On the magnetic phase diagram
of the Hubbard model

Danny ten Haaf

On the magnetic phase diagram
of the Hubbard model

On the magnetic phase diagram
of the Hubbard model

PROEFSCHRIFT

DOOR

Daniël Frederik Bastiaan ten Haaf

GEBOREN TE MOOK EN MIDDELAAR OP 11 SEPTEMBER 1968
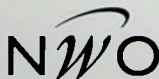
Promotiecommissie:

Promotor:	Prof.dr. J.M.J. van Leeuwen
Co-promotor:	Dr. P.J.H. Denteneer
Overige leden:	Prof.dr. J. Amesz
	Prof.dr. C.W.J. Beenakker
	Dr. H.W.J. Blöte
	Prof.dr. D. Frenkel
	Prof.dr. H. de Raedt
	Prof.dr.ir. W. van Saarloos
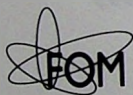
Contents

To you and to me
To us

Contents

Introduction and outline

Interacting electrons are key ingredients for understanding the properties of various classes of materials, ranging from the energetically most favorable shape of small molecules to the magnetic and superconducting instabilities of lattice electron systems. The single-band Hubbard model is presumably the simplest model for describing the behavior of correlated electrons in a solid. Initially, the model was used to describe magnetism in transition metals. More recently, it plays a role in setting up theories of high-temperature superconductivity. While it seems likely that for the latter phenomenon more complex models are needed, even this simple model is not well understood. For one dimension some rigorous results are known, but in higher dimensions the main results have been obtained from Monte Carlo and finite-lattice calculations only.

The Hubbard model is a tight-binding model. Electrons are located on lattice sites and hop from site to site with a rate t/h, where t is an overlap integral between neighboring sites. Apart from t, the model comprises only one other energy parameter U, which is the penalty for having two electrons on one site. Although the model has only two energy parameters, it has a rich phase diagram, of which the low-temperature phase is still a mystery. A wealth of (approximate) calculations exists which give more or less reliable information on the possible phases which may occur. The Hubbard model is capable of producing an antiferromagnetic as well as a ferromagnetic phase, and likely a host of variations on this theme. It shows a general tendency towards antiferromagnetic arrangements. These are most convincingly demonstrated in the one-dimensional Hubbard model, which has been solved exactly by Lieb and Wu [1]. In fact, for any finite U the ground state of the half-filled system is an antiferromagnetic insulator in that case. For higher-dimensional systems the possible phases of the Hubbard model are less well understood. It is believed to be antiferromagnetic at half filling. Off half filling, the model is superconducting for negative U under the same circumstances where it is antiferromagnetic for positive U. A well known but surprising result is a theorem by Nagaoka [2], stating that a Hubbard model on a bipartite lattice with one hole and with infinite interaction strength U has a ferromagnetic ground state.

Many authors have investigated the behavior of the model at large U near half filling, to find whether the ferromagnetic ground state is limited to only one point in the phase diagram, or part of a whole region of ferromagnetic behavior. Various methods are being used for this purpose [3], including exact diagonalization of small systems, Monte Carlo simulations, mean-field, and cluster expansion methods. Over the years, many interesting results have been obtained that shed some light on the issue of magnetism. However, a conclusive answer to this question has not been given yet.

As our main interest focuses on the existence of magnetic phases, we study in this thesis, in addition to the grand potential, spin-spin correlation functions. From the grand potential one can calculate thermodynamic quantities as density and susceptibility. We use the nearest-neighbor correlation between the z-components of the spin as an indicator for the magnetic tendencies. We will concentrate on two methods by which information on the phase diagram of the Hubbard model can be obtained: cluster expansions and Monte Carlo simulations.

By means of a cluster expansion method, series expansions in terms of the parameter βt are generated. In Chapter 2 this method is described, and expressions for various quantities are given. These expressions show very well convergent behavior for high temperatures. Standard extrapolation methods fail when trying to extend these results to lower temperatures, and therefore we present a new method of extrapolation in Chapter 3. By this method an extrapolation of the series results can be obtained all the way down to zero temperature.

A different approach, using the technique of quantum Monte Carlo simulations, is described in Chapter 4. By this technique one is able to investigate sizeable systems without a priori knowledge about the physical state which is to be expected. The character of the fermion wave function, which must be antisymmetric, causes that straightforward application of the Monte Carlo technique (which works well in the case of bosons) leads to a very bad signal-to-noise ratio, such that the noise overwhelms any properties of the system. We adapt this technique to make it suitable for lattice fermions, and use it to calculate ground-state properties. The implementation and results of these calculations are shown in Chapter 5.

Finally, we draw our conclusions for the magnetic phase diagram from the different approaches, and we compare our results to those available in the literature.

Acknowledgments

---

[1] See the list of publications on page 109.

1 The Hubbard model

A description of the single-band Hubbard model, as it is used in this thesis, is given. The parameters in this model, as well as some symmetries and limiting cases, are discussed, and the notions of configuration space and antisymmetry are clarified.

1.1 The model

The Hubbard model is a most simple description of the interactions between electrons in a metal, using only a few basic parameters. The electrons are localized on lattice sites, representing the ions in a crystalline material. They can hop from site to site, only restricted by the Pauli principle which states that two electrons of the same spin cannot occupy the same site, and by a penalty for two electrons of opposite spin that occupy the same site. This on-site interaction is considered to be a reasonable approximation for the Coulomb force between the electrons, because the range of this force is usually very short due to shielding.

The Hubbard Hamiltonian is commonly expressed in terms of the creation and annihilation operators $c_{j\sigma}^{\dagger}$, $c_{j\sigma}$ for an electron with spin σ at site j. The Pauli principle is obeyed through anticommutation relations for these operators:

$$\{c_{j\sigma}^{\dagger}, c_{j'\sigma'}\} = \delta_{jj'}\delta_{\sigma\sigma'}. \tag{1.1.1}$$

We denote the Hamiltonian as follows:

$$\mathcal{H} = \mathcal{H}_{\text{kin}} + \mathcal{H}_{\text{int}}. \tag{1.1.2}$$

The first term is the kinetic or *hopping* term, which reads

$$\mathcal{H}_{\text{kin}} = -t \sum_{\langle i,j\rangle,\sigma} c_{i\sigma}^{\dagger}c_{j\sigma}. \tag{1.1.3}$$

Here, the parameter t is the one-electron transfer integral between a nearest-neighbor pair of sites $\langle i, j\rangle$. By this term, an electron can hop from one lattice site to the other, as it is annihilated on site j and created on a neighboring site i. The other term is the on-site interaction

$$\mathcal{H}_{\text{int}} = U \sum_{i} n_{i\uparrow}n_{i\downarrow}, \tag{1.1.4}$$

where U is the interaction strength, and $n_{i\sigma}$ is the number operator $c_{i\sigma}^{\dagger}c_{i\sigma}$, counting the number of particles with spin σ at site i. Although one can very well take U negative, we will restrict ourselves to the case of positive U.

For convenience, we will use two more terms in the Hamiltonian. In Chapters 2 and 3 we will be working in the grand canonical ensemble, which makes it necessary to use a term

by which the number of particles can be varied. The particle density n can be controlled by the chemical potential μ, if a term

$$\mathcal{H}_{\text{chem}} = -\mu \sum_{i,\sigma} n_{i\sigma} \tag{1.1.5}$$

is added to the Hamiltonian. The symmetry of the Hubbard model is such that $\mu = U/2$ yields a half-filled lattice with one electron per site ($n = 1$). We will calculate various expressions in terms of μ, and use thermodynamic relations to re-express them in terms of the density.

The other term, which is useful when studying magnetic properties, introduces a uniform external magnetic field h in the z-direction, acting on the electron spin:

$$\mathcal{H}_{\text{mag}} = -\frac{h}{2} \sum_i \left(n_{i\uparrow} - n_{i\downarrow} \right). \tag{1.1.6}$$

Similarly, one can include a staggered magnetic field h_s. In that case, $\mathcal{H}_{\text{mag}}$ obtains the form

$$\mathcal{H}_{\text{mag}} = -\frac{h_s}{2} \sum_i (-1)^i \left(n_{i\uparrow} - n_{i\downarrow} \right), \tag{1.1.7}$$

where the factor $(-1)^i$ accounts for a field pointing in opposite directions on the two sublattices of a bipartite lattice. The full Hamiltonian used in some of the calculations then is

$$\mathcal{H} = \mathcal{H}_{\text{kin}} + \mathcal{H}_{\text{int}} + \mathcal{H}_{\text{chem}} + \mathcal{H}_{\text{mag}}. \tag{1.1.8}$$

1.2   Parameters, properties, and symmetries

The parameters of the full Hubbard Hamiltonian thus are t, U, μ, and h, which we will nearly always use in combination with the inverse temperature

$$\beta = \frac{1}{kT}, \tag{1.2.1}$$

where k is Boltzmann's constant. There are no restrictions on the values of these parameters, but one does not have to consider the full range of values for each parameter to study all possible situations. By simple transformations, one can map several parts of the parameter space onto other parts. A few examples are particle-hole interchange ($\bar{c}_{j\sigma} = c_{j\sigma}^{\dagger}$), which causes a change of sign of t and h and maps μ onto $U - \mu$, or spin interchange ($\bar{c}_{j\sigma} = c_{j\bar{\sigma}}$), causing only a change of sign of h. For bipartite lattices one easily finds that the sign of t is irrelevant, by changing the sign of the creation and annihilation operators on one sublattice. One can also find a relation between positive and negative values of U, under a simultaneous interchange of the roles of the chemical potential and the magnetic field (see, e.g., Ref. 4). This reveals a relation between superconductivity for $U < 0$ and antiferromagnetism for $U > 0$, and indicates how the study of magnetism in the Hubbard model for large positive

U may be of interest for understanding superconducting behavior in the case of large negative U.

From the Hamiltonian, the properties of the model can be derived via standard thermodynamic relations, starting from the grand canonical partition function

$$Z_{gr} = \text{tr } e^{-\beta \mathcal{H}}, \tag{1.2.2}$$

with $\mathcal{H}$ as given in (1.1.8). From this partition function, the basic quantity in the grand canonical ensemble can be derived: the grand potential

$$\Omega = -\frac{1}{\beta} \ln Z_{gr}. \tag{1.2.3}$$

By differentiation with respect to μ or h, relevant properties such as the particle density, the magnetization, or the susceptibility, can be derived from the grand potential.

In the limit of very large positive U, the Hubbard model is equivalent to the t–J-model for small positive J, with $J = 4t^2/U$. In the t–J-model, double occupancy of a site is forbidden, but two neighboring particles of opposite spin can be directly interchanged with energy J. The model has a configuration space which is of the order 3^N, N being the number of sites, which makes calculations of its properties much less involved than for the Hubbard model in which the possible number of configurations is 4^N. For negative J there is no equivalence between the t–J-model and the Hubbard model. Finally, at half filling, the large-U Hubbard model is equivalent to an antiferromagnetic Heisenberg spin model.

1.3   The Nagaoka state

A very interesting result, which is an important part of the motivation for this work, has been derived by Nagaoka [2]. He considered the Hubbard model on a bipartite lattice (with coordination number z), for infinite interaction strength U, and with one hole in an otherwise half-filled system. For such a system, he constructed a state with maximum total spin, and he showed that it has the lowest energy possible. One can understand this in the following way: if one has a half-filled system at infinite U, each site is occupied by one particle, and no particle can move. If one of the particles is removed, a hole is created, which has z possibilities to be interchanged with a neighboring particle. In such a system, the energy can never be below $-zt$, as one easily finds by applying the Hamiltonian to an eigenstate, and considering the site in which the wave function has its largest absolute value in this state. An eigenstate that has precisely this eigenvalue is the state

$$|\Psi_{FM}\rangle = \sum_i |\Psi_i\rangle, \tag{1.3.1}$$

where $|\Psi_i\rangle$ is the configuration in which all sites are occupied by a particle of spin up, except for site i, which is unoccupied. Finally, one can prove that, in more than one dimension, there is no other eigenstate with the lowest eigenvalue. Thus, the state with highest total spin is the unique ground state of this system.

At half filling, the Hubbard model can be mapped onto a Heisenberg antiferromagnet for large U. The result of Nagaoka, which in fact only holds for merely one point in the phase diagram, indicates the presence of ferromagnetism off half filling. A question which has become of interest to many authors is whether this ferromagnetic tendency will be stable for finite hole densities and finite interaction. This problem is far from settled yet, and we will come back to it several times, to finally discuss it again in the last paragraphs of this thesis.

1.4 Configurations and antisymmetry

An important point, which the reader should understand well before studying the following chapters, is the fact that all calculations in this thesis are performed in configuration space, which is also sometimes called the *superlattice*. One is easily confused by the difference between a hop of an electron on the lattice, and the related change of location in configuration space.

The easiest way to clarify this point is to give a simple example. Let us consider four sites connected in a loop with three particles, of which two carry spin up and one has spin down. The number of possibilities to place these particles on the sites is

$$N_c = \begin{pmatrix} 4 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 2 \end{pmatrix} \cdot 2 \qquad (1.4.1a)$$

$$= 48, \qquad (1.4.1b)$$

where the first binomial factor is the number of possibilities to choose one site for the down spin, the second factor is the number of possibilities to choose two sites for the up spins, and the factor 2 accounts for the possible number of permutations of the up spins. Thus, the configuration space belonging to this system consists of $N_c = 48$ configurations.

In Figure 1-1, three of these configurations, R_1, R_2, and R_3, are shown. Each R_i is defined by the positions of *all* particles in the system. Note that R_2 and R_3 can be obtained from each other by permutation of the up spins, and that R_1 and R_2 can be obtained from each other by a hop of one particle from a site to one of its nearest neighbors. The matrix elements of the Hamiltonian $\mathcal{H} = \mathcal{H}_{kin} + \mathcal{H}_{int}$ between these configurations are as follows:

$$\langle R_1|H|R_2 \rangle = -t, \qquad (1.4.2a)$$

$$\langle R_1|H|R_3 \rangle = \langle R_2|H|R_3 \rangle = 0, \qquad (1.4.2b)$$

$$\langle R_1|H|R_1 \rangle = U, \qquad (1.4.2c)$$

$$\langle R_2|H|R_2 \rangle = \langle R_3|H|R_3 \rangle = 0. \qquad (1.4.2d)$$

One can depict the superlattice as a set of points $\{R\}$, that are connected by bonds $\langle R, R' \rangle$ representing nonzero off-diagonal matrix elements $\langle R|H|R' \rangle = -t$. We stress again that this space has to be distinguished from the set of sites $\{i\}$ and nearest-neighbor bonds $\langle i, j \rangle$ in real space, which looks rather similar though it is of much lower dimensionality.

$$R_1 \qquad\qquad R_2 \qquad\qquad R_3$$
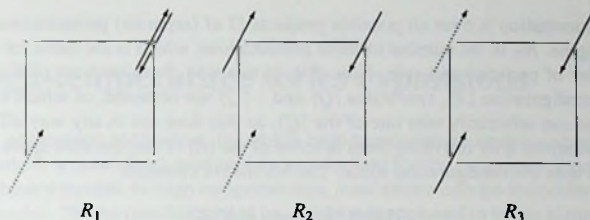
Figure 1-1. Examples of configurations for a system consisting of four sites connected in a loop, with three particles of which two carry spin up and one spin down. The up spins can be distinguished by the solid and dotted lines. Each of these configurations corresponds to one point of the superlattice.

Application of the kinetic term $\mathcal{H}_{kin}$ on a configuration R causes the electrons to hop around in the system. In configuration space, this comes down to *stepping*[1] from one point to the other along the bonds, thus performing a walk *on the superlattice*. For instance, applying one of the terms $c^\dagger c$ to the configuration R_1 makes the up spin in the upper right corner of the system hop to the upper left corner, thus resulting in the configuration R_2. Simultaneously, on the superlattice, one steps from the point R_1 to the point R_2. In this way, each electron hop is directly and uniquely related to a step in configuration space. This is used both in the series expansion method and in the Monte Carlo simulations.

Antisymmetry

For fermions, a wave function Ψ must be antisymmetric to be valid. This means that the wave function must change sign if two fermions of the same spin are interchanged by a permutation Π:

$$\langle R|\Psi\rangle = -\langle \Pi R|\Psi\rangle. \tag{1.4.3}$$

In our example, this implies $\langle R_2|\Psi\rangle = -\langle R_3|\Psi\rangle$. When solving the Schrödinger equation in the configuration space, without any restrictions, one will in general find a symmetric ground state. In order to solve the fermion problem, one has to construct antisymmetric wave functions from the eigenfunctions of the Hamiltonian. One can, however, also restrict the equations to yield antisymmetric wave functions only, thus solving only the fermion problem. This is done by considering the space of *antisymmetrized combinations* of the configurations.[2] We define the *state* $|Q\rangle$ as

$$|Q\rangle = \frac{1}{\sqrt{N_\Pi}} \sum_\Pi (-1)^{\text{sgn}(\Pi)}|\Pi R\rangle, \tag{1.4.4}$$

---

[1] In this thesis, we will always speak of *hopping* if an electron move in the system is meant, in contrast to *stepping* whenever a change of position on the superlattice takes place.

[2] In the appendix of Chapter 4, an explicit example is given in which a complete configuration space and the space of antisymmetrized combinations are depicted for an even simpler system.

where the summation is over all possible products Π of (separate) permutations of the up and down spins. N_Π is the number of these permutations, which is the same for all R with fixed number of particles and spin. Note that, in this way, from each set of N_Π permutations of a configuration $|R\rangle$, two states $|Q\rangle$ and $-|Q\rangle$ are obtained, of which only one is needed. One can arbitrarily take one of the $|Q\rangle$, as this does not in any way affect the resulting expressions after rewriting them in terms of the $|R\rangle$. Thus, the state space is a factor N_Π smaller than the configuration space. The N_Π matrix elements

$$\langle \Pi R | H | \Pi R' \rangle = \langle R | H | R' \rangle \tag{1.4.5}$$

in configuration space are all mapped onto the same matrix element

$$\langle Q | H | Q' \rangle = \pm \langle R | H | R' \rangle \tag{1.4.6}$$

in the state space, where the sign depends on the Q and Q' chosen. One can now use this Hamiltonian for the states to solve the Schrödinger equation for the fermion problem, as it automatically yields all antisymmetric solutions.

2 High-temperature series expansions

A cluster expansion technique, by which high-temperature series expansions can be calculated, is described, and results are derived for the magnetic phase diagram of the Hubbard model. At high temperatures, indications can be found for ferromagnetic behavior. These results cannot be easily extrapolated to lower temperatures.

