Analytic calculation of ground-state properties of correlated fermions with the Gutzwiller wave function

Walter Metzner, Dieter Vollhardt

Analytic calculation of ground-state properties of correlated fermions with the Gutzwiller wave function

Walter Metzner* and Dieter Vollhardt*

*Max-Planck-Institut für Physik und Astrophysik, Werner-Heisenberg-Institut für Physik, Postfach 40 12 12,
D-8000 München 40, Federal Republic of Germany*

A new analytic approach to the evaluation of ground-state properties of Hubbard-type models in terms of the Gutzwiller variational wave function is presented. It is based on the observation that expectation values in terms of this wave function may be expressed by sums over different lattice sites. This makes the application of Wick's theorem and the resulting contractions extremely simple, since the latter involve only anticommuting numbers as in a Grassmann algebra. Expressions for the momentum distribution $n_{k\sigma}$ and the Hubbard interaction in terms of a power series in a particular correlation parameter are derived which are valid for all dimensions. An explicit diagrammatic evaluation of the coefficients is described. In one dimension these coefficients may be determined to all orders which yields an approximation-free calculation of n_k and ground-state energy E for arbitrary density n and interaction strength U. In the case of a half-filled band and large U, hitherto unexpected nonanalyticities are found. The results allow for the first approximation-free assessment of the properties of the Gutzwiller wave function. It is shown that the well-known Gutzwiller approximation may be derived diagrammatically, too. The approximation is seen to yield the exact results for expectation values in terms of this wave function in the limit of infinite dimensionality.

I. INTRODUCTION

Interactions in Fermi systems lead to a particularly complex quantum-mechanical many-body problem. Even in the noninteracting case, the Pauli principle implies spatial correlations between fermions. Interactions, especially those of short range, further enhance these complications since they introduce additional subtle phase relations among the fermions in real space. In condensed-matter physics these difficulties are well known from the investigations of narrow-band metals[1] or liquid[2] and solid ^{3}He.[3] The theoretical developments in new areas such as heavy-fermion systems,[4] disordered electronic systems close to the metal-insulator transition,[5,6] and, most recently, high-temperature superconductivity[7,8] have once again drawn attention to these problems and have further challenged our understanding of strongly correlated fermions.

Theoretical studies of systems with short-range, repulsive interactions face the problem of how to incorporate an interaction that is simple only in real space into a noninteracting Fermi system that is simple (if at all) in **k** space. In this situation variational-type methods (as, for example, those developed for microscopic investigations of liquid helium[2]) are especially useful. Starting with an appropriate many-body trial wave function, the energy expectation value is calculated, which then has to be minimized with respect to some variational function. The simplest such wave function is an antisymmetrized product of one-particle functions; the next step of refinement involves two-particle functions, etc.[9]

In the case of electrons with a strongly screened Coulomb interaction a similar problem arises—but it now involves a lattice. Since electrons are pointlike, this interaction may be approximated by an on-site interaction between particles of opposite spin. Combined with a kinetic energy this lattice model becomes the "Hubbard model."[10-12]

In spite of the existence of an analytic solution in $d=1$ dimension[13] and an impressive research activity in the past, the properties of this seemingly simple model are far from being understood. In fact, even $d=1$ correlation functions are only known in very special limits.[14] In this situation numerical (Monte Carlo) methods have been particularly valuable and have yielded important insight into the properties of this model in different dimensions.[15,16] On the other hand, this approach is necessarily confined to the study of finite samples.

For an analytic investigation of the model, Gutzwiller[10,17,18] has proposed a very simple variational wave function [the "Gutzwiller wave function" (GWF)], similar to the general type described above: it introduces correlations into the noninteracting wave function via a local correlation factor in real space. He then introduced a further approximation [the "Gutzwiller approximation" (GA)] to calculate the ground-state energy.[18] In the GA, spatial correlations are neglected. It has been shown to be equivalent to an evaluation of matrix elements which calculates the classical statistical weights of different spin configurations in the noninteracting wave function.[19,20] Brinkman and Rice[21] observed that in the case of a half-filled band the results of the GA describe a transition—at a finite interaction strength U_c—to a localized state where lattice sites are singly occupied (metal-insulator transition). Although the transition itself is a consequence of the GA at half-filling, at least in one dimension, the results for the not yet localized regime ($U \lesssim U_c$) have been shown to describe static properties of normal liquid ^{3}He (Ref. 22) ("almost localized" Fermi liquid[20]). In this case the short-range interaction is be-

tween spherical hard cores and may be thought to take place within a lattice *cell* rather than on a lattice *site*. In fact, as recently discussed by Vollhardt, Wölfle, and Anderson[23] in the context of a generalized lattice model, these properties are also obtained without the actual transition. As shown by Kotliar and Ruckenstein[24] the results of the GA may be derived by a saddle-point approximation to a suitably chosen functional-integral representation of the Hubbard model, i.e., without using a variational wave function.

The result for the Hubbard model obtained by Gutzwiller[18] is a consequence of two independent approximations: (i) the choice of the GWF and (ii) the GA itself. So it is not clear to what extent this result depends on either one of the approximations. To decide on this point, one has to calculate expectation values in terms of the GWF by going beyond the GA. This was first done by Kaplan, Horsch, and Fulde[25] and Horsch and Kaplan,[26] who used numerical techniques to calculate the ground-state energy and spin correlations for finite, one-dimensional rings. They found that spin correlations in the atomic limit were in very good agreement with exact results,[27,14] while in this limit the ground-state energy deviated considerably from the Lieb-Wu result.[13] Similarly, using a combination of analytic and numerical methods, Hashimoto[28] obtained improved approximations for one- and higher-dimensional systems in the thermodynamic limit. On the other hand, employing a perturbational approach Horsch[29] and Baeriswyl and Maki[30] and Baeriswyl, Carmelo, and Maki[31] analytically calculated the ground-state energy E in terms of the GWF for a small interaction strength U. The momentum distribution n_k was thereby obtained to second order in U.[31] More recently, Gros, Joynt, and Rice[32] presented detailed numerical results, e.g., for spin-spin and hole-hole correlations, energies and excited states in terms of the GWF in $d=1$ in the infinite-U limit. Independently, Yokoyama and Shiba[33] computed the q-dependent spin-spin and density-density correlation functions, the momentum distribution, and also the ground-state energy for general U, the latter even in $d=2$ and 3 dimensions.

In spite of the above results and the simplicity of the GWF, an *exact* analytic evaluation of expectation values in terms of the GWF for arbitrary interaction strengths did not exist so far—not even in $d=1$ dimension, where complications due to lattice structures do not enter. Most recently, however, the present authors showed that an analytic diagonalization is indeed possible, at least in $d=1$.[34] This calculation is made feasible by the observation that in the case of the GWF lattice summations may be written so as to involve only *different* sites. Thereby the application of Wick's theorem for the evaluation of expectation values is greatly simplified. In this way one is able to calculate expectation values in $d=1$ without approximation. The approach was then used by Gebhard and Vollhardt[35] to evaluate correlation functions for Hubbard-type models and to obtain the exact results for the GWF in $d=1$.

Since our approach[34] consists of a sequence of steps which seem to deserve a rather detailed discussion, this paper is supposed to provide an explicit presentation of

the method. Although it may be applied in all space dimensions d, we will mainly discuss $d=1$ where approximation-free results have already been obtained, as well as the case $d=\infty$ which may be shown to yield the results of the GA.

This paper is structured as follows. In Sec. II we introduce the variational approach itself, while in Secs. III and IV general expressions for the expectation values of the Hubbard interaction and of the momentum distribution are derived which are evaluated analytically in one dimension. The resulting ground-state energy of the Hubbard model in $d=1$ is discussed in Sec. V. Based on the formalism described in Secs. III and IV, a diagrammatic derivation of the Gutzwiller approximation is given in Sec. VI. A summary in Sec. VII closes the presentation.

II. LOCAL VARIATIONAL APPROACH FOR HUBBARD-TYPE MODELS

We will discuss a variational approach to a model of fermions with spin up or down on a lattice, whose Hamiltonian is composed of a kinetic energy $\hat{H}_{\text{kin}}$ and an interaction $\hat{H}_I$:

$$\hat{H}_{\text{kin}} = \sum_{i,j,\sigma} t_{ij} \hat{c}_{i\sigma}^{\dagger} \hat{c}_{j\sigma} \tag{1a}$$

$$= \sum_{k,\sigma} \varepsilon(k) \hat{n}_{k\sigma} , \tag{1b}$$

$$\hat{H}_I = U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} . \tag{2}$$

Here $\hat{c}_{i\sigma}^{\dagger}$ and $\hat{a}_{k\sigma}^{\dagger}$ are creation operators of a fermion at site i and of momentum k, respectively, with

$$\hat{a}_{k\sigma}^{\dagger} = L^{-1/2} \sum_i \hat{c}_{i\sigma}^{\dagger} \exp(i k \cdot R_i) ,$$

$\hat{n}_{i\sigma} = \hat{c}_{i\sigma}^{\dagger} \hat{c}_{i\sigma}$ and $\hat{n}_{k\sigma} = \hat{a}_{k\sigma}^{\dagger} \hat{a}_{k\sigma}$ are the corresponding number operators, t_{ij} is a general hopping matrix element between sites R_i and R_j, and

$$\varepsilon(k) = L^{-1} \sum_{ij} t_{ij} \exp[-i k \cdot (R_i - R_j)]$$

is the corresponding dispersion. The "Hubbard interaction" $\hat{H}_I$ only takes place on the same lattice site. For nearest-neighbor hopping,

$$\hat{H} = \hat{H}_{\text{kin}} + \hat{H}_I \tag{3}$$

is the Hubbard model. More complicated interactions of the form

$$\hat{H}_V = \sum_{\sigma,\sigma'} \sum_{i,j} V_{ij,\sigma\sigma'} \hat{n}_{i\sigma} \hat{n}_{j\sigma'} , \tag{4}$$

which are determined by correlation functions, may in principle also be included, since our approach is applicable in this situation, also.[35] However, in this paper they will not be addressed for the sake of brevity.

We consider a fixed number $N_\uparrow, N_\downarrow$ of up and down spin particles; hence $N = N_\uparrow + N_\downarrow$ and $S_z = N_\uparrow - N_\downarrow$ are also fixed; L is the number of lattice sites and $n_\sigma = N_\sigma / L$, $n = n_\uparrow + n_\downarrow$. The interaction $\hat{H}_I$ counts the number of doubly occupied sites $\hat{D}_j = \hat{n}_{j\uparrow} \hat{n}_{j\downarrow}$, i.e.,

$$\hat{H}_I = U\hat{D} \ , \tag{5}$$

where $\hat{D} = \sum_i \hat{D}_i$.

To diagonalize (3) we introduce a local variational wave function[18,36]

$$| \psi \rangle = \prod_f [1 - (1-g)\hat{O}_f] | \psi_0 \rangle \ , \tag{6}$$

where $\hat{O}_f$ is a local number operator at lattice site f, g is a variational parameter, and $| \psi_0 \rangle$ is a suitable starting wave function onto which the correlation factor in (6) acts. By "local" we imply that $\hat{O}_f$ only acts on site f; more complicated forms, which include neighboring lattice sites, e.g., density-density-like correlations, have been addressed by Stollhoff and Fulde[36] and Kaplan et al.[25] In the lattice model under consideration, the only relevant local correlation operator of this type is $\hat{O}_f = \hat{n}_{f\uparrow}\hat{n}_{f\downarrow} = \hat{D}_f$. Thus we are led to the wave function proposed by Gutzwiller,[18]

$$| \psi_G \rangle = \prod_f [1 - (1-g)\hat{D}_f] | \psi_0 \rangle \ , \tag{7}$$

where we take $| \psi_0 \rangle$ to be the normalized wave function of the noninteracting, paramagnetic ground state and $0 \le g \le 1$. The correlation factor reduces the amplitude of spin configurations in $| \psi_0 \rangle$ with too many doubly occupied sites, i.e., interactions, and thus controls local density fluctuations. We note that since $\hat{D}_f$ only has eigenvalues 0 and 1, (7) may equally be written as

$$| \psi_G \rangle = g^{\hat{D}} | \psi_0 \rangle \ . \tag{8}$$

The GWF (7) will be used to calculate the expectation value of some operator $\hat{A}$ as

$$A \equiv \langle \hat{A} \rangle \equiv \frac{\langle \psi_G | \hat{A} | \psi_G \rangle}{\langle \psi_G | \psi_G \rangle} \ . \tag{9}$$

Furthermore, we introduce

$$A_0 \equiv \langle \hat{A} \rangle_0 \equiv \langle \psi_0 | \hat{A} | \psi_0 \rangle \ . \tag{10}$$

In the following we will diagonalize the Hamiltonian $\hat{H}$ in (3).

