Effect of Electron Correlation in a Narrow Band

K. A. Chao*
IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598
and Institute of Theoretical Physics, Chalmers University of Technology, S-402 20 Göteborg 5, Sweden
(Received 16 October 1972)

Gutzwiller's variational method is used to investigate the Pauli spin susceptibility and the spin waves in a degenerate narrow band. It is found that the susceptibility tends to become negative for sufficiently strong correlation if the number of holes is small and if large density of states occurs at the top of the band near the Fermi energy. For Ni, a numerical analysis is performed using the simplified density-of-states curve proposed by Kanamori. The conditions for the occurrence of ferromagnetic Ni agree with those obtained by Kanamori. The energies of spin waves in ferromagnetic metals are obtained by examining the normal modes of spin excitations in the correlated ground state. Only the improvement upon the random-phase-approximation (RPA) result which is due to correlated electron hoppings is considered. In the long-wavelength limit, the coefficient C in the expression $\hbar\omega_q = Cq^2$ is reduced from the RPA value by this correlation effect, in agreement with the predictions of other theories.

I. INTRODUCTION

The characteristic features of the d electrons in transition metals may be deduced from a number of experimental facts. One of the most important conclusions obtained from analyzing the experimental data is that the d electrons exhibit both itinerant and localized properties.[1] This led Hubbard,[2] Gutzwiller,[3] and Kanamori[4] to the construction of a model Hamiltonian, which is now commonly known as the Hubbard Hamiltonian.

The Hamiltonian consists of a tight-binding band and a very short-range interaction among electrons. Hubbard, Gutzwiller, and Kanamori used different approaches to explore the ground-state properties and obtained qualitatively similar criteria for ferromagnetism. In the past few years, the Hubbard model has been extensively investigated by many authors[5] with various assumptions. In order to be mathematically manageable, many of these treatments are effectively restricted to a nondegenerate band.

Recently, Gutzwiller's variational method was applied to the case of a doubly degenerate band by Chao and Gutzwiller.[6] It was found that ferromagnetism is favored under strong correlations if the number of holes is small and large density of states occurs at the top of the band. Unfortunately, the resulting equation which predicts this qualitative result is too complicated for accurate numerical computation. In this paper, we rederive the criterion for ferromagnetism in an alternative way so numerical computation can be performed.

If the ground state of the Hubbard Hamiltonian is ferromagnetic, one should not be surprised to find the dispersion relation of spin waves in ferromagnetic metals being affected by the intra-atomic correlation. Since Herring and Kittel[7] proposed theoretically that spin waves could exist in ferromagnetic metals the properties of such spin waves have been investigated very thoroughly by Izuyama, Mattis, Thompson, and others.[8] These authors used the random-phase approximation (RPA) to obtain the Cq^2 dispersion relation in the long-wavelength limit, where q is the wave vector of the acoustic magnon.

The effect of electron correlation on spin waves has been discussed by Roth,[9] Yamada and Shimizu,[10] Edwards,[11] Nagaoka,[12] Young and Callaway,[13] and Chan and Young.[14] They found that when the self-energy is included in the spin-wave excitation energies, the RPA solution of the coefficient C in Cq^2 should be properly modified. This paper will be devoted to the investigation of spin waves in the Hubbard model.

Since the mathematical formulation in this paper was developed in our previous work,[6] the variational scheme will only be outlined in Sec. II. In Sec. III, we compute the paramagnetic susceptibility, using a density-of-states curve which characterizes the true one for Ni. It will be seen that under strong correlation, the susceptiblity becomes negative. Assuming a ferromagnetic ground state, we proceed to obtain the spin-wave dispersion curve in the Hubbard model in Sec. IV. In our treatment, only the effect of correlation on hopping energies is considered. In Sec. V, it is briefly illustrated how the disperion curve should be further modified when the bare intra-atomic interaction energies are replaced by the smaller effective ones.

II. VARIATIONAL METHOD

Though a number of authors believe that a nondegenerate-band model is unlikely to produce ferromagnetism, there has been little doubt about the possibility of ferromagnetism in a degenerate

band. If the narrow band is degenerate, Hund's rule can be effective locally and the ground state becomes ferromagnetic without invoking nearest-neighbor exchange coupling. Without loosing generality but avoiding mathematical complexity, Gutzwiller's variational method has been generalized by Chao and Gutzwiller[6] to the simplest degenerate case, namely, the double degeneracy. In this section we will briefly summarize this variational scheme. Only the main assumptions and the general results which are relevant to this paper will be reviewed.

In a simplified picture the d band of the transition metals consists of a triply degenerate subband of t_{2g} symmetry and a doubly degenerate subband of e_g symmetry. Here we assume that the t_{2g} subband lies below the e_g subband. We also assume that this lower subband is completely filled and can therefore be ignored. The upper subband may be approximated by two tight-binding degenerate bands which have the e_g symmetry. This simplified model appropriately describes the band structure of $Fe_xCo_{1-x}S_2$.

Consider the case where there are N electrons (or N holes if the band is more than half-filled) in a cubic lattice of L sites. Two tight-binding bands are assumed to be known in the Hartree-like approximation. The Wannier states are defined in second quantized form as

$$a^\dagger_{lg\sigma} = L^{-1/2} \sum_k e^{ikg} a^\dagger_{lk\sigma} \quad , \tag{1}$$

where $l = 1$ or 2 is the band index, g labels the lattice sites, and σ is the spin quantum number. In a mixed representation the model Hamiltonian consists of a term describing the degenerate bands in the Hartree-like approximation and three intra-atomic interaction terms[15]:

$$H_0 = \sum \epsilon(lk) n_{lk\sigma} + C_2 \sum n_{1g} n_{2g} + C_3 \sum n_{lg\uparrow} n_{lg\downarrow} - J \sum n_{1g\sigma} n_{2g\sigma} \quad , \tag{2}$$

where $n_{kl\sigma}$ and $n_{ig\sigma}$ are the number operators and $n_{lg} = n_{lg\uparrow} + n_{lg\downarrow}$. k, g, l, and σ are the dummy indices. The intra-atomic Coulomb energies are indicated by C_2 and C_3.[16] J is the intra-atomic exchange energy. For convenience the band energy is shifted to make $\sum_k \epsilon(lk) = 0$.

Let us first outline the main idea of the variational method. According to Van Vleck's model of minimum polarity,[17] the ground-state wave function should exclude the states corresponding to higher degrees of ionization. To construct the trial function for the ground state, we will start from the Fermi sea which describes the completely delocalized and uncorrelated electrons. The Fermi sea is then expressed in terms of the Wannier states by using the inverse transformation of Eq. (1). At this stage, one recognizes that the uncorrelated system allows too many atoms in ionized states. To achieve minimum polarity in the system, such energetically unfavored ionized states are partially projected out of the Fermi sea. The variational parameters, which measure the percentage of the ionized states to be projected out, are thus introduced into the trial function.

Assume that our crystal contains $N(\sigma)$ electrons with σ spin, and $N(\sigma l)$ of them are in the lth band. The uncorrelated ground-state wave function is the Fermi sea Ψ. Let $K(\sigma l)$ be the occupied region of k space for the lth band and σ spin when there is no correlation among the electrons. Then the Fermi sea can be written

$$\Psi = \prod_{l\sigma} \prod_k a^\dagger_{lk\sigma} \Psi_0 \quad , \tag{3}$$

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

In order to decompose Ψ into localized states, we will use a set of lattice sites $G(\sigma l)$ to specify a set of Wannier functions $\phi_{l\sigma}(r-g)$ through the relation $g \in G(\sigma l)$. The localized many-electron wave function $\Phi(\overline{G})$ is defined as

$$\Phi(\overline{G}) = \prod_{l\sigma} \prod_g a^\dagger_{lg\sigma} \Psi_0 \quad , \tag{4}$$

where $g \in G(\sigma l)$. The configuration $\overline{G}$ represents the four sets $G(\sigma l)$ with $\sigma = \uparrow, \downarrow$ and $l = 1, 2$. Substituting the transformation

$$a^\dagger_{lk\sigma} = L^{-1/2} \sum_g e^{-ikg} a^\dagger_{lg\sigma} \tag{5}$$

into (3), Ψ becomes

$$\Psi = \sum_{\overline{G}} A(\overline{G}) \Phi(\overline{G}) \quad . \tag{6}$$

$|A(\overline{G})|^2$ measures the probability that the Wannier functions specified by $\overline{G}$ are occupied. For the present analysis, we need not know the exact form of $A(\overline{G})$.

