Effect of band structure on ferromagnetism*

D. C. Mattis and R. E. Peña
Belfer Graduate School of Science, Yeshiva University, New York, New York 10033
(Received 5 November 1973)

We extend Nagaoka's study of the ferromagnetism of nearly half-filled bands in the infinite-repulsion limit of the Hubbard model by including next-nearest-neighbor tight-binding overlap matrix elements K_2. Particles can now get past one another, even in one dimension. We find corroboration of Nagaoka's results, viz., either possibility or impossibility of ferromagnetism depending on the relative sign and magnitude of K_2.

I. INTRODUCTION

Some years ago, Nagaoka[1] studied the occurrence of ferromagnetism in half-filled bands of strongly interacting electrons. Starting with the well-known fact that bands half-filled with strongly mutually repelling electrons tend to be *antiferromagnetic*, he showed how, in certain cases, the effect of a few extra electrons or a few missing electrons (i.e., extra holes) was to stabilize a *ferromagnetic* ground state. The analysis had some puzzling features, notably a dependence on lattice geometry. The simple cubic (sc) and body-centered-cubic (bcc) lattices behaved differently from the face-centered-cubic (fcc) and hexagonal-close-packed (hcp) lattices in their dependence on the sign of the hopping matrix element t and on the sign of the occupation-number parameter $n = N - N_{el}$.

Over the years, many of Nagaoka's results and conjectures have been confirmed or extended[2] by various means, although the dependence of ferromagnetism on the crystal structure, which he obtained, has not heretofore been systematically investigated. We have therefore studied this phenomenon and conclude it to be closely related to the electron energy-band structure, which, in the tight-binding approximation, itself strongly depends on the crystal structure. In particular, the hole-particle symmetry appears to play a nontrivial role. It is a symmetry present in so-called "bipartite lattices," i.e., lattices which can be subdivided so that any point on one sublattice has its nearest neighbors on the other, such as with sc and bcc but not with fcc and hcp.

Our procedure is based on the following idea: the prototype bipartite lattice is the linear chain with only nearest-neighbor (NN) hopping matrix elements, denoted t_1. We introduce next-nearest-neighbor (NNN) hopping matrix elements t_2, thus destroying the bipartite character. The analog of sc and bcc is $t_2 = 0$. Correspondingly, the analog to fcc and hcp is $t_2 \approx t_1$. The ferromagnetism (or lack thereof) will be seen to depend on the relative signs of the t parameter and the n parameter. (We actually study the more general case of t_2 having arbitrary sign and magnitude with respect to t_1 in the calculation outlined below.)

II. STRONG-COUPLING HAMILTONIAN

For a less-than-half-filled band ($n \geq 0$) we have a Hamiltonian[1,3]

$$\mathcal{H}_a = \sum_{ij} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} + I \sum_i n_{i\uparrow} n_{i\downarrow} , \tag{1a}$$

in which we proceed to the limit $I \to \infty$, in which no atomic site can be doubly occupied. For a more-than-half-filled band, $|n|$ atomic sites are perforce doubly occupied, and it is desired to limit the double occupancy to that irreducible value. Thus, for $n < 0$ we adopt a Hamiltonian

$$\mathcal{H}_b = \sum_{ij} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} + I \sum_i n_{i\uparrow} n_{i\downarrow} + nI , \tag{1b}$$

in which we can once again proceed to the limit $I \to \infty$.

We can collapse the above into a unique Hamiltonian of type (1a). Explicitly for the linear chain, we have

$$\mathcal{H} = -K_1 \sum_{m\sigma} (c_{m\sigma}^\dagger c_{m+1\sigma} + \text{H.c.})(1 - n_{m,-\sigma})(1 - n_{m+1,-\sigma}) - K_2 \sum_{m\sigma} (c_{m\sigma}^\dagger c_{m+2\sigma} + \text{H.c.})(1 - n_{m,-\sigma})(1 - n_{m+2,-\sigma}) , \tag{2}$$

in which we have proceeded to the $I = \infty$ limit by use of projection operators. The K's are related to the t's as follows:

$$K_1 = |t_1| \quad \text{and} \quad K_2 = -(n/|n|)\, t_2 . \tag{3}$$

The ratio parameter $r \equiv K_2/K_1$ will be important in the calculation.

