Extension of Nagaoka's theorem on the large-U Hubbard model

Hal Tasaki
Department of Physics, Gakushuin University, Mejiro, Toshima-ku, Tokyo 171, Japan
(Received 21 June 1989)

An extension is given of Nagaoka's theorem on the existence of ferromagnetism in the large-U Hubbard model with precisely one hole. The present extension covers a large class of models with arbitrary non-negative hopping matrix elements and arbitrary spin-independent interactions.

I. INTRODUCTION AND MAIN RESULTS

In 1965 Nagaoka proved a theorem[1] on the almost-half-filled Hubbard model with infinitely large Coulomb repulsion U. In certain models he proved that the ferromagnetic state becomes the ground state when there exists precisely one hole.[2] Although Nagaoka proved the theorem only in regular systems where the ground-state energy can be explicitly calculated, we found that only the following four conditions are necessary to state the theorem: (i) Each hopping matrix element t_{ij} is non-negative; (ii) there exists precisely one hole; (iii) Coulomb repulsion is infinitely large; (iv) the configuration space is connected. Therefore we can treat the models with arbitrary hopping matrix elements and arbitrary spin-independent interactions, where both can be irregular and long range.

Let Λ be a set of N sites. For convenience we order the elements in Λ to identify each site with an integer $i\in\{1,2,\ldots,N\}$. We study the Hubbard model defined by the following Hamiltonian:

$$H=\sum_{i,j}t_{ij}(c_{i\uparrow}^{\dagger}c_{j\uparrow}+c_{i\downarrow}^{\dagger}c_{j\downarrow})+V(\{n_{i\uparrow}+n_{i\downarrow}\})+U\sum_{i}n_{i\uparrow}n_{i\downarrow}\ . \tag{1}$$

Here $c_{i\sigma}^{\dagger}$ and $c_{i\sigma}$ are the standard fermion operators which create and annihilate, respectively, the electron at site i with spin σ, and $n_{i\sigma}=c_{i\sigma}^{\dagger}c_{i\sigma}$ is the number operator. We consider the subspace where there are $N_e=N-1$ electrons on the lattice. The hopping matrix elements are arbitrary except for the condition[3] that $t_{ij}\geq 0$. The interaction V is an arbitrary real valued function of the occupation number $\{n_{i\uparrow}+n_{i\downarrow}\}$. A typical choice is to consider an on-site potential and nonlocal two-body interactions

$$V(\{n_{i\uparrow}n_{i\downarrow}\})=\sum_{i}V_i(n_{i\uparrow}+n_{i\downarrow})+\sum_{i,j}W_{ij}(n_{i\uparrow}+n_{i\downarrow})(n_{j\uparrow}+n_{j\downarrow})\ ,$$

where V_i, W_{ij} are arbitrary real coefficients. We assume that the Coulomb repulsion U is infinitely large, thus requiring each site to be occupied by at most one electron. Finally we introduce the spin operators by

$$S^{+}=(S^{-})^{\dagger}=\sum_{i}c_{i\uparrow}^{\dagger}c_{i\downarrow}\ ,$$

$$S^{z}=\tfrac{1}{2}\sum_{i}(n_{i\uparrow}+n_{i\downarrow})\ ,$$

$$\mathbf{S}^2=(S^z)^2+\tfrac{1}{2}(S^{+}S^{-}+S^{-}S^{+})\ ,$$

where the eigenvalues of $\mathbf{S}^2$ are denoted as $S(S+1)$.

Here we prove two versions of extended Nagaoka's theorem. The first one is weaker but does not require the condition (iv) about the lattice structure. The second one is a strict extension of the original Nagaoka's theorem.

Theorem 1: Consider the Hubbard model (1) with $t_{ij}\geq 0$, V arbitrary, $U=\infty$, and $N_e=N-1$. We make no assumptions on the lattice structure. Then among the ground states there exist at least N states with $S=S_{\max}\equiv(N-1)/2$.

Theorem 2: Consider the Hubbard model (1) with $t_{ij}\geq 0$, V arbitrary, $U=\infty$, and $N_e=N-1$. We further assume that the lattice Λ satisfies the connectivity condition stated in the following. Then the ground state has $S=S_{\max}\equiv(N-1)/2$ and is unique up to the trivial N-fold degeneracy.