The configuration $\overline{G}$ is characterized by the number of multiple occupancies it contains. This number is reduced by correlation effects. In other words, electron correlation modifies the probability amplitudes $|A(\overline{G})|^2$ according to the geometry of $\overline{G}$. In our variational-method approach to the correlation problem, $A(\overline{G})$ is reduced by a factor $\eta(l_1\sigma_1, l_2\sigma_2)$ whenever two Wannier states $\phi_{l_1\sigma_1}(r-g)$ and $\phi_{l_2\sigma_2}(r-g)$ with $g \in \overline{G}$ are occupied. The weights $\eta(l_1\sigma_1, l_2\sigma_2)$ have the values between zero and unity. Small η $(l_1\sigma_1, l_2\sigma_2)$ corresponds to strong correlation. By symmetry properties the six weight factors $\eta(l_1\sigma_1, l_2\sigma_2)$ are reduced to three independent ones. They are defined as $\eta(1) = \eta(1\uparrow, 2\uparrow) = \eta(1\downarrow, 2\downarrow)$, $\eta(2) = \eta(1\uparrow, 2\downarrow) = \eta(1\downarrow, 2\uparrow)$, and $\eta(3) = \eta(2\uparrow, 2\downarrow) = \eta(1\uparrow, 1\downarrow)$.

The trial function of the correlated ground state can then be expressed as

$$\Psi_c = \sum_{\overline{G}} \eta(1)^{\nu(1\uparrow,2\uparrow)+\nu(1\downarrow,2\downarrow)} \eta(2)^{\nu(1\uparrow,2\downarrow)+\nu(1\downarrow,2\uparrow)}$$

$$\times\eta(3)^{\nu(1\uparrow,1\downarrow)+\nu(2\uparrow,2\downarrow)}B(\bar{G})A(\bar{G})\Phi(\bar{G}) \quad , \qquad (7)$$

where $\nu(l_1\sigma_1,\ l_2\sigma_2)$ is the number of sites g such that the Wannier states $\phi_{l_1\sigma_1}(r-g)$ and $\phi_{l_2\sigma_2}(r-g)$ are occupied. The factor $B(\bar{G})$ is introduced to count for the less important interatomic correlation.

The effect of interatomic correlation is implicitly treated in the quasichemical approximation.[18] Instead of determining the exact form of $B(\bar{G})$, we let

$$\sum_{\bar{G}}{}' \left|B(\bar{G})A(\bar{G})\right|^2 = \gamma \sum_{\text{all } \bar{G}} \left|A(\bar{G})\right|^2 = \gamma \quad , \qquad (8)$$

where the primed sum over $\bar{G}$ runs through the configurations which have the same numbers of multiple occupancies. γ is the ratio of the number of such configurations to the total number of $\bar{G}$. Equation (8) has been proved to be a reasonable approximation for the narrow band, and yields exact solution at the atomic limit.[3,6]

For each set of the parameters $\eta(1)$, $\eta(2)$, and $\eta(3)$, the numbers $\nu(l_1\sigma_1,\ l_2\sigma_2)$ have certain well-defined probability distributions of a binominal type. Since they are sharply peaked we can approximate the mean numbers of multiple occupancies by the peak values, which are determined by

$$\eta(l_1\sigma_1,\ l_2\sigma_2)^2\left(N(\sigma_1 l_1) - \sum_{l\sigma}\nu(l_1\sigma_1,\ l\sigma)\right)\times\left(N(\sigma_2 l_2) - \sum_{l\sigma}\nu(l_2\sigma_2,\ l\sigma)\right)$$

$$= \nu(l_1\sigma_1,\ l_2\sigma_2)(L - N + \nu_t) \quad , \qquad (9)$$

where $\nu_t = \frac{1}{2}\sum_{l\,l'\sigma\sigma'}\nu(l\sigma,\ l'\sigma')$. From now on $\nu(l_1\sigma_1,\ l_2\sigma_2)$ refer to these most probable numbers.

In order to obtain the criterion for ferromagnetism, it is necessary to compute the occupation probabilities in reciprocal space for the strongly interacting electrons

$$\langle n_{lk\sigma}\rangle = \langle\Psi_c\left|n_{lk\sigma}\right|\Psi_c\rangle/\langle\Psi_c\left|\Psi_c\right\rangle$$

$$= L^{-1}\sum_{gh} e^{-ik(g-h)}\langle\Psi_c\left|a^\dagger_{lg\sigma}a_{lh\sigma}\right|\Psi_c\rangle/\langle\Psi_c\left|\Psi_c\right\rangle \quad . \qquad (10)$$

Using the expression Eq. (7) and the quasichemical approximation Eq. (8), the first-order density matrix is given by

$$\rho_1(l\sigma:gh) = \langle\Psi_c\left|a^\dagger_{lg\sigma}a_{lh\sigma}\right|\Psi_c\rangle/\langle\Psi_c\left|\Psi_c\right\rangle$$

$$= \langle\Psi\left|a^\dagger_{lg\sigma}a_{lh\sigma}\right|\Psi\rangle \qquad \text{if } g = h \ ,$$

$$= D(\sigma l)\langle\Psi\left|a^\dagger_{lg\sigma}a_{lh\sigma}\right|\Psi\rangle \quad \text{if } g \neq h \quad , \qquad (11)$$

where Ψ is the Fermi sea, and

$$D(\sigma l) = \{N(\sigma l)[L - N(\sigma l)]\}^{-1}\left(N(\sigma l) - \sum_{l'\sigma'}\nu(l\sigma,\ l'\sigma')\right)(L - N + \nu_t)^{-1}$$

$$\times\left[L - N + \nu_t + \sum_{l_1\sigma_1,(l_1\sigma_1)\neq(l\sigma)}\eta(l\sigma,\ l_1\sigma_1)\left(N(\sigma_1 l_1) - \sum_{l_2\sigma_2}\nu(l_1\sigma_1,\ l_2\sigma_2)\right)\right]^2 \quad . \qquad (12)$$

Substituting Eq. (11) into Eq. (10), we have

$$\langle n_{lk\sigma}\rangle = [1 - D(\sigma l)]n(\sigma l) + D(\sigma l)f(lk\sigma) \quad , \qquad (13)$$

where $n(\sigma l) = N(\sigma l)/L$ is the number density of electron and $f(lk\sigma)$ is the Fermi distribution function for the lth band and σ spin.

We should point out that in quasichemical approximation the difference between the correlation effect on long-range hoppings and that on short-range hoppings is neglected. This results in a constant reduction factor $D(\sigma l)$ in Eq. (11), regardless of the relative locations of the sites g and h. A direct consequence of this approximation is the preservation of the sharp Fermi surface, as indicated by the finite discontinuity $D(\sigma l)$ in the occupation number $\langle n_{lk\sigma}\rangle$. In other words, the Fermi surface is built in from the start to give a appealing description of a metallic state. Since the author has not been able to fully take into account the dynamics of the long-range hoppings, one cannot rigorously answer the question how sharp the Fermi surface should be in the real system. Nevertheless, the possibility of having long-range hoppings is small in the case of a narrow band. Therefore, the exact treatment of the hopping dynamics should not qualitatively change the results obtained from the quasichemical approximation.

It can be shown that the discontinuity $D(\sigma l)$ in the occupation number is a slowly varying function of the electron density and/or the correlation strength. Its value is one when there is no correlation. For the special case of one electron per atom, $D(\sigma l)$ decreases quadraticly to zero in the correlation strength. It should be noted that in the present theory $D(\sigma l) = 0$ marks a metal-to-insulator transition, and the corresponding critical correlation energy depends only on the band structure. On the other hand, if the density of electron is not equal to one, $D(\sigma l)$ is always finite. For instance, for ferromagnetic Ni, $1 \geq D(\sigma l) \geq 0.6$. In this paper we will not consider the case of one electron per atom, i.e., $D(\sigma l)$ is always finite.

