PHYSICAL REVIEW A VOLUME 8, NUMBER 5 NOVEMBER 1973

Variational Multistate-Ket Derivation of the Heisenberg Spin Hamiltonian*

D. J. Klein
Department of Physics, The University of Texas at Austin, Austin, Texas 78712

David C. Foyt†
Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712

(Received 26 March 1973)

A new variational derivation scheme for the Heisenberg exchange Hamiltonian is presented. A many-body variational *Ansatz* for the multistate ket appearing in the derivation is described and is applied to a simple model problem containing the qualitatively important different types of exchange contributions. Advantages of the present derivation over others are briefly indicated.

I. INTRODUCTION

It is often convenient to replace the Schrödinger Hamiltonian by an effective Hamiltonian which, while different in form and frequently acting on a different space, nevertheless reproduces to within some degree of approximation the effect of the Schrödinger Hamiltonian on a given space. The usual Heisenberg spin Hamiltonian describing the interaction between paramagnetic sites provides an example of just such an effective Hamiltonian. In the original Heisenberg spin-Hamiltonian derivations[1-3] the representations of the Schrödinger Hamiltonian on a space generated by permutations of the electron indices acting on a single primitive ket, taken as the simple product of site kets, is to be reproduced. However, as pointed out by Herring,[4] this restricted representation of the Schrödinger Hamiltonian does not yield the correct exchange splittings, even for weakly interacting sites. Evidently this simple product *Ansatz* for the primitive ket is primarily of use in obtaining a suggestive parametric form for an effective Hamiltonian.

More realistic derivations[4-7] of the Heisenberg spin Hamiltonian typically obtain it as a degenerate perturbation result, the space spanned by the kets generated from the simple product primitive ket being treated as a zero-order eigenspace. These perturbative derivations yield[5] not only "direct" exchange contributions to the effective Heisenberg exchange parameter but also so-called "kinetic" exchange terms, which may be of comparable magnitude but opposite in sign. Indeed in many cases the kinetic exchange term is estimated to be more important, and hence a qualitative correction in the form of the effective exchange parameter is obtained.

Another method of obtaining a Heisenberg spin Hamiltonian is to perform an *ab initio* calculation and fit the resulting energy levels to a parametrized Heisenberg spin Hamiltonian. Such computations[8] often yield results qualitatively different from the simplest derivation. This *ab initio* method is apparently only readily applicable to systems of only a few paramagnetic sites and the simple form of the resulting approximate spin Hamiltonian is obscured, and not put to use, in the computations.

Here we describe a variational method by which a more accurate single primitive ket, called a multistate ket, is derived. Using this single multistate ket in place of the simple product ket then leads, via the original derivations[1-3] to a more accurate result which, in principle, can[4,9] be exact. This variational alternative to the perturbative derivations takes into account the same "direct" and "kinetic" exchange effects, although with some additional corrections through higher orders. Indeed, the variational approach is not limited to small values of the perturbation parameter but rather by the particular variational form chosen for the multistate ket.

In the multistate-ket variational scheme, described in the preceding paper,[9] the expectation value of the Hamiltonian is minimized subject to the constraints that the multistate ket contain prescribed amounts of all the permutational symmetries arising from a given separated-atom limit. For each of these symmetry constraints one may introduce a corresponding Lagrange multiplier. One then minimizes the Hamiltonian minus an element of the group algebra which is linear in the Lagrange multipliers. When the Lagrange multipliers are chosen to correspond to (stabilized) matric basis elements of the group algebra, they become[9] estimates of the projected single-state energies. Here we interpret this Lagrange-multiplier group-algebraic element as an effective Hamiltonian, and it is expressed in terms of permutations interchanging electrons among sites in different ways. This expression allows us to identify some useful approximations applicable to the multistate scheme, and also leads directly

to a Heisenberg spin Hamiltonian.

The different patterns of interchanging electrons among sites are in one-to-one correspondence with double cosets of the symmetric group. Thus this natural group-theoretic tool plays an important part in the general development of Sec. II. In Sec. III we develop a general cluster-expansion *Ansatz* for the multistate ket; simplifying assumptions restrict our attention to one electron and orbital per site, although the development applies to infinite systems as well as finite. In Sec. IV this development is applied to a simple model example embodying both "direct" and "kinetic" exchange effects.

II. THE EFFECTIVE HAMILTONIAN

We follow the notation in preceding work.[9] Then $\mathcal{S}^0$ is the group of permutations interchanging no electrons among sites, and the multistate ket $|\psi\rangle$ is to be of a given symmetry, say $\alpha^0\gamma^0$, of $\mathcal{S}^0$. Further, $|\psi\rangle$ is to have components of all symmetries α of the full symmetric group $\mathcal{S}_N$ which arise from $\alpha^0\gamma^0$ of $\mathcal{S}^0$. To determine $|\psi\rangle$ one minimizes

$$\langle\psi|H-\mathcal{H}|\psi\rangle, \tag{2.1}$$

where $\mathcal{H}$ is a group-algebraic element

$$\mathcal{H}\equiv\sum_{\alpha\rho\sigma}\mathcal{E}_{\alpha\rho\sigma}e^{\alpha}_{\rho\sigma}, \tag{2.2}$$

with the Lagrange multipliers $\mathcal{E}_{\alpha\rho\sigma}$ chosen to satisfy the symmetry constraints

$$\langle\psi|e^{\alpha}_{\rho\sigma}|\psi\rangle=a_{\alpha}\delta_{\rho\sigma}. \tag{2.3}$$