III. EVALUATION OF THE HUBBARD INTERACTION

We now calculate the expectation value of the Hubbard interaction term, i.e., $\langle \hat{D} \rangle = \langle \hat{H}_I \rangle / U$ in (5) in arbitrary dimensions d. We first consider the numerator $\langle \psi_G | \hat{D} | \psi_G \rangle$ in (9). Inserting (7) yields

$$\langle \psi_G | \hat{D} | \psi_G \rangle = \sum_f \left\langle \prod_h [1 - (1-g)\hat{D}_h]^2 \hat{D}_f \right\rangle_0 , \tag{11}$$

where $[1 - (1-g)\hat{D}_f]^2 = 1 + (g^2 - 1)\hat{D}_f$, since $\hat{D}_f^2 = \hat{D}_f$. The product may be transformed into a sum over lattice sites by using the identity

$$\prod_h [1 + (g^2 - 1)\hat{D}_h] = 1 + \sum_{m=1}^{L} \frac{(g^2 - 1)^m}{m!}$$

$$\times \sum_{f_1, \dots, f_m}' \hat{D}_{f_1} \cdots \hat{D}_{f_m} \ , \tag{12}$$

where the prime on the sum means $f_i \ne f_j$ for $i \ne j$. In the sum over f in (11) one has to distinguish whether $f \ne f_i$ for all i or $f = f_i$ for one of the i. In the first case we may set $f = f_{m+1}$ and shift the summation; in the second we have $\hat{D}_{f_i}^2 = \hat{D}_{f_i}$. Combination of the two parts yields

$$\langle \psi_G | \hat{D} | \psi_G \rangle = g^2 \sum_{m=1}^{L} \frac{(g^2 - 1)^{m-1}}{(m-1)!}$$

$$\times \sum_{f_1, \dots, f_m}' \langle \hat{D}_{f_1} \cdots \hat{D}_{f_m} \rangle_0 \ . \tag{13}$$

In the thermodynamic limit $(L \to \infty)$ this is a power series in $g^2 - 1$.

To evaluate (13) we define the mth-order coefficient

$$\tilde{c}_m = \frac{1}{L} \frac{1}{(m-1)!} \sum_{f_1, \dots, f_m}' \langle \hat{D}_{f_1} \cdots \hat{D}_{f_m} \rangle_0 \ , \tag{14}$$

which is calculated as usual by Wick's theorem,[37] yielding the sum over all contracted terms. There are only two nonvanishing contractions: $\langle \hat{c}_{f_i\sigma}^\dagger \hat{c}_{f_j\sigma} \rangle_0$ and $\langle \hat{c}_{f_i\sigma} \hat{c}_{f_j\sigma}^\dagger \rangle_0$, where the former is the one-particle density matrix[29,30]

$$P_{f_i f_j, \sigma} \equiv \langle \hat{c}_{f_i\sigma}^\dagger \hat{c}_{f_j\sigma} \rangle_0 \tag{15}$$

$$= \int d\mathbf{k} \, e^{-2\pi i \mathbf{k} (f_i - f_j)} n_{\mathbf{k}\sigma}^0 \tag{16}$$

with $n_{\mathbf{k}\sigma}^0$ the momentum-distribution of σ spins in the noninteracting system. Because of the summation restriction in (14) ($f_i \ne f_j$) the second contraction yields

$$\langle \hat{c}_{f_j\sigma} \hat{c}_{f_i\sigma}^\dagger \rangle_0 = -P_{f_i f_j, \sigma} \ , \tag{17}$$

i.e., the usual δ term does not enter. Here and in the following, the lattice spacing is chosen such that the volume of the primitive unit cell is $V_c = 1$. In $\mathbf{k}$ space the scale is set so as to yield the volume V_c^* of a primitive unit cell in reciprocal space also as $V_c^* = 1$; hence, in the thermodynamic limit, $L^{-1} \sum_{\mathbf{k}} \to \int d\mathbf{k}$.

We now define the sum over all possible contracted terms in (14) by a curly bracket. Wick's theorem then reads

$$\langle \hat{D}_{f_1} \cdots \hat{D}_{f_m} \rangle_0 = \{ D_{f_1} \cdots D_{f_m} \}_0, \quad f_i \ne f_j \tag{18}$$

where

$$\{ c_{f_i\sigma}^\dagger c_{f_j\sigma} \}_0 \equiv P_{f_i f_j, \sigma} \ , \tag{19a}$$

$$\{ c_{f_j\sigma} c_{f_i\sigma}^\dagger \}_0 \equiv -P_{f_i f_j, \sigma} \ . \tag{19b}$$

Equation (19) is a definition of "δ-less contractions," which is supposed to apply for all f_i, f_j, i.e., even for $f_i = f_j$ although in this case the usual $\delta_{f_i f_j}$ term does not occur in the definition. This may be done for the following reason.

The expression $\{ D_{f_1} \cdots D_{f_m} \}_0$ can be written as a product of two determinants,

$$\{D_{f_1} \cdots D_{f_m}\}_0 = \{n_{f_1\uparrow} \cdots n_{f_m\uparrow}\}_0 \{n_{f_1\downarrow} \cdots n_{f_m\downarrow}\}_0 \, ,$$

(20)

where

$$\{n_{f_1\sigma} \cdots n_{f_m\sigma}\}_0 = \begin{vmatrix} P_{11,\sigma} & \cdots & P_{1m,\sigma} \\ \vdots & & \vdots \\ P_{m1,\sigma} & \cdots & P_{mm,\sigma} \end{vmatrix}$$

(21)

and $f_i \equiv i$, etc., as originally used by Gutzwiller.[10,18] We note that for $f_i = f_j$ two columns (or rows) are equal and the determinant vanishes. Hence the summation restriction $f_i \neq f_j$ may be dropped without creating new contributions. Therefore $\tilde{c}_m$, (14), is equal to

$$\tilde{c}_m = \frac{1}{L} \frac{1}{(m-1)!} \sum_{f_1, \ldots, f_m} \{D_{f_1} \cdots D_{f_m}\}_0 \, ,$$

(22)

where $\{ \cdots \}_0$ is still given in terms of the δ-less contractions in (19), although now the f_i may also be equal.

We should like to stress that the last few steps imply significant differences between our approach and earlier ones[29,30,33] although it employs the same techniques. In particular, the unrestricted sum in (22) can only be obtained if $\{ \cdots \}_0$ does not contain $\delta_{f_i f_j}$ terms. Furthermore, the objects in $\{ \cdots \}_0$ are no longer operators. In fact, the creation and annihilation operators now become *anticommuting numbers* and may be looked upon as Grassmann variables.

A. Diagrammatic representation

The contributions to $\tilde{c}_m$ in (22) may be represented diagrammatically. To this end one draws m points (corresponding to sites $f_1, \ldots, f_m$) and connects these points by lines according to the $P_{f_i f_j,\sigma}$ expressions obtained when the determinants (21) are multiplied out. Up-spin and down-spin lines may be distinguished by drawing full and broken lines, respectively. For example, for $m=2$ one has

$$\{D_{f_1} D_{f_2}\}_0 = P_{11,\uparrow} P_{11,\downarrow} P_{22,\uparrow} P_{22,\downarrow} + P_{12,\uparrow} P_{21,\downarrow} P_{12,\downarrow} P_{21,\uparrow}$$

$$- P_{11,\downarrow} P_{12,\uparrow} P_{21,\uparrow} P_{22,\downarrow}$$

$$- P_{11,\uparrow} P_{12,\downarrow} P_{21,\downarrow} P_{22,\uparrow} \, ,$$

(23)

where $f_1 \equiv 1$, etc. Equation (23) has to be summed over according to (22). The corresponding graphical representation of $\tilde{c}_2$ is shown in Fig. 1. In general, different contracted terms may lead to the same graph. We thus define the "weight" $w(G)$ of a graph G:

FIG. 1. Diagrammatic representation of the contractions of $\tilde{c}_2$.

$$w(G) \equiv \frac{1}{(m-1)!} c_G \, ,$$

(24a)

where c_G is the number of contracted terms with graph G. Furthermore, the "value" $v(G)$ of a graph is defined as

$$v(G) = \frac{1}{L} \sum_{f_1, \ldots, f_m} (\text{graph } G) \, .$$

(24b)

The evaluation of $v(G)$ is most conveniently done in k space by associating $n^0_{k\sigma}$ with every line and demanding momentum conservation at every vertex. Last, every graph carries a prefactor $(-1)^{f(G)}$, where $f(G)$ is the total number of loops of up-spin and down-spin lines in graph G.

We now turn to the norm $\langle \psi_G | \psi_G \rangle$ in (9), which is similarly obtained as

$$\langle \psi_G | \psi_G \rangle = 1 + \sum_{m=1}^{\infty} \frac{(g^2-1)^m}{m!} \sum_{f_1, \ldots, f_m} \{D_{f_1} \cdots D_{f_m}\}_0 \, .$$

(25)

As usual[37] the disconnected graphs of the numerator $\langle \psi_G | \hat{A} | \psi_G \rangle$ cancel the norm.

In the context of a variational approach this "linked cluster theorem" has been generally proved by Horsch and Fulde.[38] For the theorem to work, the lattice summation has to be unrestricted as in (22) and this is only achieved after the δ-less contractions have been introduced via Wick's theorem. Indicating the contribution of the *connected* graphs to $\tilde{c}_m$ by omitting the tilde, i.e., c_m, we define

$$c_m \equiv \frac{1}{L} \frac{1}{(m-1)!} \sum_{f_1, \ldots, f_m} \{D_{f_1} \cdots D_{f_m}\}_0^c \, ,$$

(26)

where $\{ \cdots \}_0^c$ is now given by all contractions corresponding to connected diagrams. Hence we obtain the expectation value of $\hat{D}$, including the norm, as

$$\langle \hat{D} \rangle = L g^2 \sum_{m=1}^{\infty} (g^2-1)^{m-1} c_m \, .$$

(27)

The diagrammatic evaluation of the c_m is summarized by the following Feynman rules.

(i) Draw all topologically different, connected graphs with m point vertices, m up-lines, and m down-lines such that every vertex is crossed by one up line and one down line.

(ii) Associate a momentum with every line, observing momentum conservation at each vertex (there are $m + 1$ internal momenta).

(iii) Every σ spin line with momentum k is associated with a factor $n^0_{k\sigma}$.

(iv) Integrate over all momenta to obtain the value $v(G)$, (24b), according to

$$\frac{1}{L} \sum_{f_1, \ldots, f_m} \to \int d\mathbf{k}_1 \cdots d\mathbf{k}_{m+1} \cdot$$

(v) Determine the number of loops $f(G)$, and associate

a prefactor $(-1)^{f(G)}$ with the graph.

(vi) Determine the weight $w(G)$, (24a), of every graph G by Wick's theorem or by combinatorics.

This yields

$$c_m = \sum_G (-1)^{f(G)} w(G) v(G) \ . \tag{28}$$

In Fig. 2 we show all graphs contributing to c_m up to order $m=4$; they are denoted by C_{mi}. In Table I we list the corresponding values of $f(G)$ and $w(G)$ for $m \leq 3$. The entry for $v(G)$ in this table is the result for the one-dimensional chain $(d=1)$ with $n_\uparrow = n_\downarrow = n/2$, $n \leq 1$. In this case the scaling in **k** space discussed below (17) implies that the first Brillouin zone (BZ) is in the interval $[-\frac{1}{2}, \frac{1}{2}]$ and the Fermi momentum is $k_F = n/4$. The momentum distribution of the noninteracting system is given by $n_{k\sigma}^0 = \theta(k_F - |k|)$, i.e., a step function. We see that in $d=1$ and for $n_\uparrow = n_\downarrow = n/2 \leq \frac{1}{2}$, $v(C_{mi}) \propto (n/2)^{m+1}$ as is easily understood from the diagrammatic rules defined above (see also Appendix A). Hence, in this case,

$$c_m \propto n^{m+1} \ . \tag{29}$$

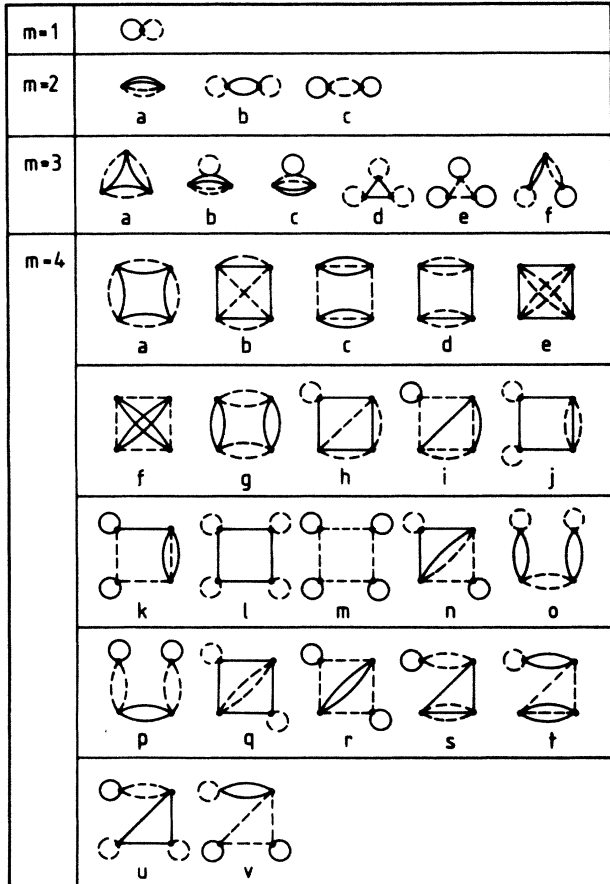We note that this result is different from those obtained by earlier investigations,[29,30,33] although the corresponding diagrams are identical. The difference originates solely from our definition of "δ-less contractions" used to evaluate the graphs. This has the great advantage that in $d=1$ we have a simple relation (29), i.e., for a given m different orders in n are not mixed as in the previous work.[29,30,33] The different definition of the contractions is accompanied by a different expansion for $\langle \hat{D} \rangle$, (27). While we use a power series in $g^2 - 1$, in Refs. 29, 30, and 33 the quantity $y = \ln(1/g)$ is used, whereby different orders of n are mixed.