2.1 Introduction

One of the methods that may be used to obtain information on the properties of a large system is the series expansion technique. The idea is to choose a region of the phase space where one of the parameters is small, and create an expansion in terms of that parameter. The other parameters are included exactly. By extrapolation of the results so obtained, one may hope to be able to derive information for regions that are outside the range of convergence of the series. The prescription for generating such an expansion is usually rather simple and straightforward: one needs a fairly simple enumeration algorithm, and one has to perform the calculations on relatively small systems only. The possibilities for obtaining accurate results are then limited by the amount of computer power available, as the time that is needed for calculating one more term in the expansion increases very fast with the power of the expansion parameter. Furthermore, as we will see, the interpretation of the results obtained is not at all trivial.

In Section 2.3 we will introduce the concept of cluster expansions, starting from a standard perturbation technique which is described in Section 2.2. In Sections 2.4 and 2.5, we show how the partition function and various correlation functions can be calculated for the Hubbard model, using the cluster expansion scheme. Some features of the resulting expressions are discussed in Section 2.6. We calculate nearest-neighbor correlations for the square and simple cubic lattices in Section 2.7. Finally, in Section 2.8, we calculate Curie and Néel temperatures for these lattices, and we show why the extrapolated results of the series expansions are unreliable. A way to deal with this problem will be the topic of Chapter 3.

2.2 Perturbation theory

In order to be able to calculate the partition function (1.2.2) for the Hubbard model by means of a cluster expansion, we need a standard perturbation technique to express the exponential operator

$$W(\beta) = e^{-\beta \mathcal{H}} \tag{2.2.1}$$

in terms of powers of the small parameter t. We write our standard lattice Hamiltonian (1.1.2) as

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1, \tag{2.2.2}$$

where we consider the local potential $\mathcal{H}_{\text{int}}$ to be the unperturbed part $\mathcal{H}_0$ of the Hamiltonian, and the kinetic term $\mathcal{H}_{\text{kin}}$ to be a perturbation $\mathcal{H}_1$. The latter operator is thus proportional to t.

By writing

$$\frac{\partial}{\partial \beta}\left[W(\beta)e^{\beta\mathcal{H}_0}\right] = -e^{-\beta\mathcal{H}}\mathcal{H}e^{\beta\mathcal{H}_0} + e^{-\beta\mathcal{H}}\mathcal{H}_0e^{\beta\mathcal{H}_0} \tag{2.2.3a}$$

$$= -W(\beta)\mathcal{H}_1e^{\beta\mathcal{H}_0}, \tag{2.2.3b}$$

with $W(\beta)$ from (2.2.1), and integrating, we obtain

$$W(\beta) = e^{-\beta\mathcal{H}_0} - \int_0^\beta d\beta_1 W(\beta_1)\mathcal{H}_1 e^{-(\beta-\beta_1)\mathcal{H}_0}. \tag{2.2.4}$$

Note that $\mathcal{H}_0$ and $\mathcal{H}_1$ need not commute. Iteration of this expression leads to

$$W(\beta) = \sum_{n=0}^{\infty} W_n(\beta), \tag{2.2.5}$$

where

$$W_n(\beta) = (-1)^n \int_0^\beta d\beta_1 \int_0^{\beta_1} d\beta_2 \cdots \int_0^{\beta_{n-1}} d\beta_n e^{-\beta_n\mathcal{H}_0}\mathcal{H}_1 e^{-(\beta_{n-1}-\beta_n)\mathcal{H}_0}\mathcal{H}_1 \cdots e^{-(\beta-\beta_1)\mathcal{H}_0}. \tag{2.2.6}$$

Each term in this summation can be obtained from the previous one by recursion.

The $W_n(\beta)$ can be evaluated straightforwardly in the state space $\{|Q\rangle\}$, where the Q are the antisymmetrized combinations of the configurations R as described in Section 1.4. The $|Q\rangle$ are eigenstates of $\mathcal{H}_0$, with eigenvalues

$$E_Q = \langle Q|\mathcal{H}_0|Q\rangle. \tag{2.2.7}$$

As we are interested in the partition function, we have to calculate the trace of $W(\beta)$:

$$Z_{\text{gr}} = \sum_Q \sum_n \langle Q|W_n(\beta)|Q\rangle, \tag{2.2.8}$$

Using (2.2.6) and inserting $n+1$ complete sets of states between the operators $\mathcal{H}_1$, we can express the matrix elements of $W_n(\beta)$ as a summation in the state space over paths

$$Q = \{Q = Q_0, Q_1, Q_2, \ldots, Q_n = Q'\}, \tag{2.2.9}$$

as follows:

$$\langle Q'|W_n(\beta)|Q\rangle = (-1)^n e^{-\beta E_Q} \sum_Q \langle Q'|\mathcal{H}_1|Q_{n-1}\rangle\langle Q_{n-1}|\mathcal{H}_1|Q_{n-2}\rangle \cdots \langle Q_1|\mathcal{H}_1|Q\rangle \times$$
$$\int_0^\beta d\beta_1 \cdots \int_0^{\beta_{n-1}} d\beta_n e^{\beta_n(E_{Q_{n-1}}-E_{Q'})}e^{\beta_{n-1}(E_{Q_{n-2}}-E_{Q_{n-1}})}\ldots e^{\beta_1(E_Q-E_{Q_1})}. \tag{2.2.10}$$

The matrix elements $\langle Q'|\mathcal{H}_1|Q\rangle$ equal $\pm t$ if the states $|Q'\rangle$ and $|Q\rangle$ are related by one electron hop to a neighboring site, or zero otherwise. The sign is determined by convention from a standard sign, for instance by looking at the sequence of creation operators necessary to create $|Q\rangle$ and $|Q'\rangle$ from the vacuum $|0\rangle$.

In the trace, the first and last states of each path are the same. A path Q can be embedded in configuration space as $\mathcal{R} = \{R_0, R_1, \ldots, R_n\}$; the sign for the whole path is then uniquely determined by the permutation Π necessary to restore the first configuration R_0 from the configuration $R_n = \Pi R_0$ obtained at the end. The convention for determining the sign along the path should obey this rule, and if it does, it gives the correct sign for the path. Thus, the partition function can be calculated as a power series in t, where the coefficient of t^n is obtained by performing the summation (2.2.10) over all closed paths of length n in state space. The multiple integrals are completely determined by the energy differences between the subsequent states in each path.

Let us consider this expression for a given system of N sites. Although the state space is usually very large, the eigenvalues E_Q of $\mathcal{H}_0$ can take a relatively small number of possible values, as they only depend on the number of up and down electrons in state $|Q\rangle$. For the Hamiltonian (1.1.8) with a chemical potential term and a magnetic field term included, they can be expressed as

$$E_Q = U n_d(Q) - (\mu + \frac{h}{2}) n_\uparrow(Q) - (\mu - \frac{h}{2}) n_\downarrow(Q), \qquad (2.2.11)$$

where $n_\uparrow(Q)$ and $n_\downarrow(Q)$ are the number of electrons in state $|Q\rangle$ with spin up and down, respectively, and $n_d(Q)$ is the number of doubly occupied sites. Using the obvious restrictions on the values of $n_\uparrow$, $n_\downarrow$, and n_d, it is easy to find that E_Q can take $(N+1)(N+2)(N+3)/6$ different values, which is a very small number compared to the 4^N possible states.

The operator $\mathcal{H}_1$ conserves the number of particles $(n_\uparrow + n_\downarrow)$ as well as the total spin $(n_\uparrow - n_\downarrow)/2$. Thus, we can make a division of the state space in a disjoint set of subspaces containing states with a fixed number of particles with spin up and with spin down, and perform the calculation of the trace in each subspace separately. Note that this reduces the set of values to consider for E_Q in one calculation even further. The energy difference between subsequent states in the paths only depends on the parameter U:

$$E_Q - E_{Q'} = U \left[n_d(Q) - n_d(Q') \right], \qquad (2.2.12)$$

and it can only be $\pm U$ or 0. We can use this to formally evaluate the integrals in (2.2.10). The matrix elements can be expressed as

$$\langle Q'|W_n(\beta)|Q\rangle = (\beta t)^n \sum_E \sum_{s=0}^n R_{E,s}^{n,Q,Q'} (\beta U)^{-s} e^{-\beta E}, \qquad (2.2.13)$$

where the summation over E involves all possible eigenvalues E_Q in the concerning subspace. The signs have been incorporated in the coefficients $R_{E,s}^{n,Q,Q'}$, which are rational numbers. With

$$R_{E,s}^n = \sum_Q R_{E,s}^{n,Q,Q} \qquad (2.2.14)$$

we find for the partition function

$$Z_{gr} = \sum_n Z_n (\beta t)^n, \tag{2.2.15}$$

where the coefficients are given by

$$Z_n = \sum_E \sum_{s=0}^n R_{E,s}^n (\beta U)^{-s} e^{-\beta E}. \tag{2.2.16}$$

The summation over E now involves all possible eigenvalues of $\mathcal{H}_0$. From this expression, we notice that we have in fact generated an expansion in the parameter βt. One can define an energy scale by taking the parameter t fixed at some value, and then it is obvious that the resulting power series is a high-temperature expansion.

2.3   Cluster expansion

The perturbation technique described in the previous section is generally applicable. In principle, the partition function can be calculated for lattices of any size, up to the power of t desired. However, as one may expect, the number of terms that have to be calculated increases rapidly with n. Furthermore, for large lattice size N, it becomes increasingly difficult to keep track of all terms, as generating a complete set of paths of given length in the configuration space is far from trivial.

To enumerate all terms in the expansions for the grand potential and the correlation functions correctly and efficiently, up to a given power of t, we use the cluster expansion method in the form presented by Gelfand *et al.* [5]. Following their, more general, discussion, we consider the Hamiltonian of the Hubbard model as the sum of an unperturbed part containing all local terms and a perturbation containing the hopping terms, as indicated in the previous section. We can formally write an expansion in the hopping integral t for any property $\mathcal{P}$ of our system. As we have seen, the factors t result from (repeated) multiplication by $\mathcal{H}_1$. We can keep track of the bonds from which all factors t arise by counting how often each bond $\langle i, j \rangle$ is used along the paths Q (i.e., how many matrix elements of $\mathcal{H}_1$ appear in each term, that make an electron hop from site i to site j or inversely):

$$\mathcal{P} = \sum_{\{n_{ij}\}} p\{n_{ij}\} \prod_{\langle i, j \rangle} t^{n_{ij}}. \tag{2.3.1}$$

Here n_{ij} denotes the number of factors t arising from the bond $\langle i, j \rangle$; by $\{n_{ij}\}$ we denote the set of these numbers for the whole system. The summation is over all possible $\{n_{ij}\}$, i.e., each n_{ij} runs from zero up to infinity. The coefficients $p\{n_{ij}\}$ are specific for each property.

A *cluster expansion* comes about by considering for each term the set of bonds $\{\langle i, j \rangle\}$ for which the n_{ij} are not equal to zero. We denote such a set, called *cluster*, by C. The terms in (2.3.1) can be rearranged to obtain a summation over clusters:

$$\mathcal{P} = \sum_C \sum_{\{n_{ij} \neq 0 | \langle i, j \rangle \in C\}} p\{n_{ij}\} \prod_{\langle i, j \rangle} t^{n_{ij}}. \tag{2.3.2}$$

We define the *weight* of a cluster by

$$\mathcal{W}_\mathcal{P}(C) \equiv \sum_{\{n_{ij} \neq 0 | (i,j) \in C\}} p\{n_{ij}\} \prod_{(i,j)} t^{n_{ij}}, \tag{2.3.3}$$

such that

$$\mathcal{P} = \sum_C \mathcal{W}_\mathcal{P}(C). \tag{2.3.4}$$

This summation is over *all* possible clusters C in our lattice system. Note that one of these clusters is in fact the whole system, and thus we may rewrite (2.3.4) to give the property for a (finite) system C:

$$\mathcal{P}(C) = \sum_{C' \subseteq C} \mathcal{W}_\mathcal{P}(C'). \tag{2.3.5}$$

This equation can be inverted to determine the weight of a cluster in terms of the weights of smaller (sub)clusters (the *subcluster subtraction*):

$$\mathcal{W}_\mathcal{P}(C) = \mathcal{P}(C) - \sum_{C' \subset C} \mathcal{W}_\mathcal{P}(C'). \tag{2.3.6}$$

From the definition of the weights (2.3.3) it follows that the terms in $\mathcal{W}_\mathcal{P}(C)$ contain at least as many factors t as there are bonds in C. Thus, to obtain an expansion for $\mathcal{P}$ up to mth order in t, one only has to identify and evaluate the weights of clusters containing up to m bonds. This is the most essential feature of the expansion.

Another important element in the expansion concerns the fact that we will consider physical quantities that have the *cluster property*. A property $\mathcal{P}$ is said to have the cluster property if for the union C of two *disjoint* clusters A and B the following holds:

$$\mathcal{P}(C) = \mathcal{P}(A) + \mathcal{P}(B). \tag{2.3.7}$$

This means that in the summation (2.3.2) for the cluster C, only terms in which all factors t originate from A *or* from B can be present. It then follows directly from (2.3.3) and (2.3.5) that

$$\mathcal{W}_\mathcal{P}(C) = 0 \tag{2.3.8}$$

for the *disconnected* cluster C. Thus, for properties obeying (2.3.7), we do not have to take the disconnected clusters into account, which is also an important feature of the expansion.

When one considers properties that do not depend on the geometry of a cluster, but merely on its topology, many clusters will give the same contribution to $\mathcal{P}$. Therefore we denote a set of *topologically equivalent* clusters by a *graph* G representing the set, and rewrite (2.3.4):

$$\mathcal{P} = \sum_G L(G) \mathcal{W}_\mathcal{P}(G). \tag{2.3.9}$$

The *lattice constant* $L(G)$ denotes the number of clusters, occurring in the system, that have the same topological structure as G. If the property is extensive, as for instance the grand potential, it is convenient to consider the property per lattice site, to avoid that this expression becomes arbitrarily large. In that case, $L(G)$ accordingly represents the number of clusters

per lattice site. In the appendix, a more accurate mathematical definition of a graph and its weight is given, and the precise meaning of connectedness is explained for some specific properties.

The set of graphs to be used does not only depend on the lattice considered, but also on the specific property $\mathcal{P}$. This will become clear in Section 2.5, where we will show how to calculate the nearest-neighbor correlation of the z-component of the spin. First we will discuss the calculation of the weights $\mathcal{W}_\Omega(G)$ for the grand potential.

2.4   Grand potential

The grand potential Ω has the cluster property, as is indicated in the appendix, such that we can apply the cluster expansion technique. Let us see how we can use it together with the perturbation technique, to calculate an expansion of the grand potential, for the Hubbard model on a square or simple cubic lattice. We can denote the grand potential for a finite system G in the following way [Cf. Eq. (1.2.3)]:

$$\Omega(G) = -\frac{1}{\beta} \ln Z_{\text{gr}}(G), \qquad (2.4.1)$$

where $Z_{\text{gr}}(G)$ obviously refers to the grand canonical partition function for this system. Substituting the expression (2.2.15) that we derived for the partition function in Section 2.2, and expanding the right-hand side of (2.4.1) with respect to the parameter βt, we find the following series expansion for the grand potential:

$$\Omega(G) = -\frac{1}{\beta} \ln Z_0(G) + \frac{1}{\beta} \left[\frac{Z_1(G)}{Z_0(G)}(\beta t) + \frac{Z_2(G)Z_0(G) - \frac{1}{2}Z_1(G)^2}{Z_0(G)^2}(\beta t)^2 + \cdots \right] \quad (2.4.2)$$

Thus, we have to enumerate all the appropriate graphs G for the thermodynamic system; calculate the Z_n for each of these graphs, up to the n desired, to obtain the $\Omega(G)$; use the subcluster subtraction to obtain the weights $\mathcal{W}_\Omega(G)$; and finally perform the summation (2.3.9). To accomplish this we must start with the smallest graph, and subsequently work ourselves through the larger graphs using (2.3.6).

Before we give the graphs G that are necessary for this summation, with their lattice constants $L(G)$, we mention some further restrictions on the calculations to be performed. From the fact that we have to consider closed paths in the state space to obtain (2.2.15), it directly follows that for bipartite lattices the coefficients $R_{E,s}^{n, Q, Q}$ must be identically zero for odd n. Also, more generally, the lowest order in βt in which a graph contributes to (2.3.9) is usually much higher than the number of bonds it contains, as only paths which use all bonds of the graph contribute to its weight. Many bonds have to be used twice to generate closed paths, unless they are part of a loop. Therefore, when calculating the grand potential up to order m, most of the larger graphs with up to m bonds, especially those that do not contain loops, can be omitted. We have seen before that graphs with more than m bonds do not have to be considered. The graphs that remain to be evaluated for obtaining expansions

Table 2-1. Graphs needed to calculate the grand potential Ω up to eighth order in βt. The numbers in the upper left corners have been assigned to the graphs for reference purposes. In the lower left and right corners we give the lattice constants (i.e. the number of topologically equivalent clusters per site that can be embedded on the lattice) for the square and simple cubic lattices, respectively. The lowest order in which a graph contributes is indicated in the upper right corner.

1 … 2	2 … 4	3 … 6	4 … 6	5 … 8	6 … 8
2 … 3	6 … 15	18 … 75	4 … 20	50 … 363	36 … 300
7 … 8	8 … 4	9 … 6	10 … 8	11 … 8	12 … 8
1 … 15	1 … 3	8 … 48	24 … 240	16 … 192	4 … 72
13 … 8	14 … 6	15 … 8	16 … 8	17 … 8	18 … 8
8 … 96	2 … 22	20 … 492	2 … 18	7 … 207	2 … 30

up to eighth order for the square and simple cubic lattices are presented, with their lattice constants,[1] in Table 2-1.

2.5 Correlation functions

In order to obtain a similar series expansion for the two-site correlation function C_{ij} of the z-component of the spin, we can use the same procedure as before, with some small adaptations. The correlation function is usually defined as

$$C_{ij} \equiv \left\langle (s_i^z - \langle s_i^z \rangle)(s_j^z - \langle s_j^z \rangle) \right\rangle \tag{2.5.1a}$$

$$= \langle s_i^z s_j^z \rangle - \langle s_i^z \rangle \langle s_j^z \rangle, \tag{2.5.1b}$$

where

$$s_i^z = \frac{1}{2}(n_{i\uparrow} - n_{i\downarrow}) \tag{2.5.2}$$

is the z-component of the spin operator, and the average is defined for any operator $\mathcal{A}$ by

$$\langle \mathcal{A} \rangle \equiv \frac{\mathrm{tr}\,(e^{-\beta\mathcal{H}}\mathcal{A})}{\mathrm{tr}\,e^{-\beta\mathcal{H}}}. \tag{2.5.3}$$

---

[1] Most of these graphs have been presented with lattice constants for various lattices in Ref. 6.

This correlation function has the cluster property (see the appendix).

When calculating these correlation functions, we have to deal with a somewhat different lattice than in the case of the grand potential. The two sites i and j are special. This must be reflected in the set of graphs that is needed for the calculations. We introduce a *labeled site* in a graph, which is a site for which a spin operator appears. We denote a graph G which contains one labeled site i by G_i, and similarly if it contains two labeled sites i and j we denote it by G_{ij}. In the appendix, this is explained in more detail.