Using Eq. (13), the expectation value of the energy $\langle H_0\rangle$ with respect to the trial function Ψ_c can

be readily obtained as

$$\langle H_0\rangle = \langle\Psi_c|H_0|\Psi_c\rangle/\langle\Psi_c|\Psi_c\rangle$$
$$= \sum D(\sigma l)\bar{\epsilon}(\sigma l) + C_2\sum\nu(1\sigma_1,\ 2\sigma_2)$$
$$+ C_3\sum\nu(l\uparrow,\ l\downarrow) - J\sum\nu(1\sigma,\ 2\sigma)\ , \quad (14)$$

where $\bar{\epsilon}(\sigma l)=\sum_k f(lk\sigma)\epsilon(lk)$. To get Eq. (14) we have used the normalization of the band energy. $\langle H_0\rangle$ is then minimized by varying $\nu(l_1\sigma_1,\ l_2\sigma_2)$. When this is done, we can compute $\langle H_0\rangle$ for various values of the electron density $n=N/L$ as functions of the magnetization $\xi=[N(\downarrow)-N(\uparrow)]/L$. It is found that for a small number of holes and large density of states near the top of the band, the ferromagnetic ground state is stable if the electron correlation is sufficiently strong.

Let us use the subscripts f and p to denote the ferromagnetic and the paramagnetic states, respectively. The energy difference $\Delta E=\langle H_0\rangle_f - \langle H_0\rangle_p$ is then schematically shown in Fig. 1. Region I corresponds to weak correlation while in region III the correlation is so strong that there is almost no double occupancy. In this paper, we are mainly interested in the region II where the correlation is fairly strong but a finite fraction of the sites are doubly occupied.

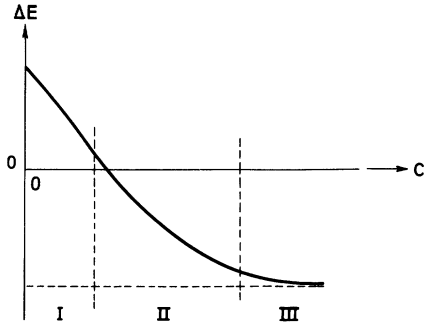

FIG. 1. Schematic plot of the energy difference $\Delta E = \langle H_0\rangle_f - \langle H_0\rangle_p$ between the ferromagnetic state and the paramagnetic state as a function of the correlation energy C. C can be along any direction in the three-dimensional mainfold spanned by $C_1=C_2-J$, C_2, and C_3.

III. PARAMAGNETIC SUSCEPTIBILITY

The criterion for ferromagnetism given in Sec. II manifests itself in the Pauli spin susceptibility. For uncorrelated electrons the magnetic properties of the ground state depend only on the average exchange interaction and the free hoppings, but not on the actual spatial distribution of the electrons. This argument does not apply to a correlated system. Equation (14) clearly indicates that the exact number of double occupancies is as important as the strength of the intra-atomic interactions in determining the magnetic properties of the ground state.

In a paramagnetic state, $N(\uparrow l)=N(\downarrow l)=\frac{1}{2}N(l)$. From the symmetry properties it is easy to see that $D(\uparrow l)=D(\downarrow l)=D(l)$, $\nu(1\uparrow,\ 2\uparrow)=\nu(1\downarrow,\ 2\downarrow)=\nu(1)$, and $\nu(1\uparrow,\ 2\downarrow)=\nu(1\downarrow,2\uparrow)=\nu(2)$. Let $\nu(3,\ l)=\nu(l\uparrow,\ l\downarrow)$, $\alpha=L-N+2\nu(1)+2\nu(2)+\nu(3,\ 1)+\nu(3,2)$, and $\beta(l)=\frac{1}{2}N(l)-\nu(1)-\nu(2)-\nu(3,\ l)$. Then from Eqs. (7) and (12) we have for the paramagnetic state

$$\eta(i)^2\beta(1)\beta(2)=\nu(i)\alpha\ , \qquad i=1,\ 2$$
$$\eta(3)^2\beta(l)^2=\nu(3,\ l)\alpha\ , \qquad l=1,\ 2$$
$$D(l)=\beta(l)\{\alpha+[\eta(1)+\eta(2)]\beta(l')+\eta(3)\beta(l)]^2$$
$$\times\{\tfrac{1}{2}\alpha N(l)[L-\tfrac{1}{2}N(l)]\}^{-1}\ , \quad (15)$$

where in Eq. (15) $l'\neq l$.

To compute the paramagnetic susceptiblity we may take $\xi(l)$ electrons from near the Fermi surface of the down-spin lth band and change their spin to up, keeping them in the same band. Such an operation changes the numbers of double occupancies. It is shown in Appendix A that for small $\xi(l)$ and strong correlation

$$\frac{\delta\nu(1)}{\nu(1)}=\frac{-\delta\nu(2)}{\nu(2)}=\frac{\xi(1)\xi(2)}{\beta(1)\beta(2)}\ , \quad (16)$$

$$\frac{\delta\nu(3)}{\nu(3)}=-\frac{\xi(1)^2+\xi(2)^2}{\beta(1)^2+\beta(2)^2}\ , \quad (17)$$

$$D(\sigma l)=D(l)\left(1\pm\frac{2[1-2\eta(3)]}{2L-N(l)}\,\xi(l)\right.$$
$$\left.+\frac{4[1+\eta(3)^2]}{[2L-N(l)]^2}\right)\xi(l)^2\ , \quad (18)$$

where $2\nu(3)=\nu(3,\ 1)+\nu(3,\ 2)$. In Eq. (18), plus and minus signs correspond to up- and down-spin bands, respectively.

Using Eqs. (14), (16), (17), and (18), the total energy increase over the energy of the paramagnetic state is obtained to the second order in $\xi(l)$,

$$\delta\langle H_0\rangle=\sum_l\frac{1}{L}\left(\frac{1}{\rho(l\mu)}+\frac{4n(l)[1+\eta(3)^2]\,\bar{\epsilon}(l)}{[2-n(l)]^2}\right.$$
$$\left.\pm\frac{4[1-2\eta(3)]\,\mu}{2-n(l)}\right)D(l)\xi(l)^2$$
$$+2C_2[\delta\nu(1)+\delta\nu(2)]+2C_3\delta\nu(3)-2J\delta\nu(1)$$
$$=\sum_l\frac{D(l)}{L\rho_{\rm eff}(l\mu)}\,\xi(l)^2-\frac{2LJ_{\rm eff}}{\beta(1)\beta(2)}\,\xi(1)\xi(2)\ , \quad (19)$$

where $n(l)=N(l)/L$, $\bar{\epsilon}(l)=[\bar{\epsilon}(\uparrow l)+\bar{\epsilon}(\downarrow l)]/N$, and μ is the Fermi energy. The effective exchange energy $J_{\rm eff}$ is defined as

$$J_{\rm eff}=[(J-C_2)\nu(1)+C_2\nu(2)]/L\ . \quad (20)$$

$\rho_{\rm eff}(l\mu)$ is the effective density of states of the lth band per atom per eV per spin. It is related to

the density of state $\rho(l\mu)$ of the uncorrelated system by

$$\rho_{eff}(l\mu) = R(l\mu)^{-1}\rho(l\mu) \ , \tag{21}$$

where

$$R(l\mu) = 1 - \rho(l\mu)\left(\frac{4n(l)[1+\eta(3)^2] \mid \bar{\epsilon}(l)\mid}{[2-n(l)]^2} \pm \frac{4\mu[1-2\eta(3)]}{2-n(l)} + \frac{2LC_3\nu(3)}{D(l)[\beta(1)^2+\beta(2)^2]}\right) \tag{22}$$

is the correlated Stoner enhancement factor. In the last equation, the minus sign applies if the particles are electrons while the plus sign corresponds to the holes. Note that $\delta\langle H_0\rangle$ includes the change of energy due to the readjustment of $D(l)$. This is equivalent to the band-shift effect discussed by Roth and by Harris and Lange.[19]

Magnetism becomes possible when $\delta\langle H_0\rangle$ is negative. It is shown in Appendix B that $\delta\langle H_0\rangle \leq 0$ if

$$J_{eff}^2 \geq \prod_{l=1,2}\left(\frac{\beta(l)}{L}\right)^2 \frac{D(l)R(l\mu)}{\rho(l\mu)} \ . \tag{23}$$

This condition is obviously satisfied when $R(l\mu) \leq 0$. The last term in the bracket of Eq. (22) makes it very difficult to compute $R(l\mu)$. However, this term is positive definite. Therefore $R(l\mu)$ is less than zero if

$$r(l\mu) = 1 - \rho(l\mu)\left(\frac{4n(l)[1+\eta(3)^2] \mid \bar{\epsilon}(l)\mid}{[2-n(l)]^2} \pm \frac{4[1-2\eta(3)]\mu}{2-n(l)}\right) \tag{24}$$

is negative.