TABLE I. The graphs C (see Fig. 2) contributing to the coefficients c_m in the Hubbard interaction, (28), are listed for $m \leq 3$. The value $v(G)$ is that for the one-dimensional case.

G	$f(G)$	$w(G)$	$v(G)$
C_1	2	1	$\left[\dfrac{n}{2}\right]^2$
C_{2a}	2	1	$\dfrac{2}{3}\left[\dfrac{n}{2}\right]^3$
C_{2b}	3	1	$\left[\dfrac{n}{2}\right]^3$
C_{2c}	3	1	$\left[\dfrac{n}{2}\right]^3$
C_{3a}	2	2	$\dfrac{1}{2}\left[\dfrac{n}{2}\right]^4$
C_{3b}	3	3	$\dfrac{2}{3}\left[\dfrac{n}{2}\right]^4$
C_{3c}	3	3	$\dfrac{2}{3}\left[\dfrac{n}{2}\right]^4$
C_{3d}	4	1	$\left[\dfrac{n}{2}\right]^4$
C_{3e}	4	1	$\left[\dfrac{n}{2}\right]^4$
C_{3f}	4	3	$\left[\dfrac{n}{2}\right]^4$

B. Particle-hole symmetry

To determine the prefactor in (29) we employ particle-hole (ph) symmetry. To this end we introduce a canonical transformation (see Ref. 13) valid for AB lattices:

$$\hat{c}'_{f\sigma} = (-1)^f \hat{c}^\dagger_{f\sigma}, \quad \hat{c}'^\dagger_{f\sigma} = (-1)^f \hat{c}_{f\sigma} \tag{30}$$

with $(-1)^f = +1$ or -1 for f on A or B, respectively. The Hamiltonian (3) and hence the exact ground-state energy $E_{ex}(n_\uparrow, n_\downarrow)$ is thereby left unchanged up to a constant

FIG. 2. The diagrams C_{mi} contributing to c_m in (26) up to order $m=4$.

$$E_{\rm ex}(1-n_\uparrow,1-n_\downarrow)=E_{\rm ex}(n_\uparrow,n_\downarrow)+(1-n_\uparrow-n_\downarrow)LU \ .$$

$$(31)$$

The exact expectation value of double occupancy $d_{\rm ex}(n_\uparrow,n_\downarrow)=\langle \hat{D}\rangle_{\rm ex}/L$ is changed according to

$$d_{\rm ex}(1-n_\uparrow,1-n_\downarrow)=d_{\rm ex}(n_\uparrow,n_\downarrow)+1-n_\uparrow-n_\downarrow \ .$$

$$(32)$$

In the case of an approximate evaluation of the ground-state energy by a variational ansatz (7), the relation (32) nevertheless still holds, as can be verified by explicit calculation. We insert (27) into (32) and, by equating the (g^2-1) coefficients, find

$$c_m(n_\uparrow,n_\downarrow)+c_{m+1}(n_\uparrow,n_\downarrow)=c_m(1-n_\uparrow,1-n_\downarrow)$$

$$+c_{m+1}(1-n_\uparrow,1-n_\downarrow)$$

$$(33)$$

for all $m \geq 1$ and arbitrary $n_\uparrow,n_\downarrow$.

C. Exact calculation of c_m in $d=1$ dimension

In the one-dimensional case we found $c_m=a_m n^{m+1}$ for $n_\uparrow=n_\downarrow=n/2$, $n \leq 1$. To make use of (33) for determining the prefactor a_m we need to know the analytic behavior of c_m for $1 \leq n \leq 2$. For this we have to reinvestigate the k summations entering in the evaluation of the corresponding graphs. For $n>1$, k sums go beyond the first BZ which makes Umklapp-processes necessary. This is done in Appendix A, where it is shown that the first two derivatives of c_m with respect to the density n are continuous. Setting $n_\uparrow=n_\downarrow=n/2$ in (33) we may then differentiate by n. At $n=1$ we find

$$c_m'(\tfrac{1}{2},\tfrac{1}{2})+c_{m+1}'(\tfrac{1}{2},\tfrac{1}{2})=0 \ ,$$

$$(34)$$

which leads to the recursion formula

$$a_{m+1}=-[(m+1)/(m+2)]a_m \ .$$

With $a_1=\tfrac{1}{4}$ this leads to the general expression

$$c_m=(-1)^{m+1}\frac{n^{m+1}}{2(m+1)}, \quad n \leq 1$$

$$(35)$$

for all $m \geq 1$.

The power series in (27) for $\langle \hat{D}\rangle$, i.e., the expectation value of the Hubbard-interaction term (2), Eq. (3), can thus be summed explicitly and yields ($n \leq 1$)

$$\langle \hat{H}_I\rangle=\frac{ULn^2}{2}\left[\frac{g}{1-G^2}\right]^2\left[\ln\frac{1}{G^2}+G^2-1\right], \quad (36)$$

where $G^2=1-(1-g^2)n$; for $n \geq 1$, (33) applies. The double occupancy $d(g)=\langle \hat{H}_I\rangle/UL$ as a function of g is shown in Fig. 3 for three densities and is compared with the corresponding results obtained within the Gutzwiller
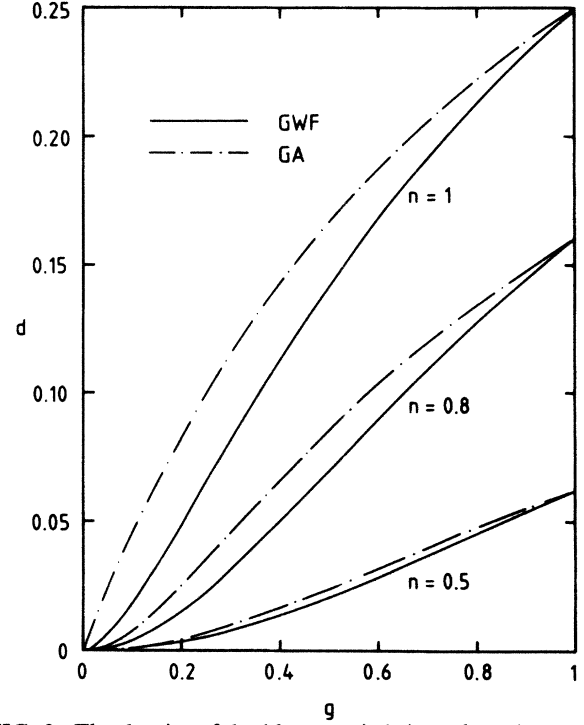

FIG. 3. The density of doubly occupied sites, d, vs the correlation parameter g for three densities in one dimension. The approximation-free results (GWF) are compared with those of the Gutzwiller approximation (GA).

approximation.[18] For strong interactions ($g \to 0$) we find $\langle \hat{H}_I\rangle$ to be *nonanalytic* in g and n:

$$\langle \hat{H}_I\rangle \simeq UL\begin{cases} g^2\ln(1/g), & n=1 \ , \\ \tfrac{1}{2}g^2\ln\dfrac{1}{1-n}, & n<1 \ . \end{cases}$$

$$(37)$$

The approximation-free result in (37) valid for small g in the thermodynamic limit is in contrast to the numerical findings[28,33] obtained earlier on the basis of *finite* systems, where a g^2 dependence for $n=1$ was concluded.

IV. EVALUATION OF THE MOMENTUM DISTRIBUTION

As in the case of $\langle \hat{H}_I\rangle$ the expectation value of $\hat{n}_{k\sigma}$ is expanded in a power series in g^2-1, where lattice sums are written so as to involve only different sites, then Wick's theorem is applied leading to δ-less contractions and, finally, the linked cluster theorem is used.

Instead of $\hat{n}_{k\sigma}$ we investigate the one-particle density matrix

$$\mathcal{P}_{gh,\sigma}=\langle \hat{c}^{\dagger}_{g\sigma}\hat{c}_{h\sigma}\rangle$$

$$(38)$$

in real space. In the noninteracting case $\mathcal{P}^{(0)}_{gh\sigma}=P_{gh\sigma}$, (16a). We first consider the numerator

$$\tilde{\mathcal{P}}_{gh,\sigma}=\langle \psi_G | \hat{c}^{\dagger}_{g\sigma}\hat{c}_{h\sigma} | \psi_G\rangle \ .$$

$$(39)$$

Assuming $g \neq h$, $| \psi_G\rangle$ as given in (7) is inserted. Then (12) is used to convert the products into sums, where g,h are excluded from the f_i sums:

$$\tilde{P}_{gh,\sigma} = \left\langle \hat{c}^\dagger_{g\sigma} \hat{c}_{h\sigma} \hat{Q}_{g-\sigma} \hat{Q}_{h-\sigma} \left[1 + \sum_{m=1}^{\infty} \frac{(g^2-1)^m}{m!} \sum'_{f_1,\dots,f_m \neq g,h} \hat{D}_{f_1} \cdots \hat{D}_{f_m} \right] \right\rangle_0, \tag{40}$$

where

$$\hat{Q}_{g\sigma} = 1 - (1-g)\hat{n}_{g\sigma} . \tag{41}$$

Since all sites in the lattice sum in (40) are different we may apply Wick's theorem to replace $\langle \cdots \rangle_0$ in (40) by $\{ \cdots \}_0$, the sum of all contracted terms with δ-less contractions. *Afterwards* we can again drop all restrictions on the sums in (40) since $\{ \cdots \}_0 = 0$ for $f_i = f_j$ or $f_i = g,h$ as is seen from writing one of the four contributions explicitly as

$$\{c^\dagger_{1\uparrow} n_{1\downarrow} c_{2\uparrow} n_{2\downarrow} D_3 \cdots D_m\}_0 = \begin{vmatrix} P_{12,\uparrow} & P_{13,\uparrow} & \cdots & P_{1m,\uparrow} \\ P_{32,\uparrow} & P_{33,\uparrow} & \cdots & P_{3m,\uparrow} \\ \vdots & & & \vdots \\ P_{m2,\uparrow} & P_{m3,\uparrow} & \cdots & P_{mm,\uparrow} \end{vmatrix} \begin{vmatrix} P_{11,\downarrow} & \cdots & P_{1m,\downarrow} \\ \vdots & & \vdots \\ P_{m1,\downarrow} & \cdots & P_{mm,\downarrow} \end{vmatrix}, \tag{42}$$

where $f_i \equiv i$, etc.

