PHYSICAL REVIEW B VOLUME 4, NUMBER 11 1 DECEMBER 1971

Magnetic Transitions in a Degenerate Band*

K. A. Chao†
IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598

(Received 27 May 1971)

The variational scheme of Gutzwiller has been reexamined so as to apply it to the case of degenerate narrow bands. The ground-state wave function for electrons is investigated for an arbitrary density of electrons and an arbitrary strength of interaction. The typical features of a metal, particularly the Fermi surface, are preserved in this approach as opposed to some other methods. The Coulomb repulsion among electrons scatters the electrons from below to above the Fermi surface of uncorrelated bands. Under the assumption that only the intra-atomic Coulomb interaction is important, the probability of electrons being scattered is larger for the paramagnetic state than for the ferromagnetic state. Therefore the ferromagnetic ground state is favored if the density of states is large at the band edges. The degenerate-band-model Hamiltonian is spin dependent, hence the results obtained from it do not have to obey the theorem of Lieb and Mattis.

I. INTRODUCTION

In 1936, Slater[1] concluded from analyzing the properties of Ni that the origin of ferromagnetism is in the intra-atomic interaction. The criticism of Slater's theory by many authors that electron correlation was not properly included in his Hartree-Fock calculation has stimulated much work trying to construct a theory for the correlation effect in narrow bands.

In accord with other theories for transition metals,[2] Van Vleck[3] has described the main features of a ground-state wave function which explicitly takes into account this correlation effect in a d band. His model of minimum polarity[4] excludes the states corresponding to higher degrees of ionization.

The idea of minimum polarity was taken up later and explored in some detail with the help of modern many-body theory by Gutzwiller,[5] Hubbard,[6] Kanamori,[7] and others. Different approaches were used to work on the Hamiltonian consisting of a tight-binding band and very short-range interactions among electrons. They obtained qualitatively different results concerning the occupation probability for electrons in reciprocal space and the comparison of the energy for a ferromagnetic and for a paramagnetic ground state.

Recently Brinkman and Rice[8] used Gutzwiller's variational method to predict a Mott transition[9] in a narrow s band. Hubbard's Green's-function decoupling approximation also indicated an insulator-conductor transition due to band splitting.[10] However the split subband does not contain exactly one electron per atom. The variational method yields a Fermi surface which contains the correct volume, and predicts the magnetic properties of an almost half-filled s band in good agreement with Nagaoka's[11] exact solution. Nevertheless, in the s-band model the magnetic properties have to be predicted without considering Hund's rule. The purpose of the present paper is to apply the variational scheme of Gutzwiller to investigate how and to what extent the electron correlation causes the transitions between the paramagnetic and the ferromagnetic states in a doubly degenerate band, which is more realistic for transition metals.

Section II states the problem and investigates the kinematics of an uncorrelated many-electron system. Pseudo-wave-functions are introduced to derive the conditions which are used in Sec. III to define the correlated ground-state wave function. This trial function contains some parameters which specify the number of lattice sites which are occupied by more than one electron. Those parameters are varied in our calculation to minimize the expectation value of the ground-state energy. In Sec. IV we introduce the quasichemical approximation to compute the various density functions. The results obtained in this section are good for all electron densities. In Sec. V we derive the condition, in terms of correlation energy and band structures, under which ferromagnetism will occur. We have restricted ourselves in Sec. V to the cases of less than one or more than three electrons per atom in order to make it mathematically manageable. In Sec. VI we point out that our model is outside the scope of applicability of the theorem of Lieb and Mattis.[12]

II. KINEMATICS OF ELECTRONS IN A DEGENERATE BAND

The atomic d level splits into two sublevels in a crystal field with cubic symmetry. One sublevel is triply degenerate and is denoted by t_{2g}. The other sublevel is doubly degenerate and transforms as $2z^2 - x^2 - y^2$ and $x^2 - y^2$ under symmetry operations. This is usually called e_g. In a simplified picture, the d band thus consists of a triply degenerate subband of t_{2g} symmetry and a doubly degenerate subband of e_g symmetry. In this paper we only consider those transition metals for which the t_{2g} subband lies below the e_g subband. The lower subband is completely filled and can therefore be ignored. The upper subband will be approximated by two tight-binding degenerate bands which have the e_g symmetry.

Consider N electrons in a simple lattice of L sites. The coordinates of the lattice sites are labeled by small italic letters g or h (not to be confused with the subscript in e_g). Four localized (Wannier) orbitals $\phi_{i\sigma}(x-g)$ of e_g symmetry are attached to each site g. The index $i=1$ or 2 designates the orbital character and σ is the spin quantum number.

Two tight-binding degenerate bands are constructed from these localized orbitals. The Bloch state of wave vector k and band index l is represented by $\psi_{kl\sigma}(x)$, where $l=\text{I}$ or II. In second quantized form, we have

$$a^\dagger_{kl\sigma}=L^{-1/2}\sum_{j,g} C_{lj}(k)e^{ikg}a^\dagger_{gj\sigma} \,. \tag{1}$$

The coefficients $C_{lj}(k)$ give the required mixture of e_g orbitals to yield the Bloch states.

Assume at first that there are $N(\sigma)$ uncorrelated electrons with spin σ, and $m(l,\sigma)$ of them are in the l band. Let $K(l,\sigma)$ be the occupied region in the l band with σ spin. These regions may all differ from one another. Their boundaries make up the Fermi surface. The uncorrelated ground-state wave function Ψ is

$$\Psi=\prod_{l,\sigma}\prod_{k} a^\dagger_{kl\sigma}\Psi_0 \,, \tag{2}$$

where $k\in K(l,\sigma)$ and Ψ_0 is the vacuum state.

The wave function Ψ describes completely delocalized and uncorrelated electrons. In order to describe the effect of correlation, the concept of a "configuration" of electrons will be used frequently. Instead of distributing the electrons over states in reciprocal space as in Ψ, consider a set $G(i,\sigma)$ of $M(i,\sigma)$ lattice sites. At each site g of $G(i,\sigma)$ insert an electron into the localized state $\phi_{i\sigma}(x-g)$. Thus, we get an uncorrelated configuration of electrons, i.e., a wave function given by

$$\Phi(\bar{G})=\prod_{i,\sigma}\prod_{g} a^\dagger_{gi\sigma}\Psi_0 \,, \tag{3}$$

where $\bar{G}$ is the collection of four $G(i\sigma)$ with $i=1$, 2 and $\sigma=\uparrow,\downarrow$. The index g runs through the set $G(i\sigma)$ for each combination of σ and i.

If (1) is inserted into (2), the wave function Ψ of delocalized and uncorrelated electrons is decomposed into uncorrelated configurations Φ. In order to write the resulting formula, the sets $K(l,\sigma)$ have to be further analyzed. Let $K(\sigma)$ be the region in reciprocal space which is common to both $K(\text{I},\sigma)$ and $K(\text{II},\sigma)$. Half of the $2m(\sigma)$ electrons in the common region $K(\sigma)$ are found in localized states of type 1σ, the other half in localized states of type 2σ. There is no such simple rule for the electrons of $K(\text{I},\sigma)$ and $K(\text{II},\sigma)$ outside $K(\sigma)$. Therefore subsets $K(li\sigma)$ have to be introduced which divide the elements of $K(l,\sigma)$ outside $K(\sigma)$ into two classes. Electrons in $K(li\sigma)$ belong to the subband $l\sigma$, but are found in the localized states of type $i\sigma$. There are obviously many ways of choosing the subsets $K(li\sigma)$ once the $K(l,\sigma)$ are known. Ψ can now be written as a double sum:

$$\Psi=\sum_{\bar{K}} D(\bar{K})\sum_{\bar{G}} A(\bar{K},\bar{G})\Phi(\bar{G}) \,, \tag{4}$$

where $\bar{K}$ represents a particular choice of the subsets $K(li\sigma)$. The sum over $\bar{K}$ runs through all the possible choices of $K(li\sigma)$ and the sum over $\bar{G}$ exhausts all the configurations.

$D(\bar{K})$ depends only on the band structure through the coefficients $C_{lj}(k)$. On the other hand, $A(\bar{K},\bar{G})$ are independent of the topology of the band but are related to the geometry of the configuration $\bar{G}$. To illustrate this, let us introduce the operators

$$\alpha^\dagger_{kj\sigma}=L^{-1/2}\sum_{g} e^{ikg}a^\dagger_{gj\sigma} \,. \tag{5}$$

They generate two artificial uncoupled bands for which $j=1$ and 2. When coupled by $C_{lj}(k)$ they reproduce the bands labeled by $l=\text{I}$ and II, as given by (1):

$$a^\dagger_{kl\sigma}=\sum_{j} C_{lj}(k)\alpha^\dagger_{kj\sigma} \,. \tag{6}$$

If we insert (5) and (6) into (2), it is easy to see that the product of $C_{lj}(k)$ and the product of e^{ikg} are separated. The former is represented by $D(\bar{K})$, and the latter gives the $A(\bar{K},\bar{G})$. The exact details concerning the quantities $D(\bar{K})$ and $A(\bar{K},\bar{G})$ are found in Appendix A.

Before introducing the correlation of electrons, we would like to investigate (4) more thoroughly. This will be vital for our later analysis. We first define the pseudodelocalized wave functions. They are constructed from the Bloch states of the uncoupled bands as

$$X(\bar{J})=\prod_{j,\sigma}\prod_{k} \alpha^\dagger_{kj\sigma}\Psi_0 \,, \tag{7}$$

where k runs through the set $J(j,\sigma)$ for fixed values of j and σ. These sets $J(j,\sigma)$ are defined as follows: We first pick a particular $K(li\sigma)$, and then let $J(j,\sigma)$ be the union of the three subsets $K(\sigma)$, $K(\text{I},j,\sigma)$, and $K(\text{II},j,\sigma)$. This relation is diagrammatically shown in Fig. 1. Note that $J(j,\sigma)$ are defined as the sets which have their images in $K(\text{I},\sigma)\cup K(\text{II},\sigma)$. To a given $J(j,\sigma)$, corresponds a particular choice of $K(li\sigma)$. As the definition of all barred quantities, $\bar{J}$ represents the combination

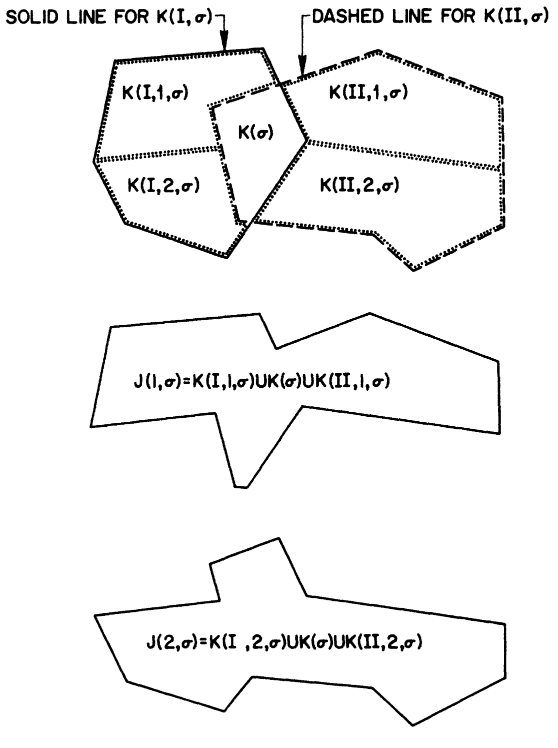

FIG. 1. The occupied Bloch states in $K(\mathrm{I}\sigma)$ and $K(\mathrm{II}\sigma)$ are divided into $J(1\sigma)$ and $J(2\sigma)$. The electrons originally occupying the states in $J(i\sigma)$ are localized in the i-type atomic orbitals.

of four $J(j, \sigma)$ where $j=1$ and 2, $\sigma = \uparrow$ and $\downarrow$.