Here the elements $e^{\alpha}_{\rho\sigma}$ of the group algebra $\mathfrak{A}(\mathcal{S}_N)$ of $\mathcal{S}_N$ form a matric basis

$$\{e^{\alpha}_{\rho\sigma};\ \alpha\rho\sigma \text{ ranging}\} \tag{2.4}$$

for the subalgebra of $\mathfrak{A}(\mathcal{S}_N)$, which is of the zero-order symmetry $\alpha^0\gamma^0$ on both the left and right. If we consider the vector space of symmetry $\alpha^0\gamma^0$ induced from a single product

$$|0\rangle\equiv|\alpha_A\gamma_A\rangle\otimes|\alpha_B\gamma_B\rangle\otimes|\alpha_C\gamma_C\rangle\otimes\cdots \tag{2.5}$$

of strongly orthogonal site kets, then it is seen that $\mathcal{H}$ leaves this space invariant. Indeed noting that a basis for this space is

$$\{e^{\alpha}_{\rho\sigma}|0\rangle;\ \alpha\rho\sigma \text{ ranging}\}, \tag{2.6}$$

we see that

$$\mathcal{H}e^{\alpha}_{\rho\sigma}|0\rangle=\sum_{\tau}\mathcal{E}_{\alpha\tau\rho}e^{\alpha}_{\tau\sigma}|0\rangle. \tag{2.7}$$

Transforming to the stabilized matric basis,[9] $\mathcal{H}$ becomes

$$\mathcal{H}=\sum_{\alpha\beta}\mathcal{E}_{\alpha\beta}e^{\alpha}_{\beta\beta}, \tag{2.8}$$

and the $\mathcal{E}_{\alpha\beta}$ and $e^{\alpha}_{\beta\beta}|0\rangle$ are seen to be eigenvalues and eigenkets to $\mathcal{H}$. Since these stabilized Lagrange multipliers $\mathcal{E}_{\alpha\beta}$ correspond to the single-state energies, $\mathcal{H}$ is seen to be an effective Hamiltonian.

Since $\mathcal{H}$ is an element of the group algebra $\mathfrak{A}(\mathcal{S}_N)$, it may be expressed as a linear combination of permutations. To accomplish this expression of $\mathcal{H}$ in the desired manner, we first consider the *double-coset* (DC) decomposition of $\mathcal{S}_N$,

$$\mathcal{S}_N=\sum_{q}\oplus\,\mathcal{S}^0G_q\mathcal{S}^0, \tag{2.9}$$

where the qth DC, $\mathcal{S}^0G_q\mathcal{S}^0$, is disjoint from every other DC. The $G_q\in\mathcal{S}_N$ are termed DC generators. The DC *symbol* of a permutation $P\in\mathcal{S}_N$ is defined to be the array with (I,J)th element D^P_{IJ} equal to the number of electrons transferred from site J to site I. Two permutations are in the same DC if and only if[10] they have the same DC symbol. Hence we see that the DC's are in unique correspondence with the different patterns of transferring electrons among sites. Further, we see that matrix elements of permutations between kets in which the electrons are localized, in accordance with the assignment of electrons to the site groups of $\mathcal{S}^0$, decrease in magnitude as the number of electrons, and the distance, increases. Indeed since differential overlap between electrons on different centers decreases exponentially with intersite separation, a similar exponential decrease is expected in these matrix elements. Hence so long as localized kets are employed, some of the matrix elements over higher DC generators are very small, and might be neglected with little compromise in a calculation.

The DC decomposition (2.9) implies[11]

$$e^{\alpha}_{\rho\sigma}=\frac{f^{\alpha}}{N!}\left(\frac{g^0}{f^{\alpha^0}}\right)^2\sum_{q}d_q^{-1}\times\sum_{s^0t^0}[G_q^{-1}]^{\alpha}_{(\sigma\alpha^0t^0)(\rho\alpha^0s^0)}e^{\alpha^0}_{r^0s^0}G_qe^{\alpha^0}_{t^0r^0}, \tag{2.10}$$

where the $e^{\alpha^0}_{r^0s^0}$ are matric-basis elements for the group algebra of the subgroup $\mathcal{S}^0$ and f^{α}, g^0, f^{α^0}, d_q, and $[G_q^{-1}]^{\alpha}_{(\sigma\alpha^0t^0)(\rho\alpha^0s^0)}$ are group-theoretic numbers identified elsewhere.[9,11] Hence the set of elements

$$\{e^{\alpha^0}_{r^0s^0}G_qe^{\alpha^0}_{t^0r^0};\ s^0qt^0 \text{ ranging}\} \tag{2.11}$$