In the case $g=h$, the factor in front of the parentheses in (40) reduces to $\hat{n}_{g\sigma} + (g^2-1)\hat{D}_g$. Again Wick's theorem is applied, transforming $\langle \cdots \rangle_0$ into $\{ \cdots \}_0$ where the quantities $\hat{n}_{g\sigma}$, $\hat{Q}_{g\sigma}$, etc., are no longer operators. We note that

$$\{[n_{g\sigma} + (g^2-1)D_g] \cdots \}_0 = \{[n_{g\sigma} Q_{g\sigma} Q_{g-\sigma} + (1-g)^2 n_{g-\sigma}] \cdots \}_0 , \tag{43}$$

where the last term has to be added since $Q_{g-\sigma} Q_{g-\sigma}$ leads to a term $n_{g-\sigma} n_{g-\sigma}$ which vanishes inside the contraction because of the anticommuting character of the n_g [see Eq. (21)]. Hence for $g=h$, (43) replaces $\langle \cdots \rangle_0$ in (40). For general g,h, (39) then takes the form

$$\tilde{P}_{gh,\sigma} = \left[c^\dagger_{g\sigma} c_{h\sigma} [1 - (1-g)(n_{g-\sigma} + n_{h-\sigma}) + (1-g)^2 n_{g-\sigma}(n_{h-\sigma} + \delta_{gh})] + \left[1 + \sum_{m=1}^{\infty} \frac{(g^2-1)^m}{m!} \sum_{f_1,\dots,f_m} D_{f_1} \cdots D_{f_m} \right] \right]_0 . \tag{44}$$

The contributions to (44) in different orders of m may again be represented diagrammatically. Straightforward application of the linked cluster theorem[37,38] shows that the norm $\langle \psi_G | \psi_G \rangle$ again cancels the disconnected diagrams. Hence the momentum distribution in k space, i.e., the Fourier transform of $\tilde{P}_{gh,\sigma}$, is given by

$$n_{k\sigma} = \frac{1}{L} \sum_{g,h} e^{2\pi i k(g-h)} [\text{Eq. (44)}]_0^c , \tag{45}$$

where only the connected diagrams ($\{ \cdots \}_0 \to \{ \cdots \}_0^c$) contribute. To simplify (45) we introduce the functions $f_{m\sigma}(k)$, $m \geq 2$,

$$f_{m\sigma}(k) \equiv \frac{1}{L} \frac{1}{(m-2)!} \sum_{f_1,\dots,f_m} e^{2\pi i k(f_1-f_2)} \{(c^\dagger_{f_1\sigma} n_{f_1-\sigma} c_{f_2\sigma} n_{f_2-\sigma} + \delta_{f_1 f_2} D_{f_2}) D_{f_3} \cdots D_{f_m}\}_0^c . \tag{46}$$

Making use of the identities

$$\sum_{g,f_1,\dots,f_m} \{n_{g\sigma} D_{f_1} \cdots D_{f_m}\}_0^c = -m \sum_{f_1,\dots,f_m} \{D_{f_1} \cdots D_{f_m}\}_0^c , \tag{47a}$$

$$\sum_{g,h,f_1,\dots,f_m} e^{2\pi i k(g-h)} \{c^\dagger_{g\sigma} c_{h\sigma} D_{f_1} \cdots D_{f_m}\}_0^c = L m! n^0_{k\sigma} f_{m\sigma}(k), \quad m \geq 2 , \tag{47b}$$

$$\sum_{g,h,f_1,\dots,f_m} e^{2\pi i k(g-h)} \{c^\dagger_{g\sigma} c_{h\sigma} n_{g-\sigma} D_{f_1} \cdots D_{f_m}\}_0^c = -L m! n^0_{k\sigma} f_{m+1,\sigma}(k), \quad m \geq 1 , \tag{47c}$$

all terms in $n_{k\sigma}$, (45), may be expressed in terms of $f_{m\sigma}(k)$, leading to

$$n_{k\sigma} = n^0_{k\sigma} - (1-g)^2 n_{-\sigma} n^0_{k\sigma} + \frac{1}{(1+g)^2} \sum_{m=2}^{\infty} (g^2-1)^m [1 - (1-g^2) n^0_{k\sigma}] f_{m\sigma}(k) . \tag{48}$$

So $n_{k\sigma}$ has been expressed as a power series in g^2-1 which converges rapidly for $g \lesssim 1$. For weak correlations or small densities the series gives accurate results even if terminated at some finite order.

A. Diagrammatic representation

The evaluation of $n_{k\sigma}$ involves the functions $f_{m\sigma}(k)$, which may be represented diagrammatically. To this end we note that the second term in (46) contributing to $f_{m\sigma}(k)$ is simply given by the coefficient c_{m-1} defined in (26); hence we define

$$h_{m\sigma}(k) \equiv \frac{1}{L} \frac{1}{(m-2)!} \sum_{f_1, \ldots, f_m} e^{2\pi i k (f_1 - f_2)} \{ c_{f_1\sigma}^\dagger n_{f_1-\sigma} c_{f_2\sigma} n_{f_2-\sigma} D_{f_3} \cdots D_{f_m} \}_0^c ,$$ (49)

such that

$$f_{m\sigma}(k) = h_{m\sigma}(k) + c_{m-1} .$$ (50)

This way of writing is convenient since the resulting graphs are identical to the ones for the two-point functions of a ϕ^4 theory[37] (without external lines). This we illustrate in the case $m=2$, $\sigma = \uparrow$:

$$f_{2\uparrow}(k) = \frac{1}{L} \sum_{f_1, f_2} e^{2\pi i k (f_1 - f_2)} \{ c_{f_1\uparrow}^\dagger n_{f_1\downarrow} c_{f_2\uparrow} n_{f_2\downarrow} \}_0^c + C_1$$ (51a)

$$= \frac{1}{L} \sum_{f_1, f_2} e^{2\pi i k (f_1 - f_2)} [-P_{12\uparrow} P_{12\downarrow} P_{21\downarrow} + P_{12\uparrow} P_{11\downarrow} P_{22\downarrow}] + c_1$$ (51b)

$$= -F_{2a\uparrow}(k) + F_{2b\uparrow}(k) + F_{2c\uparrow} ,$$ (51c)

where the graphs $F_{2i\uparrow}(k)$, $i = a,b,c$, are shown in Fig. 4. Full (broken) lines again correspond to $\uparrow$ ($\downarrow$) lines. The arrows indicate where the external momentum k enters; clearly, $F_{2c\uparrow} = C_1$ is k independent. For general $m \geq 2$ the diagrams to $f_{m\sigma}(k)$ are constructed by the following Feynman rules as in the case of the c_m: (i) draw all topologically different connected graphs G with m vertices, $m-1$ σ lines and m ($-\sigma$) lines; (ii) every line is given a momentum q with momentum conservation at the vertices and is associated with a factor $n_{q\sigma}^0$; (iii) integrate over internal momenta; this yields the value $v(G)$ for each graph G; (iv) determine the weight $w(G)$,

$$w(G) = \frac{1}{(m-2)!} c_G ;$$

(v) determine the prefactor $(-1)^{f(G)}$, where $f(G)$ is the number of loops.

This yields

$$f_{m\sigma}(k) = \sum_G (-1)^{f(G)} w(G) v(G) .$$ (52)

For $m = 1$ we define

$$f_{1\sigma}(k) = -\frac{1}{L} \sum_f \{ n_{f-\sigma} \}_0^c .$$ (53)

The graphs $F_{mi\uparrow}(k)$ contributing to $f_{m\uparrow}(k)$ for $m \leq 3$ are
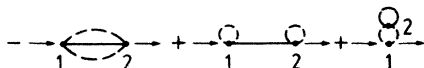
FIG. 4. Diagrammatic representation of $f_{2\uparrow}(k)$.

shown in Fig. 5; for $\sigma = \downarrow$ full and broken lines have to be interchanged.

Next we introduce the concept of one-particle irreducible (or proper) graphs[37] which do not contain single lines; otherwise they are called reducible (or improper). In Fig. 5, $F_{2b\uparrow}(k)$ and $F_{3def\uparrow}(k)$ are reducible, the others are irreducible. Reducible diagrams are given by their irreducible components times a factor $n_{k\sigma}^0$. For $|k| > k_F$ reducible graphs therefore do not contribute.

Introducing $f_{m\sigma}^{irr}(k)$, the sum of all irreducible graphs to $f_{m\sigma}(k)$, we have

FIG. 5. The diagrams $F_{mi\uparrow}(k)$ contributing to the coefficients $f_{m\uparrow}(k)$ in (46) up to order $m=3$.

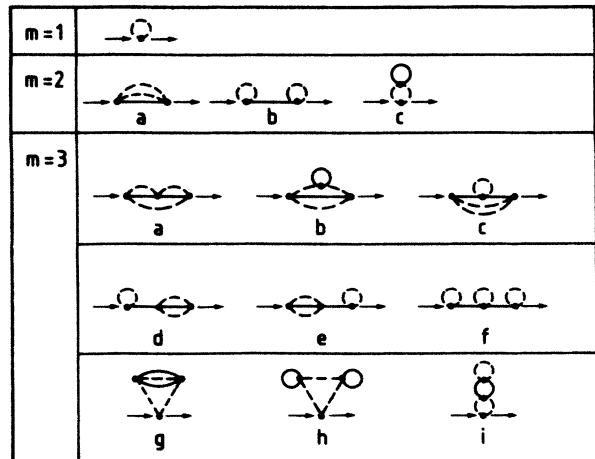$$f_{m\sigma}(\mathbf{k}) = \begin{cases} \displaystyle\sum_{\substack{v_1 \cdots v_m \geq 0 \\ \sum_n n v_n = m}} \frac{\left[\sum_n v_n\right]!}{v_1! \cdots v_m!} \prod_{l=1}^{m} [f_{l\sigma}^{\mathrm{irr}}(\mathbf{k})]^{v_l}, & |\mathbf{k}| < k_F , \\[20pt] f_{m\sigma}^{\mathrm{irr}}(\mathbf{k}), & |\mathbf{k}| > k_F . \end{cases} \tag{54}$$

Finally, we calculate $f_{2\sigma}(\mathbf{k})$ in the one-dimensional case for $n_\uparrow = n_\downarrow = n/2$, $n \leq 1$. There are three graphs as shown in Fig. 5. The graphs $F_{2b\sigma}$ and $F_{2c\sigma}$ are trivial, i.e., $v(F_{2b\sigma}) = (n/2)^2 n_{\mathbf{k}\sigma}^0$, $v(F_{2c\sigma}) = (n/2)^2$, for all k; in both cases $f(G) = 2$ and $w(G) = 1$. To evaluate $F_{2a\sigma}$, four different regimes in k and n have to be distinguished as shown in Fig. 6:

(I) $|k| < k_F = \dfrac{n}{4}$, (II) $k_F < |k| < 1 - 3k_F$, (III) $|k| > 3k_F$, (IV) $|k| > 1 - 3k_F$;

this applies to other graphs $F_{mi\sigma}$ as well. Since $k \in [-\frac{1}{2}, \frac{1}{2}]$, $k > 1 - 3k_F$ is only possible for $n > \frac{2}{3}$ and $k > 3k_F$ only for $n < \frac{2}{3}$. With $f(G) = 1$ and $w(G) = 1$ for this graph we find

$$f_{2\sigma}(k) = -v(F_{2a\sigma}) + v(F_{2b\sigma}) + v(F_{2c\sigma}) \tag{55}$$

$$= \begin{cases} \dfrac{5}{16}n^2 + k^2 & \text{(I)} , \\[14pt] \left[\dfrac{n}{2}\right]^2 - \dfrac{1}{2}\left[|k| - \dfrac{3n}{4}\right]^2 & \text{(II)} , \\[14pt] \left[\dfrac{n}{2}\right]^2 & \text{(III)} , \\[14pt] \left[\dfrac{n}{2}\right]^2 - \dfrac{1}{2}\left[|k| - \dfrac{3n}{4}\right]^2 - \dfrac{1}{2}\left[1 - |k| - \dfrac{3n}{4}\right]^2 & \text{(IV)} . \end{cases} \tag{56}$$

We see that for general n the coefficients $f_{m\sigma}(\mathbf{k})$ become rather complicated in detail even in $d=1$ and for small orders m. This is due to the role of Umklapp processes (UP) which enter for $|k| > 1 - 3k_F$. On the other hand, their structure in $d=1$ is still simple, being given by a polynomial in $|k|$ and n. The existence of UP makes the behavior of $f_{m\sigma}(k)$ in k and n nontrivial; it therefore needs to be investigated in detail to make an exact evaluation in $d=1$ possible; this is done in Appendix A.

The $\mathbf{k}$ integral over $f_{m\sigma}(\mathbf{k})$ is easily seen to be related with the coefficients c_m, i.e., the double occupancy itself. This is clear from their diagrammatic structure: integration closes the external vertices of the $F_{mi\sigma}(\mathbf{k})$ in Fig. 5, thus leading to the closed graphs C_{r^i} in Fig. 2. Indeed, integrating (48) over $k \leq k_{F\sigma}$ leads to

$$n_\sigma^< \equiv \int_{|\mathbf{k}| \leq k_F} d\mathbf{k} n_{\mathbf{k}\sigma}$$

$$= n_\sigma - (1-g)^2 n_\sigma n_{-\sigma}$$

$$- \frac{g^2}{(1+g)^2} \sum_{m=2}^{\infty} (g^2-1)^m c_m \tag{57}$$

$$= n_\sigma - \frac{1-g}{1+g}(d_0 - d) , \tag{58}$$

where $d_0 = n_\sigma n_{-\sigma}$. Hence the density of particles inside the Fermi surface is related to the density of doubly occupied sites d, i.e., to the expectation value $\langle \hat{H}_I \rangle$ itself.