To state the connectivity condition (and to prepare for the proof) we specify the basis we work with. We define the basis state as

$$|i,\sigma\rangle=(-1)^i c_{1\sigma_1}^{\dagger}c_{2\sigma_2}^{\dagger}\cdots c_{i-1\sigma_{i-1}}^{\dagger}\otimes c_{i+1\sigma_{i+1}}^{\dagger}\cdots c_{N\sigma_N}^{\dagger}|o\rangle\ , \tag{2}$$

where i denotes the position of the unique hole, and $\sigma=\{\sigma_j\}_{j\neq i}$ is a multi-index describing the spin of each electron. The vacuum $|o\rangle$ is the state which satisfies $c_{i\sigma}|o\rangle=0$ for any i,σ. Two states $|i,\sigma\rangle$ and $|j,\tau\rangle$ are said to be directly connected to each other if

$$\langle j,\tau|t_{ij}(c_{i\uparrow}^{\dagger}c_{j\uparrow}+c_{i\downarrow}^{\dagger}c_{j\downarrow})|i,\sigma\rangle\neq 0\ ,$$

and we express this fact by writing $(i,\sigma)\leftrightarrow(j,\tau)$. The relation $(i,\sigma)\leftrightarrow(j,\tau)$ naturally introduces a notion of connectivity into our basis space. Also note that whenever $(i,\sigma)\leftrightarrow(j,\tau)$ holds we have

$$\langle j,\tau|t_{ij}(c_{i\uparrow}^{\dagger}c_{j\uparrow}+c_{i\downarrow}^{\dagger}c_{j\downarrow})|i,\sigma\rangle=-t_{ij}$$

because of the sign convention in (2).

Definition: A finite lattice Λ is said to satisfy the connectivity condition if all the states $|i,\sigma\rangle$ with the same value of S^z are connected to each other in the foregoing sense.

Remarks: (1) Nagaoka has demonstrated that, when t_{ij} is nonvanishing only for the nearest neighbor (i,j), the connectivity condition is satisfied in the square lattice and in some regular three-dimensional lattices (such as the simple-cubic, the body-centered-cubic, and the face-centered-cubic lattices), but not in the one-dimensional chain. By examining his proof one finds that a sufficient condition for the connectivity condition is that each site i in Λ is contained in a loop (formed by nonvanishing t_{ij}) of length three or four and at least one of the sites (except for i) in the loop is connected (via nonvanishing t_{ij}) to all the other sites in Λ without passing through i. For example, the triangular lattice and the d-dimensional hypercubic lattice ($d \geq 2$) satisfy this criterion when t_{ij} is nonzero for nearest neighbor (i,j).

(2) It is obvious (and can be stated rigorously) that, when the conditions of theorem 2 (except that $U=\infty$) are satisfied, the statement of the theorem is true for finite but sufficiently large U. However, a crude estimate does not provide any information about how large U should be for a given N.

(3) Although it is very interesting to extend the theorem to the cases where there are more than one hole, our proof can hardly be generalized. An artificial way to make the present proof applicable to such models is to impose a constraint which makes the holes distinguishable from each other. (For example, one declares that when holes hop, they must preserve their relative ordering induced by the ordering of the lattice sites.)

II. PROOF

Proof of theorem 1: Let $|\Psi\rangle=\sum_{(i,\sigma)}\psi_{i,\sigma}|i,\sigma\rangle$ be an arbitrary normalized state. Let us define a state $|\Phi\rangle$ with $S=S_{\max}$ as $|\Phi\rangle=\sum_i\phi_i|i,\{\uparrow\}\rangle$ where $\phi_i=(\sum_\sigma|\psi_{i,\sigma}|^2)^{1/2}$ and the multi-index $\{\uparrow\}$ represents that all the electrons have upward spin. Then it is easy to see that

$$\langle\Psi|V(\{n_{i\uparrow}+n_{i\downarrow}\})|\Psi\rangle=\sum_{(i,\sigma)}|\psi_{i,\sigma}|^2\langle i,\sigma|V|i,\sigma\rangle=\sum_i|\phi_i|^2\langle i,\{\uparrow\}|V|i,\{\uparrow\}\rangle=\langle\Phi|V(\{n_{i\uparrow}+n_{i\downarrow}\})|\Phi\rangle\ ,$$