III. PROCEDURE

Following Nagaoka's procedure, we estimate the trend to ferromagnetism by obtaining an exact solution for $n = \pm 1$. The case of $n = -1$ is related to $n = +1$ by Eq. (3); therefore, $n = +1$ is all we study. The maximum spin in that case is $S = S_{max} = \frac{1}{2}(N-1)$. In the subspace of $S_z = S_{max} = \frac{1}{2}(N-1)$, in which the spin-up band is filled except for one hole and the spin-down band is empty, the exact eigenvalues are given by the usual tight-binding band structure of a *hole*:

$$\epsilon(k) = 2K_1 \cos k + 2K_2 \cos 2k \ . \tag{4}$$

The eigenfunctions are plane waves representing a traveling hole. In the subspace of $S_z = S_{\max} - 1 = \frac{1}{2} \times (N-3)$, we expect to find representatives of all the eigenfunctions and eigenvalues (4) as well as eigenfunctions belonging to total spin $S = S_{\max} - 1$. If any eigenvalue in this subspace lies lower than the lowest value of (4), then we have shown that the energy can be lowered by lowering S, and ferromagnetism is then presumed to be unstable. If, however, it is found that no state belonging to $S = S_{\max} - 1$ lies lower than the lowest value of (4), then the presumption is that ferromagnetism is stable. We need not repeat the arguments adduced by Nagaoka to justify this procedure, but it is convenient to rephrase them in terms of scattering theory.

We can visualize the $S_z = S_{\max} - 1$ subspace for the case of $n = 1$ as a three-body scattering problem, of two holes in the up-spin band and one particle in the down-spin band. The $I = \infty$ limit obliges the particle to be located on one of the holes. A plane wave constituted out of such a particle-hole pair is a *spin wave*, and it is expected that there be a continuum of energy levels describing the scattering between this spin wave and the second hole. The spin waves carry no energy *per se* (as we can see for $n = 0$; in this case of a half-filled band, states of all possible values of S are degenerate, which also implies that spin waves have no intrinsic energy). Thus, in the case $n \neq 0$ the energy of spin waves comes from the scattering with the charge-carrying holes. Because of the lack of dispersion, wave packets can be constructed out of spin waves that are localized to an arbitrary degree or that travel near the moving hole.

If the scattering of the hole on the spin waves is *repulsive*, there is no energetic advantage in the creation of spin-waves, and the ferromagnetic state cannot be ruled out. The true ground state in that case does not necessarily have total spin as large as $S_{\max}$, but it is plausible that $S \neq 0$ in the ground state. If, however, the scattering is attractive, then there is the additional possibility of a bound state with energy below the continuum, hence below any value of (4). The existence of such a bound state signifies the instability of the totally ferromagnetic state and makes a nonmagnetic $S = 0$ ground state created by the reversal of a large number of electron spins a most plausible occurrence.

Thus, the state which precludes ferromagnetism can be visualized as a stable bound complex formed by a charge-carrying hole with a spin wave localized around it. Stability is the consequence of a finite binding energy relative to the continuum of scattering states.