The density of particles outside the Fermi surface is $n_\sigma^> = n_\sigma - n_\sigma^<$. The above connection was already obtained by Hashimoto.[28]

B. Particle-hole symmetry

As in the case of the coefficients c_m ph symmetry may be used to obtain exact relations among the coefficients $f_{m\sigma}(n,k)$, where we now write the n dependence explicitly $(n_\uparrow = n_\downarrow = n/2)$. Using the canonical transformations
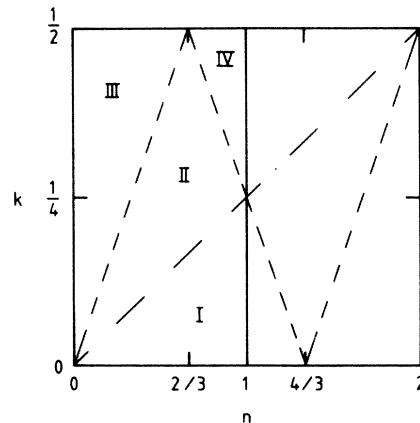

FIG. 6. Different regions in momentum k vs n due to Umklapp processes.

for AB lattices in (30), the momentum distribution is found to obey

$$n_{k_0-k,\sigma}(2-n)+n_{k\sigma}(n)=1 \; , \tag{59}$$

where k_0 is a vector in k space with $k_0(f_i-f_j)=$ half-integer if f_i,f_j are on different sublattices. Using (48) this translates into

$$f_{m\sigma}(n,k)+f_{m+1,\sigma}(n,k)=-f_{m+1,\sigma}(2-n,k_0-k) \; , \tag{60}$$

valid for all n and k with $n^0_{k\sigma}=1$.

C. Exact calculation of $f_{m\sigma}(k)$ in $d=1$ dimension

Equation (60) will now be used to derive an exact recursion relation for the $f_{m\sigma}(k)$ in $d=1$ and

$n_\uparrow=n_\downarrow=n/2\leq1$. In this case, $k_0=\frac{1}{2}$ and the spin index may be omitted. Let $\delta=1-n$ be small and $k<k_F$. According to the results in Appendix A, (n,k) lies in region (I) in Fig. 6 where $f_m(n,k)$ is given by a polynomial R_m of order $\leq m$:

$$f_m(n,k)=n^mP_m(K)+c_{m-1} \tag{61}$$

$$\equiv n^mR_m(K) \; , \tag{62}$$

where $K=k/n$. Since the first two derivatives of f_m by n at $n=1$ are continuous we may express $f_{m+1}(\bar{n},\frac{1}{2}-k)$, $\bar{n}=2-n$ by the expression it takes in region (IV), plus corrections of order $O(\delta^3)$, i.e.,

$$f_{m+1}(\bar{n},\tfrac{1}{2}-k)=\frac{(-1)^{m+1}}{2(m+1)}\bar{n}^{m+1}+\left[Q_{m+1}\left[\frac{\frac{1}{2}-Kn}{n}\right]+Q_{m+1}\left[\frac{\frac{1}{2}+Kn}{n}\right]\right]\bar{n}^{m+1}+O(\delta^3) \; . \tag{63}$$

Inserting (62) and (63) into (60) yields an equation which, according to Appendix A, may be differentiated *twice* with respect to n. Hence (60) and its two derivatives taken at $n=1$ yield the three equations

$$R_m(k)+R_{m+1}(k)=-\frac{(-1)^{m+1}}{2(m+1)}-X^{(0)}_{m+1}(k) \; , \tag{64a}$$

$$mR_m(k)+(m+1)R_{m+1}(k)=\frac{(-1)^{m+1}}{2}+(m+1)X^{(0)}_{m+1}(k)-X^{(1)}_{m+1}(k) \; , \tag{64b}$$

$$m(m-1)R_m(k)+(m+1)mR_{m+1}(k)=-\frac{m}{2}(-1)^{m+1}-m(m+1)X^{(0)}_{m+1}(k)+2mX^{(1)}_{m+1}(k)-X^{(2)}_{m+1}(k) \; , \tag{64c}$$

where we define

$$X^{(l)}_m(k)=(\tfrac{1}{2}-2k)^lQ^{(l)}_m(\tfrac{1}{2}-2k)+(\tfrac{1}{2}+2k)^lQ^{(l)}_m(\tfrac{1}{2}+2k) \tag{65}$$

and $Q^{(l)}_m(y_0)$ is the lth-derivative of $Q_m(y)$ by y at $y=y_0$. The unknown polynomials R_m and Q_m are functions of k and may be written as Taylor series around $k=\frac{1}{4}$ and $k=\frac{3}{4}$, respectively:

$$R_m(k)=\sum_{j=0}^{m}\frac{R^{(j)}_{m\,\frac{1}{4}}}{j!}(k-\tfrac{1}{4})^j \; , \tag{66a}$$

$$Q_m(k)=\sum_{j=0}^{m}\frac{Q^{(j)}_{m\,\frac{3}{4}}}{j!}(k-\tfrac{3}{4})^j \; . \tag{66b}$$

As is shown in Appendix B the coefficients $R^{(j)}_m(\frac{1}{4})$ and $Q^{(j)}_m(\frac{3}{4})$ may be calculated exactly (although not in closed form) via a recursion relation. This allows one to determine the functions $f_m(n,k)$ for arbitrary m. For the explicit evaluation of the momentum distribution n_k (or of the kinetic energy) it is usually sufficient to know f_m for $m\lesssim100$ to determine all relevant quantities to an excellent accuracy. The computation is easily done on a home computer. In Fig. 7 the calculated k dependence of $f_m(k)$ is shown for several values of m.

D. The discontinuity of n_k at the Fermi surface in $d=1$

The values of the momentum distribution at $k=k_F^\pm\equiv k_F\pm0$, as well as the slope of n_k at k_F, may now be expressed by elementary expressions. Using (61) together with (66) and the results in Appendix B one finds $(n\leq1)$

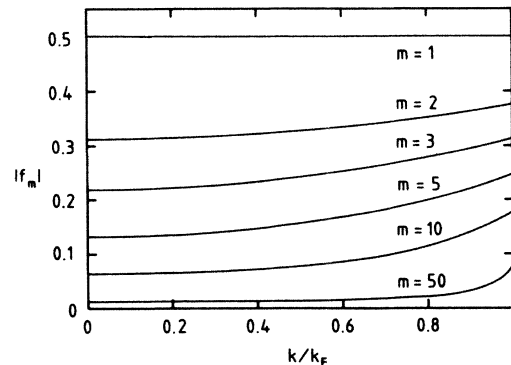$$f_m(k_F^-)=(-1)^m\frac{(2m-1)!!}{(2m)!!}n^m \tag{67}$$

FIG. 7. Coefficients $f_m(k)$ vs k/k_F for different orders of m.

and $f_m(k_F^+)=f_m(k_F^-)/(2m-1)$. The sum in (48) is easily performed and yields

$$n_{k_F^-}=1-(1-g)^2\frac{n}{2}+\frac{g^2}{(1+g)^2}\left[\frac{1}{G}-\frac{n}{2}(1-g^2)-1\right],$$

(68a)

$$n_{k_F^+}=\frac{1}{2}\left[\frac{1-G}{1+g}\right]^2,$$

(68b)

where $G^2=1-(1-n)g^2$. The discontinuity $q=n_{k_F^-}$ $-n_{k_F^+}$ of n_k at the Fermi surface is thus obtained as

$$q=\frac{1}{G}\left[\frac{G+g}{1+g}\right]^2$$

(69)

for $n\leq 1$, $0\leq g\leq 1$. For infinitely strong correlation $(g=0)$ one finds

$$\lim_{g\to 0}\begin{cases}n_{k_F^-}=1-n/2\,,\\ n_{k_F^+}=1-n/2-\sqrt{1-n}\,,\\ q=\sqrt{1-n}\,.\end{cases}$$

(70)

For a half-filled band, (69) reads

$$q=\frac{4g}{(1+g)^2}.$$

(71)

We have found this result to hold in *arbitrary* dimensions for lattices with hypercubic Fermi surfaces. It is surprising to see that this approximation-free result for the Gutzwiller wave function was already obtained within the Gutzwiller *approximation*.[18] In Fig. 8 the discontinuity q is shown as a function of $1-g$ for several densities n in comparison with the results obtained by the GA.

The slope of n_k at $k_F\pm 0$ is found by taking the derivative of $f_m(k)$, (62). Since

$$f_m'(k_F^-)=(-1)^m(2m-3)!!n^{m-1}/2(2m-4)!!$$

and

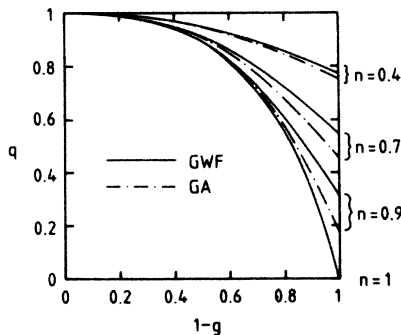$$f_m'(k_F^+)=(-1)^m(2m-5)!!n^{m-1}/2(2m-4)!!\,,$$

one obtains

$$n_{k_F^-}'=g^2\frac{n(1-g)^2}{2G^3}\,,$$

(72a)

$$n_{k_F^+}'=\frac{n(1-g)^2}{2G}\,.$$

(72b)

At $n=1$ the slopes are equal: $n_{k_F^\pm}'=(1-g)^2/2g$; for $g\to 0$ it diverges. We observe that for any n and $g\neq 0,1$ the slope is always *positive*. This is not what is expected to occur in an interacting Fermi system (see Ref. 39 for calculations of n_k in three-dimensional systems, where a monotonic decrease with increasing $|k|$ is found).

The full k dependence of n_k in $d=1$ is obtained from (48) by employing the recursion relations derived above to determine $f_m(k)$ in (62). In Fig. 9, n_k is shown for densities $n=0.8$ and $n=1$, respectively. For $n=0.8$ [Fig. 9(a)] and $|k|\leq k_F$ we see that n_k is essentially constant for arbitrary correlation strength. This behavior also holds for $|k|\geq k_F$ and weak correlations $(g\lesssim 1)$, while for strong correlations an increase for increasing $|k|$ is found. In the half-filled band case $(n=1)$, showing Fig. 9(b), n_k is hardly k dependent at all, except close to k_F as discussed above. At $|k|=k_F$, n_k is nonanalytic in the limits $n\to 1$ and $g\to 0$, i.e., these limits are not interchangeable. For $n=1$ the limits $k\to k_F$ and $g=0$ also may not be interchanged. This is seen from (48): for $k<k_F$ and small g one finds

$$n_k=\tfrac{1}{2}+g+g^2\sum_{m=2}^{\infty}(g^2-1)^mf_m(k)+O(g^2)\,.$$

(73)

FIG. 8. The discontinuity q at the Fermi surface vs $1-g$ for different densities n.
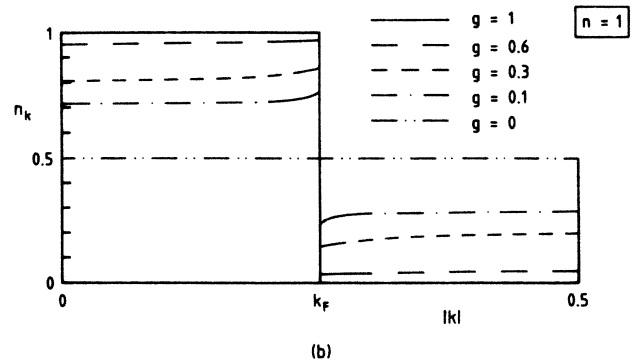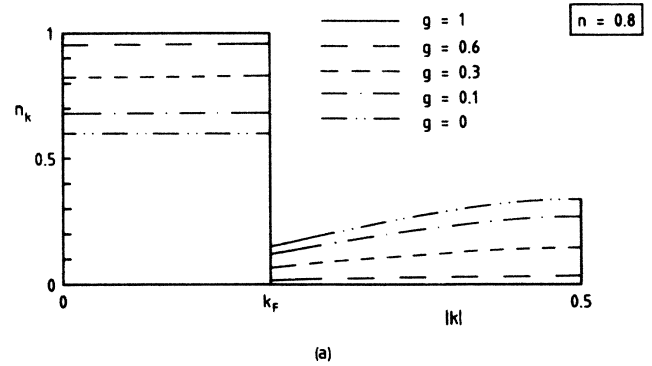

FIG. 9. The momentum distribution n_k vs k for different correlation parameters g: (a) $n=0.8$, (b) $n=1$.

A straightforward estimation of the sum in (73) shows that for $|k|$ strictly smaller than k_F the sum goes as $\ln(1/g)$. Hence $n_k = \frac{1}{2} + g$ in this limit; similarly $n_k = \frac{1}{2} - g$ for $|k| > k_F$. This was already noted in Ref. 10. On the other hand, for $k = k_F \pm 0$, (68) yields $n_{k_F^\pm} = \frac{1}{2} \mp 2g$.