The definition of $J(j, \sigma)$ is then generalized to define the sets $\tau(j, \sigma)$ of the same type which do not necessarily have images in $K(\mathrm{I}, \sigma) \cup K(\mathrm{II}, \sigma)$. The elements of every combination $\bar{\tau}$ of four $\tau(j, \sigma)$ label N Bloch states in the two uncoupled bands (5). There are $\binom{2L}{N(\uparrow)}\binom{2L}{N(\downarrow)}$ such sets with associated pseudo-wave-functions

$$X(\bar{\tau}) = \prod_{j,\sigma} \prod_{k \in \tau(j\sigma)} \alpha^{\dagger}_{kj\sigma} \Psi_0 \,. \tag{8}$$

Inserting (5) into (8), it becomes

$$X(\bar{\tau}) = \sum_{\bar{G}} A(\bar{\tau}, \bar{G}) \Phi(\bar{G}) \,, \tag{9}$$

and (4) can be rewritten as

$$\Psi = \sum_{\bar{J}} D(\bar{J}) X(\bar{J}) \,. \tag{10}$$

The coefficients $A(\bar{\tau}, \bar{G})$ are the matrix elements of a unitary transformation connecting the two bases of a $\binom{2L}{N(\uparrow)}\binom{2L}{N(\downarrow)}$-dimensional vector space. They satisfy the conditions

$$\sum_{\bar{G}} A^*(\bar{\tau}, \bar{G}) A(\bar{\tau}', \bar{G}) = \delta_{\bar{\tau}, \bar{\tau}'} \,, \tag{11}$$

$$\sum_{\bar{\tau}} A^*(\bar{\tau}, \bar{G}) A(\bar{\tau}, \bar{G}') = \delta_{\bar{G}, \bar{G}'} \,. \tag{12}$$

With the expression of $A(\bar{\tau}, \bar{G})$ given in Appendix A, (11) and (12) can be proved by direct computation. Equations (11) and (12) are the key conditions in our treatment of the electron correlation problem.

III. CORRELATED WAVE FUNCTIONS FOR A DEGENERATE BAND

Since $|A(\bar{K}, \bar{G})|^2$ in (4) measures the probability that the configuration $\bar{G}$ is being occupied by electrons, it will be altered when the correlation of electrons is taken into account. The main contribution to the correlation comes from the intra-atomic Coulomb repulsion. In order to simplify the notation, let us replace the double index $1\uparrow$ by 1, $1\downarrow$ by 2, $2\uparrow$ by 3, and $2\downarrow$ by 4. Then define ν_{ij} as the number of sites which are simultaneously occupied by electrons in orbitals of type i and j. Similarly ν_{ijm} and ν_{ijmn} are defined as the numbers of triply and quadruply occupied sites.

The main point of our approach to the correlation problem is to attach a weight to each $A(\bar{K}, \bar{G})$ according to the number and type of multiple occupancies in the corresponding configuration $\bar{G}$. We introduce the parameters η_{ij} $(0 \le \eta_{ij} \le 1)$, called the correlation strength not to be confused with the correlation energy, as the weight corresponding to a single site doubly occupied by electrons in states of types i and j. The correlation strength assigned to a triply occupied site is approximated by the product of three such parameters, each of which corresponds to one of the three couples of double occupancies. That is $\eta_{ijm} = \eta_{ij}\eta_{jm}\eta_{mi}$. The same approximation is used for the quadruply occupied sites. The symmetry property reduces six η_{ij} to three independent ones. We define them as follows: η_1 is the correlation strength between two electrons in different orbitals but with the same spin; η_2 is the correlation strength between two electrons in different orbitals and with different spins; and η_3 is the correlation strength between two electrons in the same orbital but with different spins.

The weight given to the coefficient $A(\bar{K}, \bar{G})$ by the intra-atomic Coulomb interaction is expressed in the form of $\Theta(\bar{\eta}, \bar{\nu}) = \prod_i \eta_i^{\mu_i}$, where $\bar{\eta}$ and $\bar{\nu}$ are the barred quantities conventionally defined in this paper as the combinations of the appropriate quantities η_i and ν_{ij}. μ_i is the total number of double occupations corresponding to η_i including the decomposed parts from the triply and quadruply occupied sites. If the correlation is so strong that the number of sites occupied by more than two electrons is negligibly small, we have

$$\Theta(\bar{\eta}, \bar{\nu}) = \eta_1^{\nu_{13}+\nu_{24}} \eta_2^{\nu_{14}+\nu_{23}} \eta_3^{\nu_{12}+\nu_{34}} \,. \tag{13}$$

The correlation effect in two configurations which have the same feature of multiple occupancies but are otherwise entirely different obviously cannot be described only by the intra-atomic interaction mentioned above. There is an interatomic correlation which affects the dynamics of the electron system. It depends on the over-all space density of electrons in a particular configuration. Therefore, an extra factor $B(\bar{K}, \bar{G})$ is introduced to correct both the amplitude and the phase of $A(\bar{K}, \bar{G})$. The trial function of the correlated ground state can then be written as

$$\Psi_c=\sum_{\bar{K}} D(\bar{K})\sum_{\bar{G}} A_c(\bar{K}, \bar{G})\Phi(\bar{G}) , \tag{14}$$

$$A_c(\bar{K}, \bar{G})=\Theta(\bar{\eta}, \bar{\nu})B(\bar{K}, \bar{G})A(\bar{K}, \bar{G}) . \tag{15}$$

Comparing (4), (9), (10), and (14), we can rewrite (14) as

$$\Psi_c=\sum_{\bar{J}} D(\bar{J})X_c(\bar{J}) , \tag{16}$$

$$X_c(\bar{J})=\sum_{\bar{G}} A_c(\bar{J}, \bar{G})\Phi(\bar{G}) . \tag{17}$$

Since it is impossible to precisely determine $A_c(\bar{K}, \bar{G})$, we impose conditions on $A_c(\bar{K}, \bar{G})$ in order to pick a physically acceptable one. When the correlation is turned on we like to retain (11) by proper choice of $B(\bar{K}, \bar{G})$. That is, $A_c(\bar{K}, \bar{G})$ satisfy

$$\sum_{\bar{G}} A_c^*(\bar{J}, \bar{G})A_c(\bar{J}', \bar{G})=\delta_{\bar{J},\bar{J}'} . \tag{18}$$

For convenience, let us rewrite (18) as

$$\sum_{\bar{G}} \Theta^2(\bar{\eta}, \bar{\nu})B^*(\bar{K}, \bar{G})A^*(\bar{K}, \bar{G}) \times B(\bar{K}', \bar{G})A(\bar{K}', \bar{G})\propto\delta_{\bar{K},\bar{K}'} . \tag{19}$$

Instead of an equals sign, a proportional sign is used in (19). This is because very frequently in our later calculation we will deal with quantities like $\sum_{\bar{G}}|B(\bar{K}, \bar{G})A(\bar{K}, \bar{G})|^2$. It will then be of great advantage to work with the unnormalized Ψ_c. It is easy to show that (19) is not unreasonable. Let the number of sets $\bar{K}$ be $\mathfrak{M}$. Clearly $\mathfrak{M}$ is less than $\binom{2L}{N(\uparrow)}\binom{2L}{N(\downarrow)}$. Since we are interested in using a variational scheme to obtain the ground-state wave function, we only need to work out the $\mathfrak{M}^2$ equations in (19). Nevertheless, there are $\binom{2L}{N(\uparrow)}\binom{2L}{N(\downarrow)}(\mathfrak{M})$ unknown variables $B(\bar{K}, \bar{G})$, largely exceeding the number of equations to be satisfied. With these extra degrees of freedom, we shall make certain additional assumptions concerning the coefficients $B(\bar{K}, \bar{G})$ which are suggested by (19).

The resulting approximation is called the quasichemical approximation (QCA), because the main idea is similar to the quasichemical approximation used in the theory of mixtures.[13] Instead of determining the exact form of $B(\bar{K}, \bar{G})$, we let

$$\begin{aligned}\sum_{\bar{G}}{}' B^*(\bar{K}, \bar{G})A^*(\bar{K}, \bar{G})B(\bar{K}', \bar{G})A(\bar{K}', \bar{G}) &=C(\bar{\nu})\sum_{\text{all }\bar{G}} A^*(\bar{K}, \bar{G})A(\bar{K}', \bar{G})\\ &=C(\bar{\nu})\delta_{\bar{K},\bar{K}'} ,\end{aligned} \tag{20}$$

where the primed sum over $\bar{G}$ runs through the configurations which have the same value of $\bar{\nu}$. $C(\bar{\nu})$ is the ratio of the number of such configurations to the total number of configurations. We have used (11) to get the Kronecker δ function in (20). The evaluation of $C(\bar{\nu})$ involves no more than counting configurations and this is done in Appendix B.

The reasonableness of the QCA for the s band was discussed in considerable detail by Gutzwiller in GIII.[5] The main points presented there are also valid for the d band. Nevertheless, in order to make the present paper self-contained, the validity of QCA will be investigated in the present context following Gutzwiller's argument.