We shall apply our result to Ni. Consider the simplified density of states $\rho(l\epsilon)$ shown in Fig. 2. It contains two parameters x and y. x is the total numbers of states in the high-density region. y is the product of the bandwidth W and the density of states in the same region. According to Kanamori,[4] this density-of-states curve characterizes the true one of Ni. For Ni, $y = 3$.[20] Kanamori found that the ferromagnetic ground state is favored if $x \leq 0.28$ and the Hubbard parameter is $2W$. For infinitely strong correlation, the condition relaxes to $x \leq 0.55$.

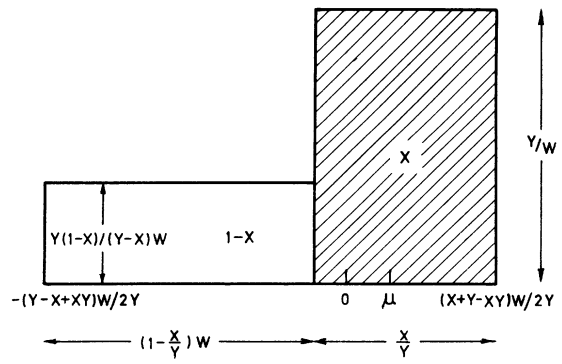

FIG. 2. Density-of-states–vs–energy curve which favors the ferromagnetic Ni. The hatched region contains X states per atom per eV per spin. The energy is so normalized that a fully occupied band has zero energy.

Since the d band of Ni contains 0.6 holes, the plus sign should be used in Eq. (24). $r(l\mu)$ is computed in Appendix B for different values of x and $\eta(3)$. For a given value of $\eta(3)$, there is a critical value x_c such that $r(l\mu) \leq 0$ if $x \leq x_c$. The computed values of $\eta(3)$ and x_c are listed in the first two columns of Table I. Recall that in Fig. 2 the Fermi energy μ is restricted in the high-density region. For 0.6 holes per atom, this condition is satisfied only if $x \geq 0.15$.

In order to have a meaningful comparison of the present results with Kanamori's, one must know the relations between $\eta(3)$ and the intra-atomic Coulomb energies C_2, C_3, and $C_1 = C_2 - J$. When we minimize $\langle H_0\rangle$ of Eq. (14) by varying the numbers of double occupancies $\nu(l_1\sigma_1, l_2\sigma_2)$, we obtain the relations between the $\nu(l_1\sigma_1, l_2\sigma_2)$ and the C_i's. Then, by using Eq. (9), the relations between C_i's and $\eta(i)$'s can be established. However, the resulting equations are too complicated to solve numerically. An alternative, which is much simpler although less accurate, is to use the relation between the corresponding quantities in s-band model. In nondegenerate-band model, there is only one Hubbard parameter C_0 and a single variational parameter η. The relation between C_0 and η has been given by Gutzwiller.[3] In this way, we obtained the values of C_0 listed in the third column of Table I.

Comparing our results with those of Kanamori, we see that ferromagnetism is more favored by the present theory than by the Kanamori's if the correlation is very strong. On the other hand, when the correlation gets weaker, Kanamori's criterion becomes less stringent.

If the Stoner enhancement factor $R(l\mu)$ is positive, one must examine the inequality Eq. (23) to check the stability of the paramagnetic state. It is then important to understand the physical meaning of the effective exchange energy J_{eff} defined by Eq. (20). Let us indicate by $(l\sigma: l'\sigma')$ the two Wannier states $\phi_{l\sigma}(r-g)$ and $\phi_{l'\sigma'}(r-g)$ localized on the same atom. We recall that C_2 and $C_2 - J$ are the intra-atomic interaction energies between two electrons in $(1, \sigma: 2, -\sigma)$ and $(1, \sigma: 2, \sigma)$, respectively. So from its definition, LJ_{eff} measures the difference in total intra-atomic interaction energy between all the electron pairs in $(1, \sigma: 2, \sigma)$ and all the electron pairs in $(1, \sigma: 2, -\sigma)$. Therefore it is the intra-atomic interaction energy rather than the pure exchange energy which is essential

TABLE I. For a given value of $\eta(3)$, x_c is the critical value such that $r(l\mu) \leqslant 0$ if $x \leqslant x_c$, where $r(l\mu)$ is given by Eq. (24) and x is indicated in Fig. 2. C_0 is the corresponding intra-atomic Coulomb energy computed from the s-band theory. W is the bandwidth.

$\eta(3)$	x_c	C_0/W
0.1	0.97	∞
0.2	0.91	...
0.3	0.82	4.2
0.4	0.62	3.4
0.47	0.28	2.7
0.5	0.12	2.3

to the magnetic properties of the narrow band.

Since both sides of Eq. (23) are reduced by the correlation effect, it becomes crucial to know which side is more affected by electron correlation. The calculation for Ni indicates that under favorable condition the Stoner enhancement factor $R(l\mu)$ can be negative even though the numbers of double occupancies $\nu(i)$'s are nonzero. Therefore under similar condition it is not impossible for a small positive $R(l\mu)$ to satisfy Eq. (23). Nevertheless, the condition $R(l\mu) \leqslant 0$ is much simpler for computation although more stringent than Eq. (23). In fact $\nu(i)$ always approaches zero faster than the discontinuity in the occupation number $D(l)$, except for one electron per atom where $\nu(i)$ and $D(l)$ go to zero simultaneously. So the appearance of magnetic ordering is mainly due to the possibility of negative $R(l\mu)$. From Eq. (22), we can make the following conclusion. Strong correlation, small number of holes, and a large density of states at the top of the band where the Fermi energy lies favor the occurrence of magnetically ordered phase.

IV. SPIN WAVES

In the following we assume that Eq. (23) is satisfied, and that the ground state Ψ_c of H_0 is ferromagnetic with the majority of electrons having down spin.[21] To treat the spin waves in correlated itinerant electrons, we add the spin-flip exchange interaction to H_0. The suitable Hamiltonian is

$$H = \sum \epsilon(lk) n_{lk\sigma} + C_2 \sum n_{lg} n_{2g} + C_3 \sum n_{lg\uparrow} n_{lg\downarrow} + \frac{J}{2} \sum (1 - \delta_{l'l}) a^\dagger_{lg\sigma} a^\dagger_{l'g\sigma'} a_{lg\sigma'} a_{l'g\sigma} \quad . \tag{25}$$

The spin-wave creation operator is assumed to be a linear combination of the elementary excitations:

$$S^\dagger(\hbar\omega_q) = \sum_{lk} \lambda(lkq) a^\dagger_{lk+q\uparrow} a_{lk\downarrow} \quad . \tag{26}$$

The coefficients $\lambda(lkq)$ and the energy $\hbar\omega_q$ are obtained from solving the equation of motion

$$[H,\ S^\dagger(\hbar\omega_q)] = \hbar\omega_q S^\dagger(\hbar\omega_q) \quad . \tag{27}$$

It is generally known that electron correlation suppresses the random excitations and thus lowers the excitation energy. For instance, the plasmon energy is less than the excitation energies associated with the random electron density fluctuations. To illustrate how correlation improves the RPA solutions of the magnon energies, we consider as an example the simpler case of a nondegenerate band. In this case, the Hubbard Hamiltonian becomes

$$H_s = \sum \epsilon(k) n_{k\sigma} + C_0 \sum n_{g\uparrow} n_{g\downarrow} \quad .$$

The RPA solution of the magnon energy E_q in this system is determined by[22]

$$1 = \frac{C_0}{L} \sum_k \frac{n_{k\downarrow} - n_{k+q\uparrow}}{C_0[n(\downarrow) - n(\uparrow)] + \epsilon(k+q) - \epsilon(k) - E_q} \quad , \tag{28}$$

where the number operator $n_{k\sigma}$ is replaced by the Fermi distribution function. However, we have shown in Eq. (13) that correlated electron hoppings cause the mean value of $n_{k\sigma}$ to deviate from the Fermi function. Therefore, as a result of this effect $n_{k\sigma}$ should be replaced by its correlated mean value. Furthermore, the bare Coulomb energy C_0 in Eq. (28) is reduced by correlation to a smaller effective one.