V. THE GROUND-STATE ENERGY OF THE HUBBARD MODEL IN $d = 1$

The ground-state energy E of the Hubbard model with general kinetic energy $\varepsilon(k)$ in terms of the GWF is obtained from the expectation value $E(g)$ of $\hat{H}$ in (3),

$$E(g) = L \sum_\sigma \int dk \, \varepsilon(k) n_{k\sigma}(g) + U L d(g) , \tag{74}$$

by minimizing with respect to the variational correlation parameter g. Since $n_k(g)$ is not given by an elementary function this has to be done numerically in general. However, for a strong correlation it may also be performed analytically. For $g \to 0$ the kinetic energy in (74) is found as

$$E_{\rm kin}(g) = \tilde{E}_\infty + 2g(\bar{\varepsilon}_0 - \tilde{E}_\infty) , \tag{75}$$

where $\tilde{E}_\infty \equiv E(U = \infty)/L = E(g = 0)/L$, and $\bar{\varepsilon}_0 = 2 \int dk \, \varepsilon(k) n_k^0$ is the average energy of the uncorrelated fermions. For general n, $\tilde{E}_\infty$ has to be determined numerically using (48) and (1b) (see below). For small n, $\tilde{E}_\infty = (-2n + n^2)t$, while for $n = 1$, $\tilde{E}_\infty = 0$ [see (73) and below] and hence $E_{\rm kin}(g) = 2g\bar{\varepsilon}_0$ as was already observed earlier.[28,33] The potential energy in this limit is given by (37). Hence, for $g \to 0$,

$$\frac{E(g)}{L} = \begin{cases} \tilde{E}_\infty + 2g(\bar{\varepsilon}_0 - \tilde{E}_\infty) + U\dfrac{g^2}{2}\ln\dfrac{1}{1-n}, & n < 1 , \tag{76a} \\[2ex] 2g\bar{\varepsilon}_0 + Ug^2\ln\dfrac{1}{g}, & n = 1 . \tag{76b} \end{cases}$$

The minimum is found at

$$g = 2(1 - \tilde{E}_\infty/\bar{\varepsilon}_0)/\bar{U}\ln[1/(1-n)]$$

and $g\ln(1/g) = 1/\bar{U}$ for $n < 1$ and $n = 1$, respectively; here $\bar{U} \equiv U/|\bar{\varepsilon}_0|$. For $U \to \infty$ the ground-state energy is thus given by

$$\frac{E}{L} = \begin{cases} -|\tilde{E}_\infty| - \dfrac{\alpha t^2}{U}, & n < 1 , \tag{77a} \\[2ex] -\beta\dfrac{t^2}{U}\dfrac{1}{\ln\bar{U}}, & n = 1 , \tag{77b} \end{cases}$$

where $\alpha = 2[(\bar{\varepsilon}_0 - \tilde{E}_\infty)/t]^2/\ln[1/(1-n)]$ and $\beta = (\bar{\varepsilon}_0/t)^2$. For less than half-filling the correction to $\tilde{E}_\infty$ goes as $-t^2/U$. However, the prefactor α has a logarithmic n dependence and goes to zero for $n \to 1$. This indicates a crossover from an n dependent to a U-dependent logarithmic correction to the $(-t^2/U)$ behavior in the limit of an exactly half-filled band. The result (77a) should be contrasted with that obtained by the GA (Ref. 18) for

$n < 1$ where no Brinkman-Rice transition occurs:

$$\tilde{E}_\infty^{\rm (GA)} = -\frac{1-n}{1-n/2}|\bar{\varepsilon}_0| ,$$

$$\alpha_{\rm GA} = \frac{1-n}{(1-n/2)^2}\left[\frac{\bar{\varepsilon}_0}{t}\right]^2 , \tag{78}$$

which is valid for

$$\bar{U} \gg 1/\sqrt{1-n}\left[1 - \frac{n}{2}\right] .$$

For $n < 1$ both the result of the GA and the approximation-free one derived here have a $(-t^2/U)$ dependence at large U. Furthermore, for small densities n the two results are seen to coincide; i.e., in this limit the results of the GA approach the correct results for the GWF, (77a).

For next-neighbor hopping, where

$$\bar{\varepsilon}_0 = -(4t/\pi)\sin(\pi n/2) ,$$

both (77) and (78) may be compared with the exact result[13,40]

$$\tilde{E}_\infty^{\rm (ex)} = -\frac{2t}{\pi}\sin\pi n ,$$

$$\alpha_{\rm ex} = (4\ln 2)n^2\left[1 - \frac{\sin(2\pi n)}{2\pi n}\right] . \tag{79}$$

For small densities, $\alpha_{\rm ex} \propto n^4$ while $\alpha_{\rm GA} \propto n^2$; this shows that for large U the GWF never approaches the exact result. In the half-filled-band case and next-neighbor hopping the ground-state energy (77b) obtained with the GWF is nonanalytical in U:

$$\frac{E}{L} = -\left[\frac{4}{\pi}\right]^2\frac{t^2}{U}\frac{1}{\ln\bar{U}} , \tag{80}$$

owing to the logarithmic correction to the $(-t^2/U)$ behavior of the exact result.[41,13] In numerical calculations,[25,28,33] where only *finite* systems are considered, the potential energy comes out proportional to g^2, since a logarithmic correction can hardly be discerned (particularly, if it is not expected). Thus, they suggest a (t^2/U) dependence of E with a seemingly much too small numerical prefactor[25] when compared with the exact result.[13] Apparently, systems considerably larger than the one studied so far $(L = 90$ in Ref. 33) have to be used to identify the logarithmic correction found here. A Brinkman-Rice transition[21] does not occur for any finite U.

The fact that for $n = 1$ and large U the ground-state energy obtained with the GWF is considerably higher than the exact result was already observed by Kaplan et al.[25] in their investigation of finite rings. This they attributed to the missing spatial correlation between empty sites and doubly occupied sites in the GWF. Indeed, for strong interactions the probability for finding a hole and a (energetically costly) doubly occupied site close to each other should be enhanced over that in the noninteracting case since this would allow the latter to decay more easily. This is, indeed, not the case in the GWF as recently

shown analytically by Gebhard and Vollhardt[35] using the present approach. Including such a correlation in the variational ansatz, Kaplan et al.[25] found the energy to be in much better agreement with the Lieb-Wu result.[13] A similar finding is due to Baeriswyl.[42] Whether the nonanalyticity in (80) is thereby removed or shifted to a higher order in t/U is, however, not clear at present.

For general interaction strength U the minimization of E has to be done numerically. The resulting U dependence of the correlation parameter g is shown in Fig. 10 for two different densities. It is compared with the results by the GA (Ref. 18) for next-neighbor hopping, where—in the special case of half-filling—a localization (Brinkman-Rice) transition ($g=0$, $E=0$) occurs at $U = 8\,|\bar{\varepsilon}_0| = 32t/\pi$. The result for the ground-state energy as a function of U/t is shown in Fig. 11 for several densities. It is compared with the exact results[13,40] and with those of the GA.[18] We observe the following features: (i) for not too strong interactions $(U/t \lesssim 2)$ both the GA and the approximation-free results reproduce the exact results very well;[31] (ii) for $n=1$ and large U all three results differ most strongly; (iii) for small densities, the results of the GA approach the correct results for the GWF (already at $n=0.3$ they cannot be distinguished in Fig. 11), but both do not approach the exact results; (iv) the GA sometimes yields energies lower than the approximation-free result and hence does not always obey the variational principle.

In Fig. 12 the ground-state energy is plotted as a function of density n for different interaction U/t and is compared with the exact result. This presentation illuminates an essential difference between the two results: while in the exact solution[13,40] the chemical potential $\mu=\partial E/\partial n$ is discontinuous at $n=1$ for all $U>0$ (absence of a Mott transition at *nonzero* U), the result for the GWF is continuous for all finite U and only shows a discontinuity at $U = \infty$ (absence of a metal-insulator transition at *finite* U). Mathematically this is expressed by the fact that for $n=1$ the exact result[13] is nonanalytic in U at $U=0$ while the result for the GWF is nonanalytic at $U = \infty$. Hence, in contrast to the exact result, the GWF always leads to a conducting state except for $U = \infty$. This is also evident
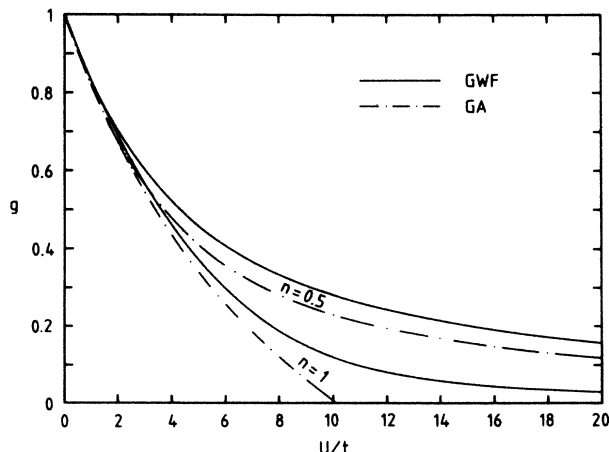

FIG. 11. The ground-state energy of the one-dimensional Hubbard model as a function of U for three different densities.

from the existence of a finite discontinuity q at the Fermi surface for all $U < \infty$ even at $n=1$ as shown in Fig. 13 for various densities. Clearly, in one dimension the exact momentum distribution for $n=1$ is not expected to have a discontinuity for $U>0$; for large U this has been shown explicitly by Takahashi.[43] On the other hand, the absence of a discontinuity is an inherent feature of any perfect nesting lattice at $n=1$. For $n<1$ even the exact solution in $d=1$ may well have a sharp Fermi surface as exhibited by the GWF. In particular, in higher dimensions and for lattices without perfect nesting a discontinuity should exist for small U even at half-filling. In this case, a Brinkman-Rice-type transition,[21] where q vanishes at a *finite* U, would have to occur.

FIG. 10. The U dependence of g for the one-dimensional Hubbard model at two different densities.

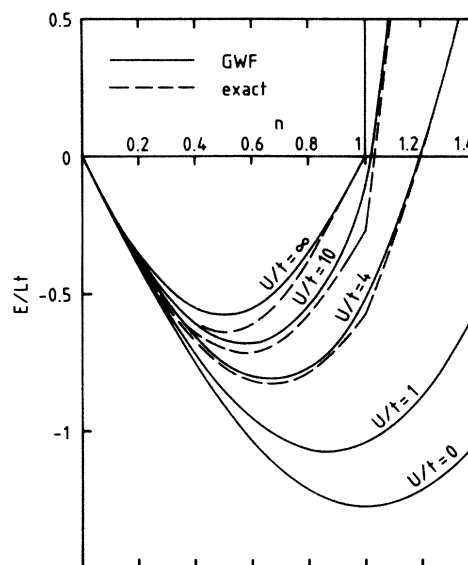FIG. 12. The density dependence of the ground-state energy for different U as compared with the exact result (Ref. 13).
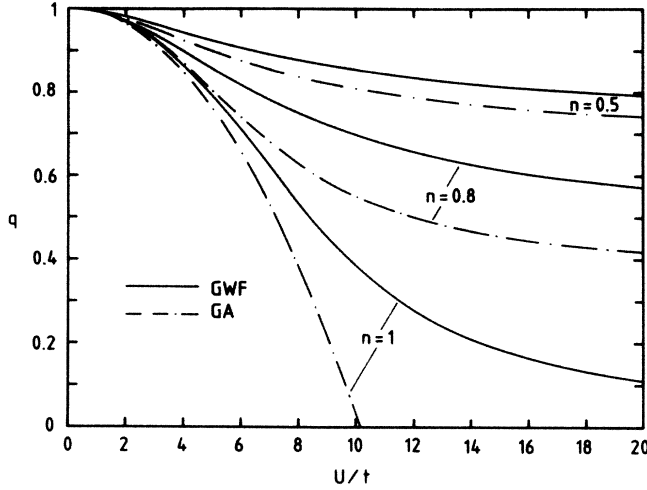
FIG. 13. The U dependence of the discontinuity q for three different densities.

VI. DIAGRAMMATIC DERIVATION OF THE GUTZWILLER APPROXIMATION

The formalism described in the preceding sections may also be used to derive the results of the *Gutzwiller approximation*[18] in terms of diagrams. For this one has to make simple approximations for the values $v(G)$ of the graphs. These approximations will then be argued to be correct in the case of *infinite* dimensionality.

