# Ferromagnetism in the Hubbard model



1991 J. Phys.: Condens. Matter 3 4917

(http://iopscience.iop.org/0953-8984/3/26/014)

J. Phys.: Condens. Matter 3 (1991) 4917–4940. Printed in the UK

# Ferromagnetism in the Hubbard model

W von der Linden† and D M Edwards
Department of Mathematics, Imperial College of Science, Technology and Medicine, London SW7 2BZ, UK

Received 22 February 1991

**Abstract.** The stability of the ferromagnetic state with complete spin alignment against a single spin reversal is studied in the square-lattice Hubbard model with nearest-neighbour hopping. A variational *ansatz* is used, which is exact in one dimension and by comparison with accurate results for small clusters yields almost exact results in the two-dimensional case. The ferromagnetic region in the $(W/U, \delta)$ phase diagram is mapped out, where $W$ is the bandwidth, $U$ the on-site electron interaction and $\delta$ is the number of holes per atom. This region is considerably smaller than in previous variational calculations and it is rigorously concluded that the state of complete spin alignment is unstable when $\delta > 0.29$, for all $U$, and when $W/U > 0.19$, for all $\delta$. The nature of the instability, and of the low-lying excitations in the ferromagnetic state for varying $\delta$ and $W/V$, is discussed.

## 1. Introduction

The Hubbard model [1] is the simplest model of strongly correlated electrons in narrow bands. It is once again being studied intensively due to its postulated relevance to the high-temperature superconductors. However, little is known for certain about the phase diagram of the model, even at $T = 0$ K. One of the first applications of the model was to the study of itinerant electron magnetism. The Hartree–Fock approximation led to the Stoner criterion, which states that the ground state is ferromagnetic if $U\rho(E_F) > 1$, where $U$ is the on-site interaction strength and $\rho(E_F)$ the density of states at the Fermi energy. It was soon clear [1, 2] that the introduction of correlation effects leads to a much more stringent condition for the ferromagnetic phase. Improved mean-field approximations, within the slave-boson approach [3], led to a very small ferromagnetic region in the phase diagram. Although the accuracy of the method is uncertain it does appear that very large values of $U$ are required to stabilize the ferromagnetic state. This agrees with quantum Monte Carlo results [4] for intermediate $U$ where no ferromagnetic phase has been found. However, in view of the small clusters and relatively high simulation temperature used, these results could be misleading.

There are a few rigorous results for the Hubbard model but these are restricted to the one-dimensional model or, in higher dimensions, to the half-filled-band case of one electron per atom. For the one-dimensional case many properties are accessible via the exact Bethe *ansatz* [5, 6]. Lieb and Mattis [7] have proved that due to the special topology of the one-dimensional problem the ground state for any even number of electrons is

† Present address: Max-Planck-Institut für Plasmaphysik, D-8046 Garching bei München, Federal Republic of Germany.

0953-8984/91/264917 + 24 $03.50 

always a non-magnetic singlet independent of $U$. For higher dimensions and half filling it has been shown recently [8] that for a bipartite lattice with $N_A$ ($N_B$) sites on sublattice A and B, respectively, the ground state is non-degenerate and has total spin $S = \frac{1}{2}|N_A - N_B|$, which is valid as long as $U > 0$ independently of the dimensionality. In particular, for $N_A = N_B$ (1D, square lattice and SC) the ground state is therefore a singlet.

Away from half filling the situation is less clear. In the large-$U$ limit, the Hubbard model can be mapped onto the $t$–$J$ model (for a review see [9]), which contains a restricted hopping term, allowing no doubly occupied sites, together with an antiferromagnetic Heisenberg Hamiltonian for the spin degrees of freedom. The pure Heisenberg model of the half-filled case leads to antiferromagnetism with algebraic long-range order in 1D [10] and long-range order in 2D [11, 12]. The kinetic energy, however, destroys the antiferromagnetism rapidly with increasing hole concentration [9, 13].

As far as the ferromagnetic part of the phase diagram is concerned very little is known for sure. Nagaoka has shown in his pioneering work [14] that for most lattices, with nearest-neighbour hopping and infinite on-site interaction $U$, the strongly ferromagnetic state, with complete spin alignment, is stable for the case of one hole in an otherwise half-filled band [15]. For two holes on a square lattice, again with $U = \infty$, Doucot and Wen found that a long-wavelength twisted spin state has lower energy than the strongly ferromagnetic state [16]. This is in agreement with results obtained by exact diagonalization for small systems up to $8 \times 8$ [17–19]. Fang *et al* [18] find for 2–4 holes that the strongly ferromagnetic state is unstable and that the energy can be lowered by flipping more and more spins. On this basis it is argued that the true ground state is presumably a singlet. In the thermodynamic limit, however, the ground state for these cases is degenerate with the Nagaoka state. This is reasonable as the hole concentration tends to zero as the number of sites tends to infinity. It is important to note that finite-size effects are crucial in the Nagaoka problem and that one can easily be deceived by the results for such a small number of holes, as we will discuss later on. Even the choice of boundary conditions changes the results in these small systems completely [19, 20]. We will see later that closed-shell configurations ($N_h = 1, 5, 9, 13, \ldots$ in 2D) favour ferromagnetism, whereas open-shell configurations ($N_h = 2, 3, 4, 6, 7, 8, \ldots$ in 2D) tend to destabilize it. This point has been emphasized previously by Barbieri *et al* [21]. It explains the discrepancy between the Nagaoka result for $N_h = 1$ and those for $N_h = 2$, 3, 4. Here $N_h$ is the number of holes, which equals $N - N_e$, where $N_e$ is the number of electrons and $N$ is the number of sites.

Barbieri *et al* [21] reported analytic results for $N_h$ large, but still less than a finite fraction of $N$. More precisely they show that for $N_h \ll \ln(N)$ in 2D, and $N_h \ll N^{1/3}$ in 3D, the strongly ferromagnetic state in the infinite-$U$ Hubbard model is locally stable with respect to a single spin flip. This is a considerable extension of Nagaoka's results [14] but is still restricted to a vanishing hole concentration $\delta = N_h/N$ in the thermodynamic limit $N \to \infty$. Thus none of the rigorous arguments addresses the question of whether the Hubbard model in two or three dimensions has a ferromagnetic phase.

In this paper we base a numerical analysis of the stability of the strongly ferromagnetic state in the 2D Hubbard model on a variational wave function proposed by Edwards [22] for the state with a single spin reversal. The wave function is exact for the one-dimensional gas with $\delta$-function interaction as well as for the 1D Hubbard model. Since it is of a rather general form, based on a physical idea that is not restricted to 1D, it is a very good wave function in higher dimensions as well. In section 2 we recall the most important previous variational approaches and discuss the physical mechanisms that may drive the instability of the ferromagnetic state. The variational *ansatz* is the subject

of section 3 and we derive all the expressions needed for the numerical analysis. Section 4 contains some brief technical remarks which may be of general interest. The results are detailed and discussed in section 5.

## 2. Variational approaches

To consider the stability of the strongly ferromagnetic state with finite hole density in the thermodynamic limit, approximate methods have to be used. We will concentrate the discussion on rigorous variational methods for states with just one reversed spin. A negative spin-reversal energy definitely indicates instability of the strongly ferromagnetic state. We will denote the ground state energy for a given trial wavefunction $\psi(\boldsymbol{q})$ by $E(\boldsymbol{q})$. To establish the notation we define the Hamiltonian of the Hubbard model:

$$H = -t \sum_{\langle i,j\rangle\sigma} a_{i\sigma}^{\dagger} a_{j\sigma} + U \sum_{i} n_{i\downarrow} n_{i\uparrow} \tag{2.1}$$

where $a_{i\sigma}^{\dagger}$ ($a_{i\sigma}$) creates (annihilates) a fermion of spin $\sigma = \uparrow, \downarrow$ at site $i$, $t > 0$ is the hopping matrix element, $U$ represents the on-site Coulomb interaction, and $\langle i, j\rangle$ indicates that only nearest-neighbour hopping is allowed. In the strongly ferromagnetic (Nagaoka) state, for which all electron spins are parallel, the particles do not experience the on-site repulsion $U$. The electrons in this state therefore occupy single-particle Bloch states $a_{k\uparrow}^{\dagger}$ with corresponding single-particle energies

$$\varepsilon(\boldsymbol{k}) = -t \sum_{\boldsymbol{\Delta}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{\Delta}} \tag{2.2}$$

where summation is over nearest-neighbour vectors $\boldsymbol{\Delta}$. Throughout this paper we set the lattice constant equal to one. The Nagaoka state is

$$|K\rangle = \prod_{k\in K} a_{k\uparrow}^{\dagger} |0\rangle \tag{2.3}$$

where $K$ is the set of wave vectors with lowest single-particle energies $\varepsilon(\boldsymbol{k})$ and $|0\rangle$ is the vacuum state. The ground state energy of the Nagaoka state is

$$E_0 = \sum_{k\in K} \varepsilon(\boldsymbol{k}). \tag{2.4}$$

The wave vectors $\boldsymbol{k}$ may be divided into shells, all wave vectors within a given shell having the same value of $\varepsilon(\boldsymbol{k})$. The non-interacting ground state (Nagaoka state) $|K\rangle$ is unique only for, what we will call, closed-shell configurations. For an electron number corresponding to an open-shell configuration the non-interacting ground state is degenerate and several Nagaoka states are possible.

There are two possible mechanisms driving the instability of the ferromagnetic state, which we will discuss separately: single-particle excitations and spin waves. We will discuss these cases in some detail to clarify the underlying physical ideas and to establish notation, which we will need in the discussion of our results.