Let us first consider a system G with one labeled site i. We can write its *magnetization* as

$$\mathcal{M}_i(G_i) = \langle s_i^z \rangle = \frac{\text{tr } (e^{-\beta\mathcal{H}} s_i^z)}{\text{tr } e^{-\beta\mathcal{H}}}, \tag{2.5.4}$$

or,

$$\mathcal{M}_i(G_i) = \frac{Z_{gr}(G_i)}{Z_{gr}(G)}, \tag{2.5.5}$$

where $Z_{gr}(G_i)$ is defined as

$$Z_{gr}(G_i) = \text{tr } (e^{-\beta\mathcal{H}} s_i^z). \tag{2.5.6}$$

The denominator on the right hand side of (2.5.5) is the partition function of G, which has already been calculated in the previous section. We only have to calculate the numerator, and subsequently expand this expression in powers of βt.

$Z_{gr}(G_i)$ can be calculated in a way very similar to the calculation of $Z_{gr}(G)$. As s_i^z is a diagonal matrix in state space, we have

$$\langle Q'|e^{-\beta\mathcal{H}} s_i^z|Q \rangle = \frac{1}{2} \left[n_{i\uparrow}(Q) - n_{i\downarrow}(Q) \right] \langle Q'|e^{-\beta\mathcal{H}}|Q \rangle. \tag{2.5.7}$$

From this expression, we see that the same equations (2.2.8) and (2.2.10) we used to calculate $Z_{gr}(G)$ can be used to calculate $Z_{gr}(G_i)$. We merely have to add a factor $-1/2$, 0, or $+1/2$ to each of the matrix elements $\langle Q'|W(\beta)|Q \rangle$ in the summation (2.2.8), depending only on the occupation of site i in state $|Q\rangle$. In this way we can calculate the magnetization $\mathcal{M}_i(G_i)$ for each graph that contains the site i. Note that for a finite system G, $\mathcal{M}_i(G_i)$ generally depends on which site in G is site i. Therefore the site has to be added to the graph as a relevant index. In Table 2-2 on page 37 we list all graphs with one labeled site, necessary to calculate the magnetization for the square or simple cubic lattice. One can perform this calculation via the subcluster subtraction (2.3.6) and the weight summation (2.3.9) for the property $\mathcal{M}_i$, using the set of graphs G_i.

The next step is to calculate the average of the product of two spin operators

$$C'_{ij}(G_{ij}) = \langle s_i^z s_j^z \rangle. \tag{2.5.8}$$

For this we have to consider

$$Z_{gr}(G_{ij}) = \text{tr } (e^{-\beta\mathcal{H}} s_i^z s_j^z). \tag{2.5.9}$$

Again, it is straightforward to find the matrix elements

$$\langle Q'|e^{-\beta\mathcal{H}} s_i^z s_j^z|Q \rangle = \frac{1}{4} \left[n_{i\uparrow}(Q) - n_{i\downarrow}(Q) \right] \left[n_{j\uparrow}(Q) - n_{j\downarrow}(Q) \right] \langle Q'|e^{-\beta\mathcal{H}}|Q \rangle \tag{2.5.10}$$

as well as the graphs G_{ij} for which the C'_{ij} are to be calculated. For the nearest-neighbor correlations these graphs are listed in Tables 2-3 and 2-4 (pp. 38, 39). A distinction has to be made between graphs consisting of one part (Table 2-3) and graphs consisting of two parts (Table 2-4). Graphs consisting of two parts are *not* disconnected in this case, if each part contains a labeled site. This slightly different form of the cluster property is explained in the appendix.

There are two possible ways to finally calculate the correlation functions C_{ij} for the system, starting from the $\mathcal{M}_i(G_i)$ and $C'_{ij}(G_{ij})$ discussed above. One way is to calculate the property

$$C_{ij}(G_{ij}) = C'_{ij}(G_{ij}) - \mathcal{M}_i(G_i)\mathcal{M}_j(G_j) \tag{2.5.11}$$

for each graph G with labeled sites i and j. By using the subcluster subtraction (2.3.6), the $\mathcal{W}_C(G_{ij})$ can then be calculated, and subsequently C_{ij} for the thermodynamic system by using the weight summation (2.3.9) with the lattice constants given in Table 2-3. We note that in this case the weights of the graphs in Table 2-4 are identically zero, unlike the weights for the property C'_{ij} (see the appendix). Alternatively, one can calculate the magnetization $\mathcal{M}$ and the correlation function C'_{ij} for the thermodynamic system by generating all weights $\mathcal{W}_{\mathcal{M}}(G_i)$ and $\mathcal{W}_{C'}(G_{ij})$ and summing those, using the respective sets of graphs and lattice constants. We can then use

$$C_{ij} = C'_{ij} - \mathcal{M}^2 \tag{2.5.12}$$

to find C_{ij} for the thermodynamic system. As both ways must obviously yield the same result, we have an independent check on the correctness of the lattice constants and the weights.

2.6   Series expansions and derivatives

By the method described in the previous sections we have generated the weights with respect to the grand potential up to eighth order in βt, for all graphs in Table 2-1. Using (2.3.9), we have subsequently calculated the expansion for the grand potential per lattice site Ω on the square and simple cubic lattices. The resulting expression can be written as

$$\beta\Omega = \sum_n (\beta t)^n (Z_{\mathrm{gr},0})^{-n} \sum_E \sum_{s=0}^n \Omega^n_{E,s} (\beta U)^{-s} e^{-\beta E}, \tag{2.6.1}$$

where

$$Z_{\mathrm{gr},0} = 1 + e^{\beta(\mu+h/2)} + e^{\beta(\mu-h/2)} + e^{2\beta\mu - \beta U}. \tag{2.6.2}$$

The coefficients $\Omega^n_{E,s}$ are rational numbers. They are related to the coefficients $R^n_{E,s}$, evaluated for all different graphs, in the expressions (2.2.15) and (2.2.16) for the partition function. The summation over E again involves a limited set of values, similar to (2.2.11):

$$E = aU - b\mu \pm ch/2, \tag{2.6.3}$$

were a, b and c are nonnegative integers. In fact, if we write

$$(Z_{\mathrm{gr},0})^n = \sum_E n_E e^{-\beta E} \tag{2.6.4}$$

then the summations over E in (2.6.1) and (2.6.4) involve the same set of values, although some of the coefficients for these E in (2.6.1) are zero.

In a similar way, we calculate the magnetization and the correlation functions. The resulting expression for a correlation function C_{ij} looks very much like the one for the grand potential:

$$C_{ij} = \sum_n (\beta t)^n (Z_{gr,0})^{-(n+2)} \sum_E \sum_{s=0}^n C_{E,s}^n (\beta U)^{-s} e^{-\beta E}. \tag{2.6.5}$$

Again, the coefficients are rational numbers.

In Tables 2-5 and 2-6 (pp. 40, 41), we present the coefficients for the grand potential Ω and for the nearest-neighbor correlation function C_{ij} up to sixth order for the square lattice. For simplicity, we have taken h equal to zero [thus we can omit c in Eq. (2.6.3)]. It would take several hundreds of pages to present the eighth-order and $h \neq 0$ coefficients in a similar fashion.[2]

In methods like these, where a very large number of numbers is obtained that can impossibly be verified by physical intuition, it is important to check the results by any means available. Some results for lower orders (2nd, 4th, 6th) or limiting cases (infinite U) can be found in the literature [7–10]. In the case of infinite U, our coefficients have been compared by Putikka [11] up to 8th order to coefficients for the t–J-model, which should be the same for $J = 0$. The full series has been checked to agree with the results of Henderson *et al.* [12, 13], who performed basically the same calculation, but did not generate the coefficients in integer form. It is also important to realize that a severe internal check exists in the cluster method: as larger graphs should not contribute in lower orders, the graphs that do contribute in lower orders can be checked by means of the subcluster subtraction. Any error in the weight of a smaller graph ruins the result for the larger graphs of which it is a subgraph. In this way all graphs contributing to sixth order are automatically checked by the eighth order calculations. To check the eighth order results we have additionally calculated, up to eighth order, the weights of some larger graphs which should only contribute from tenth or higher order, thus checking all subgraphs they contain. Only the largest eighth-order graphs could not be checked in this way, because of computer time limitations. Finally, the two methods for calculating correlation functions, as described at the end of the previous section, must yield the same result, thus also providing a check on the expansions. Having applied all these checks, we are confident of the correctness of all results presented here.

Calculating the series expansions to eighth order, as described above, takes a few weeks of CPU on a simple workstation (Sun SPARC station 1), while the sixth order can be obtained in merely a few hours. When considering extending these expansions to tenth order, one has to expect an increase in computer time by a multiplicative factor which is still much larger than between sixth and eighth order. Also, one may run into storage problems, and it may be difficult to maintain the integer form of the coefficients, since the number of possible paths also grows tremendously. As we have already invested a fair amount of time and inventiveness in making the programs more efficient, we expect that it is not possible to significantly speed up the calculations by smart tricks. For our purposes, it seemed not to be

---

[2] All these coefficients have been calculated, and they are available on request.

worthwhile to try to perform the calculation of the tenth-order coefficients on a workstation, and even on a supercomputer it would currently take too much time. Therefore, we consider the eighth-order expansion as the practical limit to what is possible at the moment.

As we want to consider various quantities as a function of the particle density n (i.e., the average number of particles per site), using

$$n = -\frac{\partial(\beta\Omega)}{\partial(\beta\mu)}, \tag{2.6.6}$$

the density is calculated by differentiating each term in the expansion for Ω separately. In a similar way one can calculate other quantities, e.g. the uniform (or ferromagnetic) susceptibility

$$\chi_{FM} = \beta\frac{\partial^2(\beta\Omega)}{\partial(\beta h)^2}\bigg|_{h=0}. \tag{2.6.7}$$

Note that the uniform susceptibility can also be obtained directly from the series for the correlation functions, as

$$\chi_{FM} = S(\mathbf{q} = \mathbf{0}), \tag{2.6.8}$$

where the *structure factor* $S(\mathbf{q})$ is defined as

$$S(\mathbf{q}) = \frac{1}{N}\sum_{ij}e^{i\mathbf{q}\cdot\mathbf{r}_{ij}}C_{ij}. \tag{2.6.9}$$

Another interesting quantity, when considering magnetic behavior, is the staggered (or antiferromagnetic) susceptibility

$$\chi_{AF} = \beta\frac{\partial^2(\beta\Omega)}{\partial(\beta h_s)^2}\bigg|_{h_s=0}, \tag{2.6.10}$$

with h_s now being a staggered magnetic field included in the Hamiltonian. This quantity cannot easily be expressed in terms of correlation functions. This is due to the fact that the z-component of the staggered total spin $\sum_i(-1)^i s_i^z$ is not a conserved quantity because this operator does not commute with the Hamiltonian, whereas the total spin is. The staggered susceptibility can still be calculated by means of the cluster expansion, by calculating the partition function explicitly in terms of the staggered field, and differentiating it twice with respect to this field. However, in order to do this, one has to include extra terms in the path integrals, as the energy difference along a path in configuration space [Cf. Eq. (2.2.12)] can now be

$$\Delta E = n_U U + 2n_h h_s, \tag{2.6.11}$$

where n_U can be 0 or ± 1, and n_h can be ± 1 only. We have included this term in the Hamiltonian, and calculated the staggered magnetization and susceptibility also to eighth order for the square and simple cubic lattices.

Let us examine some features of the series expansions more closely. They are given in terms of the variables βU, $\beta\mu$ and βh, and the perturbation parameter βt. For simplicity

we restrict ourselves to $h = 0$ and $U > 0$ here. From the fact that factors U appear in the denominator of certain terms in the expansions one could infer that they cannot be used for the case of small U, because of divergence problems. This is however not true; we can use any value for U, with approximately the same range of convergence in βt. This is due to the fact that all factors U in the denominator cancel against factors in the numerator, as can be seen by expanding the exponentials for small βU. By directly expanding the operator $e^{-\beta \mathcal{H}}$ in a power series:

$$e^{-\beta \mathcal{H}} = 1 - \beta \mathcal{H} + \frac{1}{2}(\beta \mathcal{H})^2 + \cdots \tag{2.6.12}$$

one can immediately see that no negative powers of U can occur for small U. This once more provides a check on the coefficients in the expansions.

We are most interested in the case of large U, in view of the possible appearance of a ferromagnetic state. We see that βU appears in the exponentials with negative sign only. If βU is large enough, we can neglect all terms in the expansion which contain an exponential of βU, i.e., for which the parameter a in (2.6.3) is nonzero. For still larger U (i.e., U goes to infinity) we can also neglect all of the remaining terms which contain a factor U in the denominator. This means that we only have to consider terms in (2.6.1) and (2.6.5) for which $s = 0$, or, on the level of the path summation (2.2.10), that only paths which do not contain states with doubly occupied sites, contribute. This clearly reflects the fact that this is the limit of strong repulsion.

2.7 Nearest-neighbor correlations

In order to get a good indication of the convergence of the series expansions, we consider subsequent orders of approximation for various functions. As an example we plot some of these functions in Figures 2-1 and 2-2. In Figure 2-1, we show approximations of the density as a function of $\beta \mu$ to subsequent orders, at $\beta U = 1500$, $h = 0$ and $\beta t = 0.8$, for the square lattice. Clearly for this value of βt the convergence is very good, as the difference between subsequent approximations decreases and the difference between the sixth and the eighth order approximations is only about 5% or less. For $\beta t = 0.5$ (not shown here), the difference between the subsequent approximations is much smaller, as expected. Thus, we conclude that the series for the density can be used in the region where, roughly, $\beta t < 1$. We have checked this for various values of βU, with similar results. In Figure 2-1, we only present the part of the range in $\beta \mu$ where $n < 1$. From the condition $\mu = U/2$ at half filling, we know that $n = 1$ at $\beta \mu = 750$. Because of particle-hole symmetry (see Section 1.2) the density is symmetric about $\mu = U/2$. Thus we see that the density varies very rapidly as a function of the chemical potential only in the very narrow regions $-8 < \beta \mu < +8$ and $\beta U - 8 < \beta \mu < \beta U + 8$, and that we practically have half filling in the whole range of $\beta \mu$ between these regions. Moreo *et al.* [14] find similar results for a finite (4×4) cluster using a quantum Monte Carlo method, although for much lower temperatures ($\beta t = 4$ and 8) and smaller βU.

In Figure 2-2, we show in a similar way the correlation function C_{ij} for nearest-neighbor sites i and j, as a function of the density, again in zero field and with $\beta U = 1500$ and $\beta t = 0.5$. This function is obtained by calculating both C_{ij} and n as a function of the chemical

potential, and plotting it as a one-parameter curve in the n–C_{ij}-plane. For $\beta t < 0.5$ the convergence is very good; the difference between the sixth and the eighth order approximations is already very small. At $\beta t \gtrsim 0.5$ convergence becomes much slower, and it is destroyed completely for $\beta t \approx 1$. Thus convergence is limited to smaller values of βt for this case than for the density. We have also looked at the simple cubic lattice, where we find that already at $\beta t \approx 0.4$ the difference between subsequent approximations becomes significant.

At half filling and for large U, the model can be mapped onto a $s = 1/2$ Heisenberg antiferromagnet (see, e.g., Ref. 4), which is generally believed to have an antiferromagnetically ordered ground state (see, e.g., Ref. 15). Nagaoka [2] has shown that in the limit of infinitely strong coupling the ground state of a system doped with one hole is ferromagnetically ordered. In the region where our series expansion exhibits good convergence we have examined the nearest-neighbor correlation function. We expect to find a region in the phase diagram where these correlations are ferromagnetic, indicating the existence of a Nagaoka-like phase. In order to find an indication for the onset to this behavior, we have explored the correlation function for high temperatures and strong interaction ($\beta U \gg \beta t$). In Figure 2-3 we show it as a function of the density, for various values of βU, at a fixed temperature. It can be seen that for densities above roughly $2/3$, for the square lattice as well as for the simple cubic lattice, one always finds ferromagnetic correlations when making U sufficiently large. At half filling, the correlations are antiferromagnetic for all values of U. This is consistent with the results due to the mapping on the Heisenberg model.
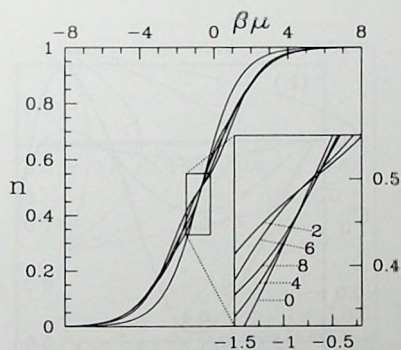

Figure 2-1. Approximations up to 0th, 2nd, 4th, 6th and 8th order of the density as a function of the chemical potential. Square lattice: $\beta t = 0.8$; $\beta U = 1500$; $h = 0$.
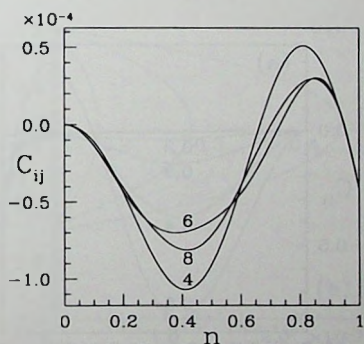
Figure 2-2. Approximations up to 4th, 6th, and 8th order of the nearest-neighbor correlation function as a function of the density. Square lattice: $\beta t = 0.5$; $\beta U = 1500$; $h = 0$.

Figure 2-3. Nearest-neighbor correlation function as a function of the density, at fixed βt for various values of βU (8th order approximation). (a) Square lattice: $\beta t = 0.5$; $\beta U = 600, 700, 800, 1000, 1500, \infty$; (b) Simple cubic lattice: $\beta t = 0.4$; $\beta U = 500, 600, 700, 800, 1000, 1500, \infty$.
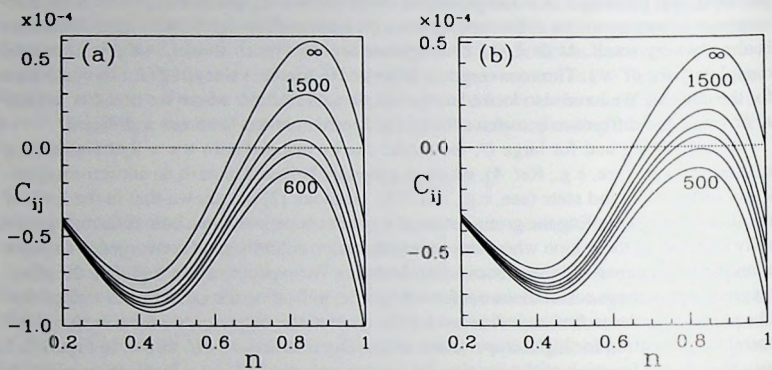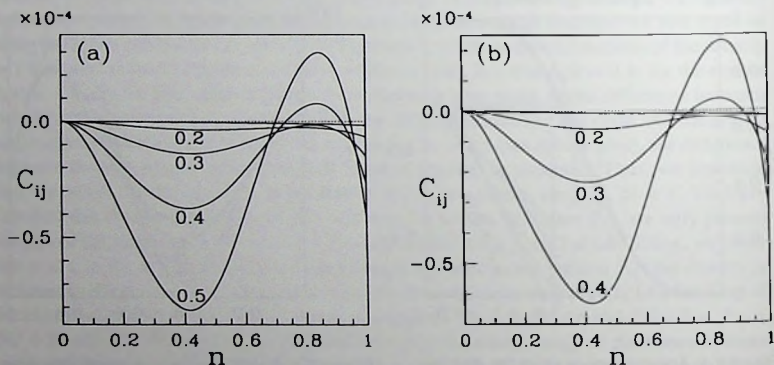

Figure 2-4. Nearest-neighbor correlation function as a function of the density, at fixed βU for various βt (8th order approximation). (a) Square lattice: $\beta U = 1500$; $\beta t = 0.1, 0.2, 0.3, 0.4, 0.5$; (b) Simple cubic lattice: $\beta U = 1250$; $\beta t = 0.1, 0.2, 0.3, 0.4$.