To this end we classify all graphs contributing to c_m and $f_{m\sigma}(\mathbf{k})$ by their "effective" number of lines, λ_σ, as follows. We consider a general graph with l_σ σ lines carrying momenta $\mathbf{k}_1, \ldots, \mathbf{k}_{l\sigma}$. The topological structure of graphs sometimes implies that different lines carry the *same* momentum [as do, for example, the two internal lines in $C_{2b,c}$ (see Fig. 2)] or, in the case of improper graphs $F_{m i \sigma}(\mathbf{k})$ as in Fig. 5, carry the external momentum $\mathbf{k}$ [e.g., the middle line in $F_{2b\uparrow}(\mathbf{k})$]. This corresponds to a constraint $\mathbf{k}_i = \mathbf{k}_j$ or $\mathbf{k}_i = \mathbf{k}$, respectively. Let the number of such independent constraints be indicated by z_σ and z'_σ, respectively. The "effective number of lines" is then defined as

$$\lambda_\sigma \equiv l_\sigma - (z_\sigma + z'_\sigma) \ . \tag{81}$$

For example, $\lambda_\uparrow(C_{2a}) = \lambda_\downarrow(C_{2a}) = 2$, $\lambda_\uparrow(C_{2b}) = 1$, $\lambda_\downarrow(C_{2b}) = 2$, $\lambda_\uparrow(F_{2a\uparrow}) = 1$, $\lambda_\downarrow(F_{2a\uparrow}) = 2$, $\lambda_\uparrow(F_{2b\uparrow}) = 0$, and $\lambda_\downarrow(F_{2b\downarrow}) = 2$.

We will now show that the following assignment of values $v(G)$ to any graph G [i.e., C_{mi} and $F_{mi\sigma}(\mathbf{k})$] is equivalent to the GA:

$$v(C_{mi}) = n_\uparrow^{\lambda_\uparrow} n_\downarrow^{\lambda_\downarrow} \ , \tag{82a}$$

$$v(F_{mi\sigma}(\mathbf{k})) = n_\uparrow^{\lambda_\uparrow} n_\downarrow^{\lambda_\downarrow} \quad \text{for proper graphs} \ , \tag{82b}$$

$$v(F_{mi\sigma}(\mathbf{k})) = n_\uparrow^{\lambda_\uparrow} n_\downarrow^{\lambda_\downarrow} n_{\mathbf{k}\sigma}^0 \quad \text{for improper graphs} \ . \tag{82c}$$

Clearly, in this approximation $n_{\mathbf{k}\sigma}$ is given by a step function, i.e., by two $\mathbf{k}$-independent values $n^{(-)}$ and $n^{(+)}$ on either side of the Fermi surface (FS), since the values

of graphs for $|\mathbf{k}| > k_F$ or $|\mathbf{k}| < k_F$ are constant. Restricting ourselves to $n_\uparrow = n_\downarrow = n/2$ for convenience we now show that (82) lead to the results of the GA.

For $n = 1$, (59) and (60) imply $n^{(+)} + n^{(-)} = 1$ and $f_m^{(-)} + f_{m+1}^{(-)} = -f_{m+1}^{(+)}$, respectively. Then we find

$$f_m^{(-)} = (-1)^m \frac{(2m-1)!!}{(2m)!!}, \quad |\mathbf{k}| < k_F \ , \tag{83a}$$

$$f_m^{(+)} = (-1)^m \frac{(2m-3)!!}{(2m)!!}, \quad |\mathbf{k}| > k_F \ . \tag{83b}$$

In $d = 1$ the right-hand side of (83a) and (83b) were the results for $f_m(k)$ only at points $k = k_F^\pm$, respectively. Using (57) the $f_m^{(\pm)}$ are related to the c_m yielding

$$c_m = (-1)^{m+1} \frac{(2m-1)!!}{2(2m)!!}, \quad n = 1 \ . \tag{84}$$

To obtain the general n dependence of the c_m we define

$$r_m \equiv c_m + c_{m+1} \ . \tag{85}$$

According to (85), r_m is a polynomial in $n \in [0,2]$ and since $r_m(n) = r_m(2-n)$ [see (34a)] r_m is a polynomial in $n(2-n)$, $r_m = \sum_p \alpha_p [n(2-n)]^p$, as follows from a theorem for symmetric functions. For any graph C_m, $m + 1 \le \lambda_\uparrow + \lambda_\downarrow \le 2m$. Consequently α_p is only different from zero for $p = m + 1$, and r_m may be written as $r_m = \alpha[n(2-n)]^{m+1}$. Using (84) this yields

$$r_m = (-1)^{m+1} \frac{(2m-1)!!}{2(2m+2)!!} [n(2-n)]^{m+1} \ . \tag{86}$$

The density of doubly occupied sites $\langle \hat{D} \rangle / L$ in (27) is thus found as

$$\frac{\langle \hat{D} \rangle}{L} = \frac{1 + n(g^2 - 1) - [1 + (2n - n^2)(g^2 - 1)]^{1/2}}{2(g^2 - 1)} \ , \tag{87}$$

as was obtained by Gutzwiller.[18]

The discontinuity q for a step-type $n_\mathbf{k}$ follows from (58) and (87) as

$$q = \frac{2(n - 2d)(\sqrt{1-n+d} + \sqrt{d})}{n(2-n)} \ , \tag{88}$$

again in accordance with the results of the GA.

We now discuss the meaning of the approximation for the graphs as formulated in (82). To this end we consider a typical Monte Carlo integration of a general graph G contributing to c_m. In this method internal momenta $\mathbf{k}_1 \cdots \mathbf{k}_{m+1}$ are generated at random (within the BZ), the value of the graph being given by the probability that every line carries a momentum lying inside the Fermi surface (FS) (which is a random event). In the case that these random events are *independent*, except of course for lines l_i and l_j with $\mathbf{k}_i = \mathbf{k}_j$, the total probability for the momenta of the different lines to lie within the FS is the *product* of the corresponding probabilities for the λ_σ effective lines. This then leads to the assignment in (82) and hence to the GA.

We have calculated a number of graphs for d-dimensional lattices with hypercubic Fermi surfaces. For high dimensions d the value $v(G)$ of a graph G always
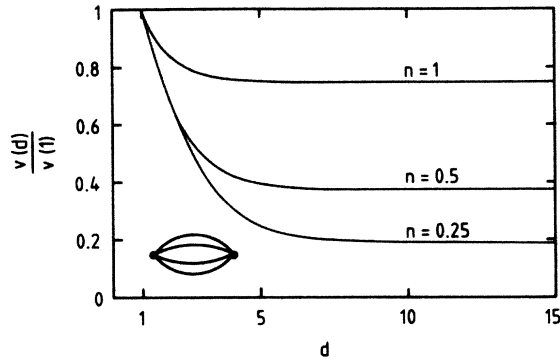
FIG. 14. Value of the graph C_{2a} for different dimensions, normalized by its value in one dimension.

converged to the expressions given in (82). In Fig. 14 we show $v(C_{2a})$, the value of graph C_{2a}, as a function of dimensionality of the system, normalized by $v(1) = \frac{2}{3}(n/2)^3$, the value of C_{2a} in one dimension. The limiting value clearly agrees with the expression obtained from (82), i.e., $\frac{1}{4}n$. Thus the GA turns out to be the exact result for the GWF in the limit $d \to \infty$. The mathematical proof for this intuitively plausible result will be published elsewhere. The assumption of independence in k space formulated above is also consistent with the approximation on which the GA is based, i.e., the neglect of spatial correlations between different lattice sites.[19,20] The GA is seen to involve all diagrams in all orders, i.e., it is not given by a particular subclass of diagrams.

We note that Hashimoto[28] derived an expression for the expectation value of the Hubbard interaction of the form

$$\langle H_I \rangle = U g^2 \mathcal{F}''(g)/\langle \psi_G | \psi_G \rangle ,$$

where $\mathcal{F}''(g)$ is an even function of g. Based on his results for $n_{k\sigma}$ and the relation (58) be concluded that—for small g—he was able to prove the absence of a Brinkman-Rice transition in the exact solution for the GWF in any dimension. However, he seems to have excluded the possibility that $\mathcal{F}''(g)$ diverges for $g \to 0$. Indeed, while in $d = 1$ we find only a weak divergence $[\mathcal{F}''(g)/\langle \psi_G | \psi_G \rangle \propto \ln(1/g^2)]$, our results for $d = \infty$ show that a Brinkman-Rice transition does occur, indicating that $\mathcal{F}''(g)/\langle \psi_G | \psi_G \rangle$ diverges as $1/g$ in this case. The behavior for $2 \leq d < \infty$ is not yet known.

VII. SUMMARY

A new diagrammatic approach to the evaluation of ground-state properties of correlated fermions in terms of the GWF has been described. The resulting expressions are simple enough to yield analytic, approximation-free results in one dimension, e.g., of the momentum distribution n_k and the ground-state energy E of the Hubbard model. This allows for an unambiguous assessment of the properties of the GWF in one dimension as discussed in Sec. V. The GWF is seen to lead to a sharp Fermi surface for all densities and finite interaction strengths. In the special case of a half-filled band, logarithmic corrections to the usual $(-t^2/U)$ behavior of E at large U are

found in contrast to the exact result.[13]

The same approach has also been applied to the calculation of correlation functions for Hubbard-type models,[35,44,45] i.e., spin-spin, density-density, hole-hole, hole doubly occupied site, and superconducting correlations. Again, analytic, approximation-free results have been obtained in one dimension. Details will be reported separately.[44] The comparison with exact results for spin-spin[43] and hole-hole correlations[36] in the half-filled band case and $U = \infty$ show very good agreement (see also Refs. 25 and 26). On the other hand, for strong interactions spatial correlations between empty and doubly occupied sites are not sufficiently described by the GWF,[25] which only accounts for the average behavior of these entities.

The case where doubly occupied sites in the ground-state wave function are *favored*, as in the case for attractive on-site interactions, is described by correlation parameters $g > 1$ rather than $0 \leq g \leq 1$ used so far in the GWF, (7). Using canonical transformations for spins on AB lattices as in (30), the above mentioned correlation functions may be shown to be related under a replacement of g by $1/g$; this allows for a simple extension to the case $g > 1$.[44,45]

In higher dimensions $(d > 1)$ analytic evaluations to arbitrary order have not yet been possible. This is due to the fact that now the shape of the Fermi surface is no longer scale invariant but depends on the density n. In addition, the contribution of Umklapp-processes becomes even more complex. Hence the general structure of the coefficients c_m and $f_m(\mathbf{k})$ is no longer given by simple polynomials as in $d = 1$ but is more complicated. This makes general evaluations much more difficult. Finite orders can of course easily be calculated numerically and in most cases already yield sufficient accuracy unless one is interested in the limit $U \to \infty$. For very large dimensionality $(d \to \infty)$ we find that the Gutzwiller *approximation* gives the correct results for expectation values in terms of the GWF. In this limit the evaluation of graphs is simple.

ACKNOWLEDGMENT

We are grateful to F. Gebhard for valuable discussions.

APPENDIX A: POLYNOMIAL STRUCTURE OF THE COEFFICIENTS $c_m, f_m(k)$ AND UMKLAPP-PROCESSES FOR $d=1$ DIMENSION

We discuss the structure of c_m and $f_m(\mathbf{k})$ as functions of n and n and k, respectively. This is necessary before we may employ the derivatives of the ph symmetric relations (33) and (60) by n, to determine the polynomial structure of $c_m, f_m(k)$.

Coefficients c_m. Momentum conservation at the vertices leads to combinations of momenta on the lines of a graph. Since a vertex is a termination point of four lines, three of them may carry different moments k_1, k_2, k_3 while the fourth carries the momentum $K = k_1 + k_2 + k_3$. K is defined up to a reciprocal lattice vector, i.e., if K is

not in the first BZ it is shifted to lie in the first BZ ["Umklapp-process" (UP)]. For a UP to occur, K must lie outside $[-1+k_F, 1-k_F]$. Since $k_F = n/4$, this is only possible for $n > 1$.

For $n \leq 1$ no UP occur. The value $v(G)$ of a graph contributing to c_m (which has $2m$ lines) is then simply the volume of a polyhedron in the $(m+1)$-dimensional space of the free momenta $k_1, \ldots, k_{m+1}$. This polyhedron is defined by $2m$ inequalities of the form

$$\left| \sum_{j=1}^{m+1} u_{ij} k_j \right| \leq k_F = \frac{n}{4} , \qquad (A1)$$

where $i = 1, \ldots, 2m$ counts the lines and $u_{ij} \in \{0, 1, -1\}$. Since they are linear in the k_j the corresponding volume is proportional to n^{m+1}. Hence $c_m \propto n^{m+1}$. This implies that for $n < 1$ and $n > 1$ (by ph symmetry) all derivatives of c_m are continuous. Only at $n = 1$, where the UP set in, the differentiability is limited. Setting $n = 1 + \delta$, $0 < \delta \ll 1$, one has $k_F = (1+\delta)/4$. The contribution of the UP is then seen to be of order δ^3. Since the UP do not contribute for $\delta < 0$ there is a discontinuity in the third derivative of the c_m by n at $n = 1$; the first two derivatives are continuous.

Coefficients $f_{m\sigma}(k)$ for $n_\uparrow = n_\downarrow = n/2$. The $f_m(k)$ (we drop the spin index) depend on $k \in [-\frac{1}{2}, \frac{1}{2}]$ and $n \in [0,2]$. We only need to discuss their structure for $k \geq 0$ and $n \leq 1$ since we have inversion symmetry (dependence on $|k|$) and because the behavior for $n > 1$ follows from ph symmetry [see (60)]. Since $f_m = h_m + c_{m-1}$, (50), we may restrict the discussion to the graphs contributing to h_m only. These consist of two external vertices, where the momentum k enters and exits, and $2m - 1$ lines; there are m free internal moments $k_i \in [-k_F, k_F]$, $k_F = n/4$, the others are determined by momentum conservation.