### *2.1. Single-particle excitations*

To create a single-particle excitation with momentum $\boldsymbol{q}$, an $\uparrow$-spin electron is removed from the occupied states with momentum $\boldsymbol{k}$ and is placed with reversed spin into a $\downarrow$-spin quasi-particle state with momentum $\boldsymbol{k} + \boldsymbol{q}$. In the Hartree–Fock approximation the corresponding wave function reads:

$$|\Psi^{\mathrm{HFA}}(\boldsymbol{q})\rangle = a_{(k+q)\downarrow}^{\dagger} a_{k\uparrow} |K\rangle \tag{2.5}$$

leading to an excitation energy $\omega_k(\boldsymbol{q}) = \varepsilon(\boldsymbol{k} + \boldsymbol{q}) - \varepsilon(\boldsymbol{k}) + U\rho$, where $\rho = N_e/N$ is the

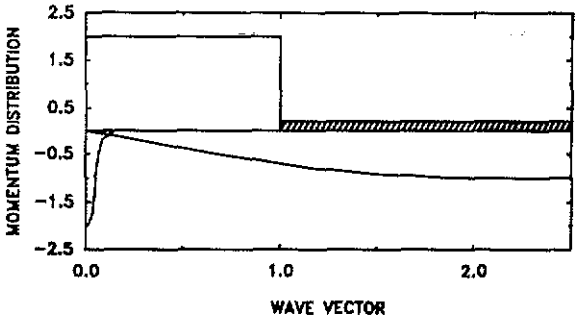


**Figure 1.** Schematic spin-dependent momentum distribution functions predicted by the Gutzwiller *ansatz* $\psi^{GA}$ for $k = k_F$ and $k + q = 0$. The upper distribution is for the ↑-spin electrons and the lower one for the ↓-spin electron.

density of electrons. The lowest excitation energy is obviously achieved if the ↑-spin electron is removed from the Fermi surface $k = k_F$ and the ↓-spin electron is placed at the bottom of the band $k + q = 0$. Thus the momentum change $q$ corresponds to a wave vector on the Fermi surface. More precisely, the total momentum is actually $q + K$. The momentum $K$ of the Nagaoka state can, however, be considered as the origin in $k$-space. We measure henceforth all wave vectors relative to the respective Nagaoka state, if not stated otherwise. For closed shells $K$ is $(\pi, \pi)$ for $\delta < 1/2$ and $(0, 0)$ for $\delta > 1/2$. The sole influence of the interaction in the Hartree–Fock approximation is a rigid upward shift of the ↓-spin electron energy by $\rho U$, which is incorrect for two reasons. It certainly overestimates the cost in energy to create a ↓-spin quasi-particle and it misses the band narrowing arising from the strongly restricted mobility of the ↓-spin electron. For large $U$ (2.5) is certainly not an appropriate wave function. Real-space configurations with many doubly occupied sites have too large a weight in this mean-field wave function. This weight is reduced in the Gutzwiller wave function [23, 24]:

$$|\Psi^{GA}\rangle = \prod_i (1 - \eta n_{i\uparrow} n_{i\downarrow})|\Psi^{HFA}\rangle. \tag{2.6}$$

$\eta$ is a variational parameter monotonically increasing from 0 to 1 with increasing $U$. Within the framework of the Gutzwiller wave function one can study nicely how the ↓-spin electron spreads out in $k$-space with increasing $U$. The spin-dependent momentum distribution $n_\sigma(p) = \langle\Psi|n_{p,\sigma}|\Psi\rangle/\langle\Psi|\Psi\rangle$ is

$$n_\uparrow(p) = n^0(p)\left(1 - \eta^2 \frac{1-\rho}{N(1-2\eta\rho+\eta^2\rho)}\right) + (1-n^0(p))\eta^2 \frac{\rho}{N(1-2\eta\rho+\eta^2\rho)}$$

$$n_\downarrow(p) = \frac{\delta(p,k+q)(1-\eta\rho)^2 + \dfrac{\eta^2}{N}\left(\rho - \dfrac{1}{N}\displaystyle\sum_{k'} n^0(k'+k+q-p)n^0(k')\right)}{(1-2\eta\rho+\eta^2\rho)}. \tag{2.7}$$

Here $n^0(p)$ is the momentum distribution of the ↑-spin electrons in the state $a_{k\uparrow}|K\rangle$ and is 1 if $p \in \{K/k\}$ and 0 otherwise. $\delta(k, k')$ is the Kronecker delta symbol and the ↑-spin electron density $\rho = N_\uparrow/N$. The momentum distribution is schematically depicted in figure 1. The number of ↑-spin electrons scattered out of the occupied states (shaded area in upper panel of figure 1) increases from 0 to $\rho$ as $U$ increases from 0 to $\infty$. At the same time the weight of the delta function at the bottom of the ↓-spin band declines

gradually from 1 to $\delta$ and the momentum distribution of the $\downarrow$-spin electron gains more and more weight outside the Fermi surface of the $\uparrow$-spin electrons. The expectation value of the energy for $U = \infty$ and one reversed spin has been studied recently by Shastry *et al* [25]. The excitation spectrum consists of a continuum of scattering states with minimum at total momentum $q = k_F$:

$$\omega_k(q) = (|E_0|/N\delta - \varepsilon(k)) + \varepsilon(k+q)\delta(1 - |E_0/N\delta z|^2). \tag{2.8}$$

$z$ is the number of nearest-neighbour sites. The first term is the kinetic energy of the $\uparrow$-spin electrons and the second that of the $\downarrow$-spin electron, which correctly includes a band-narrowing factor. It was found in [25] that the excitation energy is positive up to a critical hole concentration $\delta_{cr} \simeq 0.49$ for the square lattice and $\delta_{cr} \simeq 0.32$ for SC. The instability of the ferromagnetic state for $\delta > \delta_{cr}$ is driven by single-particle excitations at $q = k_F$.

Roth [26] first discussed this instability in detail within an approximation scheme equivalent to one used earlier by Edwards [27, 28] for discussing spin waves (see also [29]). It has been shown [30, 31] that Roth's method is equivalent to a variational *ansatz* with a wave function superior to the Gutzwiller *ansatz*, as it has more variational flexibility. The trial wave function implicitly used by Roth reads

$$|\Psi^R(q)\rangle = \left(a^\dagger_{(q+k_F)\downarrow} + \sum_k B(k) a^\dagger_{k\uparrow} S^-_{q+k_F-k}\right) a_{k_F\uparrow}\,|K\rangle \tag{2.9}$$

where $S^-_q$ is the approximate magnon creation operator:

$$S^-_q = \sum_j e^{iq\cdot x_j} a^\dagger_{j\downarrow} a_{j\uparrow} = \sum_k a^\dagger_{(k+q)\downarrow} a_{k\uparrow} \tag{2.10}$$

where $x_j$ is the position vector of site $j$. In the limit $U = \infty$, equation (2.9) simplifies to

$$|\Psi^R_\infty(q)\rangle = \sum_k B(k) S^-_{q+k_F-k} a^\dagger_{k\uparrow} a_{k_F\uparrow}\,|K\rangle \tag{2.11}$$

which has no double occupancies because according to (2.10) the $\downarrow$-spin electron is only created at empty sites. This reduces to the Gutzwiller *ansatz* used [25] if $B(k)$ is taken as a constant independent of $k$. The *ansatz* (2.11) leads to lower energies than the Gutzwiller *ansatz* and also to a critical hole concentration about 20% smaller, namely $\delta_{cr} = 0.41$ [32] for the square lattice and $\delta_{cr} = 0.24$ for the SC [26], respectively.

On minimizing the ground state energy with respect to $B(k)$ a self-consistency equation for the excitation energies $\omega(q) = E(q) - E_0$ is obtained [26]. Roth's wave function is a good approximation for the single-particle excitations and is therefore well suited to study the finite-size effects present in the Nagaoka problem. This wave function has the advantage that the ground state energy is easily accessible by numerical techniques, even for very large systems. The lowest excitation energy, which again appears at the Fermi wave vector, is depicted in figure 2. It shows typical finite-size effects, which we will encounter later in the numerical evaluation of the wave function proposed by Edwards [22]. The maxima belong to closed-shell situations, as we have mentioned earlier. Qualitatively, a closed-shell situation corresponds to a much too low density of states at $k_F$ due to the presence of the finite-size gap, whereas the opposite is true for the open shells. The picture is qualitatively the same for the Gutzwiller *ansatz* [19]. Finite-size effects are more severe in the accurate calculations of section 5 as the energies are

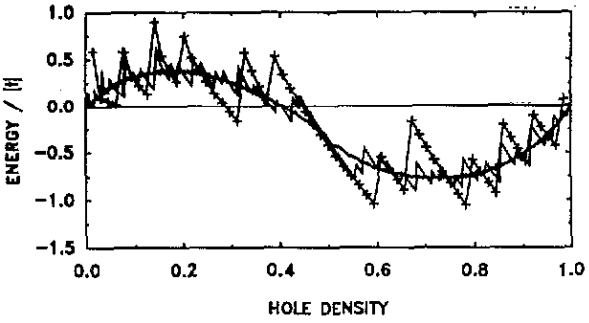


**Figure 2.** Minimum excitation energy $\omega(q = k_F)$ obtained from Roth's wave function for cluster sizes $8 \times 8$ (full curve plus crosses), $16 \times 16$ (full curve) and $100 \times 100$ (thick full curve) for all possible numbers of holes.

somewhat lower and the shell effects depicted in figure 2 lead to sign changes. Closed-shell configurations up to a critical concentration all have a positive excitation energy whereas most of the open-shell configuration would indicate an instability of the ferromagnetic state. This is the origin of the discrepancy between the Nagaoka result for one hole and the results obtained for 2–4 holes. Finite-size scaling for the open-shell results is, however, complicated due to the degeneracy of the Nagaoka states. The situation is much more transparent for the closed shells and in section 5 we focus our discussion mostly on closed shells.

### *2.2. Spin waves*

We have just seen that single-particle excitations make the ferromagnetic state unstable above a critical hole concentration. In addition, there is also the possibility that the Goldstone modes (spin waves), which tend to restore the broken rotational symmetry anyway, may become soft and lead eventually to an instability of the ferromagnetic state. The critical hole concentration could thus be reduced to a lower value or even to zero. The simplest spin wave conceivable is given by

$$|\Psi^{\mathrm{SSW}}(q)\rangle = S_q^- |K\rangle \tag{2.12}$$

with the approximate magnon operator (2.10). For $q = 0$ this wave function is an exact eigenstate of the Hubbard Hamiltonian with maximum total spin. It is therefore degenerate with the Nagaoka state. For $U = \infty$, $\Psi^{\mathrm{SSW}}(q)$ is equivalent to the random-phase approximation [33]. The excitation spectrum of the simple spin wave *ansatz* (2.12) is given by:

$$\omega^{\mathrm{SSW}}(q) = \frac{1}{N_c} \sum_k (\varepsilon(k + q) - \varepsilon(k)) n^0(k) \tag{2.13}$$

and is shown in figure 3. For small $q$ the energy can be expanded and yields in leading order

$$\omega(q) = Dq^2 \tag{2.14}$$

which defines the spin wave stiffness constant $D$. In the present approximation $D^{\mathrm{SSW}} = |E_0|/4N\rho$ and for small hole concentration $D^{\mathrm{SSW}} \simeq \delta|t|$.

