility of error in our value of $M^\dagger/N$; a fair amount of nonautomated algebra was involved.

TABLE IV. Comparison of calculations of the energy and staggered magnetization for square Heisenberg antiferromagnet consisting of N sites. Energy unit: $\hbar^2J$; magnetization unit: $\hbar$.

	E/N	$M^\dagger/N$
Néel state	−0.5000	0.50
GFMC[a]		
6×6	−0.6789	0.40[c]
8×8	−0.6734	0.38[c]
12×12	−0.6702	0.35[c]
extrapolated	−0.6692	0.31
RVB[b]		
dimer	−0.604	0.00
best short range	−0.6682	0.00
with long range	−0.6688	0.23
Present work		
T_2	−0.6508	0.40
T_2+T_4	−0.6649	0.37
$T_2+T_4+T_6$	−0.6682	0.36
extrapolated	−0.6692	0.35

[a]Green-function Monte Carlo; Trivedi and Ceperley, Ref. 6.
[b]Resonating valence bond; Liang, Doucot, and Anderson, Ref. 8.
[c]Estimated from a graph.

One of the nice features of the coupled-cluster formalism is that it applies directly in the infinite-system limit, and is (in electronic structure language) size consistent. Thus, as we have seen, truncated calculations nevertheless apply to an infinite system and extrapolations are to account for omitted types of clusters rather than for finite system size.

Finally, we note that the current formulation is not exactly a partially summed perturbation expansion, in contrast to the usual electronic structure formalism. This fact directly manifests itself in two ways in the diagrammatic representation: (1) there are no energy denominators, and (2) one cannot repeatedly substitute the diagram equations for cluster operators to obtain a traditional linked cluster expansion.

[1]F. Coester, Nucl. Phys. **7**, 421 (1958); H. Kümmel, K. H. Lührmann, and J. G. Zabolitsky, Phys. Rep. **36**, 1 (1978); J. Cizek, J. Chem. Phys. **45**, 4256 (1966).
[2]M. Roger and J. H. Hetherington, Phys. Rev. B **41**, 200 (1990).
[3]R. F. Bishop, Theor. Chim. Acta **80**, 95 (1991).
[4]R. F. Bishop, J. B. Parkinson, and Y. Xion, Phys. Rev. B **43**, 13 782 (1991); Theor. Chim. Acta **80**, 181 (1991).
[5]For a review of diagrammatic methods in electronic structure calculations, see F. E. Harris, H. J. Monkhorst, and D. L. Freeman, *Algebraic and Diagrammatic Methods in Many-Fermion Theory* (Oxford University, New York, 1992).
[6]For a review of this problem, see N. Trivedi and D. M. Ceperley, Phys. Rev. B **41**, 4552 (1990).
[7]See, for example, D. C. Mattis, *The Theory of Magnetism* (Springer-Verlag, New York, 1981).
[8]S. Liang, B. Doucot, and P. W. Anderson, Phys. Rev. Lett. **61**, 365 (1988).