First we note that (20) implies (18) as far as the Kronecker delta function is concerned. Hence we only need to study

$$\sum_{\bar{G}}{}'|B(\bar{K}, \bar{G})|^2|A(\bar{K}, \bar{G})|^2=C(\bar{\nu}) . \tag{21}$$

The approximate equation yields an exact solution in the limit of zero bandwidth. To prove it, using the expression of A in Appendix B, we have[14]

$$|A(\bar{K}, \bar{G})|^2=\prod_{i,\sigma}\left(\tilde{\omega}_{i\sigma}(g-g')\,\middle|\,\begin{matrix}G(i\sigma)\\ G(i\sigma)\end{matrix}\right). \tag{22}$$

The right-hand side is the product of four $i\sigma$ determinants, each corresponding to a specific choice of i and σ. The rows and columns of the $i\sigma$ determinant are given by the sites in set $G(i\sigma)$, and the elements of the determinant are defined by the function

$$\tilde{\omega}_{i\sigma}(g-g')=L^{-1}\textstyle\sum_k e^{ik(g-g')} , \tag{23}$$

where $k\in J(i\sigma)=K(\mathrm{I}, i, \sigma)\cup K(\mathrm{II}, i, \sigma)\cup K(\sigma)$.

For the completely degenerate case (zero bandwidth), the set $J(i\sigma)$ appearing in our variational ground-state wave function can be replaced by an ensemble average over all the possible sets $\tau(i\sigma)$ in reciprocal space. One finds with the help of a theorem concerning the determinants of products of rectangular matrices[15] that

$$\begin{aligned}\left\langle\left(\tilde{\omega}_{i\sigma}(g-g')\,\middle|\,\begin{matrix}G(i\sigma)\\ G(i\sigma)\end{matrix}\right)\right\rangle &\propto \sum_{\text{all }\tau(i\sigma)}\left(\tilde{\omega}_{i\sigma}(g-g')\,\middle|\,\begin{matrix}G(i\sigma)\\ G(i\sigma)\end{matrix}\right)\\ &=\left(L^{-1}\sum_{\text{all }k} e^{ik(g-g')}\,\middle|\,\begin{matrix}G(i\sigma)\\ G(i\sigma)\end{matrix}\right)=1 .\end{aligned} \tag{24}$$

(24) is true for arbitrary $G(i\sigma)$. If we simply pick $|B(\bar{K}, \bar{G})|^2=|B(\bar{K})|^2$ for the case of zero bandwidth, then

$$\sum_{G}{}'|B(\bar{K}, \bar{G})|^2\langle|A(\bar{K}, \bar{G})|^2\rangle=C(\bar{\nu})|B(\bar{K})|^2 . \tag{25}$$

This is just the same result as (21) if we let the phase factor be such that $|B(\bar{K})|^2=1$.

Since the electron correlation tends to narrow the band, following (25), one naturally expects that (21) is a good approximation. There is another supporting argument: the concept of the so-called exchange hole.[16] Let us write

$$|A(\overline{K},\,\overline{G})|^2=\prod_{i,\sigma}|A_{\mathrm{ex}}(J(i,\,\sigma),G(i,\,\sigma))|^2\,, \tag{26}$$

$$|A_{\mathrm{ex}}(J(i,\,\sigma),\;G(i,\,\sigma))|^2=\left(\tilde{\omega}_{i\sigma}(g-g')\left|\begin{matrix}G(i,\,\sigma)\\G(i,\,\sigma)\end{matrix}\right.\right)\,. \tag{27}$$

The quantity $|A_{\mathrm{ex}}(J(i,\,\sigma),\,G(i,\,\sigma))|^2$ describes the exchange effect between the $(i,\,\sigma)$ electrons. Its value is smaller the closer the electrons are to one another.

Since $B(\overline{K},\,\overline{G})$ is related only to the interatomic interaction between different configurations, for the convenience of our discussion let us pretend to have only this kind of correlation. The previous discussion indicates that the correct result for the zero bandwidth is obtained if we simply set $|B(\overline{K},\,\overline{G})\times A(\overline{K},\,\overline{G})|=1$. As the electrons start hopping, resulting in wider bandwidth, the value of $|B(\overline{K},\,\overline{G})\times A(\overline{K},\,\overline{G})|$ departs from 1. $B(\overline{K},\,\overline{G})$ should no longer be constant because otherwise the interatomic correlation would not narrow the band.

Suppose the $(i,\,\sigma)$ configuration $G(i,\,\sigma)$ spreads evenly throughout the entire lattice, then it has a small exchange hole. That is, $|A_{\mathrm{ex}}(J(i,\,\sigma),\,G(i,\,\sigma))|^2$ is relatively large. However, this will give the $(i,\,\sigma)$ electrons a greater chance to be close to the electrons in other $G(j,\,\rho)$, where $i\neq j$ or $\sigma\neq\rho$ or $(i,\,\sigma)\neq(j,\,\rho)$. Therefore, they feel a stronger interatomic correlation from the electrons in $G(j,\,\rho)$ as compared to the strength of interaction sensed by the electrons in a configuration which is restricted in a small portion of the lattice. Hence the occupation probabilities for configurations having larger $|A(\overline{K},\,\overline{G})|$ are reduced more than those having smaller $|A(\overline{K},\,\overline{G})|$. In other words $|B(\overline{K},\,\overline{G})|<|B(\overline{K},\,\overline{G}\,')|$ if $|A(\overline{K},\,\overline{G})|>|A(\overline{K},\,\overline{G}\,')|$. The net effect is to compensate the exchange hole and to bring $|B(\overline{K},\,\overline{G})A(\overline{K},\,\overline{G})|$ back towards the value 1.

Thus, it seems that the QCA is a reasonable approximation for the narrow band, although there is no detailed treatment of electron dynamics. The lack of exact dynamical properties, as revealed in Fig. 4, is reflected in Sec. V where the change of occupation probability for electrons in reciprocal space due to the correlation effect is discussed. However it is found then the QCA is sufficient to describe the characteristics of a correlated electron system. Therefore we will use QCA to compute the various expectation values for the variational wave function Ψ_c. This would allow us to find the values of $\overline{\eta}$ which minimize the total energy.

IV. QCA OF ELECTRON DENSITY MATRICES

The first task is to normalize the wave function Ψ_c. This is equivalent to calculating the zeroth-order density matrix $\rho_0=\langle\Psi_c|\Psi_c\rangle$. With the help of (14) and (15), we can write

$$\rho_0=\sum_{\overline{K},\overline{K}'}D^*(\overline{K})D(\overline{K}\,')\sum_{\overline{\nu}}\Theta^2(\overline{\eta},\,\overline{\nu})\times{\sum_{\overline{G}}}'A^*(\overline{K},\,\overline{G})A(\overline{K}\,',\,\overline{G})B^*(\overline{K},\,\overline{G})B(\overline{K}\,',\,\overline{G})\,. \tag{28}$$

$\overline{K}$ and $\overline{K}\,'$ represent two different choices of subsets $K(li\sigma)$ and $K'(li\sigma)$ of the set $K(l\sigma)$. The top portion of Fig. 1 is reproduced in Fig. 2 to illustrate them. These subsets are chosen arbitrarily except for one restriction. Let $m(li\sigma)$ be the number of sites in $K(li\sigma)$ and $m'(li\sigma)$ the number of sites in $K'(li\sigma)$. Then they are related by $m(\mathrm{I},\,i,\,\sigma)+m(\mathrm{II},\,i,\,\sigma)=m'(\mathrm{I},\,i,\,\sigma)+m'(\mathrm{II},\,i,\,\sigma)$, because two configurations are orthogonal if they have different numbers of electrons in orbitals of type 1 or 2. The sum of $\overline{K}$ in (28) runs through all possible choices of the subsets $K(li\sigma)$. Similarly the sum over K' exhausts all the subsets $K'(li\sigma)$ which satisfy the above condition.

The QCA (20) largely reduces ρ_0 to the simple form

$$\rho_0^{\mathrm{QCA}}=\sum_{\overline{K}}|D(\overline{K})|^2\sum_{\overline{\nu}}\Theta^2(\overline{\eta},\,\overline{\nu})C(\overline{\nu})\,. \tag{29}$$

It gives the correct answer in the limit of weak correlation, where ρ_0^{QCA} should reduce to $\langle\Psi|\Psi\rangle=1$. In this region, since $\Theta(\overline{\eta},\,\overline{\nu})=1$, we have

$$\rho_0^{\mathrm{QCA}}(\overline{\eta}=1)=\sum_{\overline{K}}|D(\overline{K})|^2\sum_{\overline{\nu}}C(\overline{\nu})=\sum_{\overline{K}}|D(\overline{K})|^2=\prod_{k,l}[|C_{l1}(k)|^2+|C_{l2}(k)|^2]=1\,. \tag{30}$$

The advantage of (29) is that the dependence on the geometry of individual configurations disappears, and the terms involving the real and reciprocal spaces are separated. The computation of ρ_0^{QCA} becomes a simple counting problem. The counting procedure starts with picking a particular $\overline{K}$. Once $\overline{K}$ is fixed, the numbers of electrons in the four sets $G(i\sigma)$ are known. Then these electrons are inserted arbitrarily in the corresponding localized orbitals. The last step is to count the number of configurations $C(\overline{\nu})$ which have the same numbers of multiple occupations. Finally these numbers are weighted by the correlation factor $\Theta^2(\overline{\eta},\,\overline{\nu})$.

The numbers $C(\overline{\nu})$ have certain well-defined probability distributions of a binominal type. In fact, $\sum_{\overline{\nu}}\Theta^2(\overline{\eta},\,\overline{\nu})C(\overline{\nu})$ is a hypergeometric series. Since the number of electrons is large, this hypergeometric series is sharply peaked at a specific value $\overline{\nu}$, which depends on the values of correlation strength $\overline{\eta}$. Therefore, the series can be approximated by its peak term. From now on, we will use $\overline{\nu}$ to refer to the most likely numbers of multiple occupations, and $C_0(\overline{\eta})$ for the $\Theta^2(\overline{\eta},\,\overline{\nu})C(\overline{\nu})$ of the peak term. (29) then becomes

$$\rho_0^{\mathrm{QCA}}=\sum_{\bar K}|D(\bar K)|^2 C_0(\bar\eta)\ . \tag{31}$$

This formula can be further simplified. Note that $C_0(\bar\eta)$ depends on the choice of $\bar K$ only through the numbers of k vectors in the sets $J(i\sigma)=K(\mathrm{I},\ i,\ \sigma)\cup K(\mathrm{II},\ i,\ \sigma)\cup K(\sigma)$. The identity of individual k vectors does not matter at all. Let $\xi(i\sigma)$ be the number of elements in the set $J(i\sigma)$, then we can rewrite (31) as

$$\rho_0^{\mathrm{QCA}}=\sum_{\bar\xi}\left[\sum_{\bar K}{}'|D(\bar K)|^2\right]C_0(\bar\eta)\ . \tag{32}$$

The primed sum over $\bar K$ is restricted to those $\bar K$ which have the same value of $\bar\xi$. If we assume that $C_{ij}(k)$ are smooth functions of k, it is shown in Appendix B that the dominating term in the summation over $\bar\xi$ corresponds to $\xi(1\sigma)=\xi(2\sigma)=\frac{1}{2}N(\sigma)$. This is consistent with the symmetry property, because $\phi_{1\sigma}(x-g)$ and $\phi_{2\sigma}(x-g)$ are equivalent because of symmetry. Hence if we approximate the first summation in (32) by the dominating term, and use the identity $C_0(\bar\eta=1)=1$, where $\bar\eta=1$ denotes $\eta_1=\eta_2=\eta_3=1$, we obtain the final form of ρ_0^{QCA} as

$$\begin{aligned}\rho_0^{\mathrm{QCA}}&=\rho_0^{\mathrm{QCA}}(\bar\eta=1)C_0(\bar\eta)\ ,\\ \rho_0^{\mathrm{QCA}}(\bar\eta=1)&=\sum_{\bar K}{}'|D(\bar K)|^2\ ,\end{aligned} \tag{33}$$

where the primed sum is restricted to $\xi(i\sigma)=\frac{1}{2}N(\sigma)$.