spans the subalgebra of $\mathfrak{A}(\mathcal{S}_N)$ transforming as $\alpha^0\gamma^0$ on both the left and right. If this spanning set (2.11) is used in place of the matric basis (2.4), the multistate variational scheme may be formulated as minimizing (2.1), where

$$\mathcal{H}=\sum_{s^0qt^0}{}'\mathcal{J}_{s^0qt^0}e^{\alpha^0}_{r^0s^0}G_qe^{\alpha^0}_{t^0r^0} \tag{2.12}$$

and the Lagrange multipliers $\mathcal{J}_{s^0q_t0}$ are chosen to satisfy the constraints

$$\langle\psi|e^{\alpha^0}_{r^0s^0}G_q\,e^{\alpha^0}_{t^0r^0}|\psi\rangle = a_{s^0q_t0}\,. \tag{2.13}$$

The summation over s^0qt^0 in (2.12) is, in general, restricted, since the spanning set (2.11) may contain linear dependencies. The effective Hamiltonian is the desired result and may be converted to a Hamiltonian in spin space as described elsewhere.[4,11]

The relation between the constraints (2.13) and (2.3) is given by

$$\begin{aligned} a_{s^0q_t0} &= \langle\psi|e^{\alpha^0}_{r^0s^0}G_q\,e^{\alpha^0}_{t^0r^0}|\psi\rangle \\ &= \sum_{\alpha\rho\sigma}[G_q]^{\alpha}_{(\rho\alpha^0s^0\chi\sigma\alpha^0t^0)}\langle\psi|e^{\alpha}_{\rho\sigma}|\psi\rangle \\ &= \sum_{\alpha\rho}[G_q]^{\alpha}_{(\rho\alpha^0s^0\chi\rho\alpha^0t^0)}\,a_\alpha\,, \end{aligned} \tag{2.14}$$

where the expansion of $e^{\alpha^0}_{r^0s^0}G_q\,e^{\alpha^0}_{t^0r^0}$ in terms of the $e^{\alpha}_{\rho\sigma}$ as used here is given in Ref. 11. In the special case that the a_α symmetry constraints are chosen so that the weights of symmetry components of the multistate ket are the same as those in a simple product of strongly orthogonal site kets [choice (i) of Ref. 9], we obtain

$$\begin{aligned} a_{s^0q_t0} &= \frac{g^0}{N!\,f^{\alpha^0}}\sum_{\alpha\rho}[G_q]^{\alpha}_{(\rho\alpha^0s^0\chi\rho\alpha^0t^0)}f^{\alpha} \\ &= \delta_{s^0,t^0}\,\delta_{q,0}\,, \end{aligned} \tag{2.15}$$

where $q=0$ denotes the identity DC. As pointed out in the discussion on DC's, these constraints (2.15) will be very nearly fulfilled for localized kets, especially for the higher DC generators. Thus we see that to a good approximation one may neglect the higher DC constraints, so that even for infinite systems we may obtain only a finite number of different types of important DC's.

In analogy to the projected energies in the preceding paper,[9] we may define projected exchange parameters for our present effective Hamiltonian. These projected quantities may be computed via a direct generalization[4,11] of the original[1-3] Heisenberg spin-Hamiltonian derivations. These projected quantities take the form

$$\mathcal{J}^{\mathcal{P}}_{s^0q_t0} \equiv \langle\psi|e^{\alpha^0}_{r^0s^0}G_q\,e^{\alpha^0}_{t^0r^0}H|\psi\rangle \tag{2.16}$$

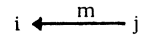
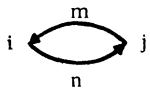


FIG. 1. Diagrams representing the terms of Eq. (3.17).



and can provide an additional check for approximate solutions, since for a good solution we expect $\mathcal{J}_{s^0q_t0}$ and $\mathcal{J}^{\mathcal{P}}_{s^0q_t0}$ to compare closely.

III. MULTISTATE-KET CLUSTER EXPANSION

In this section we describe an *Ansatz* for a multistate ket in the special case of a lattice of sites with only one orbital per site and for a separated-atom limit with one electron per site. Additional assumptions, such as a linear lattice, nearest-neighbor pairing, etc., will be introduced when convenient, although computations can be made by these same methods for less restrictive assumptions. The cluster-expansion form used here for the multistate ket is similar to those already described elsewhere[12] for antisymmetrized single-state kets.