$$\langle\Psi|t_{ij}(c_{i\uparrow}^\dagger c_{j\uparrow}+c_{i\downarrow}^\dagger c_{j\downarrow})|\Psi\rangle=\sum_{(\sigma,\tau)}(-t_{ij})\bar{\psi}_{j,\tau}\psi_{i,\sigma}\geq(-t_{ij})\bar{\phi}_j\phi_i=\langle\Phi|t_{ij}(c_{i\uparrow}^\dagger c_{j\uparrow}+c_{i\downarrow}^\dagger c_{j\downarrow})|\Phi\rangle\ ,$$

where the sum in the second equation is over σ,τ such that $(i,\sigma)\leftrightarrow(j,\tau)$. We have used the Schwarz inequality to get the second inequality. These relations imply that the expectation value of the energy of the state $|\Phi\rangle$ is always not larger than that of the original state $|\Psi\rangle$. Then the statement of the theorem follows by taking $|\Psi\rangle$ as one of the ground states and using the global SO(3) symmetry of the system.

Proof of theorem 2: One way to prove the theorem is to investigate when the foregoing Schwarz inequality is saturated. But there is a shorter (and almost trivial) proof which makes use of the Perron-Frobenius theorem.[4] Let us examine the matrix elements of H in our basis. As we have noted before the off-diagonal element $\langle j,\tau|H|i,\sigma\rangle$ is nonvanishing only when $(i,\sigma)\leftrightarrow(j,\tau)$, and is equal to $-t_{ij}$. Moreover, the connectivity condition ensures that, in a sector with fixed S^z, the matrix $\{\langle j,\tau|H|i,\sigma\rangle\}$ is indecomposable. Then, by taking $M=-H$, the Perron-Frobenius theorem implies that, in each sector with fixed S^z, there is a unique state with minimum energy, and the state is a linear combination of all $|i,\sigma\rangle$ (with the given S^z) with positive coefficients. Since such a state has $S=S_{\max}$ and the system has a global SO(3) symmetry we have proved the desired theorem.

ACKNOWLEDGMENTS

It is a pleasure to thank A. Kawabata, M. Kohmoto, and K. Kubo for useful conversations, and E. H. Lieb for correspondences.

[1]Y. Nagaoka, Solid State Commun. **3**, 409 (1965); Phys. Rev. **147**, 392 (1966). Thouless discussed the same mechanism from a combinatorial point of view in D. J. Thouless, Proc. Phys. Soc. London **86**, 893 (1965). See also E. H. Lieb, in *Phase Transitions, Proceedings of the Fourteenth Solvay Conference* (Wiley, Interscience, New York, 1971), p. 45. An extension of the theorem to the models with strong Hund couplings was given in K. Kubo, J. Phys. Soc. Jpn. **51**, 782 (1982). The (implicit) use of the Perron-Frobenius theorem is also essential in Kubo's extension.

[2]It is worth noting that, very recently, Lieb has established another very interesting mechanism of ferromagnetism. See E. H. Lieb, Phys. Rev. Lett. **62**, 1201 (1989).

[3]By using a local transformation $c_{i\sigma}^{(\dagger)}\rightarrow-c_{i\sigma}^{(\dagger)}$ (for $\sigma=\uparrow,\downarrow$), and a global electron-hole transformation, we can treat a larger class of models, as is described in Nagaoka's paper.

[4]Let $M=\{M_{ij}\}$ be an $N\times N$ matrix with $M_{ij}\geq 0$ for $i\neq j$. We assume that M is indecomposable in the sense that, for any i,j, there is a sequence $\{i_1,i_2,\ldots,i_K\}$ with $i=i_1$, $j=i_K$, and $M_{i_k i_{k+1}}\neq 0$ for all $k<K$. Then the Perron-Frobenius theorem states (among other things) that the eigenstate of M with maximum eigenvalue is unique (up to normalization), and can be written as a linear combination of all the basis vectors with strictly positive coefficients.