In Figure 2-4, we show the nearest-neighbor correlation at a constant interaction U, for some different values of the temperature. At very high temperature the correlation is antiferromagnetic for all values of the density, and increasing with decreasing temperature. In the region where n is roughly between .7 and .95, the correlations become increasingly ferromagnetic below some temperature. The correlation between nearest-neighbor sites thus changes from antiferromagnetic at high temperatures to ferromagnetic at lower temperatures.

In Figure 2-5 the region in the phase diagram where the nearest-neighbor correlation is ferromagnetic is shown for some values of the temperature. On the horizontal axis we plot $4t/U$; in the limit of large U and near half filling this equals the parameter J/t from the $t-J$-model which is then equivalent to the Hubbard model. We see that the region of ferromagnetic correlation increases very rapidly with decreasing temperature, but that the range of densities for which it occurs remains more or less the same. The density above which the ferromagnetic correlations occur is $n \approx 2/3$. At this density magnetic correlation between the nearest-neighbor sites seems to be absent. Below this density we do not expect a ferromagnetic phase, as the nearest-neighbor correlations would surely be ferromagnetic in such a phase.

The results for the simple cubic lattice are quite similar to those for the square lattice. All correlations are about 50% stronger in the three-dimensional case, which reflects the standard dimensional influence through the larger number of neighbors.
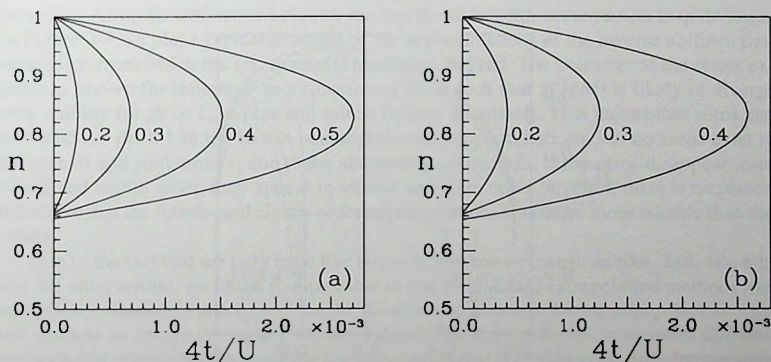

Figure 2-5. Phase diagram of the Hubbard model in the high-temperature regime. The curves enclose areas of ferromagnetic nearest-neighbor correlations (i.e., $C_{ij} > 0$) at different temperatures. (a) Square lattice: $\beta t = 0.2, 0.3, 0.4, 0.5$; (b) Simple cubic lattice: $\beta t = 0.2, 0.3, 0.4$.

2.8   Curie temperatures and convergence problems

Another interesting quantity to consider when looking for ferromagnetic behavior is the
Curie temperature. One expects a divergence in the uniform susceptibility at the transition
from a paramagnetic to a ferromagnetic phase, and so the Curie temperature T_C is defined
as the temperature at which the inverse susceptibility vanishes:

$$\chi_{PM}^{-1}(n, U, T_C) \equiv 0. \tag{2.8.1}$$

Similarly, one defines the *Néel temperature* through

$$\chi_{AF}^{-1}(n, U, T_N) \equiv 0, \tag{2.8.2}$$

indicating a possible antiferromagnetic transition. We have constructed both susceptibili-
ties from our series expansions, and we have calculated the Curie and Néel temperatures in
the eighth order approximation. In Figure 2-6 we show them as a function of the particle
density, for various values of t/U. This is an extension of the results presented by Pan and
Wang [10], who have performed the same calculations for fourth order, and Henderson *et
al.* [12], who calculated only the Curie temperature. Qualitatively, our results are very sim-
ilar to theirs. For densities above $n = .76$ a paramagnetic-ferromagnetic phase transition
is expected as a nonzero Curie temperature is found. However, as the Néel temperature is
higher than the Curie temperature in this region of the phase space, one should conclude
that the system becomes antiferromagnetic before the ferromagnetic transition is reached,
and thus one does not have clear evidence that the ferromagnetic phase exists.
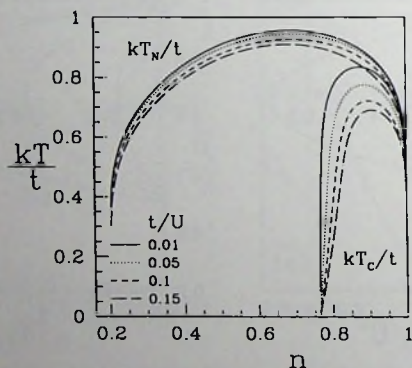

Figure 2-6. Néel and Curie temperatures, as a function of the particle density, for the Hubbard model
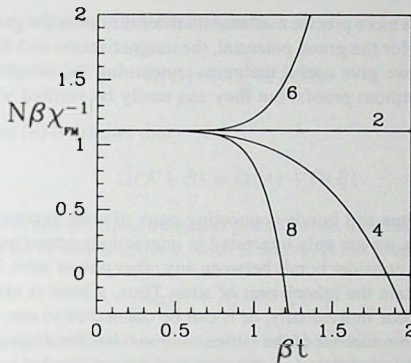on a simple cubic lattice, at constant t/U.

Figure 2-7. The inverse uniform susceptibility as a function of the parameter βt, for the Hubbard model on a simple cubic lattice, at infinite U and particle density $n = 0.9$. Approximations up to order 2, 4, 6, 8 in βt.

There are, however, several reasons why one cannot trust these results. For instance, if one considers the second- or sixth-order approximations, one does not find solutions of (2.8.1) at all, or one finds them in another range of the density ($n < .75$ for the Curie temperature). Also, the difference between the fourth- and eighth-order results is quite large. In Figure 2-7 we plot a typical example of the approximations of the inverse uniform susceptibility, from which the above results have been derived. The character of the series expansion, shown for infinite U as a function of βt, is such that $\chi_{PM}^{-1}(\beta t)$ is likely to diverge very quickly for $\beta t > 1$, to plus and minus infinity alternately. This means that zeros can be found for $\beta t \gtrsim 1$ in the fourth- and eighth-order approximations, but no zeros exist in the second and sixth orders; due to the alternating coefficients, if the zeros disappear from fourth and eighth order, they appear in second and sixth order. A priori, there is no reason to believe that the fourth- and eighth-order approximations should be more reliable than the others.

Due to the fact that we only have five terms in the series expansions (0th, 2nd, 4th, 6th, and 8th order terms), we found it impossible to rely on standard extrapolation methods like analysis by means of Padé approximants. Results obtained in different extrapolations vary too much to be able to derive any reliable values. Therefore, we have to consider different methods that may help us in revealing the information that is hidden in the high-temperature expansions. One such method will be the topic of Chapter 3.

Appendix: Graphs and weights

In this appendix we give a more precise mathematical definition for the graphs used to calculate the series expansion for the grand potential, the magnetization and the spin-correlation functions. Furthermore, we give useful theorems concerning the weights of these graphs. These theorems come without proofs, but they can easily be verified with the definitions given here.

Graphs

A graph G consists of sites and bonds connecting pairs of sites, representing the interaction between the sites. As we are only interested in interactions occurring between nearest-neighbor sites, we do not consider bonds between any other pair of sites. If a graph contains a bond, it must also contain the related pair of sites. Thus, a bond is always *connected* to two sites. A site can appear individually, or it can be connected to one, two, ..., z bonds, where z is the coordination number of the lattice we consider. We distinguish between two different kinds of sites: labeled sites ($\circ$) and sites that are not labeled ($\cdot$). A bond ($-$) is denoted by the pair of nearest-neighbor sites $\langle i, j \rangle$ it connects.

A *subgraph* $G' \subset G$ is obtained by omitting one or more bonds from G. An *empty* graph contains no bonds and only sites that are not labeled.

With a Hamiltonian of the form

$$\mathcal{H} = \sum_i \mathcal{H}_i^{\text{loc}} + \sum_{\langle i, j \rangle} \mathcal{H}_{i,j}^{\text{hop}} \tag{1}$$

we define the graph (or cluster) Hamiltonian

$$\mathcal{H}_G = \sum_{i \in G} \mathcal{H}_i^{\text{loc}} + \sum_{\langle i, j \rangle \in G} \mathcal{H}_{i,j}^{\text{hop}}, \tag{2}$$

and the following partition functions:

$$Z(G) = \text{tr } e^{-\beta \mathcal{H}_G}, \tag{3}$$

$$Z(G_i) = \text{tr } e^{-\beta \mathcal{H}_G} s_i^z, \tag{4}$$

$$Z(G_{ij}) = \text{tr } e^{-\beta \mathcal{H}_G} s_i^z s_j^z. \tag{5}$$

We use the indices with G to indicate which sites in G are labeled. For these quantities the following statements hold: If a graph G is the union of disjoint graphs A and B, then

$$Z(G) = Z(A + B) = Z(A) \cdot Z(B), \tag{6}$$

$$Z(G_i) = Z(A_i + B) = Z(A_i) \cdot Z(B), \tag{7}$$

$$Z(G_{ij}) = \begin{cases} Z(A_{ij} + B) = Z(A_{ij}) \cdot Z(B), \text{ or} \\ Z(A_i + B_j) = Z(A_i) \cdot Z(B_j). \end{cases} \tag{8}$$

Here we have chosen the labeled sites to be in some specific subgraph of G; the other possibilities are equivalent with one of these choices.

Grand potential

The grand potential is given by the expression

$$\Omega(G) = -\frac{1}{\beta} \ln Z(G). \tag{9}$$

From the statement (6) it follows that

$$\Omega(A + B) = \Omega(A) + \Omega(B) \tag{10}$$

for disjoint graphs A and B. Thus, the grand potential has the cluster property. From this, it can be shown that for A and B disjoint, with B empty, A not empty:

$$\mathcal{W}_\Omega(A + B) = \mathcal{W}_\Omega(A). \tag{11}$$

Here we see why only the sites that are directly involved in the interactions have to be included in the cluster Hamiltonian. By induction it is straightforward to prove the *linked cluster theorem* for the grand potential: If A and B are disjoint graphs, both not empty, then

$$\mathcal{W}_\Omega(A + B) = 0. \tag{12}$$

This leads to an important reduction of the set of graphs: all *disconnected* graphs, i.e., graphs consisting of two or more disjoint nonempty subgraphs, do not have to be considered in the calculation of the grand potential, as their weights are identically zero.

Magnetization

In a similar way, we can deduce a linked cluster theorem for the magnetization

$$\mathcal{M}_i(G_i) = \frac{Z(G_i)}{Z(G)}. \tag{13}$$

We have a somewhat different form of the cluster property here: for disjoint A and B,

$$\mathcal{M}_i(A_i + B) = \mathcal{M}_i(A_i). \tag{14}$$

Thus, we can simply disregard any part of the graph that is not connected to the part containing the labeled site i, when calculating the magnetization. For the weights, we can state as above: for disjoint A_i and B, with B empty, we have

$$\mathcal{W}_{\mathcal{M}}(A_i + B) = \mathcal{W}_{\mathcal{M}}(A_i). \tag{15}$$

With B nonempty, we have again the linked cluster theorem

$$\mathcal{W}_{\mathcal{M}}(A_i + B) = 0. \tag{16}$$

Correlation functions

For the correlation functions

$$C'_{ij}(G_{ij}) = \frac{Z(G_{ij})}{Z(G)} \qquad (17)$$

and

$$C_{ij}(G_{ij}) = C'_{ij}(G_{ij}) - \mathcal{M}_i(G_i)\mathcal{M}_j(G_j) \qquad (18)$$

we have

$$C'_{ij}(A_{ij} + B) = C'_{ij}(A_{ij}), \qquad (19)$$

$$C'_{ij}(A_i + B_j) = \mathcal{M}_i(A_i) \cdot \mathcal{M}_j(B_j), \qquad (20)$$

$$C_{ij}(A_{ij} + B) = C_{ij}(A_{ij}), \qquad (21)$$

$$C_{ij}(A_i + B_j) = 0 \qquad (22)$$

for disjoint A and B. Again, if B is empty, we find

$$\mathcal{W}_{C'}(A_{ij} + B) = \mathcal{W}_{C'}(A_{ij}) \qquad (23)$$

and

$$\mathcal{W}_C(A_{ij} + B) = \mathcal{W}_C(A_{ij}). \qquad (24)$$

The linked cluster theorem is different for the two correlation functions. For C' we have

$$\mathcal{W}_{C'}(A_{ij} + B) = 0, \qquad (25)$$

if A and B are disjoint and not empty. When the labeled sites i and j are not both in A or B, we find

$$\mathcal{W}_{C'}(A_i + B_j) = \mathcal{W}_{\mathcal{M}}(A_i) \cdot \mathcal{W}_{\mathcal{M}}(B_j). \qquad (26)$$

Thus we see that, for this case, the meaning of connectedness is different from what we have seen before. When a graph consists of two disjoint subgraphs, both containing one labeled site, its weight for this correlation function is the product of the weights for the magnetization of both subgraphs. Of course, when a graph consists of three or more disjoint subgraphs, its weight must be zero again. For C, the theorem is as before:

$$\mathcal{W}_C(A_{ij} + B) = 0 \qquad (27)$$

and

$$\mathcal{W}_C(A_i + B_j) = 0 \qquad (28)$$

for disjoint, nonempty A and B.

With the theorems presented here, the subcluster subtractions and the weight summations for the properties concerned can be evaluated.

Table 2-2. Graphs with one labeled site used to calculate the magnetization $\mathcal{M}$ up to eighth order in βt. The numbers in the lower left and right corners are the lattice constants for the square and simple cubic lattices, respectively. They follow by multiplication with a symmetry factor (independent of the dimension of the lattice) from the lattice constants for the corresponding graphs in Table 2-1 on page 23.

1	2	3	4	5	6	7
4 6	12 30	6 15	36 150	36 150	12 60	4 20
8	**9**	**10**	**11**	**12**	**13**	**14**
100 726	100 726	50 363	72 600	36 300	36 300	36 300
15	**16**	**17**	**18**	**19**	**20**	**21**
4 60	1 15	4 12	8 48	8 48	16 96	8 48
22	**23**	**24**	**25**	**26**	**27**	**28**
24 240	24 240	24 240	48 480	24 240	32 384	32 384
29	**30**	**31**	**32**	**33**	**34**	**35**
32 384	8 144	4 72	8 144	4 72	16 192	16 192
36	**37**	**38**	**39**	**40**	**41**	**42**
16 192	12 132	20 492	20 492	40 984	40 984	20 492
43	**44**	**45**	**46**	**47**	**48**	
8 72	4 36	56 1656	4 60	8 120	2 30	

Table 2-3. Graphs with two labeled sites, consisting of one part, needed to calculate the correlation functions C_{ij} and C'_{ij} for nearest neighbors i and j, up to eighth order in βt. The numbers in the lower left and right corners are the lattice constants for the square and simple cubic lattices, respectively. Most lattice constants follow by multiplication with a symmetry factor (depending on the dimension of the lattice) from the lattice constants in Table 2-1. Only if the two labeled sites are not directly connected, the lattice constant can not be found by a symmetry argument. Graph no. 40 does not contribute to the nearest-neighbor correlation function on the square lattice, as its labeled sites cannot be neighbors on that lattice.

1		2		3		4		5		6		7	
1	1	6	10	18	50	9	25	2	4	6	20	50	242
8		**9**		**10**		**11**		**12**		**13**		**14**	
50	242	8	32	18	100	18	100	36	200	8	32	2	20
15		**16**		**17**		**18**		**19**		**20**		**21**	
2	4	8	32	8	32	4	16	24	160	24	160	12	80
22		**23**		**24**		**25**		**26**		**27**		**28**	
12	80	4	24	8	64	16	128	8	64	16	128	2	12
29		**30**		**31**		**32**		**33**		**34**		**35**	
4	48	4	48	4	48	8	64	16	128	6	44	1	6
36		**37**		**38**		**39**		**40**		**41**		**42**	
10	164	20	328	20	328	20	328	0	32	2	36	8	96
43		**44**		**45**		**46**		**47**		**48**		**49**	
4	24	2	12	1	6	28	552	6	108	4	40	4	40

Table 2-4. Two-label graphs, consisting of two parts, necessary to calculate the correlation function C_{ij} for nearest-neighbor sites i and j, up to eighth order in βt, with their lattice constants $L(G)$ for the square (sq) and simple cubic (sc) lattices. Each part contains one labeled site out of a pair of labeled nearest-neighbor sites. The parts are indicated by numbers referring to the graphs in Table 2-2. By 0, one single labeled site is indicated. The order to which a graph contributes is the sum of the orders to which its parts contribute. Graphs 16, 31, and 48 can not be embedded on the square lattice. Some examples are ⌐⌐ (graph no. 20) and ⌐⌐ (graph no. 53).