The (unnormalized) multistate ket is taken to be of the form

$$\begin{aligned} &|\psi\rangle \equiv e^{S^+}|0\rangle, \\ &|0\rangle \equiv \varphi_1(1)\otimes\varphi_2(2)\otimes\varphi_3(3)\otimes\cdots\otimes\varphi_N(N), \\ &S^+ \equiv \sum_i S^+_i + \sum_{i<j} S^+_{ij} + \cdots, \end{aligned} \tag{3.1}$$

with S^+_i, S^+_{ij}, etc., being operators which create one-particle, two-particle, etc., excitations from the states with $\varphi_i(i)$, $\varphi_i(i)\otimes\varphi_j(j)$, etc., occupied. The excitations for each particle involved are to be orthogonal to the orbitals from which they arose. Consequently, all the excitation operators commute with one another:

$$[S^+_{i_1i_2\cdots i_m}, S^+_{j_1j_2\cdots j_n}] = 0; \tag{3.2}$$

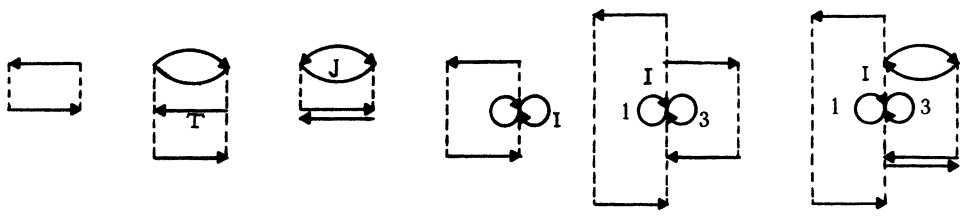


FIG. 2. Diagrams arising in the computation of the average energy $\langle\psi|H|\psi\rangle/\langle\psi|\psi\rangle$.

and also if a pair have any excitation sites in common, their product is zero:

$$S^+_{i_1 i_2 \cdots i_m} S^+_{j_1 j_2 \cdots j_n} = 0,$$
$$\{i_1, i_2, \ldots, i_m\} \cap \{j_1, j_2, \ldots, j_n\} \neq 0. \tag{3.3}$$

For our example we consider only very simple one- and two-particle excitations for a linear chain lattice

$$S^+_i \equiv x(|\varphi_{i+1}(i)\rangle\langle\varphi_i(i)| + |\varphi_{i-1}(i)\rangle\langle\varphi_i(i)|),$$
$$S^+_{ij} \equiv y\delta_{i,j\pm1}|\varphi_i(j)\otimes\varphi_j(i)\rangle\langle\varphi_i(i)\otimes\varphi_j(j)|, \tag{3.4}$$
$$S^+_{i_1 i_2 \cdots i_m} \equiv 0, \quad m \geq 3$$

where x and y are variational parameters.

In the present case each DC consists of a single permutation of $\mathcal{S}_N$, and the constraints of (2.13) and (2.15) become

$$\langle\psi|P|\psi\rangle = \langle\psi|\psi\rangle\delta_{P,1}, \quad P \in \mathcal{S}_N. \tag{3.5}$$

Because of the cyclic symmetry in $|\psi\rangle$ implied by (3.1) and (3.4) we see that the constraints for the nearest-neighbor transpositions $(i\ i+1)$, $i=1$ to N, are all satisfied if any one is satisfied. Hence all their Lagrange multipliers may be taken to be equal. Considering only the constraints for these nearest-neighbor transpositions, we are then to minimize

$$E^0 \equiv \langle\psi|H - \mathcal{J}\sum_{i=1}^{N}(ii+1)|\psi\rangle/\langle\psi|\psi\rangle, \tag{3.6}$$

subject to the single constraint

$$\langle\psi|(ii+1)|\psi\rangle = 0. \tag{3.7}$$

Matrix elements over higher-order permutations and transpositions between next-nearest neighbors may be nonzero, although we expect their magnitude to be appropriately small.

The properties (3.2) and (3.3) are especially useful in the factorization of matrix elements. To see this factorization, and the resultant cancellations, we first treat the overlap matrix element

$$\langle 0|e^S e^{S^+}|0\rangle = \langle 0|(1+S_i+S_{i\,i+1}+S_{i\,i-1})e^{S_{(i)}}e^{S^+_{(i)}}(1+S^+_i+S^+_{i\,i+1}+S^+_{i\,i-1})|0\rangle$$
$$= \langle 0|\{(1+S_i)e^{S_{(i)}}e^{S^+_{(i)}}(1+S^+_i)+S_{i\,i+1}e^{S_{(i\,i+1)}}e^{S^+_{(i\,i+1)}}S^+_{i\,i+1}+S_{i\,i-1}e^{S_{(i\,i-1)}}e^{S^+_{(i\,i-1)}}S^+_{i\,i-1}\}|0\rangle, \tag{3.8}$$

where we have defined

$$S^+_{(i_1 i_2 \cdots i_m)} \equiv \sum_{j}^{\neq i_1 i_2 \cdots i_m} S^+_j + \sum_{j<k}^{\neq i_1 i_2 \cdots i_m} S^+_{jk}. \tag{3.9}$$

Further, defining

$$A_{N-m} \equiv \langle 0|e^{S_{(12\ldots m)}}e^{S^+_{(12\ldots m)}}|0\rangle,$$
$$\xi \equiv 1 + \langle 0|S_1 S^+_1|0\rangle, \tag{3.10}$$
$$\zeta \equiv \langle 0|S_{12}S^+_{12}|0\rangle,$$

we see that (3.8) becomes

$$\langle 0|e^S e^{S^+}|0\rangle = \langle 0|1+S_i S^+_i|0\rangle A_{N-1}$$
$$+\langle 0|S_{i\,i-1}S^+_{i\,i-1}+S_{i\,i+1}S^+_{i\,i+1}|0\rangle A_{N-2}$$
$$= \xi A_{N-1} + 2\zeta A_{N-2}. \tag{3.11}$$