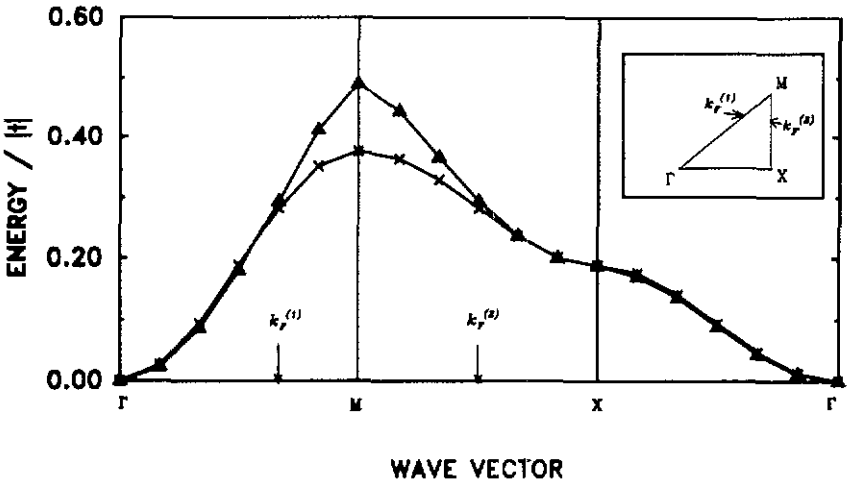


**Figure 3.** Excitation energies of the simple spin wave *ansatz* (×) multiplied by 0.347, compared with the low-lying excitations in the Hubbard model (△) as discussed in the text. The results are for the 12 × 12 cluster with 21 holes ($\delta = 0.15$). The inset shows the high-symmetry lines and the Fermi points along these lines.

The momentum distribution of the simple spin wave state can easily be derived from (2.12) and yields

$$n^{\uparrow}(\boldsymbol{k}) = (1 - 1/N_e)n^0(\boldsymbol{k}) \qquad n^{\downarrow}(\boldsymbol{k}) = (1/N_e)n^0(\boldsymbol{k} - \boldsymbol{q}). \tag{2.15}$$

The difference between the single-particle excitations and the spin waves is most obvious for the $\downarrow$-spin electrons. In the former case the electron was predominantly in one $k$-state, whereas in the latter the $\downarrow$-spin electron is homogeneously distributed over the Fermi volume of the Nagaoka state shifted by $\boldsymbol{q}$. We will use these features in the later discussion to classify qualitatively the low-lying excitations.

A wave function that contains both spin waves and single-particle excitations has recently been studied by Shastry *et al* [25]. They used the Gutzwiller-projected RPA wave function:

$$|\Psi^{\mathrm{GRPA}}(\boldsymbol{q})\rangle = \prod_i (1 - \eta n_{i\uparrow} n_{i\downarrow})|\Psi^{\mathrm{RPA}}(\boldsymbol{q})\rangle \tag{2.16}$$

with

$$|\Psi^{\mathrm{RPA}}(\boldsymbol{q})\rangle = \sum_k C(k) a^{\dagger}_{(k+q)\downarrow} a_{k\uparrow} |K\rangle. \tag{2.17}$$

In [25] the limit $U = \infty$ ($\eta = 1$) is studied and it is found that the bottom of the Stoner continuum at $k_F$ pushes the spin wave branch down and both become negative at $k_F$. The critical concentration is about the same as was obtained with the Gutzwiller *ansatz* $\Psi^{\mathrm{GA}}(\boldsymbol{q})$.

## 3. The *ansatz*

In this section we present an *ansatz* due to Edwards [22] which is exact in one dimension. As we will see, the underlying physical idea is fairly general and not restricted to one

dimension. To begin with we consider the $k$-space representation for the most general form of exact wave function with one reversed spin:

$$|\chi(q)\rangle = \sum_{Q} C(Q) a^{\dagger}_{[q-Q]\downarrow} |Q\rangle. \tag{3.1}$$

In (3.1) the summation is over all possible $N_\uparrow$-tupels $Q$ of wave vectors $k$ for the $\uparrow$-spin electrons:

$$|Q\rangle = \prod_{k\in Q} a^{\dagger}_{k\uparrow} |0\rangle \qquad Q = \sum_{k\in Q} k.$$

The *ansatz* proposed in [22] is equivalent to replacing the unknown function $C(Q)$ by a determinant of one-particle orbitals $\varphi_\alpha(k)$, $\alpha = 1, \ldots, N_\uparrow$. The number of variational parameters is therefore drastically reduced from $\binom{N}{N_e}$ to merely $N(N_e - 1)$. The underlying physical idea that led to this *ansatz* becomes immediately clear in the real-space representation of $|\chi(q)\rangle$:

$$|\chi(q)\rangle = \frac{1}{\sqrt{N}} \sum_{i} e^{iq\cdot x_i} a^{\dagger}_{i\downarrow} \prod_{\alpha} \left( \sum_{j} \varphi_\alpha(x_j - x_i) a^{\dagger}_{j\uparrow} \right) |0\rangle. \tag{3.2}$$

$\varphi_\alpha(x_i)$ are the Fourier transforms of the linearly independent one-particle orbitals $\varphi_\alpha(k)$ of the $k$-space representation. If the $\downarrow$-spin electron were to be fixed, say at position $x_i$, the exact many-body wave function for the $\uparrow$-spin electrons would be a single Slater determinant of one-particle orbitals centred at $x_i$. This is because the $\uparrow$-spin electrons interact only with the $\downarrow$-spin electron, which in turn serves as a static potential at site $x_i$. If the motion of the $\downarrow$-spin electron is taken into account, the wave function has to be a coherent wave of such states for all possible $\downarrow$-spin positions. This leads to the *ansatz* (3.2). This wave function is a special case of the one used by Wigner and Seitz [34] in their work on electron correlations in jellium. Richmond and Rickayzen [35] have elaborated upon the simplified version with the $\downarrow$-spin electron fixed at one site. This *ansatz* has the advantage that most quantities of interest can be computed analytically by means of scattering theory; however, it is not adequate for an analysis of the stability of the ferromagnetic state. It predicts, for example that for $U = \infty$ the strongly ferromagnetic state is always the ground state for arbitrary hole densities, which is in contradiction to the exact results for low electron densities [2] and to the results discussed in section 2.

The motion of the $\downarrow$-spin electron leads to a very interesting momentum-dependent interaction for the $\uparrow$-spin electrons [22], or else, after a canonical transformation, to long-range many-particle interaction. It has been shown in [22] that the wave function (3.2) is exact in one dimension. For the gas case, the one-particle orbitals $\varphi_\alpha(x_i)$ can be explicitly constructed following the ideas of McGuire [36], which started the development leading to the Bethe *ansatz*. In contrast to the assumptions leading to the exact solution in the one-dimensional problem, the physical ideas underlying the *ansatz* (3.2) are not especially geared to 1D, and it is not easy to see why it should not be a very good wave function, if not even exact, in higher dimensions. As a matter of fact, $\chi(q)$ in (3.2) is very flexible. It covers, for instance, all the variational wave functions that we discussed in the previous sections. To recover the RPA wave function $\Psi^{\mathrm{RPA}}(q)$ of (2.17) the one-particle orbitals entering (3.2) take on the form

$$\varphi_\alpha(x_i) = \frac{1}{\sqrt{(NN_e)}} \sum_{l=1}^{N_e-1} e^{i2\pi l/N_e} e^{ik_l\cdot x_i} C^{-1}(k_l). \tag{3.3}$$

The Hartree–Fock wave function $|\Psi^{\mathrm{HFA}}(q)\rangle$ and the simple spin wave *ansatz* $|\Psi^{\mathrm{SSW}}(q)\rangle$

are merely special cases of $\Psi^{\mathrm{RPA}}(q)$ with $C(k) = \delta_{k,k_0}$ for the former $C(k)$ = constant for the latter, respectively. The treatment of the Hartree–Fock wavefunction is somewhat tricky in (3.3) but it is immediately evident when using the $k$-space representation (equation (3.1)) with $C(K)$ being the Slater determinant of occupied $\uparrow$-spin orbitals with $\varphi_\alpha(k) = \delta_{k,k_\alpha}$. $\{k_\alpha\}$ is the set of wave vectors occupied by the $\uparrow$-spin electrons in (2.5). The Gutzwiller projection operator is easily incorporated. Operating with

$$\prod_i (1 - \eta n_{i\uparrow} n_{i\downarrow})$$

on the wave function (3.2) yields

$$|\chi(q)\rangle = \frac{1}{\sqrt{N}} \sum_i \mathrm{e}^{\mathrm{i}q\cdot x_i} a_{i\downarrow}^\dagger (1 - \eta n_{i\downarrow}) \prod_\alpha \left( \sum_j \varphi_\alpha(x_j - x_i) a_{j\uparrow}^\dagger \right) |0\rangle. \tag{3.4}$$

The factor $(1 - \eta n_{i\uparrow})$ leads to modified one-particle orbitals $\tilde{\varphi}_\alpha(x_i)$ in (3.2):

$$\tilde{\varphi}_\alpha(x_l) = \varphi_\alpha(x_l)(1 - \eta\delta_{l,0}). \tag{3.5}$$

Consequently for $U \to \infty$ ($\eta \to 1$) the value for the central site, which is a measure for the number of double occupancies, tends to zero: $\tilde{\varphi}_\alpha(0) \to 0$. For small hole concentrations it is more convenient to perform a particle–hole transformation for the $\uparrow$-spin electrons: $a_{i\uparrow}^\dagger \to h_{i\uparrow}$, $a_{i\uparrow} \to h_{i\uparrow}^\dagger$, leading to:

$$|\chi(q)\rangle = \frac{1}{\sqrt{N}} \sum_i \mathrm{e}^{\mathrm{i}q\cdot x_i} a_{i\downarrow}^\dagger \prod_\alpha \left( \sum_j g_\alpha(x_j - x_i) h_{j\uparrow}^\dagger \right) |0\rangle. \tag{3.6}$$

The new vacuum state $|0\rangle$ is the state where every site is occupied by an $\uparrow$-spin electron. The number of variational parameters is $N(N_{\mathrm{h}} - 1)$, which is more economical for $\delta < 0.5$. In the hole case for $U = \infty$ a possible choice of one-particle orbitals, explicitly avoiding double occupancies, is to localize one hole, say $\alpha = 1$, at the centre: $g_1(x_i) = \delta_{x_i,0}$ and to use $g_\alpha(0) = 0$ for the other orbitals. This reduces the number of variational parameters even more and stabilizes the numerical methods used to determine the variational parameters.

For the formulae that we will derive now, we stick to the electron case. The analogous expressions for the hole case are easily obtained by the particle–hole transformation. We derive the expectation value for the energy for linearly independent but not necessarily orthogonal one-particle orbitals. The derivation for the kinetic energy is outlined in the appendix. The evaluation of the interaction energy is straightforward and needs no special consideration in the appendix. The three contributions to the energy are

$$E = E_{\mathrm{kin}}^\downarrow + E_{\mathrm{kin}}^\uparrow + E_{\mathrm{int}} = t \sum_\Delta \mathrm{e}^{-\mathrm{i}q\cdot\Delta} \det(\mathbf{S}^{-1}\mathbf{S}^{(\Delta)})$$

$$+ t \sum_\Delta \mathrm{tr}(\mathbf{S}^{-1}\mathbf{S}^{(\Delta)}) + U \sum_{\alpha\beta} S_{\beta\alpha}^{-1} \varphi_\alpha^*(0)\varphi_\beta(0) \tag{3.7}$$

with the two-centre overlap matrices given by

$$S_{\alpha\beta}^{(x)} = \sum_i \varphi_\alpha^*(x_i)\varphi_\beta(x_i + x). \tag{3.8}$$

The overlap matrix $\mathbf{S}$ is a special case of (3.8) with $x = \mathbf{0}$.