As seen in the case of the c_m, UP's at an internal vertex do not occur for $n \leq 1$. Since in the present case an external vertex involves the external momentum k and two other (free) momenta UP's do not occur for $n \leq 1$ and $k \leq 1 - 3k_F$. In particular, for $n < \frac{2}{3}$ UP's are never possible for any k. Furthermore, for $k > 3k_F$ (which is only possible for $n < \frac{2}{3}$) the value of a graph has to vanish.

In Fig. 6 the different regions of $f_m(k)$ in k and n are shown. The lines $n = 1$, $k = 3n/4$, and $k = 1 - 3n/4$ and the ph-symmetric counterparts are boundaries of regions

whose analytic structures have to be discussed separately. At $k = k_F = n/4$ the h_m are discontinuous since reducible graphs vanish discontinuously.

In (I) no UP's occur. The value of a graph is given by the volume of a polyhedron in the space of free momenta $k_1 \cdots k_m$, i.e., is determined by the set of linear inequalities

$$\left| u_i \frac{k}{n} + \sum_{j=1}^{m} u_{ij} \frac{k_j}{n} \right| \leq \frac{1}{4}, \quad i = 1, \ldots, 2m - 1 , \qquad (A2)$$

where $u_i, u_{ij} \in [0, 1, -1]$. Its volume is of the form $n^m P_m(k/n)$; where P_m is a polynomial of order $\leq m$. Hence, in (I), $h_m = n^m P_m(k/n)$. The same is true for (II), i.e., $h_m = n^n Q_m(k/n)$ with a different polynomial Q_m. In (III) $(k \geq 3k_F)$, h_m vanishes. In (IV) UP's occur at the external vertices. While the usual k sums (i.e., those not involving a reciprocal lattice vector) lead to $n^m Q_m(k/n)$, the UP's yield the same contribution with k replaced by $1 - k$. Hence, in (IV) the h_m have the form

$$h_m = n^m \left[Q_m \left[\frac{k}{n} \right] + Q_m \left[\frac{1-k}{n} \right] \right] .$$

Concerning the differentiability of the h_m at the boundaries, a similar analysis as for the c_m shows[46] that $h_m(n,k)$ [and hence $f_m(n,k)$] has the following properties: (i) it is discontinuous at $k = k_F$; (ii) the first two derivatives by n at $n = 1$ are continuous for all $k \neq k_F$; (iii) the first derivative by k is continuous at $k = 3k_F$ for $n \leq \frac{2}{3}$, at $k = 1 - 3k_F$ for $\frac{2}{3} < n \leq \frac{4}{3}$, and at $k = -2 + 3k_F$ for $n \geq \frac{4}{3}$.

APPENDIX B: RECURSION RELATIONS FOR THE COEFFICIENTS $R_m^{(j)}, Q_m^{(j)}$

To determine $R_m^{(j)}$ and $Q_m^{(j)}$ we take the jth derivative of the three equations in (64) with respect to k at $k = \frac{1}{4}$; this yields

$$R_m^{(j)}(\tfrac{1}{4}) + R_{m+1}^{(j)}(\tfrac{1}{4}) = -\delta_{j0} \frac{(-1)^{m+1}}{2(m+1)} - Y_{m+1}^{(j)} , \qquad (B1a)$$

$$mR_m^{(j)}(\tfrac{1}{4}) + (m+1)R_{m+1}^{(j)}(\tfrac{1}{4}) = \delta_{j0} \frac{(-1)^{m+1}}{2} + (m+1-2j)Y_{m+1}^{(j)} - Q_{m+1}^{(j+1)}(\tfrac{1}{4}) , \qquad (B1b)$$

$$m(m-1)R_m^{(j)}(\tfrac{1}{4}) + (m+1)mR_{m+1}^{(j)}(\tfrac{1}{4}) = -\delta_{j0} \frac{m}{2}(-1)^{m+1} - [m(m+1) - 4mj + 4j(j-1)]Y_{m+1}^{(j)}$$

$$+ 2(m-2j)Q_{m+1}^{(j+1)}(\tfrac{3}{4}) - Q_{m+1}^{(j+2)}(\tfrac{3}{4}) , \qquad (B1c)$$

where

$$Y^{(j)}_{m+1} \equiv (-1)^j Q^{(j)}_{m+1}(\tfrac{1}{4}) + Q^{(j)}_{m+1}(\tfrac{3}{4}) \ . \tag{B2}$$

Together with the initial values

$$R^{(j)}_1(\tfrac{1}{4}) = -\tfrac{1}{2}\delta_{j0} \ ,$$

$$Q_{m+1}(\tfrac{3}{4}) = Q'_{m+1}(\tfrac{3}{4}) = 0 \ , \tag{B3}$$

Eqs. (B1) allow one to calculate $R^{(j)}_m(\tfrac{1}{4})$, $Q^{(j)}_m(\tfrac{1}{4})$, and $Q^{(j)}_m(\tfrac{3}{4})$ recursively for any j and m.

In particular, for $j = 0$ (B1a) and (B1b) imply

$$R_m(\tfrac{1}{4}) = (-1)^m \frac{(2m-1)!!}{(2m)!!}, \quad m \geq 1 \ , \tag{B4a}$$

$$Q_m(\tfrac{1}{4}) = -\frac{(-1)^m}{2m} + (-1)^m \frac{(2m-3)!!}{(2m)!!}, \quad m \geq 1 \ , \tag{B4b}$$

where we define $(-1)!! \equiv 1$. Furthermore,

$$R'_m(\tfrac{1}{4}) = \frac{(-1)^m}{2} \frac{(2m-3)!!}{(2m-4)!!}, \quad m \geq 2 \ , \tag{B5a}$$

$$Q'_m(\tfrac{1}{4}) = \frac{(-1)^m}{2} \frac{(2m-5)!!}{(2m-4)!!}, \quad m \geq 2 \ , \tag{B5b}$$

etc. Eliminating $Q^{(j)}_{m+1}(\tfrac{1}{4})$ from (B1) we find ($j \geq 1$)

$$R^{(j)}_{m+1}(\tfrac{1}{4}) = -\frac{m-j+\tfrac{1}{2}}{m-j+1} R^{(j)}_m(\tfrac{1}{4})$$

$$\quad - \frac{1}{2(m-j+1)} Q^{(j+1)}_{m+1}(\tfrac{3}{4}), \quad 1 \leq j \leq m \ , \tag{B6a}$$

$$R^{(m+1)}_{m+1}(\tfrac{1}{4}) = \begin{cases} 0 & \text{for } m+1 \text{ odd} \ , \\ -2Q^{(m+1)}_{m+1}(\tfrac{3}{4}) & \text{for } m+1 \text{ even} \ , \end{cases} \tag{B6b}$$

$$Q^{(j+2)}_{m+1}(\tfrac{3}{4}) = 2(m-2j)Q^{(j+1)}_{m+1}(\tfrac{3}{4}) + 2mR^{(j)}_m(\tfrac{1}{4})$$

$$\quad - 4j(m-j+1)[R^{(j)}_m(\tfrac{1}{4}) + R^{(j)}_{m+1}(\tfrac{1}{4})] \ ,$$

$$1 \leq j \leq m-1 \ . \tag{B6c}$$

Together with (B3) we have thus a simple algorithm for the recursive calculation of $R^{(j)}_m(\tfrac{1}{4})$ and $Q^{(j)}_m(\tfrac{3}{4})$ in (66) for arbitrary j and m.

*Present address: Institut für Theoretische Physik C, Rheinisch-Westfälische Technische Hochschule Aachen, Sommerfeldstrasse 26/28, D-5100 Aachen 1, Federal Republic of Germany.

[1]Electron Correlation and Magnetism in Narrow Band Systems, Vol. 29 of Solid State Sciences, edited by T. Moriya (Springer, Berlin, 1981).

[2]C.-W. Woo, in The Physics of Liquid and Solid Helium, Part I, edited by K. H. Bennemann and J. B. Ketterson (Wiley, New York, 1976), p. 349.

[3]M. Roger, J. H. Hetherington and J. M. Delrieu, Rev. Mod. Phys. 55, 1 (1983).

[4]See, for example, P. A. Lee, T. M. Rice, J. W. Serene, L. J. Sham, and J. W. Wilkins, Comments Cond. Matt. Phys. XII, 99 (1986).

[5]P. A. Lee and T. V. Ramakrishnan, Rev. Mod. Phys. 57, 287 (1985).

[6]T. M. Rice, Philos. Mag. B52, 419 (1985).

[7]P. W. Anderson, Science 235, 1196 (1987).

[8]For a preliminary review, see T. M. Rice, Z. Phys. B 67, 141 (1987).

[9]E. Feenberg, Theory of Quantum Fluids (Academic, New York, 1969).

[10]M. C. Gutzwiller, Phys. Rev. Lett. 10, 159 (1963).

[11]J. Hubbard, Proc. R. Soc. London, Ser. A 276, 238 (1963).

[12]J. Kanamori, Prog. Theor. Phys. 30, 275 (1963).

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

[14]M. Takahashi, J. Phys. 10, 1289 (1977).

[15]J. E. Hirsch, R. L. Sugar, D. J. Scalpino, and R. Blankenbecler, Phys. Rev. B 26, 5033 (1982).

[16]J. E. Hirsch, Phys. Rev. B 28, 4059 (1983); 31, 4403 (1985); 34, 3216 (1986); H. Q. Lin and J. E. Hirsch, ibid. 35, 3359 (1987).

[17]M. C. Gutzwiller, Phys. Rev. A 134, 923 (1964).

[18]M. C. Gutzwiller, Phys. Rev. A 137, 1726 (1965).

[19]T. Ogawa, K. Kanda, and T. Matsubara, Prog. Theor. Phys. 53, 614 (1975).

[20]D. Vollhardt, Rev. Mod. Phys. 56, 99 (1984).

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

[22]P. W. Anderson and W. F. Brinkman, in The Helium Liquids, edited by J. G. M. Armitage and I. E. Farqhar (Academic, New York, 1975), p. 315.

[23]D. Vollhardt, P. Wölfle, and P. W. Anderson, Phys. Rev. B 35, 6703 (1987).

[24]G. Kotliar and A. E. Ruckenstein, Phys. Rev. Lett. 57, 1362 (1986).

[25]T. A. Kaplan, P. Horsch, and P. Fulde, Phys. Rev. Lett. 49, 889 (1982).

[26]P. Horsch and T. A. Kaplan, J. Phys. C 16, L1203 (1983).

[27]H. Bethe, Z. Phys. 71, 205 (1931); L. Hulthén, Ark. Astron. Fyz. 26A, 11 (1938).

[28]K. Hashimoto, Phys. Rev. B 31, 7368 (1985).

[29]P. Horsch, Phys. Rev. B 24, 7351 (1981).

[30]D. Baeriswyl and K. Maki, Phys. Rev. B 31, 6633 (1985).

[31]D. Baeriswyl, J. Carmelo, and K. Maki (unpublished).

[32]C. Gros, R. Joynt, and T. M. Rice, Phys. Rev. B 36, 381 (1987).

[33]H. Yokoyama and H. Shiba, J. Phys. Soc. Jpn. 56, 1490 (1987).

[34]W. Metzner and D. Vollhardt, Phys. Rev. Lett. 59, 121 (1987).

[35]F. Gebhard and D. Vollhardt, Phys. Rev. Lett. 59, 1472 (1987).

[36]G. Stollhoff and P. Fulde, Z. Phys. 26, 257 (1977).

[37]See, for example, A. A. Abrikosov, L. P. Gorkov, and Y. E. Dzyaloshinskii, Quantum Field Theoretical Methods in Statistical Physics (Pergamon, New York, 1965).

[38]P. Horsch and P. Fulde, Z. Phys. B 36, 23 (1979).

[39]E. Daniel and S. H. Vosko, Phys. Rev. 120, 2041 (1960); B. I. Lundqvist and C. Lyden, Phys. Rev. B 4, 3360 (1971).

[40]H. Shiba, Phys. Rev. B 6, 930 (1972).

[41]P. W. Anderson, Solid State Phys. **14**, 166 (1963).

[42]D. Baeriswyl, in *Nonlinearity in Condensed Matter Physics,* Vol. 69 of *Solid-State Sciences,* edited by A. R. Bishop *et al.* (Springer, Berlin, 1987), p. 183.

[43]M. Takahashi, J. Phys. **10**, 1289 (1977).

[44]F. Gebhard and D. Vollhardt (unpublished).

[45]F. Gebhard, Diploma thesis, Technische Universität, München, 1987.

[46]W. Metzner, Diploma thesis, Technische Universität, München, 1987.