In this section, we will study the modification of spin-wave energies due to the correlated hoppings only. If we use the transformation [Eq. (1)] to rewrite H in Bloch representation, then we have

$$[H,\ a^\dagger_{lk+q\uparrow} a_{lk\downarrow}] = [\epsilon(l, k+q) - \epsilon(lk)] a^\dagger_{lk+q\uparrow} a_{lk\downarrow} + \frac{1}{L} \sum_{k'} [C_3(n_{lk'\downarrow} - n_{lk'\uparrow}) + J(n_{l'k'\downarrow} - n_{l'k'\uparrow})] \times a^\dagger_{lk+q\uparrow} a_{lk\downarrow} - (n_{lk\downarrow} - n_{lk+q\uparrow}) \frac{1}{L} \sum_{k'} (C_3 a^\dagger_{lk'+q\uparrow} a_{lk'\downarrow} + J a^\dagger_{l'k'+q\uparrow} a_{l'k'\downarrow}) + \cdots \quad , \tag{29}$$

where $l' \neq l$. The dots represent the terms involving more than one magnon. These higher-order terms are neglected.

We will replace the operator $n_{lk\sigma}$ by its mean value in a correlated system. The mean values $\langle n_{lk\sigma} \rangle$ are given by Eq. (13). The notations $\delta\epsilon(lkq) = \epsilon(lk+q) - \epsilon(lk)$ and $\delta\langle n(lkq) \rangle = \langle n_{lk\downarrow} \rangle - \langle n_{lk+q\uparrow} \rangle$ are used to simplify the form of our equations. Combining Eqs. (26) and (29), we obtain

$$[H, S^{\dagger}(\hbar\omega_q)] = \sum_{lk}\left(\delta\epsilon(lkq) + \frac{1}{L}\sum_{k'}[J\delta\langle n(l'k'0)\rangle + C_3\delta\langle n(lk'0)\rangle]\right)\lambda(lkq)a^{\dagger}_{lk+q\uparrow}a_{lk\downarrow}$$
$$-\sum_{lk}\frac{1}{L}\sum_{k'}[J\delta\langle n(l'k'q)\rangle\lambda(l'k'q) + C_3\delta\langle n(lk'q)\rangle\lambda(lk'q)]a^{\dagger}_{lk+q\uparrow}a_{lk\downarrow} = \hbar\omega_q\sum_{lk}\lambda(lkq)a^{\dagger}_{lk+q\uparrow}a_{lk\downarrow} \quad , \tag{30}$$

where $l' \neq l$. Equation (30) yields the following coupled equations:

$$\lambda(lkq) = X(lq)\Big/\left(\delta\epsilon(lkq) + \frac{1}{L'}\sum_{k'}[J\delta\langle n(l'k'0)\rangle + C_3\delta\langle n(lk'0)\rangle] - \hbar\omega_q\right) \quad , \tag{31}$$

$$X(lq) = \frac{1}{L}\sum_{k}[J\delta\langle n(l'kq)\rangle\lambda(l'kq) + C_3\delta\langle n(lkq)\rangle\lambda(lkq)] \quad , \tag{32}$$

where $l' \neq l$.

By either using Eq. (13) or simply applying the conservation of particles, it is easy to see that

$$\frac{1}{L}\sum_{k}\delta\langle n(lkq)\rangle = \frac{1}{L}[N(\downarrow l) - N(\uparrow l)] = \Delta(l) \quad .$$

So Eq. (31) can be rewritten

$$\lambda(lkq) = \frac{X(lq)}{\delta\epsilon(lkq) + J\Delta(l') + C_3\Delta(l) - \hbar\omega_q} \quad . \tag{33}$$

Equations (32) and (33) readily imply a determinational condition $\|A_{ij}\| = 0$. The elements A_{ij} are

$$A_{11} = 1 - \frac{C_3}{L}\sum_{k}\frac{\langle n_{1k\downarrow}\rangle - \langle n_{1k+q\uparrow}\rangle}{\delta\epsilon(1kq) + J\Delta(2) + C_3\Delta(1) - \hbar\omega_q} \quad , \tag{34a}$$

$$A_{22} = 1 - \frac{C_3}{L}\sum_{k}\frac{\langle n_{2k\downarrow}\rangle - \langle n_{2k+q\uparrow}\rangle}{\delta\epsilon(2kq) + J\Delta(1) + C_3\Delta(2) - \hbar\omega_q} \quad , \tag{34b}$$

$$A_{12} = -\frac{J}{L}\sum_{k}\frac{\langle n_{2k\downarrow}\rangle - \langle n_{2k+q\uparrow}\rangle}{\delta\epsilon(2kq) + J\Delta(1) + C_3\Delta(2) - \hbar\omega_q} \quad , \tag{34c}$$

$$A_{21} = -\frac{J}{L}\sum_{k}\frac{\langle n_{1k\downarrow}\rangle - \langle n_{1k+q\uparrow}\rangle}{\delta\epsilon(1kq) + J\Delta(2) + C_3\Delta(1) - \hbar\omega_q} \quad . \tag{34d}$$

In the case of very weak correlation $C_3 = 0$, $\langle n_{lk\sigma}\rangle$ becomes the Fermi distirbution function. The determinantal condition then reduces to exactly the one obtained by Mattis.[8]

We consider first the zero-momentum magnons. For $q = 0$, $\delta\epsilon(lk0) = 0$. $\|A_{ij}\|_{q=0} = 0$ admits the solutions $\hbar\omega_0(A) = 0$ for the acoustic mode and $\hbar\omega_0(0) = J[\Delta(1) + \Delta(2)]$ for the optical mode. These energies are the same as their RPA values. We will show in Sec. V that $\hbar\omega_0(0)$ is reduced when the bare intra-atomic interaction energies are replaced by their effective values.

For convenience but without loss of generality, we assume that the two degenerate bands lie on top of each other. We then have $\delta\epsilon(1kq) = \delta\epsilon(2kq) = \delta\epsilon(kq)$, $\langle n_{1k\sigma}\rangle = \langle n_{2k\sigma}\rangle = \langle n_{k\sigma}\rangle$, and $\Delta(1) = \Delta(2) = \Delta$. From Eqs. (34), one sees that $A_{11} = A_{22}$ and $A_{21} = A_{12}$. The condition $\|A_{ij}\| = 0$ becomes

$$\frac{\pm J + C_3}{L}\sum_{k}\frac{\langle n_{k\downarrow}\rangle - \langle n_{k+q\uparrow}\rangle}{\delta\epsilon(kq) + (J + C_3)\Delta - \hbar\omega_q} = 1 \quad . \tag{35}$$

The plus and the minus signs correspond to the acoustic and optical branches, respectively.