IV. THREE-BODY EIGENSTATES

To isolate the effective particles in the model, we transform the up-spin particles from electrons to holes, i.e., interchange $c_{m\uparrow}$ and $c_{m\uparrow}^\dagger$, bringing the Hamiltonian (2) into the form

$$\begin{aligned}\mathcal{H} = K_1 \sum_m \, &[c_{m\uparrow}^\dagger c_{m+1\uparrow}(1-n_{m\downarrow})(1-n_{m+1\downarrow}) \\ &- c_{m\downarrow}^\dagger c_{m+1\downarrow} n_{m\uparrow} n_{m+1,\uparrow}] + \text{H.c.} \\ &+ K_2 \sum_m \, [c_{m\uparrow}^\dagger c_{m+2\uparrow}(1-n_{m\downarrow})(1-n_{m+2\downarrow}) \\ &- c_{m\downarrow}^\dagger c_{m+2\downarrow} n_{m\uparrow} n_{m+2,\uparrow}] + \text{H.c.}\end{aligned} \tag{5}$$

The Schrödinger equation for our problem may be written

$$\mathcal{H} \,|\, \Psi\rangle = E \,|\, \psi\rangle \ . \tag{6}$$

An exact eigenstate $|\,\Psi\rangle$ must take the following form:

$$|\,\Psi\rangle = \sum_{l,m} F(l,m)\, c_{m\downarrow}^\dagger c_{l\uparrow}^\dagger c_{m\uparrow}^\dagger \,|\, 0\rangle \ , \tag{7}$$

in which $F(l, m)$ are amplitudes to be determined. The operator $c_{m\downarrow}^\dagger c_{m\uparrow}^\dagger$ represents a wave packet of spin waves which, in the limit $I = \infty$, can be localized at any arbitrary site m. The operator $c_{l\uparrow}^\dagger$ represents the traveling hole. It is interesting to note that the spin-wave packet $c_{m\downarrow}^\dagger c_{m\uparrow}^\dagger$ carries no charge. Making use of the following orthogonality relation:

$$\langle 0 \,|\, c_{i\uparrow} c_{j\uparrow} c_{i\downarrow} c_{m\downarrow}^\dagger c_{l\uparrow}^\dagger c_{m\uparrow}^\dagger \,|\, 0\rangle = \delta_{im}\,\delta_{lj}(1-\delta_{ij}) \ , \tag{8}$$

and of some algebra, we obtain the Schrödinger equation in the form

$$\begin{aligned} &EF(i,j)(1-\delta_{ij}) \\ &\quad = \left[\sum_{l \neq j} K_{il} F(l,j) + K_{ij} F(j,i)\right](1-\delta_{ij}) \ ,\end{aligned} \tag{9}$$

where $K_{ij} = K_1$ if $|\, j - i\,| = 1$, $K_{ij} = K_2$ if $|\, j-i\,| = 2$, and $K_{ij} = 0$ otherwise. It should be remarked that Eq. (9) makes no statement about the $F(i, i)$, and the equations for the $F(i, j \neq i)$ do not involve these diagonal terms. Thus, we may "cancel" the $(1 - \delta_{ij})$ terms in Eq. (9), making sure we retain the restriction on the summation on the right-hand side. For the form of the $F(i, j)$, we make the following *ansatz* based on momentum conservation:

$$F(i,j) = \exp[\tfrac{1}{2} i (R_i + R_j) Q] f(R_i - R_j) \ . \tag{10}$$

Equation (9) could have been derived and solved in any number of dimensions and for arbitrary K_{ij}. For simplicity, we specialize to the linear chain with NNN interactions. With an indifferent choice of origin, the equations may then be cast into the form

$$Ef(n)=K(n)f(-n)+\sum_{n'}K(n')\exp[(\tfrac{1}{2}iQn')] \times f(n+n')(1-\delta_{n',-n}) \ . \qquad (11)$$