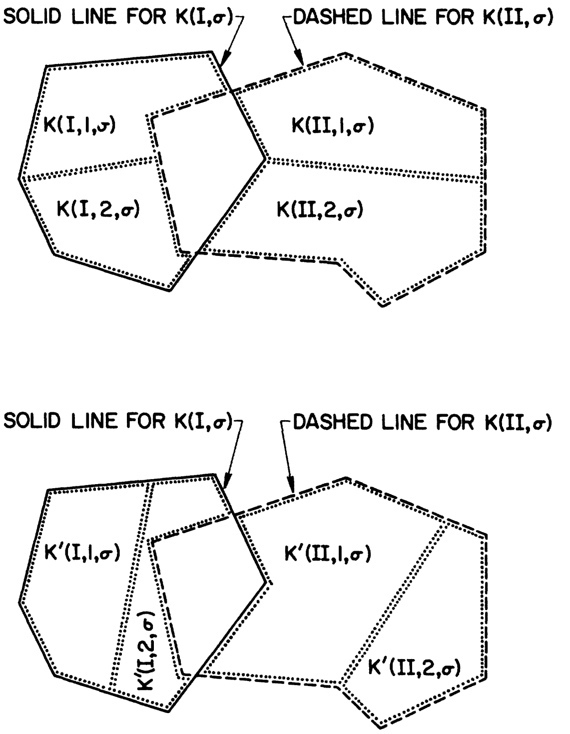

FIG. 2. Two different choices of $K(li\sigma)$ and $K'(li\sigma)$ correspond to two different localized many-electron wave functions.

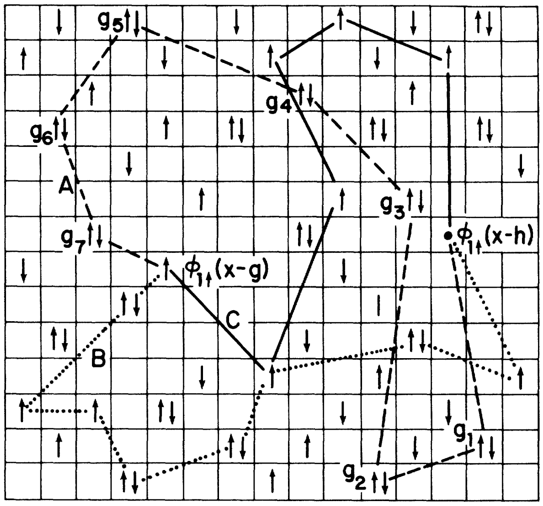

FIG. 3. Typical paths to move an electron from state $\phi_{1\uparrow}(x-g)$ to state $\phi_{1\uparrow}(x-h)$. All the electrons on the same path move simultaneously.

It is easy to see from the last formula that $\rho_0^{\mathrm{QCA}}(\bar\eta=1)=1+O(N^{-1})$, where N is the number of electrons, instead of $\rho_0^{\mathrm{QCA}}(\bar\eta=1)=1$ as in (30). This results from the approximation used to get (33) from (32). We will use the same approximation to compute the first- and second-order density matrices ρ_1 and ρ_2, and the same factor $1+O(N^{-1})$ appears there. When we normalize ρ_1 and ρ_2 this factor cancels out. Therefore, the normalized ρ_1^{QCA} and ρ_2^{QCA} reduce to the correct values when there is no correlation.

The first-order density matrix $\langle\Psi_c|a_{hi\sigma}^\dagger a_{gj\sigma}|\Psi_c\rangle$ measures the probability that an electron moves from $\phi_{j\sigma}(x-g)$ to $\phi_{i\sigma}(x-h)$. Using (14), it can be decomposed into

$$\langle\Psi_c|a_{hi\sigma}^\dagger a_{gj\sigma}|\Psi_c\rangle=\sum_{\bar K,\bar K'}D^*(\bar K)D(\bar K')\sum_{\bar G,\bar G'}A_c^*(\bar K,\ \bar G)\times A_c(\bar K',\ \bar G')\langle\Phi(\bar G)|a_{hi\sigma}^\dagger a_{gj\sigma}|\Phi(\bar G')\rangle\ . \tag{34}$$

For fixed $\bar G$ and $\bar G'$ such that $\langle\Phi(\bar G)|a_{hi\sigma}^\dagger a_{gj\sigma}|\Phi(\bar G')\rangle\neq 0$, let us first examine the uncorrelated coefficient $A^*(\bar K,\ \bar G)A(\bar K,\ \bar G')$. If the determinantal form of $A(\bar K,\ \bar G)$ given by Eq. (A3) is used, it is easy to show that $A^*(\bar K,\ \bar G)A(\bar K,\ \bar G')$ represents all the possible channels through which an electron can move from $\phi_{j\sigma}(x-g)$ to $\phi_{i\sigma}(x-h)$, while the other electrons move through various closed paths in the lattice. Figure 3 shows three typical channels in a simplified configuration which contains only one kind of localized orbitals $\phi_{1\sigma}(x-g)$. The arrows indicate the spins of the occupied Wannier orbitals. In channel A all the intermediate sites are doubly occupied. In channel B part of the intermediate sites are doubly occupied. However, in channel C

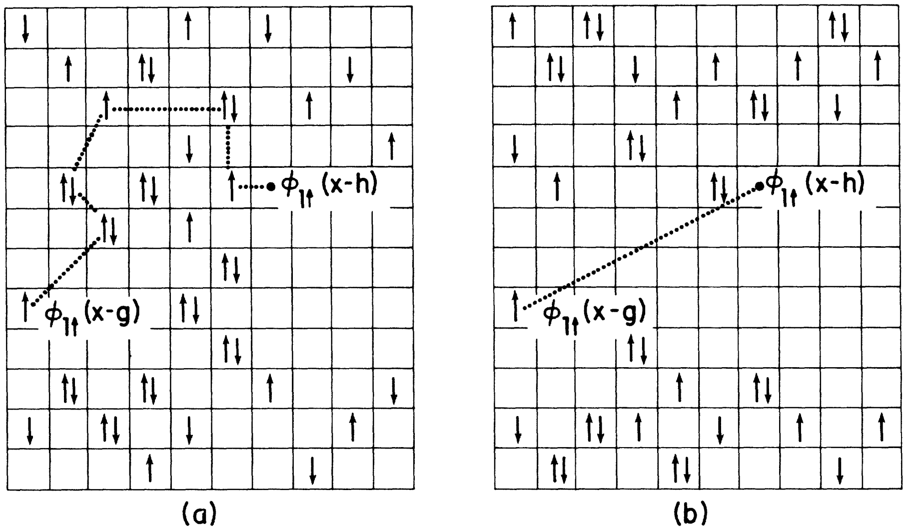

FIG. 4. Interatomic correlation makes it harder to transfer an electron from $\phi_{1\uparrow}(x-g)$ to $\phi_{1\uparrow}(x-h)$ in (a) than in (b).

there is no doubly occupied intermediate site.

The transfer of the electrons in each channel, say channel A, is characterized by moving the electron at site g_1 to site h, and the electron at g_2 moves to g_1, and so on. The electron at site g only moves to g_7 rather than to site h directly. In this manner the numbers of double occupations before and after the electron transfer can differ by no more than 1. Therefore, the QCA given by (20) is applicable to (34). The computation is very tedious, but presents no additional difficulty. Its QCA expression is given in Appendix B.

However, the computation of the first-order density matrix reveals the weak point of QCA in its present form of (20). In the following analysis we show that (20) may need modification in order to take into full consideration the dynamics of the electrons. Consider the two configurations in Fig. 4. Although the numbers of doubly occupied sites in (a) and in (b) are the same, it is more difficult for an electron in (a) to go from $\phi_{1\sigma}(x-g)$ to $\phi_{1\sigma}(x-h)$ than the electron in (b). It seems hopeless to justify rigorously that (20) has taken full account of the pseudointeratomic correlation (the interatomic correlation is actually induced by the intra-atomic correlation whenever one electron passes the other electrons).

There is one argument which favors the reasonableness of (20) with respect to this particular point. For a fixed number of doubly occupied sites, the configurations which evenly spread over the whole crystal largely outnumber those having sites concentrated in a small portion of the lattice. Therefore, the pseudointeratomic correlation cannot be a first-order effect for our model Hamiltonian (42).

We shall use $\rho_1^{\text{QCA}}(hi\sigma, gj\sigma)$ to denote the QCA expression for the normalized first-order density matrix. With the help of (33) and the results from Appendix B, we have

$$\rho_1^{\text{QCA}}(hi\sigma, gj\sigma) = \langle\Psi_c|a^\dagger_{hi\sigma}a_{gj\sigma}|\Psi_c\rangle^{\text{QCA}}/\langle\Psi_c|\Psi_c\rangle^{\text{QCA}} = \rho_1^0(hi\sigma, gj\sigma)C(hi\sigma, gj\sigma;\ \eta_1, \eta_2, \eta_3)\ . \tag{35}$$

$\rho_1^0(hi\sigma, gj\sigma)$ is the density matrix if there is no correlation, which can be easily obtained from Bloch representation as

$$\rho_1^0(hi\sigma, gj\sigma) = L^{-1}\sum_l \sum_k C^*_{li}(k)C_{lj}(k)e^{-ik(h-g)}\ , \tag{36}$$

where k belongs to the set $K(l\sigma)$. $C(hi\sigma, gj\sigma;\ \eta_1, \eta_2, \eta_3)$ is a well-defined weighted number of configurations. It is computed by first picking up all the configurations $\overline{G}$, the associated $\Phi(\overline{G})$ of which contain $\phi_{j\sigma}(x-g)$ but not $\phi_{i\sigma}(x-h)$ provided $(jg)\neq(ih)$. The next step is to filter out those $\overline{G}$ for which $M(1\sigma)\neq M(2\sigma)$, where $M(i\sigma)$ is the number of elements in $G(i\sigma)$. This step comes from the same approximation as that used to get (33) from (32). Then the numbers of configurations having the same numbers of multiply occupied sites are weighted by proper correlation effect. The sum of these weighted numbers is a hypergeometric series. $C(hi\sigma, gj\sigma;\ \eta_1, \eta_2, \eta_3)$ is the sharp peak term which approximates the series.