Similarly, we find that

$$A_M = \xi A_{M-1} + \zeta A_{M-2}, \quad M \leq N-1. \tag{3.12}$$

Thus

$$A_M/A_{M-1} = \xi + \zeta A_{M-2}/A_{M-1}$$
$$= \xi + \zeta\{\xi + \zeta\{\xi + \zeta\{\cdots\}^{-1}\}^{-1}\}^{-1}. \tag{3.13}$$

This continued fraction becomes independent of M for sufficiently large M,

$$f(\xi,\zeta) \equiv \lim_{M\to\infty}\frac{A_M}{A_{M-1}} = \xi + \frac{\zeta}{f(\xi,\zeta)}, \quad M \leq N-1. \tag{3.14}$$

Solving this quadratic equation for f yields

$$f(\xi,\zeta) = \tfrac{1}{2}\xi + [(\tfrac{1}{2}\xi)^2 + \zeta]^{1/2}, \tag{3.15}$$

where the positive root of the quadratic solution has been chosen because of its correct behavior as $S^+_i,\ S^+_{ij} \to 0$. We now have

$$\frac{A_{N-1}}{\langle\psi|\psi\rangle} = \left\{\xi + 2\zeta\frac{A_{N-2}}{A_{N-1}}\right\}^{-1} = \frac{f}{\xi f + 2\zeta},$$
$$\frac{A_{N-M}}{\langle\psi|\psi\rangle} = \frac{A_{N-1}}{\langle\psi|\psi\rangle}\prod_{j=2}^{m}\frac{A_{N-j}}{A_{N-j+1}} = \frac{f^{2-m}}{\xi f + 2\zeta}. \tag{3.16}$$

To further aid in the evaluation of the pertinent matrix elements we find a diagrammatic representation convenient. Simple one- and two-arrow diagrams of Fig. 1 are defined to represent the operators

$$|\varphi_i(m)\rangle\langle\varphi_j(m)|,$$
$$|\varphi_i(m)\otimes\varphi_j(N)\rangle\langle\varphi_j(m)\otimes\varphi_i(n)|. \tag{3.17}$$

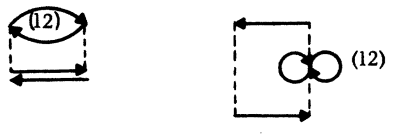


FIG. 3. Diagrams arising in the computation of the constraint matrix element $\langle\psi|(12)|\psi\rangle/\langle\psi|\psi\rangle$.

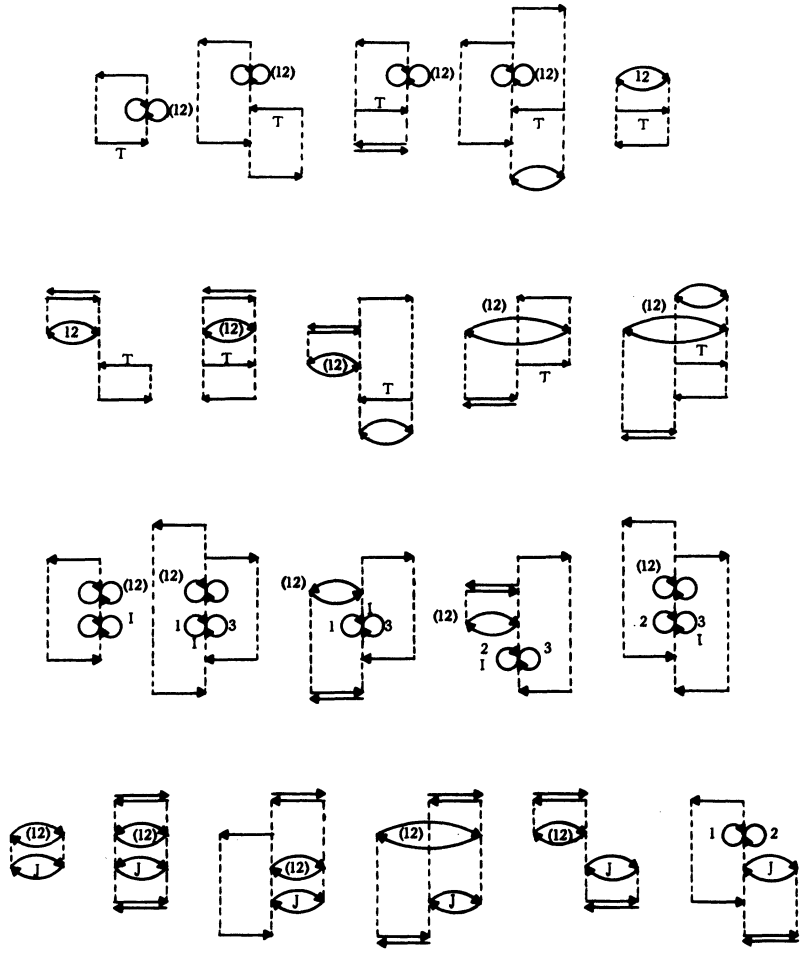


FIG. 4. Additional diagrams arising in the computation of the projected "exchange" parameter $\mathcal{J}^{\mathcal{P}}$.