By $K(n)$ we mean $K_{p,p+n}$. This is the principal equation which we propose to study below. Because $f(n)$ has a simple form (single exponential or single standing wave) *only* in the special case of NN interactions, we use the general method of Fourier transforms to effect a solution:

$$f(n)\equiv\sum_q f_q e^{iqn} \ , \quad -\pi\le q\le\pi \ , \qquad (12)$$

where we have the usual

$$\frac{1}{N}\sum_{n=-1/2N}^{1/2N} e^{i(q-q')n}=\delta_{qq'} \ .$$

These expressions in (11) yield

$$f_q=\frac{1}{N}\ \frac{\sum_{q'}[\in(q+q')-\in(q+\frac{1}{2}Q)]f_{q'}}{E-\in(q+\frac{1}{2}Q)} \ , \qquad (13)$$

where

$$\in(q)=\sum_{n=-1/2N}^{1/2N}K(n)e^{iqn}=2K_1\cos q+2K_2\cos 2q \ . \qquad (14)$$

The sign of K_1 is intrinsically positive. (If initially negative, it can *always* be made positive by the transformation $c_{n\sigma}\to c_{n\sigma}e^{i\pi n}$, without any other parameter being modified.) The sign and magnitude of K_2 relative to K_1 will, however, be an important factor in the results. By varying K_2 we are, in effect, modifying the band structure.

The inhomogeneous scattering solutions of (13) have energies that interlace the unperturbed eigenvalues, i.e., span the continuum $E=\in(k)$, where $k=q+\frac{1}{2}Q$ can take on any value in the range $-\pi$, $+\pi$. Thus, they are correctly visualized as the superposition of a charge-carrying hole [Eq. (4)] and a spin wave (the latter carrying no energy in the case of a nearly half-filled band in the limit $I=\infty$). The energy shift due to the scattering interaction never exceeds the spacing $O(1/N)$ between unperturbed energy levels, so if we wished to study these states, it is the phase shifts we should study. Such analysis would, however, distract us from the primary purpose, which is to establish the conditions for the existence of a bound state below the continuum.

We therefore now seek a self-consistent homogeneous (bound-state) solution to Eq. (13). In this process we will obtain, for general Q, five coupled homogeneous equations in the five unknowns defined as follows:

$$\langle 1\rangle\equiv\frac{1}{N}\sum_q f_q \ ,$$

$$\langle\cos nq\rangle\equiv\frac{1}{N}\sum_q\cos(nq)f_q \ , \quad n=1,2, \qquad (15)$$

$$\langle\sin nq\rangle\equiv\frac{1}{N}\sum_q\sin(nq)f_q \ , \quad n=1,2 \ .$$

The coefficients appearing in these coupled equations are of the form

$$\Delta(f_n,g_m)=\frac{1}{N}\sum_q\frac{f_n(q)g_m(q)}{E-\in(q+\frac{1}{2}Q)} \ ,$$

$$\Delta(f_n)=\frac{1}{N}\sum_q\frac{f_n(q)}{E-\in(q+\frac{1}{2}Q)} \ , \qquad (16)$$

$$\Gamma(f_n)=\frac{1}{N}\sum_q\frac{f_n(q)\in(q+\frac{1}{2}Q)}{E-\in(q+\frac{1}{2}Q)} \ ,$$

where $f_n(q)$, $g_n(q)$ may be any of the four trigonometric functions

$$C_n\equiv\cos nq \ , \quad S_n\equiv\sin nq \ , \quad n=1,2 \ .$$

A solution will have been found when the determinant of the 5×5 matrix of coefficients vanishes. We will not exhibit that matrix here explicitly, except for the two most interesting cases: $\frac{1}{2}Q=0$ and $\frac{1}{2}Q=\pi$, in which many matrix elements vanish. This follows from the fact [see (4)] that for these choices of Q,

$$\in(q+\tfrac{1}{2}Q)=\in(-q+\tfrac{1}{2}Q) \ , \qquad (17)$$

and, therefore,

$$\Delta(S_n,C_m)=\Delta(S_n)=\Gamma(S_n)=0 \ , \quad n,m=1,2 \ . \qquad (18)$$

Under these conditions, the determinant factors and the determinantal condition become