G no.	parts		$L(G)$ sq	$L(G)$ sc	G no.	parts		$L(G)$ sq	$L(G)$ sc	G no.	parts		$L(G)$ sq	$L(G)$ sc
1	0	1	6	10	25	0	25	44	616	49	1	1	9	25
2	0	2	18	50	26	0	26	24	320	50	1	2	50	242
3	0	3	6	20	27	0	27	44	616	51	1	3	18	100
4	0	4	50	242	28	0	28	16	384	52	1	4	142	1178
5	0	5	36	200	29	0	29	32	512	53	1	5	100	968
6	0	6	18	100	30	0	30	12	240	54	1	6	46	468
7	0	7	2	20	31	0	31	0	48	55	1	7	6	100
8	0	8	142	1178	32	0	32	8	192	56	1	17	12	80
9	0	9	92	936	33	0	33	4	96	57	1	18	32	376
10	0	10	50	484	34	0	34	24	320	58	1	19	12	240
11	0	11	100	968	35	0	35	8	192	59	1	20	44	616
12	0	12	18	300	36	0	36	16	256	60	1	21	24	320
13	0	13	46	468	37	0	37	10	164	61	1	37	30	788
14	0	14	36	400	38	0	38	30	788	62	2	2	71	589
15	0	15	6	100	39	0	39	8	456	63	2	3	46	468
16	0	16	0	10	40	0	40	32	1216	64	2	17	32	376
17	0	17	4	16	41	0	41	32	1184	65	3	3	9	100
18	0	18	12	80	42	0	42	18	620	66	3	17	12	160
19	0	19	4	48	43	0	43	8	96	67	17	17	4	64
20	0	20	16	128	44	0	44	2	36					
21	0	21	8	64	45	0	45	44	1992					
22	0	22	32	376	46	0	46	4	80					
23	0	23	24	320	47	0	47	8	160					
24	0	24	12	240	48	0	48	0	20					

Table 2-5. Coefficients $\Omega^n_{E,s}$ in the expansion of the grand potential Ω for the square lattice (see Eq. 2.6.1). $E = aU - b\mu$; the magnetic field h has been put to zero. Note that all $\Omega^n_{E,n}$ are identically zero, due to a coincidental symmetry.

b	a	s	$\Omega^2_{E,s}$
1	0	0	4
2	0	1	8
2	1	1	-8
3	1	0	4

b	a	s	$\Omega^4_{E,s}$
1	0	0	3
2	0	3	8
2	0	2	-48
2	0	1	44
2	0	0	-64/3
2	1	3	-8
2	1	2	40
3	0	3	-48
3	0	2	24
3	0	1	-48
3	0	0	10
3	1	3	48
3	1	2	24
3	1	1	48
3	1	0	7
4	0	3	-48
4	0	2	48
4	1	3	-16
4	1	2	96
4	1	1	-88
4	1	0	-112/3
4	2	3	64
4	2	2	-32
5	1	3	-48
5	1	2	24
5	1	1	-48
5	1	0	10
5	2	3	48
5	2	2	24
5	2	1	48
5	2	0	7
6	2	3	8
6	2	2	-48
6	2	1	44
6	2	0	-64/3
6	3	3	-8
6	3	2	40
7	3	0	3

b	a	s	$\Omega^6_{E,s}$	b	a	s	$\Omega^6_{E,s}$	b	a	s	$\Omega^6_{E,s}$
1	0	0	10/9	5	1	5	-992	7	3	5	-11040
2	0	5	-128	5	1	4	-11960	7	3	4	-4792
2	0	3	640/3	5	1	3	3664/3	7	3	3	-1208
2	0	2	-192	5	1	2	-260	7	3	2	52
2	0	1	272/3	5	1	1	718	7	3	1	208/3
2	0	0	-2312/45	5	1	0	8462/15	7	3	0	308/45
2	1	5	128	5	2	5	-11040	8	2	5	8016
2	1	4	128	5	2	4	-4792	8	2	4	2976
2	1	3	-448/3	5	2	3	-1208	8	2	3	56
3	0	5	992	5	2	2	52	8	2	2	664
3	0	4	536	5	2	1	208/3	8	2	1	2570/3
3	0	3	-3664/3	5	2	0	308/45	8	2	0	-1486/9
3	0	2	1148	6	0	5	4000	8	3	5	-7520
3	0	1	-718	6	0	4	-2304	8	3	4	-4000
3	0	0	1148/5	6	0	3	128	8	3	3	-8888/3
3	1	5	-992	6	1	5	20096	8	3	2	2032
3	1	4	-1528	6	1	4	-15808	8	3	1	-1762/3
3	1	3	568/3	6	1	3	10736/3	8	3	0	-6932/45
3	1	2	-360	6	1	2	1600	8	4	5	-496
3	1	1	50	6	1	1	868/3	8	4	4	-1536
3	1	0	181/45	6	1	0	-13024/45	8	4	3	1760/3
4	0	5	8016	6	2	5	-20832	9	3	5	992
4	0	4	-2976	6	2	4	-13760	9	3	4	536
4	0	3	56	6	2	3	1808	9	3	3	-3664/3
4	0	2	-664	6	2	2	-3408	9	3	2	1148
4	0	1	2570/3	6	2	1	-3028/3	9	3	1	-718
4	0	0	-1486/9	6	2	0	-1044/5	9	4	5	-992
4	1	5	-7520	6	3	5	-3264	9	4	4	-1528
4	1	4	-4000	6	3	4	512	9	4	3	568/3
4	1	3	-8888/3	6	3	3	-3008/3	9	4	2	-360
4	1	2	2032	7	1	5	12032	9	4	1	50
4	1	1	-1762/3	7	1	3	3056/3	9	4	0	181/45
4	1	0	-6932/45	7	1	2	-308	10	4	5	-128
4	2	5	-496	7	1	1	-358/3	10	4	3	640/3
4	2	4	-1536	7	1	0	539/45	10	4	2	-192
4	2	3	1760/3	7	2	5	-992	10	4	1	272/3
5	0	5	12032	7	2	4	-11960	10	4	0	-2312/45
5	0	4	-6320	7	2	3	3664/3	10	5	5	128
5	0	3	3056/3	7	2	2	-260	10	5	4	128
5	0	2	-308	7	2	1	718	10	5	3	-448/3
5	0	1	-358/3	7	2	0	8462/15	11	5	0	10/9
5	0	0	539/45								

Table 2-6. Coefficients $C_{E,s}^n$ in the expansion of the correlation function C_{ij} for nearest neighbors i and j, for the square lattice (see Eq. 2.6.5). $E = aU - b\mu$; the magnetic field h has been put to zero.

b a s	$C_{E,s}^2$	b a s	$C_{E,s}^2$
3 0 2	1	4 1 1	-2
2 0 1	-1	4 2 2	-2
2 1 2	-1	5 1 2	4
3 0 2	4	5 1 1	4
3 0 1	-4	5 2 2	-4
3 1 2	-4	6 2 2	1
4 0 2	4	6 2 1	-1
4 0 1	-4	6 3 2	-1
4 1 2	-2		

b a s	$C_{E,s}^4$	b a s	$C_{E,s}^4$
2 0 2	-8	6 2 4	576
2 0 3	-2	6 2 3	100
2 0 2	6	6 2 2	-76
2 0 2	-11/3	6 2 1	-22
2 0 0	-1/6	6 2 0	-1/3
2 1 4	8	6 3 4	96
2 1 3	10	6 3 3	4
3 0 4	-84	7 1 4	-400
3 0 3	36	7 1 3	416
3 0 2	2	7 1 2	-104
3 0 1	-6	7 1 1	104/3
3 0 0	-1/2	7 1 0	2/3
3 1 4	64	7 2 4	52
3 1 3	48	7 2 3	300
3 1 2	4	7 2 2	-266
4 0 4	-288	7 2 1	-18
4 0 3	228	7 2 0	-1/6
4 0 2	-68	7 3 4	348
4 0 1	20	7 3 3	32
4 1 4	232	7 3 2	12
4 1 3	104	8 2 4	-288
4 1 2	-32	8 2 3	228
4 1 1	-44/3	8 2 2	-68
4 1 0	-1/3	8 2 1	20
4 2 4	56	8 3 4	232
4 2 3	12	8 3 3	104
5 0 4	-400	8 3 2	-32
5 0 3	416	8 3 1	-44/3
5 0 2	-104	8 3 0	-1/3
5 0 1	104/3	8 4 4	56
5 0 0	2/3	8 4 3	12
5 1 4	32	9 3 3	-84
5 1 3	300	9 3 3	36
5 1 2	-266	9 3 2	2
5 1 1	-18	9 3 1	-6
5 1 0	-1/6	9 3 0	-1/2
5 2 4	348	9 4 4	84
5 2 3	32	9 4 3	48
5 2 2	12	9 4 2	4
6 0 4	-192	10 4 4	-8
6 0 3	240	10 4 3	-2
6 0 2	-16	10 4 2	6
6 1 4	-480	10 4 1	-11/3
6 1 3	416	10 4 0	-1/6
6 1 2	-424	10 5 4	8
6 1 1	40	10 5 3	10
6 1 0	4/3		

b a s	$C_{E,s}^6$	b a s	$C_{E,s}^6$
2 0 6	-102	2 1 4	-56
2 0 5	166	3 0 6	288
2 0 4	-59	3 0 5	1020
2 0 3	-7	3 0 4	-894
2 0 2	157/12	3 0 3	401
2 0 1	-349/60	3 0 2	-281/2
2 0 0	-3/10	3 0 1	441/10
2 1 6	102	3 0 0	8/3
2 1 5	-64	3 1 6	-288

b a s	$C_{E,s}^6$	b a s	$C_{E,s}^6$	b a s	$C_{E,s}^6$
3 1 3	-1308	7 1 1	-20623/30	10 3 5	-53456
3 1 4	-270	7 1 0	-965/36	10 3 4	22492
3 1 3	-65	7 2 6	-113576	10 3 3	14300/3
3 1 2	9/2	7 2 5	-57960	10 3 2	2531
4 0 6	7751	7 2 4	16600	10 3 1	4359/10
4 0 5	-1314	7 2 3	-2392	10 3 0	2453/180
4 0 4	-2048	7 2 2	565/2	10 4 6	-69898
4 0 3	3239/3	7 2 1	2236/3	10 4 5	-26874
4 0 2	-929/4	7 2 0	359/30	10 4 4	-817
4 0 1	1007/60	7 3 6	-44476	10 4 3	349/3
4 0 0	377/40	7 3 5	-12012	10 4 2	-7037/12
4 1 6	-7072	7 3 4	-2338	10 4 1	-1739/20
4 1 5	-6124	7 3 3	-80	10 4 0	-7/4
4 1 4	-1274	7 3 2	29/2	10 5 6	-4514
4 1 3	2473/3	8 0 6	6480	10 5 5	-1760
4 1 2	-1421/12	8 0 5	-15216	10 5 4	8
4 1 1	-196/5	8 0 4	4288	11 3 6	31804
4 1 0	-13/12	8 0 3	448/3	11 3 5	-20340
4 2 6	-679	8 1 6	132448	11 3 4	3394
4 2 5	-992	8 1 5	-130312	11 3 3	-4540/3
4 2 4	108	8 1 4	58496	11 3 2	3338/3
5 0 6	31804	8 1 3	-20056/3	11 3 1	-2267/10
5 0 5	-20340	8 1 2	-3122/3	11 3 0	-423/180
5 0 4	3394	8 1 1	-4272/5	11 4 6	-17672
5 0 3	-4540/3	8 1 0	-106/5	11 4 5	-17234
5 0 2	3338/3	8 2 6	-2086	11 4 4	-258
5 0 1	-2267/10	8 2 5	-107884	11 4 3	14869/3
5 0 0	-623/180	8 2 4	63656	11 4 2	-562/3
5 1 6	-17672	8 2 3	-8643	11 4 1	5977/30
5 1 5	-17236	8 2 2	34255/6	11 4 0	773/90
5 1 4	-258	8 2 1	21247/30	11 5 6	-14132
5 1 3	14869/3	8 2 0	1429/180	11 5 5	-8360
5 1 2	-562/3	8 3 6	-129376	11 5 4	8
5 1 1	5977/30	8 3 5	-34640	11 5 3	-135
5 1 0	773/90	8 3 4	-2916	11 5 2	37/3
5 2 6	-14132	8 3 3	-1430	12 4 6	7751
5 2 5	-8360	8 3 2	-5459/6	12 4 5	-1314
5 2 4	8	8 3 1	-1607/15	12 4 4	-2048
5 2 3	-135	8 3 0	-29/15	12 4 3	3239/3
5 2 2	37/3	8 4 6	-7466	12 4 2	-929/4
6 0 6	53200	8 4 5	-1664	12 4 1	1007/60
6 0 5	-50036	8 4 4	-312	12 4 0	377/40
6 0 4	15336	9 1 6	36540	12 5 6	-7072
6 0 3	-5580	9 1 5	-47840	12 5 5	-6124
6 0 2	1695	9 1 4	15272	12 5 4	-1274
6 0 1	-1295/3	9 1 3	-9104/3	12 5 3	2473/3
6 0 0	-317/18	9 1 2	-252	12 5 2	-1421/12
6 1 6	21212	9 1 1	86	12 5 1	-196/5
6 1 5	-53456	9 1 0	79/45	12 5 0	-13/12
6 1 4	22492	9 2 6	121492	12 6 6	-679
6 1 3	14300/3	9 2 5	-121276	12 6 5	-992
6 1 2	2531	9 2 4	63782	12 6 4	108
6 1 1	4359/10	9 2 3	-15724/3	13 5 6	288
6 1 0	2453/180	9 2 2	13004/3	13 5 5	1020
6 2 6	-69898	9 2 1	-20623/30	13 5 3	401
6 2 5	-26874	9 2 0	-965/36	13 5 2	-281/2
6 2 4	-817	9 3 6	-113576	13 5 1	441/10
6 2 3	349/3	9 3 5	-57960	13 5 0	8/3
6 2 2	-7037/12	9 3 4	16600	13 6 6	-288
6 2 1	-1739/20	9 3 3	-2392	13 6 4	-288
6 2 0	-7/4	9 3 2	565/2	13 6 3	-1308
6 3 6	-4514	9 3 1	2236/3	13 6 4	-270
6 3 5	-1760	9 3 0	359/30	13 6 3	-65
6 3 4	8	9 4 6	-44476	13 6 2	9/2
7 0 6	36540	9 4 5	-12012	14 6 6	-102
7 0 5	-47840	9 4 4	-2338	14 6 5	166
7 0 4	15272	9 4 3	-80	14 6 4	-59
7 0 3	-9104/3	9 4 2	29/2	14 6 3	-7
7 0 2	-252	10 2 6	53200	14 6 2	157/12
7 0 1	86	10 2 5	-50036	14 6 1	-349/60
7 0 0	79/45	10 2 4	15336	14 6 0	-3/10
7 1 6	121492	10 2 3	-5580	14 7 6	102
7 1 5	-121276	10 2 2	1695	14 7 5	-64
7 1 4	63782	10 2 1	-1295/3	14 7 3	-56
7 1 3	-15724/3	10 2 0	-317/18		
7 1 2	13004/3	10 3 6	21212		

3 Extrapolation of the series expansions

A method is proposed to extrapolate the results of high-temperature series expansions to low temperatures. An extrapolation down to zero temperature is obtained, and information on the magnetic phase diagram of the Hubbard model is extracted.

3.1 Introduction

We will consider a method that does not encounter the problems of extrapolation to low temperatures, as described in the previous chapter. In this method the density of holes is used as a small parameter. The high-temperature results are re-expressed in terms of an effective density of states for holes (as has been done before by Brinkman and Rice [16]), and extended to interactions between hole levels. With this density of states, expressions for the free energy of the thermodynamic system can be obtained in the whole range of temperatures. We will define a partition function for the holes in Section 3.2 and express it in terms of an effective chemical potential for the holes. In Section 3.3 we derive the density of states for non-interacting holes, and we determine its moments, for infinite U. We present an improvement on the non-interacting hole picture in Section 3.4, where we consider interacting holes by introducing a Fermi-liquid-like interaction in energy space. In Section 3.5 we show how to use the density of states to calculate zeros of the inverse susceptibility. Section 3.6 deals with the non-interacting hole approximation applied for finite U. Finally, in Section 3.7, we show results for the magnetic phase diagram of the Hubbard model.

3.2 Hole formulation

We use again the Hubbard Hamiltonian (1.1.8), containing the chemical potential and a magnetic field. In order to investigate the thermodynamic properties we want to calculate the grand canonical partition function (1.2.2), which we rewrite for a system consisting of N sites as

$$Z_{gr} = \sum_{N_s=0}^{2N} e^{\beta \mu N_s} Z_{N_s},$$ (3.2.1)

where Z_{N_s} is the canonical partition function for $N_s = N_\uparrow + N_\downarrow$ particles:

$$Z_{N_s} = \sum_{N_\uparrow=0}^{N_s} e^{\beta h (N_\uparrow - N_\downarrow)} \sum_j e^{-\beta \varepsilon_j^{(N, N_s, N_\uparrow)}}$$ (3.2.2)

Here, $\{\varepsilon_j^{(N, N_s, N_\uparrow)}\}$ is the set of eigenvalues of $\mathcal{H}_{kin} + \mathcal{H}_{int}$ for $N_\uparrow$ up spins and $N_\downarrow$ down spins on N sites (notice that the ε_j are functions of t and U only).

In order to approach to lower temperatures in the limit of strong interactions and near half filling, we are going to express the partition function in terms of an effective chemical potential for holes. We associate the kinetic part of the Hamiltonian with the motion of the (dilute) holes, and its magnetic part with the background of spins. Thus, we have to divide out the spin degrees of freedom to obtain the canonical partition function for the holes:

$$Z_{N_h}^h \equiv Z_{N-N_h} e^{\beta \epsilon_{hf}(N-N_h)}, \tag{3.2.3}$$

where we define the number of holes as

$$N_h \equiv N - N_s. \tag{3.2.4}$$

One could define the number of holes also as the number of sites where no particles are present, which reduces to (3.2.4) only for infinite U. Such a definition, however, would make the interpretation of Eq. (3.2.3) problematic for the case of finite U, in which it is possible to create extra holes by creating pairs of electrons on the same site. Therefore, we will use (3.2.4), and the other definitions in this section, also in the case of large but finite U. In that case, a pair of electrons located on the same site causes a very high energy, and the contribution of the corresponding 'extra' hole to the kinetic part of the Hamiltonian is some orders of magnitude smaller than the contribution of a 'real' (non-removable) hole. As we will see in Section 3.6, this leads to terms to be added to the expressions for infinite U of order $1/U$ or higher. First, we will consider the case of infinite U.

In (3.2.3), we have introduced a parameter ϵ_{hf} which can be viewed as the free energy per spin in the absence of holes (i.e., at half filling; we take $Z_0^h \equiv 1$):

$$\epsilon_{hf} \equiv -\frac{1}{N\beta} \ln Z_N, \tag{3.2.5}$$

which for infinite U can be explicitly given in terms of the magnetic field as

$$\epsilon_{hf} = -\frac{1}{\beta} \ln(2 \cosh \beta h). \tag{3.2.6}$$

The grand canonical partition function for the holes then is

$$Z_{gr}^h = Z_{gr} e^{\beta(\epsilon_{hf}-\mu)N} \tag{3.2.7a}$$

$$= \sum_{N_h} Z_{N_h}^h e^{\beta(\epsilon_{hf}-\mu)N_h}, \tag{3.2.7b}$$

suggesting the definition of an effective chemical potential for the holes:

$$\mu_h \equiv \epsilon_{hf} - \mu \tag{3.2.8}$$

[Cf. Eq. (3.2.1)]. With this definition we can rewrite (3.2.7a) as

$$\ln Z_{gr} = -\beta \mu_h N + \ln Z_{gr}^h. \tag{3.2.9}$$

3.3 Construction of a density of states

We consider a system near half filling, with, for simplicity, infinitely strong coupling U (the case of finite U will be treated in Section 3.6). We assume that the system can be described in terms of the kinetic energy of non-interacting dilute holes and the magnetic energy of the background particles. We define the spectral distribution of the energy levels of one hole in an otherwise half-filled system, $\rho(\varepsilon, \beta h)$, in terms of the one-hole partition function Z_1^h, through

$$\frac{Z_1^h}{N} \equiv \int d\varepsilon \rho(\varepsilon, \beta h) e^{-\beta t\varepsilon}, \tag{3.3.1}$$

where we write $\beta t\varepsilon$ to make the integration parameter ε dimensionless. One can see this as a Laplace transform, since Z_1^h is a function of βt. We take ρ to be normalized to one.