The most general forms of $C(hi\sigma, gj\sigma;\ \eta_1, \eta_2, \eta_3)$ as given in Appendix B, which are valid for all values of correlation strength and electron density, are extremely complicated. However, under strong correlation, the probability of having more than two particles (either electrons or holes) at the same lattice site is negligibly small if we assume that the number of particles $N\leq L$. This is the case of interest for which $C(hi\sigma,\ gj\sigma;\ \eta_1,\ \eta_2,\ \eta_3)$ is greatly simplifed to the following forms.

Since the numbers of σ spin electrons in both types of localized orbitals are equal, $\nu_{14}=\nu_{23}$ and $\nu_{12}=\nu_{34}$ as a result of the symmetry property. Let us normalize them as $\nu_\uparrow=\nu_{13}/L$, $\nu_\downarrow=\nu_{24}/L$, $\nu_2=\nu_{14}/L=\nu_{23}/L$, and $\nu_3=\nu_{12}/L=\nu_{34}/L$. Accordingly let $n=N/L$ and $n_\sigma=N(\sigma)/L$, the density of electrons in lattice. Further define $\alpha=1-n+\nu_\uparrow+\nu_\uparrow+2\nu_2+2\nu_3$ and $\beta_\sigma=n_\sigma/2-\nu_\sigma-\nu_2-\nu_3$. Then, from Appendix B, we have

$$C(hi\sigma,\ hi\sigma;\ \eta_1,\ \eta_2,\ \eta_3)=1\ , \tag{37}$$

$$C(hi\sigma, hj\sigma; \eta_1, \eta_2, \eta_3) = [(\tfrac{1}{2}n_\sigma)(1-\tfrac{1}{2}n_\sigma)]^{-1}\beta_\sigma[1+(\gamma_\uparrow+\gamma_\downarrow)\eta_2\eta_3], \quad i\neq j \tag{38}$$

$$\begin{aligned}C(hi\sigma, gi\sigma; \eta_1, \eta_2, \eta_3) &= C(hi\sigma, gj\sigma; \eta_1, \eta_2, \eta_3)\\ &= [(\tfrac{1}{2}n_\sigma)(1-\tfrac{1}{2}n_\sigma)]^{-1}\beta_\sigma\alpha[1+\gamma_\rho(\eta_2+\eta_3)+\gamma_\sigma\eta_1]^2, \quad g\neq h\end{aligned} \tag{39}$$

where $\gamma_\sigma=\beta_\sigma/\alpha$, and ρ is the spin opposite to σ. The ν_i and η_i are related by

$$\eta_1^2=\nu_\sigma\ \alpha\beta_\sigma^{-2}, \qquad \sigma=\uparrow,\downarrow \tag{40}$$

$$\eta_i^2=\nu_i\ \alpha(\beta_\uparrow\beta_\downarrow)^{-1}, \quad i=2,\ 3. \tag{41}$$

V. CRITERION ON FERROMAGNETISM

Consider the model Hamiltonian consisting of two degenerate tight-binding bands and intra-atomic Coulomb interaction:

$$H=\sum \epsilon(l,k)n_{lk\sigma}+C_1\sum n_{1g\sigma}n_{2g\sigma} + C_2\sum n_{1g\sigma}n_{2g\rho}+C_3\sum n_{ig\uparrow}n_{ig\downarrow}\,. \tag{42}$$

$n_{lk\sigma}$ and $n_{ig\sigma}$ are the number operators. l, σ, ρ, i, k, and g are the appropriate dummy indices. The band energy is normalized to $\sum_k \epsilon(l,k)=0$ for convenience.

To compute the expectation value $\langle\Psi_c|H|\Psi_c\rangle$, the two-body interaction energy turns out to be trivial. Whenever two electrons orbit around the same atom, they contribute one unit of the appropriate correlation energy. Hence the interaction energy is obtained by counting the number of double occupancies and weighting them with the corresponding correlation energy C_i.

With the help of (1), the correlated band energy can be expressed in terms of first-order density matrices. To be mathematically manageable, our computation is restricted to the cases of less than one or more than three electrons per atom. In such a situation, $\langle\Psi_c|H|\Psi_c\rangle$ can then be written as

$$\langle H\rangle=\sum_\sigma q_\sigma N(\sigma)\bar{\epsilon}_\sigma+C_1L\sum_\sigma\nu_\sigma+2C_2L\nu_2+2C_3L\nu_3, \tag{43}$$

$$q_\sigma=[\tfrac{1}{2}n_\sigma(1-\tfrac{1}{2}n_\sigma)]^{-1}\ \alpha\beta_\sigma[1+\gamma_\rho(\eta_2+\eta_3)+\gamma_\sigma\eta_1]^2. \tag{44}$$

$\bar{\epsilon}_\sigma$ is the mean band energy of the occupied uncorrelated σ-spin Bloch states. q_σ measures the discontinuity of the occupation probability for the σ-spin electrons in reciprocal space, as indicated in Fig. 5.

With the help of the definitions of α and β_σ as well as the relations (40) and (41), it is easy to see that $\langle H\rangle$ depends on ν_i and C_i. We use the variational method to minimize $\langle H\rangle$ with respect to ν_i. This yields the necessary relations between ν_i and C_i through which either C_i or ν_i can be eliminated. In the similar s-band problem, Gutzwiller[16] kept ν_i to investigate the stability of the ferromagnetic ground state. On the other hand, Brinkmann and Rice[8] worked with C_i to predict a metal-insulator transition.

We will eliminate C_i. The resulting form for the general case is very complicated. However, for the ferromagnetic state, it becomes quite simple owing to the fact that $\nu_\downarrow=\nu_2=\nu_3=0$ and $\eta_2=\eta_3=0$. Letting the subscript f denote the ferromagnetic state and $\nu_\uparrow=\nu_f$, $\bar{\epsilon}_\uparrow=\bar{\epsilon}_f$, and $q_\uparrow=q_f$, we find

$$\langle H\rangle_f=q_f\,N\bar{\epsilon}_f\left[1-\frac{\nu_f}{q_f}\ \frac{dq_f}{d\nu_f}\right], \tag{45}$$

$$q_f=[\tfrac{1}{2}n(1-\tfrac{1}{2}n)]^{-1}\ (\tfrac{1}{2}n-\nu_f) \times[(1-n+\nu_f)^{1/2}+\nu_f^{1/2}]^2. \tag{46}$$

For the paramagnetic state which is symbolized by a subscript p, we let $\nu_1=\nu_\uparrow=\nu_\downarrow$ and $q_p=q_\uparrow=q_\downarrow$. Since $n_\uparrow=n_\downarrow=\frac{1}{2}n$, (43) and (44) then reduce to

$$\langle H\rangle_p=q_pN\bar{\epsilon}_p\left[1-\sum_i\frac{\nu_i}{q_p}\ \frac{dq_p}{d\nu_i}\right], \tag{47}$$

$$\begin{aligned}q_p=[\tfrac{1}{4}n(1-\tfrac{1}{4}n)]^{-1}\ (\tfrac{1}{4}n-\nu_1-\nu_2-\nu_3)&\times[(1-n+2\gamma_2+2\nu_3+2\nu_1)^{1/2}\\ &+\nu_1^{1/2}+\nu_2^{1/2}+\nu_3^{1/2}]^2,\end{aligned} \tag{48}$$

where $N\bar{\epsilon}_p=\sum_\sigma N(\sigma)\bar{\epsilon}_\sigma$.

If the correlation is so strong that there is no doubly occupied site, the energy difference $\Delta E=\langle H\rangle_f-\langle H\rangle_p$ has the simple form

$$\Delta E=N(1-n)(1-\tfrac{1}{2}n)^{-1}\ (\bar{\epsilon}_f-\bar{\epsilon}_p) \times[1+\tfrac{1}{4}n(1-\tfrac{1}{4}n)^{-1}\bar{\epsilon}_p\ (\bar{\epsilon}_f-\bar{\epsilon}_p)^{-1}]. \tag{49}$$

Equation (49) is computed for various simple density-of-states curves which are assumed to be step functions consisting of three steps. It is found that $\Delta E<0$, and therefore the ground state is ferromagnetic if the density of states is large at the band edges. If the density of states is large at the middle of the band, $\Delta E > 0$ and so we have a stable paramagnetic ground state.

The criterion on ferromagnetism depends on the shape of the band, rather than the local properties around $\bar{\epsilon}_f$ and $\bar{\epsilon}_p$ as one might conclude from (49). This is because the band energy is so normalized that $\bar{\epsilon}_f$ and $\bar{\epsilon}_p$ are measured from a zero point which depends on the entire band structure. This dependence on the whole band structure can be demonstrated by the occupation probability for the electrons in reciprocal space.

Since the electron correlation serves as a scattering mechanism to excite the uncorrelated Bloch electrons from below to above the Fermi surface, let us first compute the correlated Fermi distribution function

$$F_c(lk\sigma) = \langle \Psi_c | n_{lk\sigma} | \Psi_c \rangle / \langle \Psi_c | \Psi_c \rangle . \qquad (50)$$

Substituting (1) into (50), it becomes

$$F_c(lk\sigma) = L^{-1} \sum_{j,j',g,g'} e^{-ik(g-g')} C_{lj}^* C_{lj'} \, \rho_1(gj\sigma, g'j'\sigma). \qquad (51)$$

Using (35)–(41), the correlated occupation probability in k space can be obtained as

$$F_c(lk\sigma) = \tfrac{1}{2} n_\sigma + [\tfrac{1}{2} n_\sigma \, (1 - \tfrac{1}{2} n_\sigma)]^{-1} \, [F(lk\sigma) - \tfrac{1}{2} n_\sigma] \times \alpha\beta_\sigma [1 + \gamma_p(\eta_2 + \eta_3) + \gamma_\sigma \eta_1]^2 \pm \tilde{F}(lk\sigma). \qquad (52)$$

$F(lk\sigma)$ is the Fermi distribution function for the uncorrelated wave function. $\tilde{F}(lk\sigma)$ is a complicated k-dependent fluctuation which depends on the band structures. The + spin goes with one band, and the – sign goes with the other band. Since $\tilde{F}(lk\sigma)$ is of the order $O(N^{-1})$ as compared to the k-independent part, it will be neglected. We can then rewrite (52) in terms of energy as

$$F_c(lE\sigma) = \tfrac{1}{2} n_\sigma + [F(lE\sigma) - \tfrac{1}{2} n_\sigma] \, [\tfrac{1}{2} n_\sigma (1 - \tfrac{1}{2} n_\sigma)]^{-1} \times \alpha\beta_\sigma [1 + \gamma_p(\eta_2 + \eta_3) + \gamma_\sigma \eta_1]^2 . \qquad (53)$$