$$\det(A)\det(B)=0 \ , \qquad (19)$$

where

$$A=\begin{pmatrix}1+\Gamma(1) & -2K_1\Delta(C_1) & -2K_2\Delta(C_2)\\ \Gamma(C_1) & 1-2K_1\Delta(C_1,C_1) & -2K_2\Delta(C_1,C_2)\\ \Gamma(C_2) & -2K_1\Delta(C_2,C_1) & 1-2K_2\Delta(C_2,C_2)\end{pmatrix}, \qquad (20)$$

and

$$B=\begin{pmatrix}1+2K_1\Delta(S_1,S_1) & 2K_2\Delta(S_1,S_2)\\ 2K_1\Delta(S_2,S_1) & 1+2K_2\Delta(S_2,S_2)\end{pmatrix} \ . \qquad (21)$$

The sums (15) and (16) occurring in these determinants can be replaced by integrations, $N^{-1}\sum_q\to(2\pi)^{-1}\int_{-\pi}^{\pi}dq$, in the case of the bound state (finite integrands).

V. NUMERICAL ANALYSIS

We used a computer to evaluate the integrals which are the limiting values as $N\to\infty$ of the sums occurring in matrices A and B. Because of time

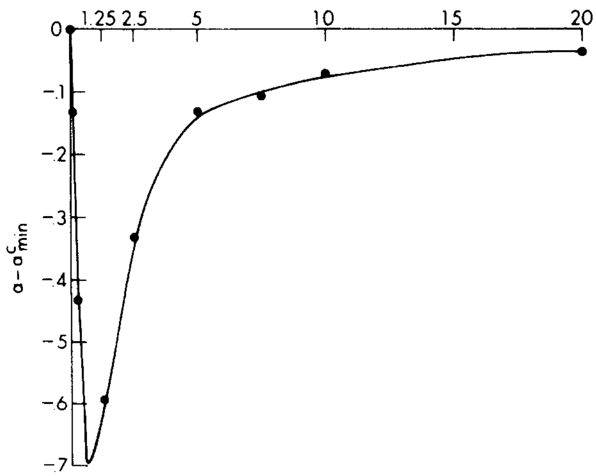

FIG. 1. Computed binding energy of an eigenstate belonging to $S_{max}-1$, as measured from the bottom of the continuum of states belonging to S_{max} and expressed in units of $|K_1|$. It is denoted $a-a^c_{min}$ and plotted as a function of $r=K_2/|K_1|$ for positive r. For $r<0$ there is no bound state, and ferromagnetism is stable. The relation of r to the hopping matrix elements t_1 and t_2 is given in Eq. (3) of the text.

considerations, the integrals held to within an accuracy of 6%. However, we did check a few scans to accuracies of up to 0.1% and found no significant discrepancies, except possibly when $K_2\to 0$. For appropriate choices of the ratio parameter $r\equiv K_2/|K_1|$, we scanned over energies $a=E/|K_1|$, which lay below the continuum minimum a^c_{min}. For each value of r, we obtained the determinants of A and B, scanned over a reasonable field of energy values. A bound state would appear as a change in the sign of either determinant between two successive values of the energy. Both the cases $\frac{1}{2}Q=0$ and $\frac{1}{2}Q=\pi$ were considered. Expected asymptotic behavior was verified. The results are as follows.

a. Matrix A. Its determinant does not vanish for any value of $a<a^c_{min}$ for all choices of r. It is finite for all $a<a^c_{min}$, although it may develop an apparent singularity in the limit $a\to a^c_{min}$.

b. Matrix B. Its determinant exhibits a zero (a single bound state) of energy $a<a^c_{min}$ for all r, as far as we have been able to determine, in the range $r>0.02$. The accompanying Fig. 1 displays the binding energy of the bound state as a function of r. For $r<0$, no bound states are found. The bound-state energy which we calculate has the same value for $\frac{1}{2}Q=\pi$ as for $Q/2=0$, a result which is a consequence of the structure of Eq. (21). While we have been unable to find a bound state in the range $0<r<0.02$, this may be an artifact of the numerical calculation. In any case, the binding is a maximum at $r\approx O(1)$, the value appropriate for comparison with three dimensions, as discussed in Sec. I.