The acoustic-branch solution of Eq. (35) can be simply written

$$\frac{J'}{L}\sum_{k}\frac{\langle n_{k\downarrow}\rangle - \langle n_{k+q\uparrow}\rangle}{\delta\epsilon(kq) + J'\Delta - \hbar\omega_q(A)} = 1 \quad , \tag{36}$$

where $J' = J + C_3$. In order to illustrate clearly the effect of correlation on the spin-wave energies, we consider the small-q magnons in a complete ferromagnetic state where $\langle n_{k\uparrow}\rangle = 0$. In this region, we can expand the left-hand side of Eq. (36) to second order in $\delta\epsilon(kq)$ to obtain

$$\hbar\omega_q(A) = \frac{1}{\Delta L}\sum_{k}\langle n_{k\downarrow}\rangle\left(\delta\epsilon(kq) - \frac{1}{J'\Delta}[\delta\epsilon(kq)]^2\right) \quad . \tag{37}$$

Substituting the mean value $\langle n_{k\downarrow}\rangle = [1 - D(\downarrow)]n(\downarrow) + D(\downarrow)f(k\downarrow)$ into this equation, we obtain

$$\hbar\omega_q(A) = D(\downarrow)\hbar\omega_q(A)^{\mathrm{RPA}} - \frac{2[1 - D(\downarrow)]}{\Delta L J'}\sum_{k}[\delta\epsilon(kq)]^2 \quad , \tag{38}$$

where

$$\hbar\omega_q(A)^{\mathrm{RPA}} = \frac{1}{\Delta L}\sum_{k}f(k\downarrow)\left(\delta\epsilon(kq) - \frac{1}{J'\Delta}[\delta\epsilon(kq)]^2\right) \quad . \tag{39}$$

$\hbar\omega_q(A)^{\mathrm{RPA}}$ consists of two competing terms: namely, a kinetic energy term and a superexchange term. Since the kinetic energy term is positive, the RPA always predicts ferromagnetism for $J' \to \infty$. Therefore, the RPA magnon dispersion curve has an increasing slope as J' gets larger.

However, it is well known that the RPA overestimates the magnon excitation energies. While most theories which improve the RPA solution assume some effective value of J', Roth[9] has computed the magnon energies using the s-band Hubbard model. She then derived from the magnon dispersion relation the effective exchange integral which one must use with the Heisenberg model to obtain the same spin-wave energy. What Roth has

obtained can be summarized as follows: Let $\hbar\omega_q^{\mathrm{RPA}} = C(J')^{\mathrm{RPA}} q^2$ and $\hbar\omega_q^{\mathrm{COR}} = C(J')^{\mathrm{COR}} q^2$ be the RPA solution and the solution including the correlation effect, respectively. Then both $C(J')^{\mathrm{RPA}}$ and $C(J')^{\mathrm{COR}}$ are monotonically increasing functions of J', and $C(J')^{\mathrm{RPA}} > C(J')^{\mathrm{COR}}$ for all values of J'.

It will be seen in the next section that Roth's conclusion follows from the fact that RPA allows too many atoms being doubly occupied. The RPA method therefore overestimates the magnon excitation energies. In this section, we will show the similar change of magnon energies due to the correlation effect on electron hoppings. We are interested in a strongly correlated ferromagnetic system which contains a finite fraction of doubly occupied sites. Such system corresponds to the ferromagnetic portion of region II in Fig. 1. In this case, the discontinuity in occupation number $D(\downarrow)$ in Eq. (38) is almost constant. For instance, $1 \geq D(\downarrow) \geq 0.6$ for ferromagnetic Ni for values of J' from zero to infinity. In region II, $D(\downarrow)$ varies roughly between 0.7 and 0.8.

Since the $D(\downarrow)$ in Eq. (38) can be approximated by a constant, $\hbar\omega_q(A)$ is then an increasing function of the intra-atomic interaction energy J'. Furthermore, for a given value of J', $\hbar\omega_q(A) \leq \hbar\omega_q(A)^{\mathrm{RPA}}$. The equal sign holds only if $D(\downarrow) = 1$, that is, if there is no correlation. Therefore, the magnon dispersion curve is flattened from its RPA solution. Note that the above conclusion is restricted to systems which have a finite discontinuity in the occupation number, $D(\sigma l)$. The model Hamiltonian [Eq. (2)] is inadequate to describe the magnon excitation as $D(\sigma l) \to 0$, where a metal-to-insulator transition occurs. This extreme case will be further discussed in Sec. V.

For the optical mode, a relation similar to but less accurate than Eq. (39) can also be derived. The $q = 0$ magnons have the striking property that in Eq. (26) the coefficients $\lambda(lk0) = \lambda(l)$ for all k. This is easily seen from Eq. (33). If q is small, we expect $\lambda(lkq) \simeq \lambda(lq)$ to be a fairly good approximation for all k. Equation (26) is then reduced to the simple form $S^\dagger(\hbar\omega_q) = \sum_l \lambda(lq) S^\dagger(lq)$, where $S^\dagger(lq) = \sum_k a^\dagger_{lk+q\uparrow} a_{lk\downarrow}$. Commuting H past the operator $S^\dagger(lq)$, we have

$$[H, S^\dagger(lq)] = \sum_k \delta\epsilon(lkq) a^\dagger_{lk+q\uparrow} a_{lk\downarrow} + J\Delta(l') S^\dagger(lq) - J\Delta(l) S^\dagger(l'q) \quad ,$$

where $l' \neq l$. Again we consider only the case of complete ferromagnetism and complete overlap of the two degenerate bands. Under this situation, the above commutator yields the following determinantal condition for $q \to 0$:

$$\begin{vmatrix} \frac{D(\downarrow)}{\Delta L} \sum_k f[\epsilon(k)]\, \delta\epsilon(kq) + J\Delta - \hbar\omega_q & -J\Delta \\ -J\Delta & \frac{D(\downarrow)}{\Delta L} \sum_k f[\epsilon(k)]\, \delta\epsilon(kq) + J\Delta - \hbar\omega_q \end{vmatrix} = 0 \quad .$$

This condition leads to the solutions

$$\hbar\omega_q(A) = \frac{D(\downarrow)}{\Delta L} \sum_k f[\epsilon(k)]\, \delta\epsilon(kq) \quad , \tag{40}$$

$$\hbar\omega_q(0) = \hbar\omega_q(A) + 2J\Delta \quad . \tag{41}$$

Equations (38) and (40) differ by a term which is inversely proportional to J'. Hence for strong correlation, the discrepancy becomes very small.

This result can be explained in physical terms as follows: If there is no correlation, $\langle n_{lk\downarrow} \rangle$ is just the Fermi distribution function $f(lk\downarrow)$. The contribution of the band energy to the magnon energy comes from the change of the number of occupied Bloch states in the hatched regions of Fig. 3. However, in the correlated system, some electrons are scattered from below to above the Fermi surface. The plotted $\langle n_{lk\downarrow} \rangle$ in Fig. 3 is the superposition of a constant $[1 - D(\downarrow l)] n(\downarrow l)$ and the $D(\downarrow l) f(lk\downarrow)$. Since the constant part of $\langle n_{lk\downarrow} \rangle$ does not contribute to the magnon energy, the only contribution to $\hbar\omega_q$ comes from the dotted region in Fig. 3. Therefore this correlation effect reduces the contribution of the band energy to $\hbar\omega_q$ by a factor $D(\downarrow l)$, and thus flattens the $\hbar\omega_q$-vs-q curve.

The coefficient C of the characteristic quadratic shape Cq^2 of the magnon dispersion curve for small q is thus roughly reduced to $D(\downarrow l)C$ by the correlation effect. For instance, in the case of a nearly half-filled band Edwards[11] has found that the reduction of C from its RPA value is more than 40%.

V. DISCUSSION

The Gutzwiller's variational method has been previously formulated in the atomic-orbital representation.[6] In this paper, we have reformulated the scheme, using the Wannier representation to achieve a simpler mathematical derivation. The qualitative features of these two alternations are quite similar. In the case of a degenerate band, the Wannier functions are less localized than the

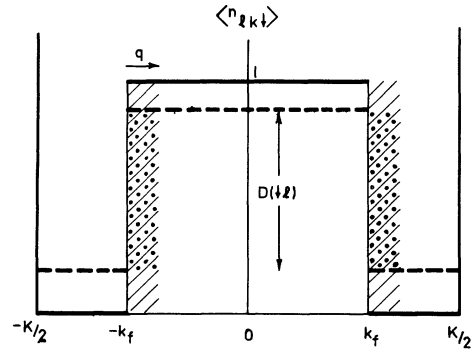

FIG. 3. Occupation probabilites for the electrons in reciprocal space along an arbitrary direction. The solid line is for uncorrelated system. k_f and k_{-f} mark the Fermi surface. $K/2$ and $-K/2$ indicate the zone boundary. q is the momentum of the magnon.

atomic orbitals. One may think then that the use of Wannier functions underestimates the intra-atomic interaction energies C_2, C_3, and J. Since there is technical difficulty to precisely determine C_2, C_3, and J, they have been treated as varying parameters in the Hubbard Hamiltonian. In this respect, the Wannier representation is adequate for the qualitative analysis.