TABLE I. Exchange parameters for $J=0$.

T/I	$-2T^2/I^2$	$\mathcal{J}/I$	$\mathcal{J}^{\mathcal{P}}/I$	x	E_0/I
0	0	0	0		
0.01	−0.0002	−0.000 199 940	−0.000 199 86	−0.009 995 0	−0.000 199 91
0.05	−0.005	−0.004 962 905	−0.004 915	−0.049 388 7	−0.004 944 94
0.1	−0.02	−0.019 425 74	−0.018 75	−0.095 407 0	−0.019 170 6
0.2	−0.08	−0.071 914 6	−0.064 703	−0.170 176	−0.069 141 2
0.3	−0.18	−0.145 933 2	−0.123 215	−0.222 139	−0.137 053
0.4	−0.36	−0.232 269 3	−0.187 044	−0.257 715	−0.214 366
0.5	−0.5	−0.325 589 2	−0.253 392	−0.282 755	−0.296 734
0.6	−0.72	−0.423 043 8	−0.321 107	−0.301 037	−0.382 006
0.7	−0.98	−0.523 083 6	−0.389 641	−0.314 854	−0.469 061
0.8	−1.28	−0.624 824 1	−0.458 708	−0.325 612	−0.557 277
0.9	−1.62	−0.727 734	−0.528 139	−0.334 201	−0.646 285
1.0	−2.0	−0.831 478 4	−0.597 832	−0.341 204	−0.735 855
1.5	−4.50	−1.357 001 3	−0.984 448	−0.362 847	−1.188 27
2.0	−8.0	−1.887 683 2	−1.300 729	−0.373 979	−1.644 12

Horizontal positions label sites, while arrows identify electrons with the initial and final states indicated by the tails and heads of the arrows. Products of operators are indicated by superpositions of the corresponding arrow diagrams, with the usual convention that arrows at higher vertical positions correspond to operators at positions further to the left in the product. For a matrix element $\langle\psi|\mathcal{O}|\psi\rangle$, with $\mathcal{O}$ a product of arrow operators, to be nonzero, we see that the diagrams corresponding to $e^{S}\mathcal{O}e^{S^{+}}$ must return every electron to its original site in $|0\rangle$. Thus every site must have arrows of type i coming in as many times as they leave; in addition, those of type i coming in and out of a given site must alternate, and S (and S^{+}) can contribute only one arrow coming into (and going out of) any site. Finally, if in constructing a diagram a point is reached in which the arrows of $\mathcal{O}$ have already been employed and all the electrons transferred have returned to their original sites, then the result of all the remaining arrows is merely the residual overlap not involving any of those sites already included in the diagram; these residual overlaps may then be evaluated by the results of (3.17).

IV. MODEL HAMILTONIAN EXAMPLE

To illustrate the general theory of Sec. II we consider the special case of a linear chain of sites with only one orbital per site. Further, we assume a simple model Hamiltonian,

$$H \equiv T\sum_{mn}\left(|\varphi_n(m)\rangle\langle\varphi_{n+1}(m)| + |\varphi_{n+1}(m)\rangle\langle\varphi_n(m)|\right) + J\sum_{lmn}|\varphi_n(l)\otimes\varphi_{n+1}(m)\rangle\langle\varphi_{n+1}(l)\otimes\varphi_n(m)| + I\sum_{lmn}|\varphi_n(l)\otimes\varphi_n(m)\rangle\langle\varphi_n(l)\otimes\varphi_n(m)|, \tag{4.1}$$

which is defined on the space of linear combinations of products of the orthonormal φ orbitals. Here T, J, and I may be interpreted as charge transfer, exchange, and intrasite Coulomb repulsion integrals. Although this is a very simple model Hamiltonian it is generally regarded[5,6] as exhibiting both "direct" and "kinetic" exchange processes associated with J and T, respectively. When $T=0$ and there is one electron per site, this model is the nearest-neighbor linear Heisenberg model. When $J=0$, this model is the nearest-neighbor linear Hubbard model.

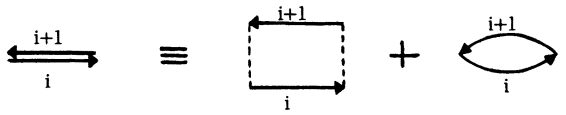


FIG. 5. A useful abbreviation.

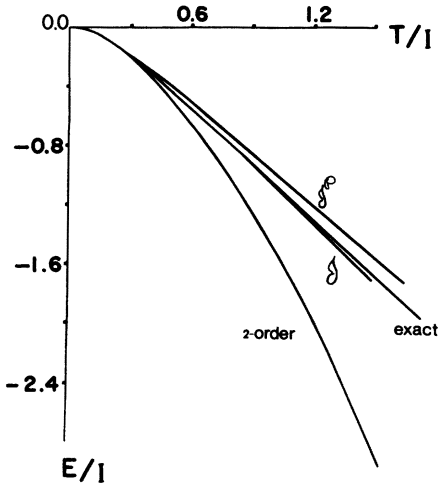


FIG. 6. Ground-state energies of the effective Heisenberg models as compared to the exact ground-state energy of the Hubbard model ($J=0$).