As mentioned earlier, it is expected that the hopping of the $\downarrow$-spin electron is reduced. To discuss this point we assume orthonormal orbitals in (3.7). The band narrowing is qualitatively obtained by averaging $\det(\mathbf{S}^{(\Delta)})$ in (3.7):

$$E_{\mathrm{kin}}^{\downarrow}(q) = \left(\frac{1}{z}\sum_{\Delta} \det(\mathbf{S}^{(\Delta)})\right) \varepsilon(q). \tag{3.9}$$

The band narrowing is therefore given by the first factor in (3.9), which is the average overlap of the many-body wave function for the $\uparrow$-spin electrons with this wave function shifted by one lattice constant. This is similar to what Doucot and Rammal [37] derived in the coherent-spin-state approximation and it is also akin to the hopping of polarons.

As in the Hartree–Fock approximation, the energy is invariant under any arbitrary regular transformation of the occupied orbitals

$$\ddot{\varphi}_{\beta} = \sum_{\gamma} \varphi_{\gamma} M_{\gamma\beta}.$$

This feature is very useful for stabilizing the numerical algorithm to obtain the stationary one-particle orbitals. To obtain an effective Hamiltonian for the $\uparrow$-spin electron orbitals, we differentiate the energy with respect to $\varphi_{\eta}(x_i)$ and equate the result to zero. We need not introduce Lagrangian parameters for the orthonormality constraint as in [22], since (3.7) is valid for arbitrary orbitals. The derivation is outlined in section 2 of the appendix and leads to the eigenvalue equation:

$$\begin{aligned} E_{\eta}\varphi_{\eta}(x_i) = {} & t\sum_{\Delta} \mathrm{e}^{-\mathrm{i}q\cdot\Delta} \det(\mathbf{S}^{(\Delta)}) \sum_{\gamma} \varphi_{\gamma}(x_i + \Delta) S_{\gamma\eta}^{(\Delta)-1} \\ & + t\sum_{\Delta} \varphi_{\eta}(x_i + \Delta) - t\sum_{\Delta\gamma} \varphi_{\gamma}(x_i) S_{\gamma\eta}^{(\Delta)} \\ & + U\left(\delta_{x_i,0} - \sum_{\gamma} \varphi_{\gamma}(x_i)\varphi_{\gamma}^{*}(0)\right)\varphi_{\eta}(0). \end{aligned} \tag{3.10}$$

We have thus derived a self-consistent-field equation for the one-particle orbitals. As described in section A2, a Löwdin orthogonalization was performed after differentiation, and so orbitals finally appearing in (3.10) are orthogonal. The overlap matrices are also understood to be evaluated in these orthonormal orbitals. If the true eigenvectors of (3.10) are inserted, the effective Hamiltonian becomes Hermitian and the eigenvalues for the occupied orbitals are all degenerate and identical to the kinetic energy of the $\downarrow$-spin electron:

$$E_{\eta} = E_{\mathrm{kin}}^{\downarrow} = t\sum_{\Delta} \mathrm{e}^{-\mathrm{i}q\cdot\Delta} \det(\mathbf{S}^{(\Delta)}). \tag{3.11}$$

This can be verified by multiplying (3.10) by $\varphi_{\eta}^{*}(x_i)$ and summing over $x_i$. As a matter of fact, for any set of orthonormal orbitals used in the definition of the effective Hamiltonian, these very orbitals are the left-eigenvectors of the effective Hamiltonian. The one-particle energies are all degenerate and given by expression (3.11). The first term on the RHS of (3.11) acts only on the occupied orbitals and is the only term that correlates

the orbitals of the ↑-spin electrons. The second and third terms describe a nearest-neighbour hopping followed by a projection into the subspace orthogonal to the occupied orbitals:

$$H_{2,3} = \left(1 - \sum_{\alpha} |\varphi_\alpha\rangle\langle\varphi_\alpha|\right) H_{\text{kin}}. \tag{3.12}$$

Similarly, the fourth and fifth terms consist of a static scattering potential for the central site, again followed by the projection operator into the space of non-occupied states. The eigenvectors of the eigenvalue problem (3.10) without the first term would be the eigenstates of a tight-binding model with nearest-neighbour hopping and an impurity at $\boldsymbol{x}_i = \mathbf{0}$.

A quantity that allows us to classify the low-lying states according to the single-particle or spin wave character is the momentum distribution. The expectation value of the spin-dependent momentum distribution function for the wave function $\chi(q)$ of (3.2) with orthonormal orbitals $\varphi_\alpha$ is derived in section A3 of the appendix:

$$n^{\uparrow}(\boldsymbol{k}) = \sum_{\alpha=1}^{N_\uparrow} |\varphi_\alpha(\boldsymbol{k})|^2 \qquad n^{\downarrow}(\boldsymbol{k}) = \frac{1}{N}\sum_{l} \mathrm{e}^{\mathrm{i}(\boldsymbol{k}-\boldsymbol{q})\cdot\boldsymbol{x}_l} \det(S^{(\boldsymbol{x}_l)}_{\alpha\beta}). \tag{3.13}$$

The expression for the ↑-spin electrons is what one expects for uncorrelated electrons. The momentum distribution for the ↓-spin electron involves the two-centre overlap matrices (3.8) for all possible centres and it depends explicitly on the total momentum $\boldsymbol{q}$ of the many-body wave function.

## 4. Some technical details

Here we briefly discuss some of the technical details, in as far as they are of general interest or necessary for the understanding of the results. We used several techniques to determine the one-particle orbitals $\varphi_\alpha(\boldsymbol{x}_i)$. The most obvious scheme for minimizing the energy functional is presumably simulated annealing [38]. There are at first glance several advantages: (i) if one has enough patience and computer time this algorithm will find the global minimum; (ii) the algorithm is numerically stable; (iii) when changing only one orbital at a time, the update technique for determinants [39] can be used. To perform one lattice sweep (to change each variational parameter once) involves $O(NN_h^3)$ operations. In the actual calculation it turned out that this scheme is still very slow, mainly for two reasons: (a) the number of operations for one step is considerable and the updated determinants become inaccurate very quickly and have to be recalculated from scratch at about every 100th step; (b) due to the random nature of the changes it was necessary to make a huge number of lattice sweeps to reach the desired accuracy. The problem is that the total energy is of order $N$ while the excitation energy is only O(1).

By making use of the gradient, implicitly given in (3.10), it is actually possible to devise an algorithm that is superior in both respects. One can obtain a much faster convergence as far as the number of lattice sweeps is concerned and the computational costs per lattice sweep are reduced to $O(NN_h^2)$. The optimal algorithm among the

possible standard minimization schemes for the present problem is the power method [40]. Formally, (3.10) can be written as

$$\mathcal{L}\varphi_\alpha = E_\alpha \varphi_\alpha \tag{4.1}$$

where the operator $\mathcal{L}$ is non-local and depends on all occupied orbitals. As for a standard eigenvalue problem we define the iteration scheme

$$\varphi_\alpha^{n+1} = (\mathcal{L} - E_s \mathbf{I})\varphi_\alpha^n \tag{4.2}$$

with a proper spectral shift $E_s$ to ensure and speed up the convergence towards the ground state. Starting from an arbitrary set of orthonormal orbitals the iteration procedure (4.2) produces a new set of orbitals, which in general is no longer orthonormal. We can orthonormalize the orbitals at each step, since the energy is invariant under this operation. Extensive numerical tests have shown that the power method yields precisely the same ground state energies and orbitals as obtained by simulated annealing. It is worth mentioning that the power method and steepest descent produce a very similar sequence of energies, when starting from the same initial set of orbitals. The former is, however, more economical per iteration step. Another standard technique, which in principle should be superior to steepest descent, is the conjugate gradient method [41]. It turned out, however, that this scheme has one crucial shortcoming, which in the end makes it less efficient than the power method. Similarly to what happens in quantum Monte Carlo algorithms for many-fermion problems, the one-particle orbitals entering the determinant have the tendency to become more and more parallel during the iterations. To keep the algorithm stable one has to reorthonormalize the orbitals after about five iteration steps. Although the energy is invariant under this transformation, the conjugate gradients are not, and the whole scheme loses its advantage. The Lanczos algorithm, which for standard eigenvalue problems is equivalent to the conjugate gradient method, cannot be applied to the non-linear eigenvalue problems (3.10). Most of the results reported in this paper are obtained with the power method. Simulated annealing is merely used from time to time as a check.

## 5. Results and discussion

In the first part of this chapter we concentrate on $U = \infty$. We begin the discussion of the results with a comparison of the ground state energies obtained with the present *ansatz* and by the Lanczos method [19], which is supposed to be exact. The biggest system accessible to exact diagonalization is the $8 \times 8$ cluster with three holes. This is an open-shell case and there exists no unique Nagaoka state. The wave vectors reported here are therefore absolute and not measured relative to a reference state. The ground state energies for this system for wave vectors along the (1, 1) direction are given in table 1. We see that the data, obtained by the present *ansatz*, are in very good agreement with the results of the Lanczos method. The excitation energies, also given in table 1, are negative for periodic boundary conditions, which we used for all our calculations. Barbieri *et al* [19] found that the sign of the excitation energies depends on the boundary conditions. As mentioned earlier, the case of three holes is an open-shell situation and several Nagaoka states with different total momenta $\boldsymbol{K}$ are possible. This is why the excitation energies for $q = (\pi, \pi)$ and $q = (3\pi/4, 3\pi/4)$ are degenerate in our calculation. They both belong to zero-energy spin wave excitations. In the following discussion we will concentrate on closed-shell situations, since they have two major advantages: (a)

**Table 1.** Ground state energies for the $8 \times 8$ cluster with three holes. Wave vectors are given in units of $\pi/4$ and energies in units of $|t|$. The exact (Lanczos) results are from [19]. The values in brackets are the excitation energies.