The condition for the negative Pauli spin susceptibility in Ni, given in Table I, is a consequence of correlation effect on hopping energy. In order to calculate the correction term for this condition, it is necessary to perform a full scale computation for the Stoner enhancement factor $R(l\mu)$. Nevertheless, we can predict qualitatively how the critical value x_c in Table I should be altered. From Eq. (22) we see that the correction term (the third term in the bracket) is proportional to $C_3\eta(3)/D(l)$. Except for one electron per atom, the discontinuity in occupation number $D(l)$ is always greater than zero. Since $C_3 \to 0$ at sufficiently weak correlation and $\eta(3) \to 0$ at extremely strong correlation, the correction term contributes only in the strong-correlation region while $\nu(i)$ is still finite. Therefore, a large increase of x_c should occur at about $\eta(3) = 0.5$. No significant modification on x_c is expected at the region of small $\eta(3)$.

Owing to the mathematical difficulty, the computation of the Hubbard parameter C_0 in the third column of Table I is based on the s-band model. Hence this result should be considered as an estimation of the average value of C_2, C_3, and $C_1 = C_2 - J$.

In Sec. IV we have only considered the effect of correlated hoppings on the spin-wave energies. The bare intra-atomic interaction energies C_2, C_3, and J are used throughout the analysis from Eqs. (30)–(41). As we mention before, in a correlated system, C_2, C_3, and J should be replaced by the smaller effective values. Here we only discuss briefly how to reduce the bare interaction energies to the effective ones. The details of a complete treatment is reported elsewhere.[23]

It is more convenient for our present analysis if we express the single-particle excitations in Wannier representation;

$$S^{\dagger}(lkq) = a^{\dagger}_{lk+q\uparrow} a_{lk\downarrow} = \frac{1}{L}\sum_{g_1 g_2} \exp[-i(k+q)g_1 + ikg_2]\, a^{\dagger}_{lg_1\uparrow} a_{lg_2\downarrow} \quad . \tag{42}$$

The commutator of $S^{\dagger}(lkq)$ and the Coulomb interaction H_c can be easily obtained:

$$[H_c, S^{\dagger}(lkq)] = \frac{1}{L}\sum_{g_1 g_2}{}' \exp[-i(k+q)g_1 + ikg_2] \times C_3(n_{lg_1\downarrow} - n_{lg_2\downarrow}) a^{\dagger}_{lg_1\uparrow} a_{lg_2\downarrow} \quad . \tag{43}$$

The primed sum is restricted to $g_1 \neq g_2$ because of the exclusion principle. The terms in the restricted double sum are the specially correlated electron hoppings. When one electron hops from g_2 to g_1, it sees another electron at either g_1 or g_2.

The nontrivial components in $n_{lg_2} a^{\dagger}_{lg_1} a_{lg_2} |\Psi_c\rangle$ are those configurations [see the definition in Eq. (7)] containing two occupied Wannier states $\phi_{l\uparrow}(r-g_2)$ and $\phi_{l\downarrow}(r-g_2)$. So the mean value of n_{lg_2} must be proportional to the number $\nu(l\uparrow, l\downarrow)$ of such doubly occupied sites. Furthermore, these specially correlated hoppings may cause the numbers of double occupancies to deviate from their optimum values, depending on whether the final site g_1 is singly occupied or empty. If g_1 is empty, then the deviation is one. According to Gutzwiller's variational scheme, the occupation probability of the corresponding configuration is reduced by a factor $\eta(l\uparrow, l\downarrow)$.

When this correlation effect is taken into account in this way, one finds that C_3 is reduced to its effective value $C_3\eta(l\uparrow, l\downarrow)\Gamma(C_3)$. $\Gamma(C_3)$ is a slowly varying function which has the value between 0 and 1. By the same token, the effective exchange energy can be expressed as $J\eta(1\uparrow, 2\downarrow)\Gamma(J)$. These effective energies should appear in the matrix elements of Eqs. (34). We immediately see that the J' in the superexchange term of Eq. (38) is lowered by the correlation effect. So the magnon spectra are further flattened through the enhancement of the superexchange energy.

The energy of the zero-momentum optical magnon now becomes $\hbar\omega_0(0) = J\eta(1\uparrow, 2\downarrow)\Gamma(J)[\Delta(1) + \Delta(2)]$. Similarly, the Stoner gap parameters can be expressed as $J\eta(1\uparrow, 2\downarrow)\Gamma(J)\Delta(l)$. In RPA, $\Gamma(J) = 1$ and $\eta(1\uparrow, 2\downarrow) = 1$ and so the Stoner gap parameters become $J\Delta(l)$. Therefore, electron correlation also reduces the Stoner gap parameters

from their RPA values. In physical terms, this can be explained as follows. The zero-momentum optical magnon corresponds to the reverse of one of the two spins which are localized on the same atom. The correlation effect decreases the number of such atoms from its RPA value. Therefore the energy of the zero-momentum optical magnon, or the Stoner gap parameters, should be less than the corresponding RPA solution.

The results obtained in Secs. IV and V concerning the spin-wave energy are valid in region II of Fig. 1, where the correlation is fairly strong but a finite fraction of sites are doubly occupied. If the correlation is so strong that there is no doubly occupied atoms (in region III of Fig. 1), then there is no mechanism to couple together the elementary spin excitations in Hubbard model which neglects the interatomic interactions. Under this situation, the Hubbard model is inadequate to describe the magnon modes. One example is the case of one electron per atom. As we mentioned before, a metal-to-insulator transition occurs in such system if the correlation is sufficiently strong. If we neglect the interatomic interaction as well as the virtual hoppings, as in the Hubbard model, there is no mechanism to generate the spin waves in the insulating state. Since the discontinuity in occupation number, $D(\sigma l)$, vanishes only at the metal-to-insulator transition, one should not apply Eq. (38) to the systems of diminishing $D(\downarrow)$.

ACKNOWLEDGMENTS

The author wishes to thank Dr. M. C. Gutzwiller for valuable suggestions and comments during his stay at IBM where this work started. He would also like to thank Professor B. Goodman and Professor A. Sjölander, as well as Dr. B. I. Lundqvist for very helpful discussions. The critical reading of the manuscript by Dr. J. Morrison is gratefully acknowledged.

APPENDIX A

The $\nu(i)$'s can be more generally defined as $2\nu(1)=\nu(1\uparrow, 2\uparrow)+\nu(1\downarrow, 2\downarrow)$, $2\nu(2)=\nu(1\uparrow, 2\downarrow)+\nu(1\downarrow, 2\uparrow)$, and $2\nu(3)=\nu(1\uparrow, 1\downarrow)+\nu(2\uparrow, 2\downarrow)$. For given values of $\eta(i)$ and $N(\sigma l)$ let $\nu_0(\sigma l, \sigma' l')$ be the solution of Eq. (9). There are six coupled equations due to the six different assignments of $(\sigma l, \sigma' l')$. For paramagnetic state $N(\uparrow l)=N(\downarrow l)=\frac{1}{2}N(l)$. Then because of the symmetry properties only four of these six equations are linearly independent.

We substitute $N(\uparrow l)=\frac{1}{2}N(l)+\zeta(l)$, $N(\downarrow l)=\frac{1}{2}N(l)-\zeta(l)$, and $\nu(\sigma l, \sigma' l')=\nu_0(\sigma l, \sigma' l')+\delta\nu(\sigma l, \sigma' l')$ into Eq. (9). Then the so obtained six equations are combined pairwisely as follows: The right-hand sides as well as the left-hand sides of the two equations involving the same $\eta(i)$ are added up, respectively. This sequence of algebraic manipulations yields the following results:

$$2\{[\tfrac{1}{4}N-\nu(1)-\nu(2)-\nu(3)]\eta(i)^2+\nu(i)\}[\delta\nu(1)+\delta\nu(2)+\delta\nu(3)]+[L-N+2\nu(1)+2\nu(2)+2\eta(3)]\delta\nu(i)=\eta(i)^2h(i)\ ,\quad i=1,\ 2,\ 3 \tag{A1}$$

where $h(1)=-h(2)=-\zeta(1)\zeta(2)$ and $h(3)=-\frac{1}{2}[\zeta(1)^2+\zeta(2)^2]$. To obtain Eq. (A1), we have assumed $|N(1)-N(2)|\ll N$ and neglected the higher-order terms in ζ.