Under extremely strong correlation such that $\nu_i = 0$, we have for the paramagnetic state

$$F_{cp}(lE\sigma) = \tfrac{1}{4} n + [F_p(lE\sigma) - \tfrac{1}{4} n] \, (1 - n)(1 - \tfrac{1}{4} n)^{-1}, \qquad (54)$$

and for the ferromagnetic state

$$F_{cf}(lE\sigma) = \tfrac{1}{2} n + [F_f(lE\sigma) - \tfrac{1}{2} n] \, (1 - n)(1 - \tfrac{1}{2} n)^{-1}. \qquad (55)$$

Let $\mathcal{E}_f$ and $\mathcal{E}_p$ be the Fermi energies of the uncorrelated ferromagnetic and paramagnetic states, respectively; then

$$\begin{aligned} \Delta F_c(lE) &= F_{cf}(lE\uparrow) - F_{cp}(lE\uparrow) - F_{cp}(lE\downarrow) \\ &= -(1-n) && \text{if } E < \mathcal{E}_p \\ &= (1-n)(1 + \tfrac{1}{4} n)(1 - \tfrac{1}{4} n)^{-1} && \text{if } \mathcal{E}_p < E < \mathcal{E}_f \\ &= -\tfrac{1}{8} n^2 (1-n)(1 - \tfrac{1}{2} n)^{-1}(1 - \tfrac{1}{4} n)^{-1} && \text{if } \mathcal{E}_f < E . \qquad (56) \end{aligned}$$

(54), (55), and (56) are shown in Fig. 5.

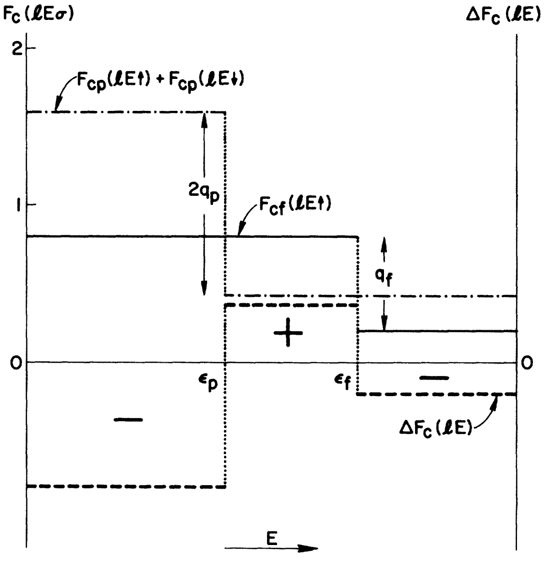

FIG. 5. Occupation probabilities of electrons vs energy. Subscripts p and f stand for paramagentic and ferromagnetic states, respectively, and $\Delta F_c(lE)$ is defined by (56).

The contribution from the band energy to the energy difference ΔE (48) can be expressed as

$$\Delta E_b = \textstyle\sum_l \int E \rho_l(E) \, \Delta F_c(lE) \, dE . \qquad (57)$$

For convenience we reset the zero point of the energy at the bottom of the band, that is, $E \geq 0$ in (57). This does not affect the value of (57) because there are equal numbers of electrons in both the ferromagnetic and the parametic states. If the density of states $\rho_l(E)$ is large at the center of the band where $F_c(lE)$ is positive, then $\Delta E_b > 0$ and we have a stable paramagnetic ground state. However if $\rho_l(E)$ is large at the band edges, especially at the top of the band where $\Delta F_c(lE)$ is negative, then $\Delta E_b < 0$. In this case the correlation of electrons favors a ferromagnetic ground state.

We must point out that the result which is presented in Fig. 5 has some obvious shortcomings. To excite a Bloch electron from the bottom to the top of the band may be energetically unfavorable. The present result comes from neglecting the dynamics of the electrons as mentioned in Secs. III and IV. The correct treatment including such dynamical properties is expected to produce a modification of the Fermi distribution function only within a finite region around the Fermi surface. However, except for the detailed numerical evaluation of (57), the exact solution does not give additional characteristics to the criterion on ferromagnetism. In this sense, QCA leads to a reasonable prediction.

VI. DISCUSSION

We have used the variational method to investigate the Hubbard model of degenerate bands. Since the number of particles is assumed to be less than 1, and there are four localized orbitals at each site, our results are valid only for less than one or more than three electrons per atom. The method can be extended to all electron densities simply by devoting the necessary effort to handle the complicated algebraic work. The results, however, may not add any new features to what we have already presented.

Only the intra-atomic correlation is emphasized in both the trial function and the model Hamiltonian. Although the interatomic correlation is considered in the trial function, it is not fully taken into account by the QCA. The ignorance of the dynamics of electrons by QCA yields the step-function-type occupation probability of electrons in reciprocal space. Hence this approximation is good only for narrow bands.

The correlation of electrons scatters the electrons from below to above the uncorrelated Fermi surface in such a way that the total occupied volume in k space is preserved. The probability that electrons are excited into the holes is larger for the paramagnetic state than for the ferromagnetic state as shown in Fig. 5. Therefore, if the density of states is large at the band edges, the ground-state energy of the paramagnetic state increases much faster than that of the ferromagnetic state as the intra-atomic Coulomb repulsion increases. Eventually this is the mechanism through which the correlation of electrons produces ferromagnetism.

The results of the present paper can be summarized by the schematic plot in Fig. 6. The top portion corresponds to the cases in which a large density of states lies at the middle of the band. The energy $\langle H\rangle_f$ for the ferromagnetic state is always greater than the energy $\langle H\rangle_p$ for the paramagnetic state, no matter how strong the correlation is. However for those bands which have a large density of states at the band edges, $\langle H\rangle_f$ and $\langle H\rangle_p$ cross at $C = C_m$, as indicated in the bottom part of Fig. 6. At this particular value of correlation energy C_m, the paramagnetic and ferromagnetic states are degenerate. Hence, around C_m the ground state of the correlated electron system is a mixture of both the magnetically ordered and disordered phases. A magnetic transition thus occurs as the intra-atomic Coulomb interaction increases and passes through C_m.

The QCA results (for d-band as well as the s-band models) are valid not only for three-dimensional but also for two- and one-dimensional lattices. Therefore the ferromagnetic state can occur in a one-dimensional case which is forbidden by the theorem of Lieb and Mattis.[12] However their theorem is true only for a spin-independent Hamiltonian like the one used in the s-band model. Since the intra-atomic Coulomb energies vary with the spin of the electrons in the localized orbital, the degenerate-band Hamiltonian (42) is spin dependent. Hence the present model is out of the scope of applicability of Lieb and Mattis's theorem. This is very important because it is very difficult to construct a solvable model Hamiltonian which predicts the existence of ferromagnetism only for three-dimensional lattices.

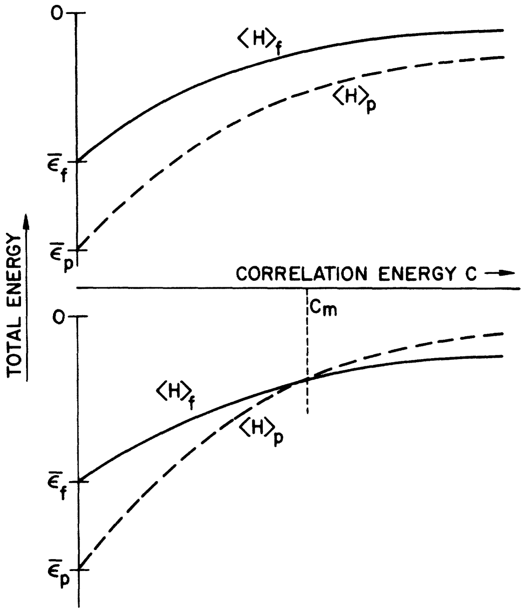

FIG. 6. Schematic plot of the total energies of the electron system in ferromagnetic state (solid curve) and in paramagnetic state (dashed curve) as functions of the electron correlation energy.

ACKNOWLEDGMENTS

The author would like to express his sincere thanks to Dr. M. C. Gutzwiller who kindly suggested the problem and continuously gave very valuable advice, and also to Dr. W. F. Brinkman for helpful discussions.

APPENDIX A

Before going into the computations, it is helpful to review the definitions of the various sets, the relations between them, as well as the relations between the numbers of elements in the sets. We use the parenthesized indices preceded by either K or G to represent the sets. The indices are $\sigma = \uparrow$ or $\downarrow$ (the spin quantum number), $l = \mathrm{I}$ or II (the band index), and $i = 1$ or 2 (the characteristics of the localized orbitals).

If the parenthesis is preceded by K, it is a set of reciprocal vectors. If preceded by G, it is a set of lattice sites. A set of k vectors is associated with a set of Bloch states, and similarly a set of sites is associated with a set of atomic orbitals localized around these sites. The natures of the Bloch and the localized states are characterized by the indices in the parentheses. Hence in our notation, $K(l\sigma)$ is a set of k vectors which are related to the Bloch states $\Psi_{lk\sigma}(x)$, where $k\in K(l\sigma)$. In the same fashion $G(i\sigma)$ is related to the atomic orbitals $\phi_{i\sigma}(x-g)$, where $g\in G(i\sigma)$.

Let the numbers of elements in $K(l\sigma)$ and $G(i\sigma)$ be $m(l\sigma)$ and $M(i\sigma)$, respectively, and $m(\mathrm{I},\sigma)+m(\mathrm{II},\sigma)=M(1\sigma)+M(2\sigma)=N(\sigma)$ (total number of electrons of spin σ). Therefore we can match the $N(\sigma)$ Bloch states associated with $K(\mathrm{I},\sigma)$ and $K(\mathrm{II},\sigma)$ and the $N(\sigma)$ atomic orbitals associated with $G(1\sigma)$ and $G(2\sigma)$ by a one-to-one mapping. We only consider the restricted mappings defined as follows: Let $K(\sigma)$ be the subset of the k vectors which are common to $K(\mathrm{I},\sigma)$ and $K(\mathrm{II},\sigma)$. Topologically define the arbitrary subsets $K(li\sigma)$ by the top portion of Fig. 1. Then we match the $J(1\sigma)$ in Fig. 1 to $G(1\sigma)$ and $J(2\sigma)$ to $G(2\sigma)$. If the number of elements in $K(li\sigma)$ is $m(li\sigma)$ and that in $K(\sigma)$ is $m(\sigma)$, it is clear that $m(l\sigma)=m(l1\sigma)+m(l2\sigma)+m(\sigma)$, and $M(i\sigma)=m(\mathrm{I},i,\sigma)+m(\mathrm{II},i\ \sigma)+m(\sigma)$.