| $q$ | Lanczos | *Ansatz* |
|---|---|---|
| (0,0) | −10.942 | −10.940 (−0.056) |
| (1,1) | −10.919 | −10.916 (−0.044) |
| (2,2) | −10.875 | −10.874 (−0.023) |
| (3,3) | −10.834 | −10.828 (−0.000) |
| (4,4) | −10.750 | −10.744 ( 0.000) |

they are unique and have a fairly regular finite-size behaviour; (b) it is possible to follow the dispersion relation of a single branch of excitation energies, because there is only one Nagaoka state defining the origin in $k$-space. In figure 3 the low-lying dispersion for a system of size $12 \times 12$ with 21 holes is given. This dispersion is typical for the case of low hole concentration. We find qualitatively similar results for hole densities $\delta \leqslant 0.15$ for all systems studied. For small $\delta$ the dispersion curves are very much like that for the simple spin wave, which for comparison is also given in figure 3. They both start off $\propto q^2$ from the $\Gamma$ point. The results are shown along high-symmetry lines connecting the points $\Gamma = (0, 0)$, $M = (\pi, \pi)$ and $X = (\pi, 0)$. The similarity between the simple spin wave and the low-lying excitation of the Hubbard model is the greater the lower the hole concentration. The only important difference between the two results is an overall renormalization factor. The simple RPA spin wave results for the case in figure 3 are about three times larger than our more accurate results. We conclude that the low-lying excitations of the Hubbard model for small $\delta$ and $U = \infty$ are renormalized spin waves. To obtain a more quantitative measure for the renormalization we analyse the spin wave stiffness constant (2.13). We obtain $D = E(q)/q^2$ from a quadratic fit of the excitation energies for the smallest $q$-values in the (1, 0) and (1, 1) directions for a given lattice size. We collected data for $8 \times 8$, $12 \times 12$, $16 \times 16$ and $20 \times 20$ clusters for all possible numbers of holes corresponding to closed-shell situations. We find a universal linear-$\delta$-dependence of $D$ for $\delta \leqslant 0.15$: $D = 0.069(5)\delta z|t|$. In particular, for $\delta = 0.1$, $D = 0.0069(5)z|t|$ as compared with $D = 0.009z|t|$ obtained by Shastry *et al* with the Gutzwiller-projected RPA wave function. Our result is smaller as a consequence of the better trial wave function. The discrepancy becomes more pronounced for $\delta = 0.2$, where the results are $0.0082(7)z|t|$ and $0.014z|t|$, respectively. The spin wave stiffness constant of the Hubbard model is therefore much smaller than the RPA value, which for small $\delta$ is $D^{\mathrm{RPA}} = 0.25\delta z|t|$.

$D$ as a function of $\delta$ is already deviating from linear behaviour at $\delta = 0.15$, and it passes through a maximum with increasing $\delta$. Eventually it becomes negative beyond some concentration $\delta^*$. For lower hole concentration than this, the overall dispersion curve already deviates qualitatively from the simple spin wave curve. It has regions of negative excitation energies indicating instability of the strongly ferromagnetic state. Figure 4 shows a typical dispersion curve slightly below the concentration $\delta^*$. For the smaller clusters (up to $12 \times 12$) the instability occurs at the X point ($q = (\pi, 0)$). This point has, however, no deeper physical significance, contrary to the speculations in [18] for the two-hole case. In our calculations it turns out that the X point is part of, or close to, the Fermi surface for these system sizes and for the concentration where the excitation energy becomes negative ($\delta_{\mathrm{cr}} \simeq 0.3$). In fact, the minimum of the excitation energy

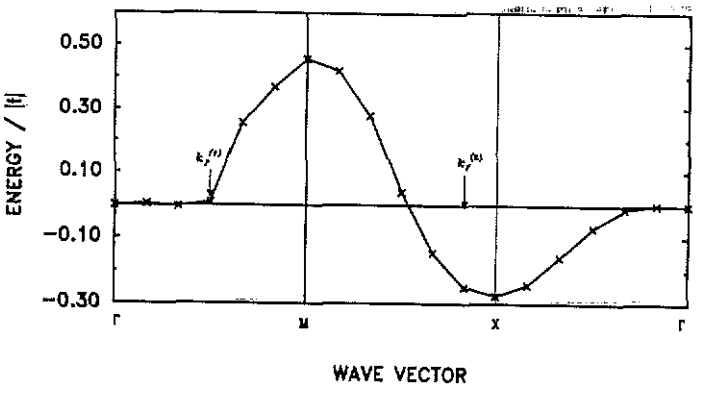


**Figure 4.** Low-lying excitations of the 2D Hubbard model for the 12 × 12 cluster with 57 holes ($\delta = 0.40$).

occurs at $k_F$. This can either be verified for a small cluster at higher hole concentration (figure 5) or for a bigger cluster. In the latter case (figure 6) we also see that the spin wave stiffness constant is still positive, while the ferromagnetic state is already unstable due to the single-particle excitation at $k_F$.

Next we study the $\delta$-dependence of the dispersion curve. In figure 7 the excitation energy for $q = (\pi, 0)$ is depicted for various cluster sizes. We find a rather universal behaviour for all system sizes: a linear increase for small $\delta$ and a transition to negative values at $\delta = 0.29$. It has been argued in [25] that the stability of the ferromagnetic phase at small $\delta$ is due to the $\delta^2$-dependence of the kinetic energy of the $\downarrow$-spin electron. This particular $\delta^2$-dependence is, however, an artifact of the Gutzwiller wave function and is not a generic argument for the stability of the ferromagnetic state. This is not surprising as the result for $U = \infty$ is obtained by merely projecting out double occupancies from the $U = 0$ solution. The wave function implicitly used by Roth already shows a linear $\delta$-dependence of $E_{\text{kin}}^{\downarrow}$. We obtain for the (20 × 20) cluster with $\delta \leq 0.1$ the following

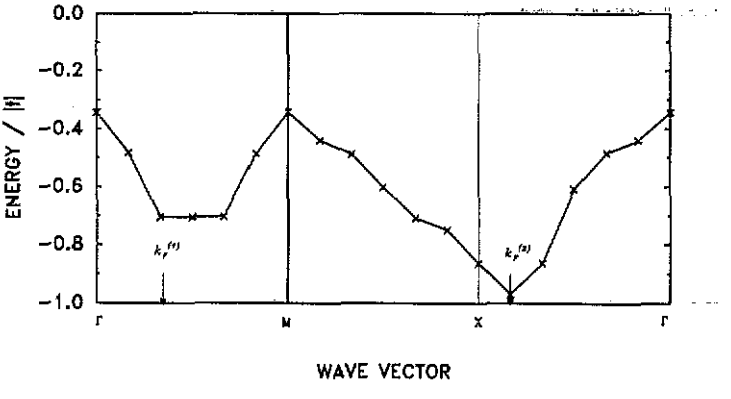


**Figure 5.** Low-lying excitations of the 2D Hubbard model for the 12 × 12 cluster with 83 holes ($\delta = 0.58$).

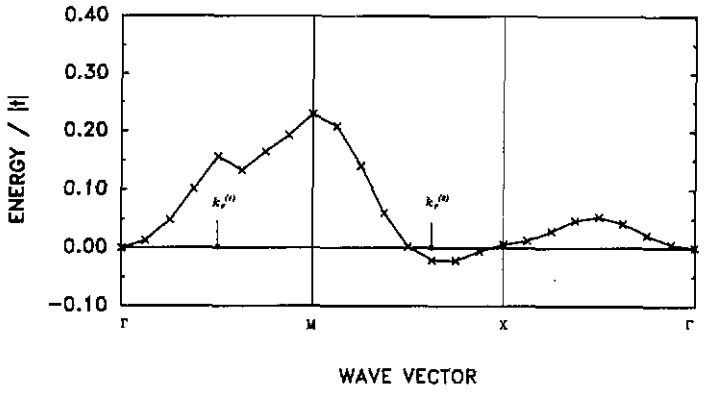


**Figure 6.** Energy dispersion of the 2D Hubbard model for the 16 × 16 cluster with 69 holes ($\delta = 0.27$).

dependence:

$$E^{\uparrow}_{\mathrm{kin}} = z|t|(0.7\delta + 3.7\delta^2) + E_0 \qquad E^{\downarrow}_{\mathrm{kin}} = -z|t|(0.2\delta + 4.0\delta^2). \tag{5.1}$$

Since $E^{\downarrow}_{\mathrm{kin}}$ corresponds to the bottom of the $\downarrow$-spin electron the band width is twice the value of the kinetic energy. The kinetic energy for the $\downarrow$-spin electron for arbitrary $\boldsymbol{k}$ is to a good approximation given by $E^{\downarrow}_{\mathrm{kin}}(\boldsymbol{k}) = q_{\downarrow}\,\varepsilon(\boldsymbol{k})$, as given in (3.9), and the narrowing factor $q_{\downarrow}$ for small $\delta$ would be, according to (5.1), $q_{\downarrow} = 0.2\delta + 4.0\delta^2$. At this point it is interesting to compare this with the result of the dominant term approximation discussed by Vollhardt [24]. This approximation has been employed by Rice and Ueda [42] for the periodic Anderson model within the framework of the heavy-fermion problem. The exact result for the energy of the Gutzwiller wave function for one reversed spin yields, according to (2.8), $q_{\downarrow} = \pi\delta^2$. The dominant term approximation leads to [24] $q_{\downarrow} = \delta$ in the thermodynamic limit. The exact evaluation of the Gutzwiller *ansatz* yields therefore

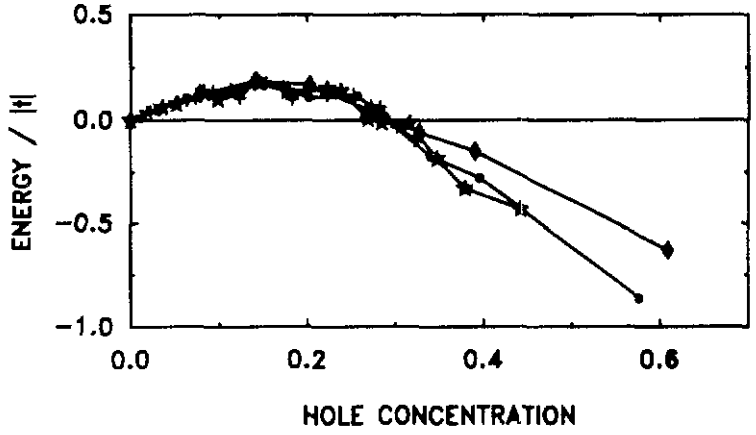


**Figure 7.** Excitation energy for $\boldsymbol{q} = (\pi, 0)$ as a function of the hole concentration (closed shells only) and for various lattice sizes: 8 × 8 (◇), 12 × 12 (●), 16 × 16 (★) and 20 × 20 (+).

the wrong $\delta$-dependence for $q_{\downarrow}$. The dominant-term approximation gives the right $\delta$-dependence, but with a prefactor that is one order of magnitude too big. This has an important impact on the question of ferromagnetism in the heavy-fermion problem.