If the simplest primitive ket $|0\rangle$ is used in a direct construction of an effective Hamiltonian, one obtains[1-3] the nearest-neighbor Heisenberg exchange Hamiltonian

$$-J\sum_{i=1}^{N}(ii+1), \tag{4.2}$$

where $(ii+1)$ is a nearest-neighbor transposition in spin-free space. The more accurate second-order perturbation result (with the T and J terms of (4.1) of first and second orders) yields[5,6] a nearest-neighbor Heisenberg exchange Hamil-

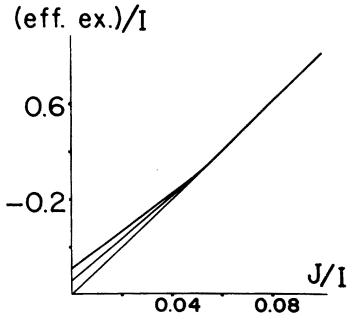


FIG. 7. Effective exchange constants for the effective Heisenberg model. The upper curve is obtained using the projected $\mathcal{J}^{P}$, the middle curve using the Lagrange multiplier $\mathcal{J}$, and the lower linear curve using the second-order perturbation result $J-2T^{2}/I$.

tonian

$$-\frac{2NT^2}{I}+\left(J-\frac{2T^2}{I}\right)\sum_{i=1}^{N}(ii+1). \tag{4.3}$$

Evidently, for $2T^2/I$ comparable to or larger than J, the simple single primitive-ket result of (4.2) is in serious error, so that these cases in which $2T^2/I$ is important shall be of special interest for our more accurate multistate-ket treatment.

We use a multistate ket of the form (3.1) with one- and two-particle excitations as in (3.4). The diagrammatic method of evaluating matrix elements is employed. In Figs. 2, 3, and 4 we list those diagrams encountered in evaluating the matrix elements over H, $(ii+1)$, and $(ii+1)H$, respectively. In these figures we have not labeled the sites but have assumed they are sites 1, 2,... in proceeding from left to right; further we have suppressed the arrow labels when they are evident. We have also used the abbreviation of Fig. 5. The matrix elements for H, $(ii+1)$, and $(ii+1)H$ are

$$\frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}=\frac{N}{\xi f+2\zeta}\left\{4x\,(f+y)\,T+2(x^2+y)J+x^2I\left(2+\frac{x^2+2x^4+2y^2+4x^2y}{f}\right)\right\}$$

$$\frac{\langle\psi|(ii+1)|\psi\rangle}{\langle\psi|\psi\rangle}=2\frac{y+2x^2}{\xi f+2\zeta}$$

$$\mathcal{J}^{\Phi}=\frac{\langle\psi|(ii+1)H|\psi\rangle}{\langle\psi|\psi\rangle}=\frac{1}{\xi f+2\zeta}\left\{2xT\left(2+5x^2+4y+\frac{3x^2y+2y^2}{f}\right)+J\left(1+x^4+2x^2y+y^2+\frac{8x^4+12x^2y+4y^2}{f}\right)\right.$$
$$\left.+2x^2I\left(1+\frac{4x^2+2y}{f}\right)\right\}.$$

Minimization of E^0 as in (3.6) gives, for the $J=0$ case, the results of Table I. For small T/I the Lagrange multiplier $\mathcal{J}$, the projected $\mathcal{J}^{\Phi}$, and the second-order perturbation result for the effective Heisenberg exchange interaction are all seen to be similar in value. For larger T/I the two multistate-ket results are further away from the perturbation result which is expected to be less valid. We thus have three different nearest-neighbor Heisenberg models. The exact ground-state energies[13] for these three Heisenberg models are compared in Fig. 6 with the exact ground-state energy[14] for the Hubbard model. As expected, both the multistate-ket variational results are seen to provide better estimates than the second-order perturbation result, particularly for larger T/I.

The variational multistate formulation has also been applied to the model Hamiltonian in the case where the "direct" exchange J is nonzero. Choosing $J+2T^2/I$ to be a constant value $\frac{1}{10}I$, we obtain the results of Fig. 7. The multistate-ket procedure again gives results similar to, and probably better than, the second-order perturbation results.

V. DISCUSSION AND CONCLUSION

The general derivation of the Heisenberg spin Hamiltonian appears to be the first of a variational nature. We have found several advantages over the more common perturbative derivation. The effective exchange parameters for the variational *Ansatz* of Secs. III and IV seemingly gave better results, at least for the ground state, than a second-order perturbation result. Indeed, this cluster-expansion variational *Ansatz* does not seem to be limited to particularly small values for the intersite perturbation.