VI. DISCUSSION AND CONCLUSIONS

First recall some well-known facts concerning the linear chain. The linear chain of electrons with just NN hopping but arbitrary two-body interactions was long ago proved to have a ground state of *minimum* spin[4]: $S=0$ or $\frac{1}{2}$, depending on whether the total number of electrons N_{el} was even or odd. Exact solutions of the linear-chain Hubbard model with finite interaction have fully confirmed this.[5] It is only in the pathologic limit $I=\infty$ that a spin degeneracy develops; in that limit, every eigenstate is the product of a determinant of plane waves and of a spin function, so that *every* eigenvalue belongs to *any* possible spin value ranging from $S=0$ or $\frac{1}{2}$ to S_{max}. This is analogous, insofar as the magnetic properties are concerned, to a number N_{el} of localized noninteracting spins-$\frac{1}{2}$, and is as close to ferromagnetism as the theorem permits in one dimension.

With NNN hopping introduced into the Hamiltonian, there are no previous calculations or exact results to fall back upon, and, moreover, the minimum-spin theorem does not apply.[6] We have been able to obtain the exact eigenstates only in the limit $I=\infty$ with $n=\pm 1$ and $S_z=S_{max}$ or $S_{max}-1$. As we shall see below, our results clarify Nagaoka's. Note that our method could be applied to two or three dimensions as well, and to arbitrary range hopping. That is, Eq. (9) with *ansatz* (10) can be Fourier transformed in any number of dimensions, leading to a set of equations not unlike (15) or (16). The main impediment is computational. We found the number of equations to rise rapidly with increasing dimensionality and range of hopping, and the integrals over two- or three-dimensional Brillouin zones are seen to present computational difficulties. Nevertheless, it is conceptually feasible, if computationally impractical, to extend the present work by, e.g., introducing NNN hopping into a sc lattice.

If, in our calculation, we require $t_2=t_1\equiv t$ to simulate the NN fcc and hcp lattice and $t_2=0$ (i.e., $r=0$) to simulate the sc and bcc lattices, then, combining the results summarized in Fig. 1 with the definitions (3), we may reasonably conclude

$$\begin{aligned}&\text{sc and bcc: not nonmagnetic,}\\&\text{fcc and hcp: nonmagnetic for } (N_{el}-N)\times t>0\\&\qquad\qquad\text{magnetic for } (N_{el}-N)\times t<0\;, \quad (22)\end{aligned}$$

which is not in disaccord with Nagaoka. We have distinguished between the cases where hole and spin wave do not interact at all (we call this "not nonmagnetic") and those in which they actually repel (this case is "magnetic" since the addition of holes inhibits the creation of spin waves, hence stabilizing a macroscopic value of S). The cases where hole and spin wave attract evidently imply a lack of ferromagnetism, as we have extensively discussed in Sec. III, and are "nonmagnetic."

*Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR-72-2153C.

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

[2]D. J. Thouless, Proc. Phys. Soc. Lond. 86, 893 (1965); L. Roth, J. Phys. Chem. Solids 28, 1549 (1966); D. M. Edwards, Proc. R. Soc. A 300, 373 (1967); L. Roth, Phys. Rev. 186, 428 (1969).

[3]J. Hubbard, Proc. R. Soc. A 276, 238 (1963).

[4]E. Lieb and D. Mattis, Phys. Rev. 125, 164 (1962); *Mathematical Physics in One Dimension* (Academic, New York, 1966), Chap. 4.

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

[6]The theorems developed in Ref. 4 are not applicable because with NNN hopping, particles of parallel spin can "get past" one another.