Although we said before that we divide out the magnetic degrees of freedom in the spin background, there is still a dependence of ρ on the magnetic field h. It is not easy to see how the hole motion depends on the field exactly, but one can easily understand why this dependence exists: a magnetic field influences the distribution of the spin background, which in turn determines the behavior of the hole. The hole motion depends on the field only indirectly, and the mechanism that governs the hole dynamics can in fact be much better described in terms of the average magnetization of the spin background than in terms of the field. It is important to understand that, in this picture, one has to treat the spin background as if it were at half filling, with the dilute holes subjected to its magnetization. Therefore we change variables at this level from βh to the magnetization per spin m:

$$m(h) = -\frac{\partial \epsilon_{\mathrm{hf}}(\beta h)}{\partial h}. \tag{3.3.2}$$

We can write ρ as a function $\bar{\rho}(\varepsilon, m)$ of m by solving βh from this equation, and then

$$\frac{Z_1^h}{N} = \int d\varepsilon \bar{\rho}(\varepsilon, m) e^{-\beta t\varepsilon}. \tag{3.3.3}$$

This change of variables could also be interpreted as a Legendre transformation

$$\bar{\epsilon}_{\mathrm{hf}}(m) = \epsilon_{\mathrm{hf}}(\beta h) + mh \tag{3.3.4}$$

at half filling, which we will use later on.

With $\bar{\rho}(\varepsilon, m)$ we can write down a first approximation for the grand canonical partition function. A one-hole level can be occupied, with a Boltzmann weight $e^{-\beta t\varepsilon}$, or it can be unoccupied, in which case there is an electron in the system with Boltzmann weight $e^{-\beta \mu_h}$ (with the magnetic energy included in μ_h). Thus, in the case of non-interacting holes we have (dropping the m-dependence of $\bar{\rho}$)

$$\ln Z_{\mathrm{gr}} = N \int d\varepsilon \bar{\rho}(\varepsilon) \ln(e^{-\beta t\varepsilon} + e^{-\beta \mu_h}), \tag{3.3.5}$$

or equivalently, using (3.2.9),

$$\ln Z_{\mathrm{gr}}^h = N \int d\varepsilon \bar{\rho}(\varepsilon) \ln \left(1 + e^{-\beta(t\varepsilon - \mu_h)}\right). \tag{3.3.6}$$

This equation becomes exact in a one-dimensional system, as in that case the holes cannot disturb the magnetic background of the particles, thus being really non-interacting, and also in a ferromagnetic system (at $m = \pm 1$), for similar reasons. In other, higher-dimensional systems (3.3.6) is only correct to first order in $e^{\beta\mu_h}$. We make an expansion of the right-hand side with respect to the small parameter $e^{\beta\mu_h}$ to obtain

$$\ln Z_{gr}^h = N \int d\varepsilon \bar{\rho}(\varepsilon) \left(e^{\beta\mu_h} e^{-\beta t\varepsilon} - \frac{1}{2} e^{2\beta\mu_h} e^{-2\beta t\varepsilon} + \cdots \right). \tag{3.3.7}$$

Comparing this to a similar expansion of the logarithm of Eq. (3.2.7b) we see that this is consistent with the definition of the density of states in the first-order term.

Now, as an illustration of the calculation, let us have a look at the form that the grand potential [Eq. (1.2.3)] actually takes when evaluating it for this system by means of the cluster expansion method. We can use the expression (2.6.1) that we gave in Chapter 2, which for infinite U reduces to

$$\frac{\ln Z_{gr}}{N} = \ln Z_{gr,0} + \sum_{n=1}^{\infty} (\beta t)^{2n} \sum_{m=0}^{2n-1} \sum_{l=-m}^{m} \frac{\Omega_{m,l}^n e^{\beta(m\mu + lh)}}{(Z_{gr,0})^{2n}}. \tag{3.3.8}$$

Here, $Z_{gr,0}$ is again the partition function for a system consisting of only one site:

$$Z_{gr,0} = 1 + 2e^{\beta\mu} \cosh(\beta h) = 1 + e^{-\beta\mu_h}. \tag{3.3.9}$$

By substituting μ_h for μ, using (3.2.8) and (3.2.6), and expanding in the small parameter $e^{\beta\mu_h}$, we can obtain an expression for the grand potential for the holes again, now in the form of a series expansion:

$$\frac{\ln Z_{gr}^h}{N} = \sum_{N_h=1}^{\infty} e^{N_h \beta\mu_h} \sum_{n=0}^{\infty} (\beta t)^{2n} \Omega(N_h, n), \tag{3.3.10}$$

where

$$\Omega(p, 0) = \frac{(-1)^{(p-1)}}{p}, \tag{3.3.11}$$

and

$$\Omega(p, n) = \sum_{m=0}^{2n-1} \sum_{l=-m}^{m} \binom{p+m-1}{p+m-2n} \frac{(-1)^p \Omega_{m,l}^n e^{\beta lh}}{[-2\cosh(\beta h)]^m} \tag{3.3.12}$$

for $n \neq 0$. Finally, we obtain a relation between the coefficients $\Omega(1, n)$ and the moments of $\rho(\varepsilon)$ [$= \rho(\varepsilon, \beta h)$] by expanding (3.3.1) in powers of βt:

$$\frac{Z_1^h}{N} = \sum_{n=0}^{\infty} (\beta t)^n \frac{(-1)^n}{n!} \int d\varepsilon \rho(\varepsilon) \varepsilon^n. \tag{3.3.13}$$

Thus, we see from Eq. (3.3.13) and the first-order term in (3.3.10) that we have

$$\int d\varepsilon \rho(\varepsilon) \varepsilon^{2n} = (2n)! \Omega(1, n) \tag{3.3.14}$$

for the even moments of ρ, all odd moments being zero. Although we have restricted our-selves to the case of infinite U here, this expression can easily be extended for finite U, as we will see in Section 3.6. U then enters the equation as a parameter at the right-hand side.

For infinite U there is another, faster way to calculate these moments. They can then be expressed directly in the number of possible paths in state space for a system with one hole. This has been done first by Brinkman and Rice [16], who calculated the first 10 moments of the density of states for ferromagnetic, antiferromagnetic, and paramagnetic spin backgrounds on a simple cubic lattice. Yang *et al.* [17] have presented a large number of moments for the same spin backgrounds on 2- to 5-dimensional hyper-cubic lattices, including 18 moments for the square lattice and 14 for the simple cubic lattice. In Appendix A, we outline a method which enables us to enumerate the paths in an efficient way, and by which we have extended their results to 22 and 16 moments, respectively. These moments are pre-sented in Table 3-1 for $m = 0$ and $m = 1$.

We now approximate $\bar{\rho}(\varepsilon)$ by a polynomial which we fit with the moments. In this way we calculate an approximation to the density of states, and we obtain an impression of the convergence of subsequent orders of approximation. In Figures 3-1 and 3-2 we show the result for a paramagnetic ($m = 0$) and a ferromagnetic ($m = 1$) system.

Table 3-1. Moments of $\bar{\rho}(\varepsilon)$ for $m = 0$ and $m = 1$, corresponding to a paramagnetic and a ferro-magnetic system, respectively, for the square and the simple cubic lattices (odd moments vanish).

square lattice			simple cubic lattice	
$m = 0$	$m = 1$	n	$m = 0$	$m = 1$
1	1	0	1	1
4	4	2	6	6
30	36	4	72	90
$269\frac{1}{2}$	400	6	$1072\frac{1}{2}$	1860
$2641\frac{3}{4}$	4900	8	$17781\frac{3}{4}$	44730
$27279\frac{15}{16}$	63504	10	$314403\frac{3}{16}$	1172556
$291718\frac{31}{32}$	853776	12	$5804323\frac{47}{64}$	32496156
$3199250\frac{73}{256}$	11778624	14	$110549185\frac{29}{64}$	936369720
$35766660\frac{57}{256}$	165636900	16	$2156004418\frac{11}{256}$	27770358330
$405989247\frac{577}{4096}$	2363904400	18		
$4665921461\frac{101}{16384}$	34134779536	20		
$54182396281\frac{55139}{65536}$	497634306624	22		
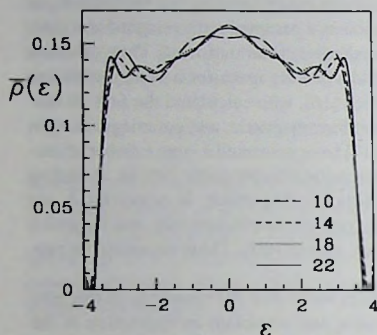
Figure 3-1. The density of states for a para-
magnetic system on a square lattice, using up
to the number of moments indicated.
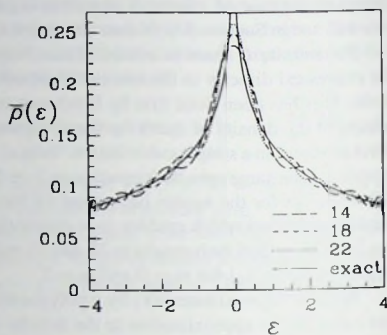
Figure 3-2. The density of states in the ferro-
magnetic regime on a square lattice. Exact re-
sult, and approximations using up to the num-
ber of moments indicated.

For $m = 1$, the system behaves like a system of free fermions, for which the exact density
of states is known [18]:

$$\bar{\rho}(\varepsilon) = \frac{1}{2\pi^2} K\left[1 - \left(\frac{\varepsilon}{4}\right)^2\right], \qquad (3.3.15)$$

with K the complete elliptic integral of the first kind. It has an integrable singularity at $\varepsilon = 0$.
This is difficult to approximate and causes some oscillations away from $\varepsilon = 0$. Convergence
towards the exact result is rather good. For $m = 0$ convergence is very good, as from 14th
order on the difference between subsequent approximations becomes very small.

Meshkov and Berkov [19] fit the density of states by postulating that the integral of $\bar{\rho}^2$ be
minimal ('smoothness' criterion), using a discretized $\bar{\rho}$. They claim that this method gives
faster convergence then a polynomial fit. Comparing their results for the ferromagnetic den-
sity of states with the exact result and the results presented here, however, one may question
that claim. We feel that the polynomial fits, when using an equal number of moments, give
similar or even better results, which are also easier to handle in further calculations.

Before calculating various quantities which can tell us something about the low-tempe-
rature properties of the system, we will in the next section consider a method to improve
the approximation of the density of states by including interactions between the holes.

3.4 Interacting holes

The crucial question to investigate is for which domain of hole densities the assumption of
independent holes is justified. This range can be determined from an estimate of the interac-

tions between the holes. Very similar to the theory of the classical dilute gas [20], the inter-
action can be deduced from the two-hole partition function as defined by (3.2.3) for $N_h = 2$.
It is a matter of choice how to represent the hole interaction. One could think of a spatial
representation, but one must realize that in this strongly quantal system the interaction is
non-local, which complicates the transparency of the spatial representation substantially.
Having the one-hole system represented by a density of states it is natural here to choose
an interaction between energy levels. First we formulate the interaction in terms of discrete
levels and then we take the continuum limit as in Eq. (3.3.5). The discrete version of this
expression can be written in terms of levels ε_i, distributed according to the density $\bar{\rho}(\varepsilon_i)$:

$$Z_{gr}^{(1)} = e^{-\beta\mu_h N} \sum_{\{n_i\}} e^{-\beta \sum_i (t\varepsilon_i - \mu_h) n_i}, \tag{3.4.1}$$

where $\{n_i\}$ with $n_i = 0, 1$ is the occupation of the levels ε_i. We have given the expression a
super-index 1 to indicate that $Z_{gr}^{(1)}$ matches Z_{gr} up to the one-hole terms. The next approxi-
mation can be of the form

$$Z_{gr}^{(2)} = e^{-\beta\mu_h N} \sum_{\{n_i\}} e^{-\beta\left[\sum_i (t\varepsilon_i - \mu_h) n_i + \sum_{(i,j)} f_{ij} n_i n_j\right]}, \tag{3.4.2}$$

where f_{ij} accounts for the interaction between the levels ε_i and ε_j. The second term in the
exponent is a sum over all pairs of levels (i, j). In the energy space a distance between lev-
els does not seem to be a measure for the strength of the interaction as in real space. There,
interactions usually decay sufficiently fast with the distance, such that the sum over pairs
does not increase with the square of the number of elements, but only linearly as is necessary
for a thermodynamic system. In order to make the exponent in (3.4.2) of the correct ther-
modynamic behavior the interaction should therefore decrease with the size of the system
as

$$f_{ij} = \frac{t}{N}\phi_{ij}, \tag{3.4.3}$$

with ϕ_{ij} of order unity. An additional advantage of (3.4.3) is the fact that interactions of this
type can be handled rigorously in the thermodynamic limit by mean-field theory [21]. Thus
we can write

$$\ln Z_{gr}^{(2)} = -\beta\mu_h N + \sum_i \ln(1 + e^{-\beta\bar{\varepsilon}_i}) + \frac{\beta t}{N} \sum_{(i,j)} \phi_{ij} n(\bar{\varepsilon}_i) n(\bar{\varepsilon}_j), \tag{3.4.4}$$

where the $\bar{\varepsilon}_i$ are the shifted energy levels

$$\bar{\varepsilon}_i = t\varepsilon_i - \mu_h + \frac{t}{N} \sum_{j\neq i} \phi_{ij} n(\bar{\varepsilon}_j), \tag{3.4.5}$$

and $n(\bar{\varepsilon})$ is the Fermi occupation number

$$n(\bar{\varepsilon}) = \frac{1}{1 + e^{\beta\bar{\varepsilon}}}. \tag{3.4.6}$$

Now the interaction ϕ_{ij} must be chosen such that $Z_{gr}^{(2)}$ produces the correct two-hole partition function. Expanding Eq. (3.4.2) with respect to the number of holes,

$$N_h = \sum_i n_i, \tag{3.4.7}$$

and using Eq. (3.2.7a), we find

$$Z_2^h = \sum_{(i,j)} e^{-\beta t(\varepsilon_i + \varepsilon_j + \frac{1}{N}\phi_{ij})}. \tag{3.4.8}$$

In our high-temperature expansion we have no direct information on Z_2^h, but we have the coefficient of the second-order term in the hole expansion of $\ln Z_{gr}^h$ [Cf. (3.2.7b)], which is

$$U_2 = Z_2^h - \frac{1}{2}(Z_1^h)^2. \tag{3.4.9}$$

Note that this expression is of order N, and not of order N^2 as are both terms on its right-hand side. Using (3.4.8) and the corresponding expression for Z_1^h we may equate

$$U_2 = \sum_{(i,j)} e^{-\beta t(\varepsilon_i + \varepsilon_j + \frac{1}{N}\phi_{ij})} - \frac{1}{2} \sum_{i,j} e^{-\beta t(\varepsilon_i + \varepsilon_j)}. \tag{3.4.10}$$

For $N \to \infty$ we may write

$$U_2 = -\frac{\beta t}{N} \sum_{(i,j)} e^{-\beta t(\varepsilon_i + \varepsilon_j)} \phi_{ij} - \frac{1}{2} \sum_i e^{-2\beta t \varepsilon_i}, \tag{3.4.11}$$

and we see that U_2 is indeed of order N by virtue of (3.4.3). Note that, even for zero interaction, the terms $i = j$ in the second term of (3.4.10) are not compensated by the first term. The second term in (3.4.11) gives the ideal-gas term of the hole system on the two-hole level.

Since we have moments of U_2 from our high-temperature expansions, and also the last term in (3.4.11) is known from our one-hole density of states, it is convenient to split U_2 into an interacting and an ideal part

$$U_2 = U_2^{int} + U_2^{id}, \tag{3.4.12}$$

with

$$U_2^{id} = -\frac{1}{2} \sum_i e^{-2\beta t \varepsilon_i} \tag{3.4.13a}$$

$$= -\frac{N}{2} \int d\varepsilon \, \bar{\rho}(\varepsilon) e^{-2\beta t \varepsilon} \tag{3.4.13b}$$

[Cf. (3.3.7)]. ϕ_{ij} must then be determined from U_2^{int}. In a continuous version the equation for $\phi(\varepsilon, \varepsilon')$ becomes

$$U_2^{int} = -\frac{\beta t}{2} N \int d\varepsilon \, \bar{\rho}(\varepsilon) \int d\varepsilon' \, \bar{\rho}(\varepsilon') e^{-\beta t(\varepsilon + \varepsilon')} \phi(\varepsilon, \varepsilon'). \tag{3.4.14}$$

This relation is not strong enough to yield a unique $\phi(\varepsilon, \varepsilon')$, in the same way as the second virial coefficient of a classical gas is not sufficient to determine the interaction potential. The freedom in choice will be reflected upon the efficiency of the program to determine the higher-order interactions. We have chosen to have the dependence of $\phi(\varepsilon, \varepsilon')$ only on the sum variable $\varepsilon + \varepsilon'$, and we approximate it by a polynomial:

$$\phi(\varepsilon, \varepsilon') = \sum_l \phi_l (\varepsilon + \varepsilon')^l. \tag{3.4.15}$$

Equating moments in (3.4.14) and expanding U_2^{int} in powers of βt,

$$U_2^{\text{int}} = \sum_k \left(U_2^{\text{int}} \right)_k (\beta t)^k, \tag{3.4.16}$$

we have

$$\left(U_2^{\text{int}} \right)_k = \sum_l \frac{(-1)^k N}{2(k-1)!} \int d\varepsilon \, \bar{\rho}(\varepsilon) \int d\varepsilon' \bar{\rho}(\varepsilon')(\varepsilon + \varepsilon')^{k-1+l} \phi_l. \tag{3.4.17}$$

Because we are working on a bipartite lattice, all odd moments of U_2^{int} are identically zero. Hence k is even, and as also $\bar{\rho}(\varepsilon)$ has only even moments, the combination $k - 1 + l$ must be even and therefore the sums in Eqs. (3.4.15) and (3.4.17) contain only odd l. The set of equations (3.4.17) for a finite number of the moments $(U_2^{\text{int}})_k$ determines an equal number of coefficients ϕ_l. We have computed $(U_2^{\text{int}})_k$ for the square lattice at $U = \infty$, up to $k = 12$. This involves 6 terms ($k = 2, 4, \ldots, 12$) and so we can determine 6 values $\phi_1, \phi_3, \ldots, \phi_{11}$. In the equations we thus need $k - 1 + l = 12 - 1 + 11 = 22$ as the highest moment of $\bar{\rho}(\varepsilon)$, which is just the number of moments we have determined. The values of $\frac{1}{N}(U_2^{\text{int}})_k$ are given in Table 3-2. For the ferromagnetic system ($m = 1$) these coefficients are zero, as the holes do not interact in that case.