We now consider what conclusions can be drawn from our results about the stability of the strongly ferromagnetic state in the thermodynamic limit $N \rightarrow \infty$. The discussion above, for example that of figure 6, indicates that this state first becomes unstable via single-particle excitation in which an electron is removed from the $\uparrow$-spin Fermi surface and placed in a state with $k = 0$ at the bottom of the $\downarrow$-spin quasi-particle band. The question is whether this excitation energy is positive or not, for a given hole concentration $\delta$. For closed-shell situations the wave vector $q$ of the excited state, relative to the unique Nagaoka state, is equal to a Fermi wave vector but for the range of $\delta$ of interest a close upper bound for the excitation energy is obtained with $q = (\pi, 0)$. This excitation energy is plotted in figure 7 as a function of $\delta$ for different system sizes. The small scatter, particularly at small $\delta$, suggests that these results are close to the thermodynamic limit and that there are no negative excitation energies for $\delta < 0.29$. It is interesting to compare figure 7 with figure 2, where the peaks in the data correspond to closed-shell results within Roth's approximation. In the latter case there is marked downward shift of the peaks from the $8 \times 8$ to the $16 \times 16$ cluster as the thermodynamic limit (essentially the curve for the $100 \times 100$ cluster) is approached. The absence of such an effect in figure 7 further supports the suggestion that the results in that figure are close to the thermodynamic limit.

In figure 8 we plotted the lowest excitation energy for any wave vector and for all numbers of holes in an $8 \times 8$ lattice. We obtain negative excitation energies even for $\delta < 0.29$. In cases where the lowest excitation energy is zero, corresponding to the trivial state of zero-wave-vector spin wave, the result plotted is for $q$ equal to a Fermi wave vector, to aid comparison with figure 2. There is an overall similarity between figures 2 and 8 but in the latter case the fluctuations between closed and open shells are reduced by a factor 2. If in the thermodynamic limit there are no negative excitation energies for $\delta < 0.29$ we expect the negative values in this range in figure 8 to tend to zero as $N \rightarrow \infty$ with increasing system size. To illustrate this tendency, figure 8 also contains results for the $16 \times 16$ cluster for a few typical hole concentrations, which exhibit the largest fluctuations. To sum up, we can conclude rigorously from figure 7 that the strongly ferromagnetic state is unstable for $\delta > 0.29$ and that it is very probably stable for $\delta < \delta_{cr}$ where $\delta_{cr}$ is somewhat smaller than 0.29 but non-zero.

We close the discussion about the infinite-$U$ limit with the results for the momentum distribution function. Tables 2 and 3 show the values of $n(k)$ for $\delta = 0.08$ and $\delta = 0.33$, respectively. Both results are for the X point in $k$-space. For the $\uparrow$-spin electrons the difference $\Delta n^{\uparrow}(k) = n^{\uparrow}(k) - n_0(k)$ from the Nagaoka state is given. In the first case of low density (table 2) we observe that $\Delta n^{\uparrow}(k)$ is uniformly distributed proportionally to $n^0(k)$. This is precisely the result we obtained for the simple spin wave *ansatz*. The same is true for the momentum distribution for the $\downarrow$-spin electron. We see in table 2($b$) that $n^{\downarrow}(k) \propto n^0(k - q)$. These findings confirm our conclusions, based on the dispersion curves, that the low-lying excitations of the Hubbard model for low hole concentrations ($\delta \leq 0.15$) are spin waves. Also, in agreement with our earlier results, the momentum distribution for $\delta = 0.33$ given in table 3 resembles very much that of a single-particle excitation. $n^{\uparrow}(k)$ is strongly reduced in a localized region near $k = (\pi, 0)$ and $n^{\downarrow}(k)$ has a clear structure near $k = (0, 0)$. The momentum distribution for infinite $U$ given in table 3 resembles much more closely the ideal case of a single-particle excitation ($U = 0$) than that of the Gutzwiller *ansatz* sketched in figure 1. However, the rather uniform $\Delta n^{\uparrow}(k)$

**Table 2.** Values for the momentum distribution in units of 0.01 for the $8 \times 8$ cluster with five holes ($\delta = 0.08$) for the state with momentum $q = (0, \pi)$. (*a*) For the $\uparrow$-spin electrons; the difference $\Delta n^{\uparrow}(k) = n^{\uparrow}(k) - n^{0}(k)$. (*b*) For the $\downarrow$-spin electron. The $k_x$ and $k_y$ coordinates, in units of $\pi/4$, are indicated at the bottom and left, respectively, of each part of the table.

(*a*)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 4 | 1 | −3 | −2 | −2 | −2 | −3 | 1 | 1 |
| 3 | −3 | −2 | −2 | −2 | −2 | −2 | −3 | 2 |
| 2 | −2 | −2 | −2 | −2 | −2 | −2 | −2 | −2 |
| 1 | −2 | −2 | −2 | −2 | −2 | −2 | −2 | −2 |
| 0 | −2 | −2 | −2 | −2 | −2 | −2 | −2 | −2 |
| −1 | −2 | −2 | −2 | −2 | −2 | −2 | −2 | −2 |
| −2 | −2 | −2 | −2 | −2 | −2 | −2 | −2 | −2 |
| −3 | −3 | −2 | −2 | −2 | −2 | −2 | −3 | 2 |
| | −3 | −2 | −1 | 0 | 1 | 2 | 3 | 4 |

(*b*)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 4 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 |
| 3 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 |
| 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 0 |
| 0 | 0 | 2 | 2 | 2 | 2 | 2 | 0 | 0 |
| −1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 0 |
| −2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| −3 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 |
| | −3 | −2 | −1 | 0 | 1 | 2 | 3 | 4 |

over much of the zone, which is spin-wave-like, indicates that the $\downarrow$-spin quasi-particle has a strong ($\uparrow$-spin particle + magnon) component as in the wave function (2.9). The physical picture is therefore as obtained by diagrammatic approaches [26–29]. For small hole concentration the single-particle excitations are much higher in energy than the Goldstone mode (spin waves), which have a $q^2$-behaviour at the $\Gamma$ point. The spin wave stiffness constant is, however, strongly reduced by correlation effects as compared with the simple spin wave or RPA result. With increasing hole density the gain in kinetic energy of the $\downarrow$-spin electron overcomes the loss in kinetic energy of the $\uparrow$-spin electrons and the energy for a spin-flip lies below the Fermi energy of the $\uparrow$-spin electrons. Above the critical hole concentration, the ferromagnetic state is unstable.

We conclude for $U = \infty$ that the strongly ferromagnetic state is the ground state for $\delta < \delta_{cr}$. On the other hand for $U = 0$ the ground state of the Hubbard model is non-magnetic. There exists therefore a critical on-site repulsion $U^*(\delta)$, below which the ferromagnetic state becomes unstble. As expected from strong-coupling perturbation theory, we find that the ground state energy behaves as follows: $E(U) = E(\infty) - \alpha/U$ ($\alpha > 0$). This illustrates the gain in energy due to short-range antiferromagnetic order. For the system sizes and hole concentrations under consideration $\alpha$ is of the order of

**Table 3.** Values for the momentum distribution in units of 0.01 for the 8 × 8 cluster with 21 holes ($\delta = 0.33$) for the state with momentum $q = (0, \pi)$. (*a*) For the ↑-spin electron; the difference $\Delta n^{\uparrow}(k) = n^{\uparrow}(k) - n^{0}(k)$ is given. (*b*) For the ↓-spin electron.

(*a*)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 4 | 1 | 2 | −2 | −2 | −2 | 2 | 1 | 0 |
| 3 | 1 | 4 | −2 | −2 | −2 | 4 | 1 | 0 |
| 2 | 2 | −2 | −2 | −2 | −2 | −2 | 2 | 1 |
| 1 | −4 | −2 | −2 | −2 | −2 | −2 | −4 | −21 |
| 0 | −4 | −2 | −2 | −2 | −2 | −2 | −4 | −9 |
| −1 | −4 | −2 | −2 | −2 | −2 | −2 | −4 | −21 |
| −2 | 2 | −2 | −2 | −2 | −2 | −2 | 2 | 1 |
| −3 | 1 | 4 | −2 | −2 | −2 | 4 | 1 | 0 |
| | −3 | −2 | −1 | 0 | 1 | 2 | 3 | 4 |

(*b*)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 4 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 0 | 1 | 2 | 2 | 2 | 1 | 0 | 0 |
| 1 | 0 | 1 | 3 | 3 | 3 | 1 | 0 | 0 |
| 0 | 0 | 1 | 10 | 6 | 10 | 1 | 0 | 0 |
| −1 | 0 | 1 | 3 | 3 | 3 | 1 | 0 | 0 |
| −2 | 1 | 1 | 2 | 2 | 2 | 1 | 0 | 0 |
| −3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | −3 | −2 | −1 | 0 | 1 | 2 | 3 | 4 |

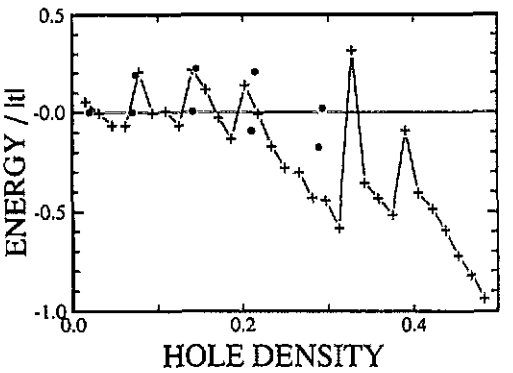


**Figure 8.** Excitation energy at $k = k_F$ (see text) for 8 × 8 (+) and all possible numbers of holes 1, . . . , 32 and for 16 × 16 (●) at a few important hole concentrations. The latter points occur in pairs, the upper one corresponding to a closed shell and the lower one to the closed shell minus one hole.

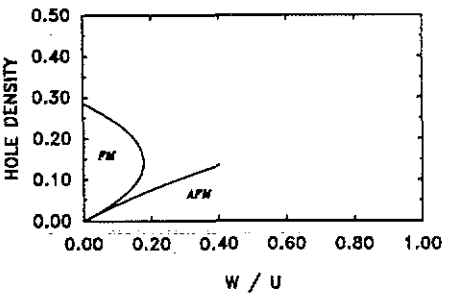


**Figure 9.** Phase diagram of the Hubbard model as function of $\delta$ and $W/U$, where $W$ is the band width.

$10|t|$. Assuming a constant $\alpha$ for small hole concentration—which is in very good agreement with numerical results—together with $E(\infty) = \gamma\delta$ (equation (5.1)) we conclude that the phase boundary in the ($\delta$, $W/U$) phase space is linear in $\delta$ as first predicted by Nagaoka [14]. In figure 9 the phase diagram obtained for the ($8 \times 8$) cluster is given. The strongly ferromagnetic state is unstable for all densities if $U < 42|t|$. These results are in qualitative agreement with those of [43], where a variational method was used exploiting a modified Gutzwiller wave function as proposed in [44]. Using a perturbative treatment, the authors of [44] found that no ferromagnetic phase exists for $U/W < 2$, which is the region for which the scheme yields reliable results.