There still are some shortcomings in the present variational treatment. Not all constraints beyond those for nearest-neighbor transpositions were explicitly treated. Expectation values of the approximate multistate ket over permutations, such as $(ii+2)$, $(ii+1\,i+2)$, $(ii+2\,i+3\,i+1)$, and $(ii+2)\,(i+3\,i+5)$, may give nonzero results, although of comparatively small order: 4th, 6th, and 8th in the four examples above. However, zero expectation values are obtained over permutations which transfer an electron more than two sites or which has a disjoint transposition between nearest neighbors. To take into account higher permutations (or DC's) a more complicated *Ansatz* for the multistate ket would be required. In principle this could easily be improved by including, for example, single and pair excitations between next-nearest-neighbor pairs; however, although some more constraints could be handled exactly, not all of the very-high-order ones could. In practice such more general variational *Ansätze* would lead to an even greater number of diagrams. An approximate alternative would be to neglect all diagrams of sufficiently high order; omission of those of 8th and higher order, say, would eliminate even a number of those already given in Fig. 3.

In the treatment of more general Hamiltonians than those of Secs. III and IV similar *Ansätze* still apply. The discussion of Sec. II already describes how the constraints may be treated.

*Supported by the National Science Foundation, Washington, D.C. and the Robert A. Welch Foundation, Houston, Tex.

†Present address: Department of Chemistry, Louisiana State University at New Orleans, New Orleans, La.

[1]W. Heisenberg, Z. Phy. 49, 619 (1928).

[2]P. A. M. Dirac, Proc. Roy. Soc. Lond. A123, 714 (1929); J. H. Van Vleck, Phys. Rev. 45, 405 (1934).

[3]R. A. Serber, Phys. Rev. 45, 461 (1934).

[4]C. Herring, Rev. Mod. Phys. 34, 631 (1962); C. Herring, in *Magnetism*, edited by G. T. Rado and H. Suhl (Academic, New York, 1963), Vol. 2B, p. 1.

[5]P. W. Anderson, Phys. Rev. 115, 2 (1959); P. W. Anderson, Solid State Phys. 14, 99 (1963).

[6]F. Keffer and T. Oguchi, Phys. Rev. 115, 1428 (1959); T. Arai, Phys. Rev. 134, A824 (1964); T. R. Waite, J. Chem. Phys. 33, 256 (1960); L. N. Buleavski, Zh. Eksp. Teor. Fiz. 51, 230 (1966) Sov. Phys.-JETP 24, 154 (1967)]; J. H. Van Vleck, *Quantum Theory of Atoms, Molecules, and the Solid State,* edited by P. O. Löwdin (Academic, New York, 1966), p. 475; D. J. Klein and W. A. Seitz, Phys. Rev. (to be published).

[7]J. Yamishita and J. Kondo, Phys. Rev. 109, 730 (1958); N. L. Huang and R. Orbach, Phys. Rev. Lett. 12, 225 (1964); K.-I. Gondiara and Y. Tanabe, J. Phys. Soc. Jap. 21, 1527 (1966); N. L. Huang and R. L. Orbach, Phys. Rev. 154, 487 (1967); J. Hubbard, D. E. Rimmer, and F. R. A. Hopgood, Proc. Phys. Soc. Lond. 88, 13 (1966); E. Lombardi, J. Jansen, and R. Ritter, Chem. Phys. Lett. 12, 529 (1972).

[8]L. F. Mattheiss, Phys. Rev. 123, 1209 (1961); R. K. Nesbet, Phys. Rev. 122, 1407 (1961); Phys. Rev. 135, A460 (1964); A. J. Freeman and R. F. Watson, Phys. Rev. 124, 1439 (1961); D. J. Klein, J. Chem. Phys. 51, 5151 (1969); D. C. Foyt, R. W. Kramling, and F. A. Matsen, J. Phys. Chem. 75, 1882 (1971).

[9]D. C. Foyt and D. J. Klein, preceding paper, Phys. Rev. A 8, 2274 (1973).

[10]P. Kramer and T. H. Seligman, Nucl. Phys. A 136, 545 (1969); Nucl. Phys. A 186, 49 (1972); B. R. Junker and D. J. Klein, J. Chem. Phys. 55, 5533 (1971).

[11]F. A. Matsen and D. J. Klein, J. Phys. Chem. 75, 1860 (1971); F. A. Matsen, D. J. Klein, and D. C. Foyt, J. Phys. Chem. 75, 1867 (1971).

[12]O. Sinanoglu, J. Chem. Phys. 36, 706, 3198 (1962); O. Sinanoglu, in *Advances in Chemical Physics,* edited by I. Prigogine (Academic, New York, 1964), Vol. 6, p. 315; R. Brout, Phys. Rev. 111, 1324 (1958); H. Primas, in *Modern Quantum Chemistry,* edited by O. Sinanoglu (Academic, New York, 1965), Vol. 2, p. 45; J. da Provedencia, Nucl. Phys. 44, 572 (1963); J. Cizek, J. Chem. Phys. 45, 4256 (1966).

[13]H. Bethe, Z. Phys. 71, 205 (1931); L. Hulthén, Ark. Mat. Astron. Fys. 26A, 1 (1938).

[14]E. H. Lieb and F. Y. Wu, Phys. Rev. Lett. 20, 1445 (1968).