Again we add up the three equations which are obtained by substituting $i=1$, 2, and 3 into Eq. (A1). Using the notation $\nu=\nu(1)+\nu(2)+\nu(3)$, the resulting equation can be written

$$\{(\tfrac{1}{2}N-2\nu)[\eta(1)^2+\eta(2)^2+\eta(3)^2]+L-N+4\nu\}\delta\nu=[\eta(1)^2-\eta(2)^2]\zeta(1)\zeta(2)-\tfrac{1}{2}\eta(3)^2[\zeta(1)^2+\zeta(2)^2]\ . \tag{A2}$$

$\delta\nu(i)$ are then obtained by substituting $\delta\nu$ from Eq. (A2) in Eq. (A1). The so obtained $\delta\nu(i)$ are valid for the entire range of correlation strength. For strong correlation, $(\frac{1}{4}N-\nu)\eta(i)^2$ and $\nu(i)$ are much less then $L-N+2\nu$. If we neglect these small quantities, then we have the simple solutions for strong correlation region as $\delta\nu(i)=\eta(i)^2h(i)(L-N+2\nu)^{-1}$.

$D(\sigma l)$ is computed to the second order in ζ by the same technique. Equation (18) is readily obtained under the condition of strong correlation.

APPENDIX B

Consider a two-variable quadratic function

$$f(x,\ y)=Ax^2-2Bxy+Cy^2\ . \tag{B1}$$

A clockwise rotation of the coordinate axis through an angle $\theta=\frac{1}{2}\tan^{-1}[2B/(A-C)]$ will transfer $f(x,\ y)$ into

$$f(u,\ v)=\tfrac{1}{2}(A+C+\alpha)u^2+\tfrac{1}{2}(A+C-\alpha)v^2\ , \tag{B2}$$

where $\alpha=[(A-C)^2+4B^2]^{1/2}$. Let us choose the particular values x_0 and y_0 such that $x_0=y_0\tan\theta$, then the corresponding u_0 vanishes. Therefore $f(u=0,\ v)\le 0$ if $A+C-\alpha\le 0$, or $AC\le B^2$. Comparing Eqs. (B1) and (19), we see that Eq. (23) is the condition for $\delta\langle H_0\rangle\le 0$.

We are particularly interested in the possibility of negative $r(l\mu)$ which is defined by Eq. (24). The negative $r(l\mu)$ will surely make $\delta\langle H_0\rangle\le 0$. Consider the density of states per atom per eV per spin

$$\rho(lE)=(1-x)/(1-x/y)W \quad \text{for } -(y-x+xy)W/2y\le E<-(x-y+xy)W/2y$$
$$=y/W \quad \text{for } -(x-y+xy)W/2y\le E\le(x+y-xy)W/2y\ , \tag{B3}$$

where x and y are adjustable parameters. The $\rho(lE)$ is plotted in Fig. 1 as a function of E. $\rho(lE)$ is so normalized that a completely filled band has

zero energy. Note that we have assumed the same density of states curve for both bands.

We restrict our discussion to small number of holes such that the Fermi energy of a paramagnetic state lies in the high-density-of-states region. The Fermi energy μ and the mean hole energy $\bar{\epsilon}(l)$ can be easily computed as

$$\mu = (x + y - xy - \tfrac{1}{2}n)W/2y \quad , \tag{B4}$$

$$\bar{\epsilon}(l) = -(x + y - xy - \tfrac{1}{4}n)W/2y \quad , \tag{B5}$$

where n is the hole density which is assumed to be less than $4x$.

For Ni, $y = 4.5$[20] and $n = 0.6$. Substituting μ, $\bar{\epsilon}$, and $\rho(l\mu)$ into Eq. (24), the condition for $r(l\mu) \leq 0$ is obtained in terms of $\eta(3)$ and the critical values x_c. The results are listed in Sec. III.

---

*Present address: Department of Physics, University of Warwick, Coventry, England.

[1]For a general discussion, see the review articles of J. H. Van Vleck, Rev. Mod. Phys. **25**, 220 (1953); E. P. Wohlfath, Rev. Mod. Phys. **25**, 211 (1953); J. C. Slater, Rev. Mod. Phys. **25**, 199 (1953); C. Zener and R. R. Heikes, Rev. Mod. Phys. **25**, 191 (1953).

[2]J. Hubbard, Proc. R. Soc. Lond. **276**, 238 (1963); Proc. R. Soc. Lond. **277**, 238 (1964); Proc. R. Soc. Lond. **285**, 542 (1965).

[3]M. C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963); Phys. Rev. **134**, 923 (1964); Phys. Rev. **137**, 1726 (1965).

[4]J. Kanamoki, Prog. Theor. Phys. **30**, 275 (1963).

[5]Y. Nagaoka, Phys. Rev. **147**, 392 (1966); A. B. Harris and R. V. Lange, Phys. Rev. **157**, 295 (1967); R. A. Bari, D. Adler, and R. V. Lange, Phys. Rev. B **2**, 2898 (1970); W. F. Brinkman and T. M. Rice, Phys. Rev. B **2**, 4302 (1970); Battelle Symposium on Critical Phenomena, 1971 (unpublished).

[6]K. A. Chao and M. C. Gutzwiller, J. Appl. Phys. **42**, 1420 (1971); K. A. Chao, Phys. Rev. B **4**, 4034 (1971).

[7]C. Herring and C. Kittel, Phys. Rev. **81**, 869 (1951); C. Herring, Phys. Rev. **85**, 1003 (1952); Phys. Rev. **87**, 60 (1952).

[8]T. Izuyama, Prog. Theor. Phys. **23**, 969 (1960); D. C. Mattis, Phys. Rev. **132**, 2521 (1963); E. D. Thompson, Adv. Phys. **14**, 213 (1965); D. M. Edwards, Proc. R. Soc. Lond. **269**, 338 (1962); T. Izuyama and R. Fubo, J. Appl. Phys. **35**, 1074 (1964); J. F. Cornwell, Proc. R. Soc. Lond. **279**, 346 (1964); Proc. R. Soc. Lond. **284**, 423 (1965).

[9]L. M. Roth, J. Phys. Chem. Solids **28**, 1549 (1967).

[10]H. Yamada and M. Shimizu, J. Phys. Soc. Jap. **22**, 1404 (1967).

[11]D. M. Edwards, Proc. R. Soc. Lond. **300**, 373 (1967).

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

[13]W. Young and J. Callaway, J. Phys. Chem. Solids **31**, 865 (1970).

[14]K. Y. Chan and W. Young, J. Phys. C **4**, 1855 (1971).

[15]In degenerate band model, the Wannier states are not as localized as the atomic orbitals. However, at least for the present work, we found that the Wannier representation and the atomic-orbital representation admit similar analytical expressions. Therefore we can use the Wannier representation to avoid algebraic complexity. For numerical analysis, one must compute the intra-atomic interaction energies from the atomic orbitals. See the discussion in Sec. V.

[16]The notations C_2 and C_3 are used here in order to be consistent with the notations C_1, C_2, and C_3 in our previous papers, Ref. 6. In fact, $J = C_2 - C_1$.

[17]H. Hurwitz, thesis (Harvard University, 1941) (unpublished); J. H. Van Vleck, Rev. Mod. Phys. **17**, 42 (1945).

[18]For its definition, see Ref. 6.

[19]L. M. Roth, J. Phys. Chem. Solids **28**, 1549 (1967), Eq. (27); A. B. Harris and R. V. Lange, Phys. Rev. **157**, 295 (1967), Eq. (6.9).

[20]Kanamori uses $y = 3$ for his computation, which is based on the triply degenerate T_{2g} subband. Since we are dealing with the doubly degenerate e_g subband, y should be 4.5 in our case.

[21]Strictly speaking, if Eq. (23) is satisfied, the magnetically ordered state may be either ferromagnetic or antiferromagnetic. The antiferromagnetic phase will not be discussed in this paper.

[22]T. Izuyama, Prog. Theor. Phys. **23**, 969 (1960).

[23]K. A. Chao, Solid State Commun. **11**, 1633 (1972).