Finally we give the continuum form of the expressions (3.4.4) and (3.4.5) for the grand potential:

$$\begin{aligned}
\ln Z_{\text{gr}}^{(2)} &= -\beta \mu_{\text{h}} N + N \int d\varepsilon \, \bar{\rho}(\varepsilon) \ln(1 + e^{-\beta \bar{\varepsilon}}) \\
&+ \frac{\beta t}{2} N \int d\varepsilon \int d\varepsilon' \bar{\rho}(\varepsilon) n(\bar{\varepsilon}) \bar{\rho}(\varepsilon') n(\bar{\varepsilon}') \phi(\varepsilon, \varepsilon').
\end{aligned} \tag{3.4.18}$$

Table 3-2. Values of $\frac{1}{N}\left(U_2^{\text{int}}\right)_k$ for $m = 0$ and $m = 1$, for the square lattice with $U = \infty$.

k	0	2	4	6	8	10	12
$m = 0$	0	0	$-\frac{1}{3}$	$-\frac{47}{80}$	$-\frac{1713}{4032}$	$-\frac{989681}{5806080}$	$-\frac{160327813}{3832012800}$
$m = 1$	0	0	0	0	0	0	0

with

$$\bar{\varepsilon} = t\varepsilon - \mu_h + t \int d\varepsilon' \bar{\rho}(\varepsilon') \phi(\varepsilon, \varepsilon') n(\bar{\varepsilon}'). \tag{3.4.19}$$

3.5   Inverse susceptibility

We return to the uniform susceptibility

$$\chi_{FM} = \left. \frac{\partial M}{\partial h} \right|_{h=0}, \tag{3.5.1}$$

with M the total magnetization of the system. As before, we try to find indications of divergences of χ_{FM}, which should be related to second-order phase transitions between a paramagnetic and a ferromagnetic state. It is usually more convenient to express this by stating that the inverse susceptibility must be zero:

$$\chi_{FM}^{-1} = 0, \tag{3.5.2}$$

and to study

$$\chi_{FM}^{-1} = \left. \frac{\partial h}{\partial M} \right|_{M=0}, \tag{3.5.3}$$

or,

$$\beta N \chi_{FM}^{-1} = \left. \frac{\partial \beta h}{n_s \partial m} \right|_{m=0}, \tag{3.5.4}$$

where m is the magnetization per spin as defined in Section 3.3. In order to find an expression for h, to be able to calculate (3.5.4), we construct a generalized (Landau-like) free energy

$$\varphi(n_s, \mu, m, h) = -\frac{1}{\beta N} \ln Z_{gr} + \mu n_s - hmn_s, \tag{3.5.5}$$

where $\ln Z_{gr}$ is given by (3.3.5). The function φ has to be minimized with respect to μ and m at fixed particle density n_s and field h, to obtain the free energy. Note that this h is *not* the same field as we used before in Section 3.3. There we interpreted h as a field that is felt only by the spins in the background, whereas now we obtain the physical external field that would be necessary to yield the given magnetization. Of course, in the case of a finite number of holes (the limit of half filling), these fields are the same, as we will see in the resulting expressions. Note also that, due to the definition of m as the magnetization per spin, its conjugated variable is hn_s, not h.

We can rewrite (3.5.5) using (3.2.8) and (3.3.4):

$$\beta \varphi = \beta \bar{\epsilon}_{hf} n_s + \beta \mu_h n_h - \frac{1}{N} \ln Z_{gr}^h, \tag{3.5.6}$$

where we can interpret the first term as the contribution of the background of spins, and the other terms as the contribution of the holes. Minimization leads to the following equations:

$$n_h = \int d\varepsilon \bar{\rho}(\varepsilon) \frac{1}{1 + e^{\beta(t\varepsilon - \mu_h)}}, \tag{3.5.7}$$

$$\beta h = \beta h_{hf} - \int d\varepsilon \frac{\partial \bar{\rho}(\varepsilon)}{n_s \partial m} \ln \left(1 + e^{-\beta(t\varepsilon - \mu_h)}\right). \tag{3.5.8}$$

with

$$h_{hf} = \frac{\partial \bar{\epsilon}_{hf}}{\partial m}. \tag{3.5.9}$$

The expression for the inverse uniform susceptibility (3.5.4) then becomes

$$\beta N \chi_{FM}^{-1} = \left. \frac{\partial \beta h_{hf}}{n_s \partial m} \right|_{m=0} - \int d\varepsilon \left. \frac{\partial^2 \bar{\rho}(\varepsilon)}{n_s^2 \partial m^2} \right|_{m=0} \ln\left(1 + e^{-\beta(\iota\varepsilon - \mu_h)}\right). \tag{3.5.10}$$

This can be rewritten in terms of $\rho(\varepsilon)$, using the Legendre transform (3.3.4) [thus $\bar{\rho}(\varepsilon, m) = \rho(\varepsilon, \beta h_{hf})$]:

$$\beta N \chi_{PM}^{-1} = \left. \frac{\partial \beta h_{hf}}{n_s \partial m} \right|_{m=0} \left(1 - \left. \frac{\partial \beta h_{hf}}{n_s \partial m} \right|_{m=0} \int d\varepsilon \left. \frac{\partial^2 \rho(\varepsilon, \beta h_{hf})}{\partial(\beta h_{hf})^2} \right|_{h_{hf}=0} \ln\left(1 + e^{-\beta(\iota\varepsilon - \mu_h)}\right) \right). \tag{3.5.11}$$

Note that $m = 0$ is equivalent to $h_{hf} = 0$, and that, for reasons of symmetry, the first derivative of ρ with respect to h_{hf} vanishes at $h_{hf} = 0$.

According to (3.5.2) we want to find values of n_h and βt for which the right-hand side of (3.5.11) is zero, with n_h fixed by Eq. (3.5.7). For infinite U we have $\beta h_{hf}(m) = \arctan(m)$, so putting (3.5.11) to zero gives

$$\int d\varepsilon \left. \frac{\partial^2 \rho(\varepsilon, \beta h_{hf})}{\partial(\beta h_{hf})^2} \right|_{h_{hf}=0} \ln\left(1 + e^{-\beta(\iota\varepsilon - \mu_h)}\right) = 1 - n_h. \tag{3.5.12}$$

This equation can be solved by an iterative procedure to calculate the value of μ_h for a given value of βt. The density of states $\rho(\varepsilon)$, necessary to calculate n_h according to (3.5.7), is determined from its moments as described in Section 3.3, and its second derivative is calculated in a similar way.

To include the interaction described in Section 3.4, one should use the grand potential as given in (3.4.18) rather than the non-interacting hole approximation of (3.3.5). The final equation, equivalent to (3.5.11), then involves one extra term which contains the second derivative with respect to βh_{hf} of the interaction ϕ. We give a derivation of this equation in Appendix B.

In Figure 3-3 we show Curie temperatures for the square-lattice Hubbard model at infinite U, in three different approximations: (a) The non-interacting hole approximation, with ρ determined by interpolation from 8 of its moments (of which 4 moments are non-zero); (b) The same but with ρ determined from 22 (11 non-zero) moments; and (c) The interacting-hole approximation, with ρ determined from 22 moments and ϕ from 12 (5 non-zero) interaction coefficients.

One can see that the difference between the 8th-order and the 22nd-order non-interacting approximations is small. In both approximations, ferromagnetism is stable against paramagnetism for $n_h \lesssim 0.27$, at low T. The interaction does not change this picture very much. It slightly enhances the stability of the ferromagnetic state, up to $n_h \lesssim 0.29$. The difference between the non-interacting and the interacting approximations becomes larger with increasing hole density, as expected. Numerically, the results agree very well for $n_h \lesssim 0.06$.
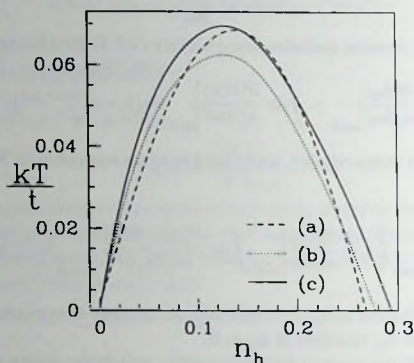
Figure 3-3. Curie temperatures (contours of zero inverse uniform susceptibility) for the square lattice at infinite U. (a) Non-interacting hole approximation, 8th order; (b) Non-interacting hole approximation, 22nd order; (c) Interacting-hole approximation, 22nd order.

In the next section we will treat the case of finite U. We have been able to calculate 8 moments of the density of states in that case, thus we can do an eighth-order approximation at the most. One can then calculate merely two coefficients ϕ_i of the interaction, resulting in an approximation of the interaction which is rather crude. We have seen that the picture in the non-interacting hole approximation is qualitatively the same as the one in the interacting-hole approximation, in eighth order already. For small n_h it agrees rather well also numerically. Therefore, we will not include the interaction in the following calculations.

3.6   Large U

As we pointed out before, at finite U, excitations in the spin background become possible due to the creation of pairs of electrons with opposite spin at the same site. This means that extra empty sites are created, and thus the number of empty sites in the system is no longer fixed. Taking U large, however, we can consider the contributions to the partition function due to these excitations to be small corrections of the infinite-U system, and we can neglect the terms that would arise from permanently present electron pairs. To do this, we consider the grand potential of the Hubbard model on a square lattice up to the second-order term:

$$-\frac{\beta\Omega}{N} = \ln\left[1 + 2e^{\beta\mu}\cosh(\beta h) + e^{2\beta\mu-\beta U}\right] +$$

$$+ \quad (\beta t)^2 \frac{4e^{\beta\mu}(1 + e^{2\beta\mu - \beta U})\cosh(\beta h) + \frac{8}{\beta U}e^{2\beta\mu}(1 - e^{-\beta U})}{\left(1 + 2e^{\beta\mu}\cosh(\beta h) + e^{2\beta\mu - \beta U}\right)^2} + \cdots. \quad (3.6.1)$$

In this expression, we will neglect the terms that contain the exponential of $-\beta U$, but we keep terms that are proportional to a power of $1/U$. This precisely distinguishes the terms that are due to permanent electron pairs, which cause an energy βU, from those due to temporary excitations in a system where otherwise no double occupancies are present. It is necessary to make this approximation, as the exponential terms cannot be treated in this method. However, it can be seen easily that these terms are always exponentially smaller than other terms in the expansion, and thus that this approximation is justified.

First we consider the case of half filling, where we have $\mu = U/2$:

$$-\frac{\beta\Omega_{hf}}{N} = \ln\left[2 + 2e^{\beta U/2}\cosh(\beta h)\right] + (\beta t)^2 \frac{8e^{\beta U/2}\cosh(\beta h) + \frac{8}{\beta U}(e^{\beta U} - 1)}{\left[2 + 2e^{\beta U/2}\cosh(\beta h)\right]^2} + \cdots.$$
$$(3.6.2)$$

Here we can neglect all but the most important terms at large U, i.e., we only take the terms containing the highest power of $e^{\beta U}$, to get

$$-\frac{\beta\Omega_{hf}}{N} = \frac{\beta U}{2} + \ln\left[2\cosh(\beta h)\right] + (\beta t)^2 \frac{2}{(\beta U)\left[\cosh(\beta h)\right]^2} + \cdots. \quad (3.6.3)$$

By definition, this expression must be equal to $\frac{\beta U}{2} + \frac{1}{N}\ln Z_N$, so using the definition (3.2.5) for ϵ_{hf} we get

$$-\beta\epsilon_{hf} = \ln\left[2\cosh(\beta h)\right] + (\beta t)^2 \frac{2}{(\beta U)\left[\cosh(\beta h)\right]^2} + \cdots, \quad (3.6.4)$$

and we see that this is indeed a correction of order $\frac{1}{U}$ on Eq. (3.2.6). Note that we obtain the same result if we first omit the $e^{-\beta U}$ terms in (3.6.1), and only then substitute $U/2$ for μ. This once more supports our statement that these terms may be neglected.

Off half filling, we have to rewrite (3.6.1) (without the $e^{-\beta U}$ terms) in terms of the effective chemical potential μ_h for the holes, as defined by Eq. (3.2.8), but now containing the corrected ϵ_{hf} as given by (3.6.4). For simplicity, we do this in a few steps. First, we substitute the chemical potential for the holes without the correction terms, as in Section 3.3. Then we expand the logarithm and the numerators with respect to the exponential of this chemical potential. Finally, we include the corrected μ_h by expanding the exponentials with respect to the correction terms. Thus, we obtain for the grand potential

$$-\frac{\beta\Omega}{N} = -\beta\mu_h + e^{\beta\mu_h}\left(1 + (\beta t)^2\left[2 - \frac{2}{(\beta U)\left[\cosh(\beta h)\right]^2}\right] + \cdots\right) + \cdots. \quad (3.6.5)$$

The coefficient of $e^{\beta\mu_h}$ in this expression again determines the moments of the distribution $\rho(\varepsilon, \beta h)$, as described in Section 3.3. Of course these are now functions of βU. In Table 3-3 we give the moments that we have been able to derive from the series expansion data, for

Table 3-3. Moments of the density of states $\left(\int d\varepsilon \rho(\varepsilon)\varepsilon^n\right)$ for the square and simple cubic lattices (odd moments vanish), for large U and $h = 0$.

n	nth moment for the square lattice
0	1
2	$2\left(2 - \frac{2}{\beta U}\right)$
4	$24\left(\frac{5}{4} - \frac{2}{\beta U} - \frac{12}{(\beta U)^2} + \frac{3}{(\beta U)^3}\right)$
6	$720\left(\frac{539}{1440} - \frac{59}{48\beta U} - \frac{93}{8(\beta U)^2} - \frac{89}{6(\beta U)^3} + \frac{1}{2(\beta U)^4} + \frac{127}{2(\beta U)^5}\right)$
8	$40320\left(\frac{10567}{161280} - \frac{271}{576\beta U} - \frac{1459}{320(\beta U)^2} - \frac{377}{32(\beta U)^3} + \frac{4531}{96(\beta U)^4} + \frac{4043}{8(\beta U)^5} - \frac{28837}{8(\beta U)^6} + \frac{78593}{8(\beta U)^7}\right)$

n	nth moment for the simple cubic lattice
0	1
2	$2\left(3 - \frac{3}{\beta U}\right)$
4	$24\left(3 - \frac{9}{2\beta U} - \frac{27}{(\beta U)^2} + \frac{3}{(\beta U)^3}\right)$
6	$720\left(\frac{143}{96} - \frac{67}{16\beta U} - \frac{315}{8(\beta U)^2} - \frac{71}{2(\beta U)^3} - \frac{633}{2(\beta U)^4} + \frac{825}{2(\beta U)^5}\right)$
8	$40320\left(\frac{1129}{2560} - \frac{869}{320\beta U} - \frac{7407}{320(\beta U)^2} - \frac{249}{32(\beta U)^3} - \frac{10551}{32(\beta U)^4} + \frac{13725}{8(\beta U)^5} - \frac{215739}{8(\beta U)^6} + \frac{279837}{4(\beta U)^7}\right)$

$h = 0$. Note that the moments for $h = \infty$ are the same as in the case of infinite U (Table 3-1), because U has no significance in a system where all spins point in the same direction.

We can apply the same method described in Section 3.5 to calculate Curie temperatures for finite U. One has to realize, though, that the inverse susceptibility at half filling then depends on the temperature, which was not the case for infinite U. Due to the excitations we get corrections of the type $\beta t^2/U$, thus we still have a series expansion in the parameter βt. The coefficients in this expansion are suppressed by large factors βU, however, and the range of convergence of the expansion is $\beta t \lesssim 30$ or further, depending on the value of βU. Therefore, we may hope that convergence is good enough in the region where we expect to find solutions of (3.5.2). We give the full expression for the inverse susceptibility at half filling, for the square lattice and up to the $(\beta t)^8$ terms:

$$
\beta N \chi_{\text{hf}}^{-1} = 1 + \frac{4(\beta t)^2}{(\beta U)} + \frac{8(-2 + \beta U)(\beta t)^4}{(\beta U)^3} + \frac{\left[1131 - 648\beta U + 32(\beta U)^2\right](\beta t)^6}{3(\beta U)^5} +
$$
$$
+ \frac{\left[-9129 + 6296\beta U - 1132(\beta U)^2 + 4(\beta U)^3\right](\beta t)^8}{(\beta U)^7}. \tag{3.6.6}
$$

We have checked that (3.6.6) does not vanish for any value of βt and βU. Therefore we ex-

pect no transition from a paramagnetic to a ferromagnetic state in the half-filled system. We only have to consider the second factor on the right-hand side of (3.5.11), which vanishes at

$$\int d\varepsilon \left. \frac{\partial^2 \rho(\varepsilon, \beta h)}{\partial(\beta h)^2} \right|_{h=0} \ln\left(1 + e^{-\beta(\iota\varepsilon - \mu_h)}\right) = (1 - n_h) \frac{\chi_M}{\beta N}. \tag{3.6.7}$$

We show the results for the square and the simple cubic lattices in the next section.

3.7   Magnetic phase diagram

We have used the theory described above to calculate Curie temperatures for the square and simple cubic lattices. For both lattices, we find a surface of Curie temperatures in the n_h–t/U–T diagram. In Figures 3-4 and 3-5 we display these results.