When (1) is inserted into (2), we have (4) as a linear combination of the localized wave functions. Equation (4) is obtained through all possible ways to map $K(l\sigma)$ into $G(i\sigma)$. The coefficients $A(\overline{K},\overline{G})$ are defined in terms of determinants with the rows labeled by the elements in $\overline{K}$ and columns by the elements in $\overline{G}$. Therefore the sign of these determinants depends on the order in which the k vectors and the lattice sites are enumerated. Of course, the final results do not depend on the particular order which was chosen. If we enumerate the k vectors and the sites by a certain order, we can write[14]

$$\Psi=\sum_{\overline{K}}\left(\prod_{jl\sigma}\prod_{k}C_{lj}(k)\right)\times\sum_{\overline{G}}\left[\prod_{j\sigma}\left(L^{-1/2}e^{ikg}\,\middle|\,\begin{matrix}J(j\sigma)\\G(j\sigma)\end{matrix}\right)\right]\Phi(\overline{G}),\tag{A1}$$

where $k\in K(lj\sigma)$, and the notation for the determinant is explained in Ref. 14. Thus we have

$$D(\overline{K})=\prod_{jl\sigma}\prod_{k}C_{lj}(k),\tag{A2}$$

$$A(\overline{K},\overline{G})=\prod_{j\sigma}\left(L^{-1/2}\ e^{ikg}\,\middle|\,\begin{matrix}J(j\sigma)\\G(j\sigma)\end{matrix}\right).\tag{A3}$$

For the convenience of later computations, we would like to write

$$\Psi=\left[\prod_{\sigma}\sum_{n(\sigma),m(\mathrm{I},1,\sigma)}\sum_{\overline{K}}{}'\prod_{il}\prod_{k}C_{li}(k)\times\sum_{\overline{G}}\prod_{j}\left(L^{-1/2}e^{ikg}\,\middle|\,\begin{matrix}J(j\sigma)\\G(j\sigma)\end{matrix}\right)\right]\Phi(\overline{G}),\tag{A4}$$

where $n(\sigma)=m(\mathrm{I},1,\sigma)+m(\mathrm{II},1,\sigma)$. Once $n(\sigma)$ and $m(\mathrm{I},1,\sigma)$ are fixed, we can determine $m(\mathrm{II},1,\sigma)=n(\sigma)-m(\mathrm{I},1,\sigma)$ and $m(l2\sigma)=m(l\sigma)-m(\sigma)-m(l1\sigma)$. The ranges of the dummy indices are $0\le m(\mathrm{I},1,\sigma)\le n(\sigma)$ and $0\le n(\sigma)\le m(\mathrm{I}\sigma)+m(\mathrm{II},\sigma)-2m(\sigma)$. The primed summation over $\overline{K}$ runs only through those $\overline{K}$ having the same values of $m(li\sigma)$.

APPENDIX B

Substituting (A2) and (A3) into (28), it becomes

$$\rho_0=\prod_{\sigma}\sum_{n(\sigma),m(\mathrm{I},1,\sigma),m'(\mathrm{I},1,\sigma)}\sum_{\overline{\nu}}\Theta^2(\overline{\eta},\overline{\nu})\times\sum_{\overline{K},\overline{K}'}D^*(\overline{K})D(\overline{K}')\tilde{\rho}_0(\overline{K},\overline{K}'),\tag{B1}$$

where

$$\tilde{\rho}_0=\prod_{\sigma}\sum_{\overline{G}}\prod_{j}\left(\omega_{j\sigma}(g-g')\ \begin{matrix}G(j\sigma)\\G(j\sigma)\end{matrix}\right)B^*(\overline{K},\overline{G})B(\overline{K}',\overline{G}).\tag{B2}$$

At this point the degenerate band turns out to be much more complicated than the s band. In the latter the function $\omega_\sigma(g-g')$ is a simple summation over the occupied region in reciprocal space, i.e., over the states in $K(\mathrm{I}\sigma)$ (if we assume that we have only the I band). However, in the present case we have to be very careful in defining the $\omega_{i\sigma}(g-g')$ in order to avoid any ambiguity.

Since the k vectors are enumerated according to a certain order, we can also label the k vectors in the union of $K(\mathrm{I},i,\sigma)$, $K(\mathrm{II},i,\sigma)$, and $K(\sigma)$ by numbers. That is

$$k_n\in K(\mathrm{I},i,\sigma)\cup K(\mathrm{II},i,\sigma)\cup K(\sigma),$$
$$n=1,2,\ldots,m(\mathrm{I},i,\sigma)+m(\mathrm{II},i,\sigma)+m(\sigma).$$

By the same token, we have another set

$$k'_n\in K'(\mathrm{I},i,\sigma)\cup K'(\mathrm{II},i,\sigma)\cup K(\sigma),$$
$$n=1,2,\ldots,m'(\mathrm{I},i,\sigma)+m'(\mathrm{II},i,\sigma)+m(\sigma).$$

The elements in the $i\sigma$ determinant can now be defined as

$$\omega_i(g-g')=L^{-1}\sum_{n}e^{i(k_ng-k'_ng')}.\tag{B3}$$

By the QCA (20), $\tilde{\rho}_0=C(\overline{\nu})\delta_{\overline{K},\overline{K}'}$. Hence (B1) is simplified to

$$\rho_0^{\mathrm{QCA}}=\left[\prod_{\sigma}\sum_{n(\sigma),m(\mathrm{I},1,\sigma)}\sum_{\overline{K}}{}'\prod_{il}\prod_{k}C_{li}(k)\right]\Theta^2(\overline{\eta\nu})C(\overline{\nu}).\tag{B4}$$

Since $C_{li}(k)$ are smooth functions of k, then the primed sum over $\overline{K}$ is proportional to the number of $\overline{K}$ to be summed up. This number is sharply peaked at $n(\sigma)=\frac{1}{2}N(\sigma)-m(\sigma)$, a result in agreement with the symmetry properties of the e_g sublevels. If we approximate the sum over $n(\sigma)$ by the peak term, we have

$$\rho_0^{\mathrm{QCA}} = \Big(\prod_{\sigma} \sum_{m(\mathrm{I},1,\sigma)} \sum_{\overline{K}}{}' \prod_{il} \prod_{k} C_{il}(k) \Big) \Theta^2(\overline{\eta}\overline{\nu})\, C(\overline{\nu}), \tag{B5}$$

where $0 \le m(\mathrm{I}, 1, \sigma) \le \frac{1}{2}N(\sigma) - m(\sigma)$.

To write down the complicated quantity $\Theta^2(\overline{\eta}\overline{\nu}) \times C(\overline{\nu})$, let us use the notation $\sum[A\,|\,a]$ to represent a summation over the integer a from 0 to A. Then we have

$$\begin{aligned}
\Theta^2(\overline{\eta}\overline{\nu})\, C(\overline{\nu}) = {}& \binom{L}{\frac{1}{2}N(\uparrow)}^{-1} \binom{L}{\frac{1}{2}N(\downarrow)}^{-2} \sum[\tfrac{1}{2}N(\uparrow)\,|\,\nu_{12}] \sum[\tfrac{1}{2}N(\uparrow) - \nu_{12}\,|\,\nu_{13}] \sum[\tfrac{1}{2}N(\uparrow) - \nu_{12} - \nu_{13}\,|\,\nu_{14}] \sum[\tfrac{1}{2}N(\uparrow) - \nu_{12} - \nu_{23}\,|\,\nu_{24}] \\
& \times \sum[\nu_{23}\,|\,\nu_{123}] \sum[\tfrac{1}{2}N(\uparrow) - \nu_{13} - \nu_{23} - \nu_{123}\,|\,\nu_{34}] \sum[\nu_{12} - \nu_{123}\,|\,\nu_{124}] \sum[\nu_{13}\,|\,\nu_{123}] \sum[\nu_{23}\,|\,\nu_{234}] \sum[\nu_{123}\,|\,\nu_{1234}] \\
& \times \binom{L - \frac{1}{2}N(\uparrow)}{\frac{1}{2}N(\uparrow) - \nu_{12}} \binom{\frac{1}{2}N(\uparrow)}{\nu_{12}} \binom{\nu_{12}}{\nu_{123}} \binom{\frac{1}{2}N(\uparrow) - \nu_{12}}{\nu_{13}} \binom{\frac{1}{2}N(\uparrow) - \nu_{12}}{\nu_{23}} \binom{\frac{1}{2}N(\uparrow) - \nu_{12} - \nu_{13}}{\nu_{14}} \binom{\frac{1}{2}N(\uparrow) - \nu_{12} - \nu_{23}}{\nu_{24}} \\
& \times \binom{\nu_{12} - \nu_{123}}{\nu_{124}} \binom{\nu_{23}}{\nu_{234}} \binom{L - N(\uparrow) + \nu_{12}}{\frac{1}{2}N(\downarrow) - \nu_{13} - \nu_{23} - \nu_{123}} \binom{\frac{1}{2}N(\downarrow) - \nu_{13} - \nu_{23} - \nu_{123}}{\nu_{34}} \\
& \times \binom{L - N(\uparrow) - \frac{1}{2}N(\downarrow) + \nu_{12} + \nu_{13} + \nu_{23} + \nu_{123}}{\frac{1}{2}N(\downarrow) - \nu_{14} - \nu_{24} - \nu_{34} - \nu_{124} - \nu_{134} - \nu_{234} - \nu_{1234}} \binom{\nu_{123}}{\nu_{1234}} \\
& \times \eta_3^{2(\nu_{12}+\nu_{34}+\nu_{124}+\nu_{234}+\nu_{1234})}\, \eta_2^{2(\nu_{23}+\nu_{14}+\nu_{124}+\nu_{234}+\nu_{134}+\nu_{1234})}\, \eta_1^{2(\nu_{13}+\nu_{24}+\nu_{123}+\nu_{134}+\nu_{234}+\nu_{1234})}\,.
\end{aligned} \tag{B6}$$

We should point out that $\Theta^2(\overline{\eta}\overline{\nu})\, C(\overline{\nu})$ is symmetric in $N(\uparrow)$ and $N(\downarrow)$, although the present form is chosen for convenience only.