A strong $U$ is thus necessary to stabilize the ferromagnetic state. The critical $U^*$ that we find is, however, significantly smaller than that obtained by recent slave-boson mean-field calculations [3]. We observe the following behaviour of the low-lying excitations on reducing $U$: the bottom of the single-particle continuum comes down in energy and pushes down the spin wave branch; the latter eventually becomes negative first at $k_F$; for small $U$ the minimum near $k_F$ is much more pronounced than has been found in the case of infinite $U$.

## 6. Summary

We have studied the low-lying states of the two-dimensional Hubbard model with one $\downarrow$-spin electron and an arbitrary number of $\uparrow$-spin electrons as a function of the on-site repulsion $U$ and the hole density $\delta$. The analysis has been done in the framework of a wave function, which was proposed by one of the authors [22]. It is known that the *ansatz* is exact in 1D. The underlying physical idea, however, is not restricted to 1D and comparison of numerical results with exact data for small clusters shows that the *ansatz* is very accurate in 2D as well. We have studied systems of size $8 \times 8$, $12 \times 12$, $16 \times 16$ and $20 \times 20$ for all possible hole concentrations. To see whether the ground state is ever ferromagnetic, we investigated the infinite-U case in great detail. It has been possible to follow a single branch of the spectrum throughout the entire Brillouin zone, by considering only closed-shell systems. For open-shell situations this is very difficult, as the non-interacting ground state is degenerate and several reference (Nagaoka) states are possible. The degeneracy of the Nagaoka state leads to a superposition of low-lying bands, each starting at a different point in the Brillouin zone. This is precisely what one expects in the thermodynamic limit. For $N \to \infty$ the Fermi volume can be rearranged slightly, which involves only energy changes of order $1/N$, to construct Nagaoka states with arbitrary total momentum $\boldsymbol{K}$. Each of these states can serve as the origin of a dispersion curve similar to the one we obtained for the closed-shell situations. Consequently, the bottom of the low-lying excitation spectrum should be dispersionless in the thermodynamic limit. However we have interpreted the dispersion curves of figures 3–6 for finite $N$ in terms of spin wave and single-particle excitations from a particular Nagaoka state and discussion of figures 7, 8 and 2 indicates that corresponding excitation energies are close to their values in the thermodynamic limit. From the dispersion curves and the spin-dependent momentum distribution it follows that for small hole density $\delta \leqslant 0.15$ the low-lying excitations are spin waves with a strongly reduced spin wave stiffness constant $D$. We find that $D$ is about four times smaller than obtained in the RPA. For larger hole concentration the spectrum near $k_F$ takes on a single-particle character and becomes soft at $k_F$. The energy becomes negative at a critical concentration $\delta_{cr} \simeq 0.29$. The critical hole concentration is considerably lower than that obtained earlier

using simpler variational wave functions. We conclude therefore that the 2D Hubbard model has a small region in the $(\delta, W/U)$ phase diagram where the strongly ferromagnetic state is the ground state. We find, however, that the generic argument for the stability of the ferromagnetic state, claimed by Shastry *et al*, is not valid. In this paper we have concentrated on the instability against a single spin flip. As discussed, the instability of the strongly ferromagnetic state indicates that the quasi-particle energies for the $\downarrow$-spin electrons dips below the Fermi energy of the $\uparrow$-spin electrons. This indicates, as found in exact diagonalization of very small clusters, that the energy can be lowered even further by flipping more than one spin. Whether the transition is directly to a non-magnetic state, as found in Hartree–Fock calculations by Oles [43], is not yet clear. Our results have, moreover, shown that the diagrammatic approaches proposed earlier [26–30], which implicitly used Roth's wave function, are rather accurate. These schemes have the advantage that they can be applied to an arbitrary number of $\downarrow$-spin electrons. We reckon that they should yield a fairly accurate basis for further studies of the Hubbard model.

**Acknowledgments**

We are very grateful to Peter Young for permission to make use of unpublished results [19]. We would also like to thank Jelle Beenen for his support in testing parts of the numerical schemes. One of us (WvdL) is grateful to the Max-Planck-Gesselschaft for the support of this project via the Otto-Hahn-Stipendium.

**Appendix**

*A1. Kinetic energy*

Here we derive the expression for the kinetic energy. To this end we define new operators:

$$b_\alpha^{(i)\dagger} = \sum_j \varphi_\alpha(x_j - x_i) a_{j\uparrow}^\dagger \tag{A1}$$

and rewrite the wave function:

$$|\chi(q)\rangle = \frac{1}{\sqrt{N}} \sum_i e^{iq\cdot x_i} a_{i\downarrow}^\dagger \prod_\alpha b_\alpha^{(i)\dagger} |0\rangle. \tag{A2}$$

The norm of the wave function is readily written down:

$$\langle\chi|\chi\rangle = \frac{1}{N} \sum_i \langle 0| \left(\prod_\alpha b_\alpha^{(i)}\right)\left(\prod_\alpha b_\alpha^{(i)\dagger}\right) |0\rangle. \tag{A3}$$

By means of Wick's theorem we obtain

$$\langle 0 | \left(\prod_\alpha b_\alpha^{(i)}\right)\left(\prod_\alpha b_\alpha^{(j)\dagger}\right) |0\rangle = \det(\mathbf{S}^{(i,j)}) \tag{A4}$$

with

$$S_{\alpha\beta}^{(i,j)} = \langle 0| b_\alpha^{(i)} b_\beta^{(j)\dagger} |0\rangle = \sum_l \varphi_\alpha^*(x_l)\varphi_\beta(x_l + x_i - x_j). \tag{A5}$$

For the norm we need only the result for $i = j$:

$$\langle\chi|\chi\rangle = \det(\mathbf{S}) \tag{A6}$$

where $\mathbf{S}$ is the overlap matrix $\mathbf{S}^{(0,0)}$.

The expectation value for the kinetic energy of the $\downarrow$-spin electron is obtained in a similar way:

$$\begin{aligned}\langle\chi|H_{\rm kin}^{\downarrow}|\chi\rangle/\langle\chi|\chi\rangle &= \frac{t}{N}\sum_{\langle i,j\rangle} e^{iq\cdot(x_j-x_i)}\det(\mathbf{S}^{(i,j)})/\det(\mathbf{S}) \\ &= t\sum_{\Delta} e^{-iq\cdot\Delta}\det(\mathbf{S}^{-1}\mathbf{S}^{(\Delta,0)})\end{aligned} \tag{A7}$$

where like in (A4) Wick's theorem has been employed.

The expectation value for the $\uparrow$-spin electrons is somewhat more complicated:

$$\langle\chi|H_{\rm kin}^{\uparrow}|\chi\rangle = \frac{t}{N}\sum_{\langle i,j\rangle}\sum_{l}\langle 0|\left(\prod_{\alpha} b_{\alpha}^{(l)}\right)a_{i\uparrow}^{\dagger}a_{j\uparrow}\left(\prod_{\alpha} b_{\alpha}^{(l)\dagger}\right)|0\rangle. \tag{A8}$$

We evaluate first the expression for the summand in a slightly generalized form, with different centres $l$, $l'$, as we will need it for the momentum distribution functions as well:

$$\begin{aligned}\langle 0|\left(\prod_{\alpha} b_{\alpha}^{(l')}\right)a_{i\uparrow}^{\dagger}a_{j\uparrow}\left(\prod_{\alpha} b_{\alpha}^{(l)\dagger}\right)|0\rangle &= \delta_{i,j}\langle 0|\left(\prod_{\alpha} b_{\alpha}^{(l')}\right)\left(\prod_{\alpha} b_{\alpha}^{(l)\dagger}\right)|0\rangle \\ &\quad - \langle 0|\left(\prod_{\alpha} b_{\alpha}^{(l')}\right)a_{i\uparrow}a_{j\uparrow}^{\dagger}\left(\prod_{\alpha} b_{\alpha}^{(l)\dagger}\right)|0\rangle.\end{aligned} \tag{A9}$$

According to (A4) the first term on the RHS is, apart from the Kronecker $\delta$,

$$\langle 0|\left(\prod_{\alpha} b_{\alpha}^{(l')}\right)\left(\prod_{\alpha} b_{\alpha}^{(l)\dagger}\right)|0\rangle = \det(\mathbf{S}^{(l',l)}). \tag{A10}$$

To evaluate the second term on the RHS of (A9) we introduce auxiliary orbitals:

$$b_0^{(l)\dagger} = a_{j\uparrow}^{\dagger} \qquad b_0^{(l')} = a_{i\uparrow}. \tag{A11}$$

Again invoking Wick's theorem, we obtain

$$\langle 0|\left(\prod_{\alpha} b_{\alpha}^{(l')}\right)a_{i\uparrow}a_{j\uparrow}^{\dagger}\left(\prod_{\alpha} b_{\alpha}^{(l)\dagger}\right)|0\rangle = \det(\mathbf{M}). \tag{A12}$$

The matrix $M_{\alpha\beta}$ is identical to $S_{\alpha\beta}^{(l',l)}$ for $\alpha,\beta = 1,\ldots,N_{\uparrow}$, but it contains one additional column and row with index 0:

$$\begin{aligned}&M_{00} = \langle 0|b_0^{(l')}b_0^{(l)\dagger}|0\rangle = \delta_{i,j} \qquad M_{0\beta} = \langle 0|b_0^{(l')}b_{\beta}^{(l)\dagger}|0\rangle = \varphi_{\beta}(x_i - x_l) \\ &M_{\alpha 0} = \langle 0|b_{\alpha}^{(l')}b_0^{(l)\dagger}|0\rangle = \varphi_{\alpha}^{*}(x_j - x_{l'}).\end{aligned} \tag{A13}$$

On adding to the 0th column of $\mathbf{M}$ a proper linear combination of the other columns, which leaves the determinant unchanged, one can make all elements of this column vanish, except the (0, 0) element. The determinant we seek is therefore

$$\det(\mathbf{M}) = \det(\mathbf{S}^{(l',l)})\left(1 - \sum_{\alpha,\beta=1}^{N_{\uparrow}}\varphi_{\alpha}(x_i - x_l)S_{\alpha\beta}^{(l',l)-1}\varphi_{\beta}^{*}(x_j - x_{l'})\right). \tag{A14}$$

The expression in brackets is the 0th element of the transformed 0th column. Collecting (A14), (A12), (A10), (A9) and (A8) we obtain

$$\begin{aligned}&\langle 0|\left(\prod_{\alpha} b_{\alpha}^{(l')}\right)a_{i\uparrow}^{\dagger}a_{j\uparrow}\left(\prod_{\alpha} b_{\alpha}^{(l)\dagger}\right)|0\rangle \\ &\quad = \det(\mathbf{S}^{(l',l)})\left(\sum_{\alpha,\beta=1}^{N_{\uparrow}}\varphi_{\alpha}(x_i - x_l)S_{\alpha\beta}^{(l',l)-1}\varphi_{\beta}^{*}(x_j - x_{l'})\right).\end{aligned} \tag{A15}$$

For the kinetic energy we only need $l = l'$. The first factor on the RHS of (A15) is therefore the norm of the wave function. Inserting (A15) into the kinetic energy expression (A8) and dividing by the norm leads to

$$\langle\chi|H^{\uparrow}_{\mathrm{kin}}|\chi\rangle/\langle\chi|\chi\rangle = t\sum_{\alpha\beta} S^{-1}_{\alpha\beta}\sum_{\langle ij\rangle} S^{(j,i)}_{\beta\alpha}. \tag{A16}$$

*A2. Eigenvalue equation for non-orthonormal orbitals*

To generate the energy expression in non-orthonormal orbitals

$$E = t\sum_{\Delta}\mathrm{tr}(\mathbf{S}^{-1}\mathbf{S}^{\Delta}) + t\sum_{\Delta}\det(\mathbf{S}^{-1}\mathbf{S}^{\Delta})\,\mathrm{e}^{-\mathrm{i}q\cdot\Delta}$$

$$+ U\sum_{\alpha\beta} S^{-1}_{\beta\alpha}\varphi^{*}_{\alpha}(0)\varphi_{\beta}(0) \tag{A17}$$

with respect to $\varphi^{*}_{\eta}(i)$ we will make use of the following results:

$$\frac{\partial}{\partial\varphi^{*}_{\eta}(i)}S^{(\Delta)}_{\alpha\beta} = \delta_{\alpha\eta}\varphi_{\beta}(x_i+\Delta) \qquad \frac{\partial}{\partial\varphi^{*}_{\eta}(i)}S^{-1}_{\alpha\beta} = -S^{-1}_{\alpha\eta}\sum_{\gamma}\varphi_{\gamma}(x_i)S^{-1}_{\gamma\beta}$$

$$\frac{\partial}{\partial\varphi^{*}_{\eta}(i)}\det(S^{(\Delta)}_{\alpha\beta}) = \det(S^{(\Delta)}_{\alpha\beta})\sum_{\gamma}\varphi_{\gamma}(x_i+\Delta)S^{(\Delta)^{-1}}_{\gamma\eta}. \tag{A18}$$

We perform the following tedious but straightforward steps:

(i) differentiate (A17) by using (A18) and equate the result to zero;
(ii) multiply both sides with $S^{-1/2}_{\eta\rho}$ and sum over $\eta$;
(iii) perform a Löwdin transformation to orthonormal orbitals:

$$\tilde{\varphi}_{\alpha} = \sum_{\beta}\varphi_{\beta}S^{-1/2}_{\beta\alpha} \tag{A19}$$

and we end up with the eigenvalue equation for the orthonormal orbitals $\tilde{\varphi}_{\alpha}$:

$$E_{\eta}\tilde{\varphi}_{\eta}(x_i) = t\sum_{\Delta\gamma} e^{-\mathrm{i}q\cdot\Delta}\det(\tilde{\mathbf{S}}^{(\Delta)})\tilde{\varphi}_{\gamma}(x_i+\Delta)\tilde{S}^{(\Delta)^{-1}}_{\gamma\eta}$$

$$+ t\sum_{\Delta}\tilde{\varphi}_{\eta}(x_i+\Delta) - t\sum_{\Delta}\tilde{\varphi}_{\gamma}(x_i)\tilde{S}^{(\Delta)}_{\gamma\eta}$$

$$+ U\left(\delta_{i,0} - \sum_{\gamma}\tilde{\varphi}_{\gamma}(x_i)\tilde{\varphi}^{*}_{\gamma}(0)\right)\tilde{\varphi}_{\eta}(0). \tag{A20}$$

The overlap matrices $\tilde{\mathbf{S}}$ are evaluated with the orthonormal orbitals.

*A3. Momentum distribution*

In this section we assume orthonormal one particle orbitals $\varphi_\alpha$. The overlap matrix $\mathbf{S}$ is therefore the unit matrix and the norm $\langle\chi|\chi\rangle = 1$. By making use of the operators $b^\dagger$, $b$ of section A1 of this appendix, the momentum distribution for the $\uparrow$-spin electrons is:

$$\langle\chi|\eta_{k\uparrow}|\chi\rangle = \frac{1}{N^2}\sum_{l,l'}\exp(\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}_l - \boldsymbol{x}_{l'}))\sum_i \langle 0|\left(\prod_\alpha b_\alpha^{(i)}\right) a_{l\uparrow}^\dagger a_{l'\uparrow}\left(\prod_\alpha b_\alpha^{(i)\dagger}\right)|0\rangle. \tag{A21}$$

Together with (A17) for orthonormal orbitals we obtain

$$\langle\chi|n_{k\uparrow}|\chi\rangle = \sum_\alpha\left|\frac{1}{\sqrt{N}}\sum_l \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}_l}\varphi_\alpha(\boldsymbol{x}_l)\right|^2 = \sum_\alpha|\varphi_\alpha(\boldsymbol{k})|^2 \tag{A22}$$

with $\varphi_\alpha(\boldsymbol{k})$ being the Fourier transform of $\varphi_\alpha(\boldsymbol{x})$, as is obvious from (A22). For the $\downarrow$-spin electron

$$\langle\chi|n_{k\downarrow}|\chi\rangle = \frac{1}{N^2}\sum_{i,j}\exp(\mathrm{i}(\boldsymbol{k}-\boldsymbol{q})\cdot(\boldsymbol{x}_i - \boldsymbol{x}_j))\langle 0|\left(\prod_\alpha b_\alpha^{(i)}\right)\left(\prod_\alpha b_\alpha^{(j)\dagger}\right)|0\rangle. \tag{A23}$$

Along with (A12) we finally obtain the desired expression:

$$\langle\chi|n_{k\downarrow}|\chi\rangle = \frac{1}{N}\sum_l \mathrm{e}^{\mathrm{i}(\boldsymbol{k}-\boldsymbol{q})\cdot\boldsymbol{x}_l}\det(\mathbf{S}^{(l,0)}). \tag{A24}$$

## References


[1] Hubbard J 1963 *Proc. R. Soc.* A **276** 238
[2] Kanamori J 1963 *Prog. Theor. Phys.* **30** 275
[3] Kotliar G and Ruckenstein A E 1986 *Phys. Rev. Lett.* **57** 1362
[4] Hirsch J 1985 *Phys. Rev.* B **31** 4403
[5] Lieb E and Wu F Y 1968 *Phys. Rev. Lett.* **20** 1445
[6] Ogata M and Shiba H 1990 *Phys. Rev.* B **41** 2326
[7] Lieb E and Mattis D C 1962 *Phys. Rev.* **125** 164
[8] Lieb E 1989 *Phys. Rev. Lett.* **62** 1201
[9] Rice T M 1990 *Helv. Phys. Acta* **63** 336
[10] Luther A and Peschel I 1975 *Phys. Rev.* B **12** 3908
[11] Horsch P and von der Linden W 1988 *Z. Phys.* B **72** 181
[12] Reger J D and Young A P 1988 *Phys. Rev.* B **37** 5978
[13] Birgeneau R J 1990 *Physical Properties of High Temperature Superconductors II* ed. D M Ginsberg (Singapore: World Scientific)
[14] Nagaoka Y 1966 *Phys. Rev.* **147** 392
[15] The same result had been found independently by Thouless D J 1965 *Proc. Phys. Soc.* **86** 893
[16] Doucot B and Wen X G 1989 *Phys. Rev.* B **40** 2719
[17] Takahashi M 1982 *J. Phys. Soc. Japan* **51** 3475
[18] Fang Y, Ruckenstein A E, Dagotto E and Schmitt-Rink S 1989 *Phys. Rev.* B **40** 7406
[19] Barbieri A, Riera J A and Young A P unpublished
[20] Riera J A and Young A P 1989 *Phys. Rev.* B **40** 5285
[21] Barbieri A, Riera J A and Young A P 1990 *Phys. Rev.* B **41** 11697
[22] Edwards D M 1990 *Prog. Theor. Phys. Suppl.* **101** 453
[23] Gutzwiller M C 1963 *Phys. Rev. Lett.* **10** 159
[24] Vollhardt D 1984 *Rev. Mod. Phys.* **56** 99
[25] Shastry B S, Krishnamurthy H R and Anderson P W 1990 *Phys. Rev.* B **41** 2375

[26] Roth L 1969 *Phys. Rev.* **186** 428
[27] Edwards D M 1968 *J. Appl. Phys.* **39** 481
[28] Edwards D M 1967 *Proc. R. Soc.* A **300** 373
[29] Edwards D M and Hertz J 1973 *J. Phys. F: Met. Phys.* **3** 2191
[30] Tan B W 1974 *PhD Thesis* London
[31] Allan S R and Edwards D M 1982 *J. Phys. F: Met. Phys.* **12** 1203
[32] Basile A G and Elser V 1990 *Phys. Rev.* B **41** 4842
[33] Mattis D C 1981 *The Theory of Magnetism I* (*Springer Series in Solid-State Sciences*) ed P Fulde (Berlin: Springer)
[34] Wigner E P and Seitz F 1934 *Phys. Rev.* **46** 509
[35] Richmond P and Rickayzen G 1969 *J. Phys. C: Solid State Phys.* **2** 528
[36] McGuire J B 1965 *J. Math. Phys.* **6** 432
[37] Doucot B and Rammal R 1990 *Phys. Rev.* B **41** 9617
[38] Kirkpatrick S, Gelatt C D and Vecchi M P 1983 *Science* **220** 671
[39] Blankenbecler R, Scalapino I J and Sugar R L 1981 *Phys. Rev.* B **24** 4295
[40] Wilkinson J H 1965 *The Algebraic Eigenvalue Problem* (Oxford: Clarendon)
[41] Preuss W H, Flannery B P, Teukolsky S A and Veterling W T 1988 *Numerical Recipe* (Cambridge: Cambridge University Press)
[42] Rice T M and Ueda K 1986 *Phys. Rev.* B **34** 6420
[43] Oles A 1982 *J. Phys. C: Solid State Phys.* **15** L1065
[44] Stollhoff G and Fulde P 1977 *Z. Phys.* B **20** 257