If we are interested in the strong correlation (sc) region, the numbers of triple and quadruple occupations are small and neglected. $\Theta^2(\overline{\eta}\overline{\nu})\, C(\overline{\nu})$ is then approximated by its peak term $C_0(\overline{\eta})$ as

$$\begin{aligned}
\Theta^2(\overline{\eta}\overline{\nu})\, C(\overline{\nu})\big|_{\mathrm{sc}} = {}& C_0(\overline{\eta}) \\
= {}& \binom{L}{\frac{1}{2}N(\uparrow)}^{-1} \binom{L}{\frac{1}{2}N(\downarrow)}^{-2} \binom{L - \frac{1}{2}N(\uparrow)}{\frac{1}{2}N(\uparrow) - \nu_{12}} \binom{\frac{1}{2}N(\uparrow)}{\nu_{12}} \binom{\frac{1}{2}N(\uparrow) - \nu_{12}}{\nu_{13}} \binom{\frac{1}{2}N(\uparrow) - \nu_{12} - \nu_{23}}{\nu_{24}} \\
& \times \binom{\frac{1}{2}N(\uparrow) - \nu_{12}}{\nu_{23}} \binom{\frac{1}{2}N(\uparrow) - \nu_{12} - \nu_{13}}{\nu_{14}} \binom{L - \frac{1}{2}N(\uparrow) - \nu_{12}}{\frac{1}{2}N(\downarrow) - \nu_{13} - \nu_{23}} \binom{N(\downarrow) - \nu_{13} - \nu_{23}}{\nu_{34}} \\
& \times \binom{L - N(\uparrow) - \frac{1}{2}N(\downarrow) + \nu_{12} + \nu_{13} + \nu_{14}}{\frac{1}{2}N(\downarrow) - \nu_{14} - \nu_{24} - \nu_{34}}\, \eta_3^{2(\nu_{12}+\nu_{34})}\, \eta_2^{2(\nu_{14}+\nu_{23})}\, \eta_1^{2(\nu_{13}+\nu_{24})}\,.
\end{aligned} \tag{B7}$$

The most likely numbers of double occupation ν_{ij} are obtained by solving $dC_0(\overline{\eta})/d\nu_{ij} = 0$. Because of the symmetry property of the e_g sublevels, we have $\nu_{12} = \nu_{34}$ and $\nu_{14} = \nu_{23}$. If we normalize the ν_{ij} as $\nu_\uparrow = \nu_{13}/L$, $\nu_\downarrow = \nu_{24}/L$, $\nu_2 = \nu_{14}/L = \nu_{23}/L$, and $\nu_3 = \nu_{12}/L = \nu_{34}/L$, then they satisfy the conditions

$$\eta_1^2 = \nu_\sigma A_\sigma^2\,, \qquad \sigma = \uparrow, \downarrow \tag{B8}$$

$$\eta_i^2 = \nu_i A_\uparrow A_\downarrow\,, \qquad i = 2, 3 \tag{B9}$$

$$\begin{aligned}
A_\sigma = {}& (1 - n + \nu_\uparrow + \nu_\downarrow + 2\nu_2 + 2\nu_3)^{1/2} \\
& \times (\tfrac{1}{2} n_\sigma - \nu_\sigma - \nu_2 - \nu_3)^{-1}\,,
\end{aligned} \tag{B10}$$

where $n = N/L$ and $n_\sigma = N(\sigma)/L$, the density of electrons in lattice [not to be confused with the dummy indices n and $n(\sigma)$ which appeared in (B3) and (B4)].

Following the same procedure and approximations, the QCA of the first-order density matrices (34) are obtained. If we define

$$\begin{aligned}
\omega_\sigma &= 1 - n_\sigma - \tfrac{1}{2}n_\rho + \nu_\sigma + \nu_2 + \nu_3\,, \\
\beta_\sigma &= \tfrac{1}{2}n_\sigma - \nu_\sigma - \nu_2 - \nu_3\,, \\
\alpha &= 1 - n + \nu_\uparrow + \nu_\downarrow + 2\nu_2 + 2\nu_3 + \nu_{124} + \nu_{134} + \nu_{234} + \nu_{1234}\,, \\
\lambda_\sigma &= (\beta_\sigma - \nu_{124} - \nu_{134} - \nu_{234} - \nu_{1234})/\alpha\,, \\
\mu_\sigma &= \omega_\rho(\beta_\sigma - \nu_{123})/\alpha(\omega_\rho + \nu_{123})\,, \\
\chi_\sigma &= \omega_\sigma \beta_\sigma/\alpha(\omega_\sigma + \nu_{123})\,,
\end{aligned}$$

where ρ is the spin opposite to σ, then we have

$$C(hi\sigma\,,\ hi\sigma\,;\ \eta_1,\ \eta_2,\ \eta_3) = 1, \quad i = 1 \text{ or } 2 \tag{B11}$$

$$C(h1\sigma, h2\sigma; \eta_1, \eta_2, \eta_3) = \beta_\sigma[\tfrac{1}{2}n_\sigma(1 - \tfrac{1}{2}n_\sigma)]^{-1}\,[1 + (\lambda_\rho + \mu_\rho)\eta_2\eta_3 + \lambda_\rho\mu_\rho\,\eta_1^2\,\eta_2^2\,\eta_3^2]\,, \tag{B12}$$

$$C(h1\sigma, g1\sigma; \eta_1, \eta_2, \eta_3) = \alpha\beta_\sigma(\omega_\sigma + \nu_{123})\,[(\tfrac{1}{2}n_\sigma)(1 - \tfrac{1}{2}n_\sigma)\,\omega_\sigma]^{-1}$$
$$\times[1 + \chi_\sigma\eta_1 + \lambda_\rho\eta_2 + \mu_\rho\,\eta_3 + (\lambda_\rho\mu_\rho + \mu_\rho\,\chi_\sigma + \chi_\sigma\lambda_\rho)\,\eta_1\eta_2\eta_3 + \lambda_\rho\mu_\rho\chi_\sigma\eta_1^2\,\eta_2^2\,\eta_3^2\,]^2, \tag{B13}$$

$$C(h1\sigma\,,\ g2\sigma\,;\ \eta_1\,,\ \eta_2\,,\ \eta_3) = \alpha\beta_\sigma(\omega_\sigma + \nu_{123})\,[\tfrac{1}{2}n_\sigma(1 - \tfrac{1}{2}n_\sigma)\,\omega_\sigma\,]^{-1}$$
$$\times\prod_{i,j}\{1 + \lambda_\rho\,\eta_2 + \mu_\rho(1 + \lambda_\rho\,\eta_1^2\,\eta_i)\,\eta_j + [1 + \lambda_\rho\,\eta_i\,\eta_j^2 + \mu_\rho\,\eta_i\,\eta_j\,(1 + \lambda_\rho\,\eta_1^2\,\eta_i\,\eta_j^2)\,]\chi_\sigma\,\eta_1\}\,. \tag{B14}$$

In (B14) $i = 2$ and $j = 3$, or $i = 3$ and $j = 2$. The other coefficients $C(h2\sigma,\ h1\sigma;\ \eta_1,\ \eta_2,\ \eta_3)$, $C(h2\sigma,\ g2\sigma;\ \eta_1,\ \eta_2,\ \eta_3)$, and $C(h2\sigma,\ g1\sigma,\ \eta_1,\ \eta_2,\ \eta_3)$ are obtained from (B12)–(B14) by exchanging 1 and 3, and 2 and 4 of the indices of $\bar{\nu}$. For strong correlation such that the triple and quadruple occupations are negligible, (B11)–(B14) reduce to (36)–(38).

---

[*] Parts of this work were reported in J. Appl. Phys. 42, 1420 (1971).

[†] Present address: Institute of Theoretical Physics, Fack, S-402 20, Göteborg 5, Sweden.

[1] J. C. Slater, Phys. Rev. 49, 537 (1936).

[2] E. P. Wohlfarth, Rev. Mod. Phys. 25, 211 (1953); J. C. Slater, *ibid.* 25, 199 (1953); C. Zener and R. R. Heikes, *ibid.* 25, 191 (1953).

[3] J. H. Van Vleck, Rev. Mod. Phys. 25, 220 (1953).

[4] H. Hurwitz, thesis (Harvard University, 1941) (unpublished); J. H. Van Vleck, Rev. Mod. Phys. 17, 42 (1945); Physica 15, 204 (1949).

[5] M. C. Gutzwiller, Phys. Rev. Letters 10, 159 (1963); Phys. Rev. 134, A923 (1964); 137, A1726 (1965), to be referred to as GIII.

[6] J. Hubbard, Proc. Roy. Soc. (London) A276, 238 (1963); 277, 238 (1964).

[7] J. Kanamori, Progr. Theoret. Phys. (Kyoto) 30, 275 (1963).

[8] W. F. Brinkman and T. M. Rice, Phys. Rev. B 2, 4302 (1970).

[9] N. F. Mott, Proc. Phys. Soc. (London) 62, 416 (1949); Phil. Mag. 6, 287 (1961); 40, 677 (1968).

[10] J. Hubbard, Proc. Roy. Soc. (London) A285, 542 (1965).

[11] Y. Nagaoka, Phys. Rev. 147, 392 (1966).

[12] E. Lieb and D. Mattis, Phys. Rev. 125, 164 (1962).

[13] See E. A. Guggenheim, *Mixtures* (Oxford U. P., New York, 1952), p. 38.

[14] Determinants are written in abbreviated form as

$$\left(f(x-y)\,\middle|\,\begin{matrix}X\\Y\end{matrix}\right) = \begin{vmatrix} f(x_1,\ y_1) & f(x_1,\ y_2) & \cdots & f(x_1,\ y_n) \\ f(x_2,\ y_1) & f(x_2,\ y_2) & \cdots & f(x_2,\ y_n) \\ \cdots & \cdots & \cdots & \cdots \\ f(x_n,\ y_1) & f(x_n,\ y_2) & \cdots & f(x_n,\ y_n) \end{vmatrix}\,,$$

where X and Y are ordered sets and $x_i \in X$, $y_i \in Y$.

[15] See R. M. Thrall and L. Tornheim, *Vector Spaces and Matrices* (Wiley, New York, 1957), p. 127.

[16] For details, see M. C. Gutzwiller, Ref. 5, GIII, Sec. 2.

PHYSICAL REVIEW B VOLUME 4, NUMBER 11 1 DECEMBER 1971

Mössbauer Study of Hyperfine Magnetic Interactions in Fe-Ga Solid Solutions[*]

L. R. Newkirk and C. C. Tsuei
W. M. Keck Laboratory of Engineering Materials,
California Institute of Technology, Pasadena, California 91109

(Received 15 March 1971)

Mössbauer spectra have been obtained for bcc Fe alloys containing up to 25-at.% Ga in solid solution. The spectra have been analyzed in a manner which permits the study of the effect of configuration on the hyperfine magnetic field and on the isomer shift with very few prior assumptions. Various assumptions commonly made in handling data of this type have been examined in detail and their validity tested. The details of 16 different configurations are presented along with an empirical formula to describe the hyperfine field as a function of configuration and concentration.

I. INTRODUCTION

The range and strength of the hyperfine magnetic interaction in nondilute solid solutions is of general interest in understanding the mechanism which causes these alloys to be ferromagnetic. Many of