In Figures 3-4(a) and 3-5(a), contours of fixed Curie temperature are plotted in the n_h–t/U plane. In the range of temperatures up to about $kT/t = 0.20$ we find a curve enclosing a region of ferromagnetism. For $kT/t \gtrsim 0.07$ these curves are closed and lie away from the $t/U = 0$ axis. Thus, at given density n_h and temperature T_C, one has to go to finite U to find a transition. In other words: allowing for excitations in the spin background enhances the ferromagnetic behavior. Furthermore, curves are generally not enclosed by all contours at lower temperatures. This would imply that, at given n_h and t/U, one would find a paramagnetic-ferromagnetic transition when lowering the temperature from a region of high temperature, but also when letting it increase from zero. This reentering of a paramag-
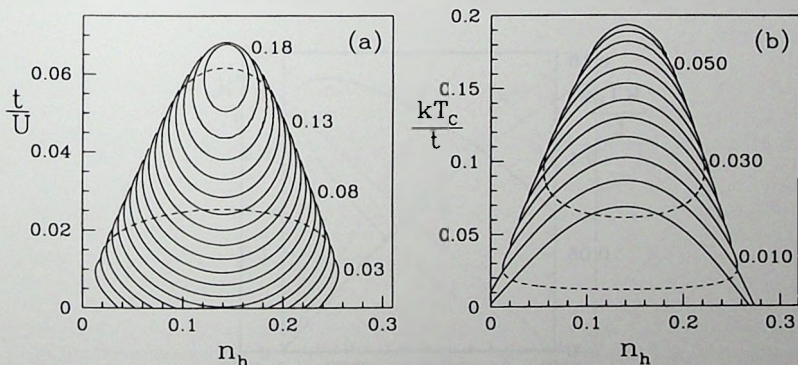
Figure 3-4. Magnetic phase diagram for the square lattice. (a) Contours of fixed Curie temperature, with $kT_C/t = 0.03, 0.04, \ldots, 0.19$ (increment 0.01); (b) Curie temperature at fixed $t/U = 0, 0.005, \ldots, 0.055$ (increment 0.005). For a few curves, the unphysical part has been indicated by a dashed line.
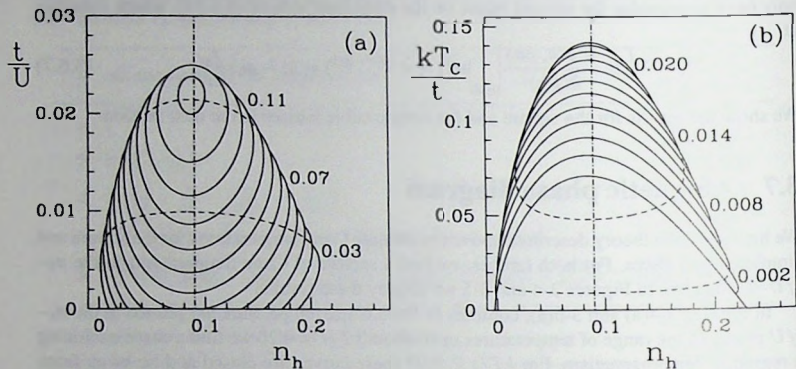
Figure 3-5. Phase diagram for the simple cubic lattice. (a) Contours of fixed Curie temperature, with $kT_C/t = 0.03, 0.04, \ldots, 0.14$ (increment 0.01); (b) Curie temperature at fixed $t/U = 0, 0.002, \ldots,$ 0.022 (increment 0.002). The unphysical part of a few curves is indicated by a dashed line. For $n_h = 0.09$, indicated by the dotted-dashed line, Curie temperatures are plotted as a function of t/U in Figure 3-6.
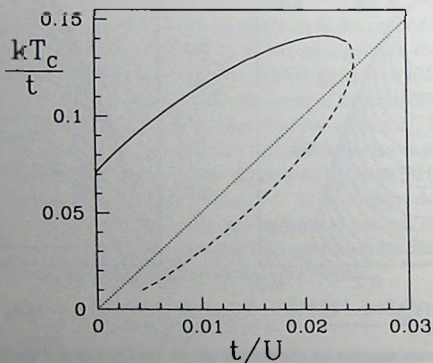
Figure 3-6. Curie temperature for the simple cubic lattice, at $n_h = 0.09$. The dashed part of the curve is unreliable, due to lack of convergence below the dotted line.

netic phase at low temperatures does not seem to be physical. It is an artefact of this method, due to convergence problems at very low temperatures. One can understand this by looking at the expression (3.6.6). If the highest-order term becomes of order one, the series is clearly too short and does not converge properly. This means that the results become unreliable for $\frac{t}{U} \gtrsim \frac{kT}{2t}$ in the case of the square lattice, and $\frac{t}{U} \gtrsim \frac{kT}{4t}$ on the simple cubic lattice. We have drawn the unphysical part of a few curves only, indicated by a dashed line. As the approximations are better for higher temperatures, we assume that the actual curve at $T_C = 0$ (for which we can only perform a calculation at infinite U) should enclose all curves shown.

In Figures 3-4(b) and 3-5(b), we show Curie temperatures in contours of fixed t/U. Again we see the non-physical behavior of curves being closed at the low-temperature side, for almost all values of t/U. Figure 3-6 shows Curie temperatures at fixed $n_h = 0.09$, for the simple cubic lattice, as indicated by the dotted-dashed lines in Figure 3-5. The dotted line in Figure 3-6 indicates the region where the series expansion becomes unreliable, according to the arguments presented above.

There is one other point we want to mention here. As we have stated in the introduction, we have also constructed the staggered susceptibility by replacing the magnetic field h by a staggered field h_s. Although it is much more complicated to calculate the high-temperature expansions for that case, as the number of terms involved increases significantly, it is not difficult to obtain expressions for the staggered susceptibility, both at half filling and in the one-hole approximation, for $h_s = 0$. Thus, one may think that it is possible to obtain similar results for the transition between a paramagnetic and an antiferromagnetic state, and conclude which transition occurs first. When putting the inverse staggered susceptibility at half filling (the equivalent of (3.6.6) for the antiferromagnetic system) to zero, one finds solutions for all values of the parameter βU. This means that the staggered susceptibil-
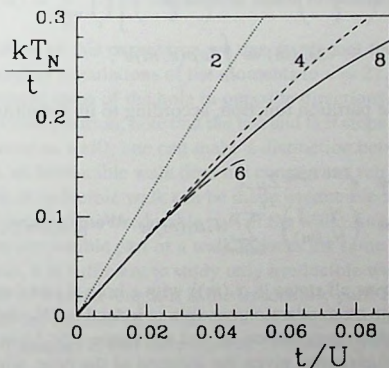
Figure 3-7. Néel temperature for the simple cubic lattice at half filling. Approximations to different orders in βt, as indicated.

ity of the half-filled system diverges at a finite temperature. Apparently, the paramagnetic-antiferromagnetic transition is driven by the background itself, and may be disturbed by a finite hole density. In our formulation, however, it is the holes that drive the system into an ordered state, and the background only indirectly contributes to the transition via its interaction with the holes.

This formulation is clearly not suitable to describe the transition to an antiferromagnetic state. Therefore we only briefly indicate what we expect for the paramagnetic-antiferromagnetic transition. In Figure 3-7 we plot Néel temperatures for the simple cubic lattice at half filling, in approximations to different orders in the parameter βt. We see that the convergence of the series expansion is very good for large U. A transition from a paramagnetic to an antiferromagnetic phase is expected for all values of U. It is at $T_N = 0$ for infinite U, and at increasing temperatures with decreasing U. For finite hole densities we expect the transition to occur at lower temperatures, and at some point cross the paramagnetic-ferromagnetic transition.

Appendix A: Enumeration of paths

In this appendix we describe an efficient way to calculate the moments of the density of states for the case of infinite U, as presented in Table 3-1. We start from Eq. (3.3.3), which we expand in terms of the parameter βt:

$$\frac{Z_1^h}{N} = \sum_{n=0}^{\infty} (\beta t)^n \frac{(-1)^n}{n!} M_n(m), \tag{A.1}$$

with the moments of the density of states defined as

$$M_n(m) \equiv \int d\varepsilon \bar{\rho}(\varepsilon, m)\varepsilon^n. \tag{A.2}$$

We can write the one-hole partition function, according to its definition [Cf. (3.2.3)], also as

$$Z_1^h \equiv e^{(N-1)\beta\epsilon_{hf}} Z_{N-1} \tag{A.3}$$

$$= \binom{N-1}{N_\uparrow(m)}^{-1} \sum_{i, \alpha_i(m)} \langle i, \alpha_i(m)| e^{-\beta\mathcal{H}_{kin}} |i, \alpha_i(m)\rangle, \tag{A.4}$$

where the summation is over all states $|i, \alpha_i(m)\rangle$ with a hole at site i and with a spin background $\alpha_i(m)$ such that the magnetization per spin is indeed m. $N_\uparrow$ denotes the number of electrons with spin up, which depends on m, and the factor $\binom{N-1}{N_\uparrow(m)}$ is the total number of possible background configurations given the location of the hole, which accounts for the spin degrees of freedom. In the thermodynamic limit, this factor is exactly equal to the exponential factor in (A.3), as one easily checks by applying Stirling's formula for the binomial, and with (3.2.6) for ϵ_{hf}. The summation over i gives a trivial (translational) factor N, and

we can expand the exponential in powers of βt to obtain

$$\frac{Z_1^h}{N}\binom{N-1}{N_\uparrow(m)} = \sum_{\alpha(m)}\sum_n \frac{(-1)^n}{n!} A_n[\alpha(m)](\beta t)^n, \tag{A.5}$$

where

$$A_n[\alpha(m)] = \langle\alpha(m)|\left(\frac{\mathcal{H}_{\text{kin}}}{t}\right)^n|\alpha(m)\rangle \tag{A.6}$$

is the number of walks of length n in the configuration space, that restore the spin background $\alpha(m)$ to its original state. Comparing (A.1) and (A.5) we see that

$$M_n(m) = \binom{N-1}{N_\uparrow(m)}^{-1}\sum_{\alpha(m)} A_n[\alpha(m)]. \tag{A.7}$$

Thus $M_n(m)$ is precisely the sum over all possible closed walks w_n of length n, summing the fraction of spin backgrounds that is restored by w_n. Such walks induce permutations $\Pi(w_n)$ of the spins, which can be written as products of disjunct cyclic permutations $\Pi_i(w_n)$ with length $|\Pi_i(w_n)| > 1$. In order to restore the spin background $\alpha(m)$, the direction of the spin on each site must remain unchanged, when applying $\Pi_i(w_n)$. Thus, all spins that are interchanged by this permutation must point in the same direction. As the number of spins involved is negligible compared to the total number of spins, we may approximate that the probability to find an individual spin pointing up or down is given by $(1+m)/2$ and $(1-m)/2$, respectively. Hence the fraction of backgrounds in which the alignment of the spins remains unchanged under the permutation $\Pi_i(w_n)$ is $[(1+m)/2]^l + [(1-m)/2]^l$, where $l = |\Pi_i(w_n)|$ is the number of spins involved in the permutation. We can then calculate M_n as

$$M_n(m) = \sum_{w_n}\prod_i\left[\left(\frac{1+m}{2}\right)^{|\Pi_i(w_n)|} + \left(\frac{1-m}{2}\right)^{|\Pi_i(w_n)|}\right]. \tag{A.8}$$

For the actual evaluation of this expression we use an elegant theorem that enables us to significantly extend earlier calculations of the moments to $n = 22$. Defining a *retracing sequence* as two subsequent steps of the hole in opposite directions (thus after two steps the hole is back in its previous position; note that the last and first steps of a closed walk are considered to be subsequent as well), one can make a distinction between *reducible* and *irreducible* closed walks: an irreducible walk does not contain any retracing sequence, whereas a reducible walk does. A reducible walk can be made irreducible by removing its retracing sequences; the result is called the irreducible part of the walk. As a retracing sequence does not permute spins, the irreducible part of a walk induces the same permutation of the spins as the walk itself. Thus, it is sufficient to study only irreducible walks if one knows of how many reducible walks of a given length it is the irreducible part. Brouwer [22] has proved the following formula: the number of closed walks of length $l + 2n$ on a hyper-cubical lattice with coordination number z, that have a given irreducible part of length $l > 0$, is

$$N_{\text{ir}}(l,n) = (z-1)^n\binom{l+2n}{n} \tag{A.9}$$

This greatly facilitates the calculation of (A.8).

Appendix B: The inverse susceptibility

In this appendix we give the formula for the inverse susceptibility in the interacting-hole approximation, using the theory given in Section 3.4. We start from Eq. (3.5.5), which has to be differentiated with respect to m in order to get the equivalent of (3.5.8), with (3.4.18) for $\ln Z_{gr}$:

$$
\begin{aligned}
\beta h &= \beta h_{hf} + n_h \frac{\partial \beta \mu_h}{n_s \partial m} - \int d\varepsilon \frac{\partial \bar{\rho}(\varepsilon)}{n_s \partial m} \ln(1 + e^{-\beta \bar{\varepsilon}}) + \int d\varepsilon \bar{\rho}(\varepsilon) n(\bar{\varepsilon}) \frac{\partial \beta \bar{\varepsilon}}{n_s \partial m} \\
&- \beta t \int d\varepsilon \int d\varepsilon' \frac{\partial \bar{\rho}(\varepsilon)}{n_s \partial m} n(\bar{\varepsilon}) \bar{\rho}(\varepsilon') n(\bar{\varepsilon}') \phi(\varepsilon, \varepsilon') \\
&- \beta t \int d\varepsilon \int d\varepsilon' \bar{\rho}(\varepsilon) \frac{\partial n(\bar{\varepsilon})}{n_s \partial m} \bar{\rho}(\varepsilon') n(\bar{\varepsilon}') \phi(\varepsilon, \varepsilon') \\
&- \frac{\beta t}{2} \int d\varepsilon \int d\varepsilon' \bar{\rho}(\varepsilon) n(\bar{\varepsilon}) \bar{\rho}(\varepsilon') n(\bar{\varepsilon}') \frac{\partial \phi(\varepsilon, \varepsilon')}{n_s \partial m},
\end{aligned}
\tag{B.1}
$$

where n_h is given by

$$
n_h = - \int d\varepsilon \bar{\rho}(\varepsilon) n(\bar{\varepsilon}) \frac{\partial \beta \bar{\varepsilon}}{\partial \beta \mu_h} + \beta t \int d\varepsilon \int d\varepsilon' \bar{\rho}(\varepsilon) n(\bar{\varepsilon}) \bar{\rho}(\varepsilon') \frac{\partial n(\bar{\varepsilon}')}{\partial \beta \mu_h} \phi(\varepsilon, \varepsilon').
\tag{B.2}
$$

This may look awkward, but if we look at the derivatives of $\bar{\varepsilon}$ [see Eq. (3.4.19)] we see that many of these terms cancel. Let us first look at the expression (B.2) for the hole density. As we are working at fixed hole density, derivatives of the Fermi factor do not play a role in these equations, and they vanish. We need the derivative of $\bar{\varepsilon}$ with respect to $\beta \mu_h$,

$$
\frac{\partial \beta \bar{\varepsilon}}{\partial \beta \mu_h} = -1 + \beta t \int d\varepsilon' \bar{\rho}(\varepsilon') \phi(\varepsilon, \varepsilon') \frac{\partial n(\bar{\varepsilon}')}{\partial \beta \mu_h},
\tag{B.3}
$$

and so we see that indeed there is a cancelation of terms, leaving us with the relation

$$
n_h = \int d\varepsilon \bar{\rho}(\varepsilon) n(\bar{\varepsilon}).
\tag{B.4}
$$

Then, we rewrite the expression for the magnetic field with

$$
\begin{aligned}
\frac{\partial \beta \bar{\varepsilon}}{n_s \partial m} &= -\frac{\partial \beta \mu_h}{n_s \partial m} + \beta t \int d\varepsilon' \frac{\partial \bar{\rho}(\varepsilon')}{n_s \partial m} n(\bar{\varepsilon}') \phi(\varepsilon, \varepsilon') + \beta t \int d\varepsilon' \bar{\rho}(\varepsilon') \frac{\partial n(\bar{\varepsilon}')}{n_s \partial m} \phi(\varepsilon, \varepsilon') \\
&+ \beta t \int d\varepsilon' \bar{\rho}(\varepsilon') n(\bar{\varepsilon}') \frac{\partial \phi(\varepsilon, \varepsilon')}{n_s \partial m}.
\end{aligned}
\tag{B.5}
$$

Using this expression it is straightforward to check that (B.1) reduces to

$$
\beta h = \beta h_{hf} - \int d\varepsilon \frac{\partial \bar{\rho}(\varepsilon)}{n_s \partial m} \ln(1 + e^{-\beta \bar{\varepsilon}}) + \frac{\beta t}{2} \int d\varepsilon \int d\varepsilon' \bar{\rho}(\varepsilon) n(\bar{\varepsilon}) \bar{\rho}(\varepsilon') n(\bar{\varepsilon}') \frac{\partial \phi(\varepsilon, \varepsilon')}{n_s \partial m}.
\tag{B.6}
$$

In order to derive the inverse susceptibility from this expression, we have to take the derivative with respect to $n_s m$ again, and put $m = 0$. For reasons of symmetry it is easy to show

that the first derivatives with respect to m of all functions appearing in the integrals vanish at $m = 0$. Thus, in the terms in (B.6) we only have to consider the derivatives of the functions that have been differentiated once already:

$$\beta N \chi_{PM}^{-1} = \left.\frac{\partial \beta h_{hf}}{n_s \partial m}\right|_{m=0} - \int d\varepsilon \left.\frac{\partial^2 \bar{\rho}(\varepsilon)}{n_s^2 \partial m^2}\right|_{m=0} \ln(1 + e^{-\beta \bar{\varepsilon}}) +$$
$$+ \frac{\beta t}{2} \int d\varepsilon \int d\varepsilon' \bar{\rho}(\varepsilon) n(\bar{\varepsilon}) \bar{\rho}(\varepsilon') n(\bar{\varepsilon}') \left.\frac{\partial^2 \phi(\varepsilon, \varepsilon')}{n_s^2 \partial m^2}\right|_{m=0}. \quad \text{(B.7)}$$

This can again be expressed in terms of $\rho(\varepsilon)$ (note that also ϕ is being Legendre transformed):

$$\beta N \chi_{PM}^{-1} = \left.\frac{\partial \beta h_{hf}}{n_s \partial m}\right|_{m=0} \left(1 - \left.\frac{\partial \beta h_{hf}}{n_s \partial m}\right|_{m=0} \left[\int d\varepsilon \frac{\partial^2 \rho(\varepsilon)}{\partial (\beta h)^2} \ln(1 + e^{-\beta \bar{\varepsilon}})\right.\right.$$
$$\left.\left.+ \frac{\beta t}{2} \int d\varepsilon \int d\varepsilon' \rho(\varepsilon) n(\bar{\varepsilon}) \rho(\varepsilon') n(\bar{\varepsilon}') \frac{\partial^2 \phi(\varepsilon, \varepsilon')}{\partial (\beta h)^2}\right]_{h=0}\right), \quad \text{(B.8)}$$

which is the modification of (3.5.11) for interacting holes.

4 Theory of quantum Monte Carlo methods

A few different forms are described of a technique called Monte Carlo simulation, which is devised to calculate properties of relatively large systems. The aim is to find a Monte Carlo method that is suitable for calculating properties of interacting fermions on a lattice.

4.1 Introduction

Monte Carlo simulations are among the few methods known to extract unbiased information from systems that are too large to apply exact calculation methods. In a Monte Carlo simulation, a sum or integral in a complete phase space is replaced by a sum or integral over a well chosen set of configurations, such that the result of this summation equals the exact result within a statistical error bar. In Quantum Monte Carlo methods, one often uses a trial function, on the basis of which one chooses the configurations. In these configurations one samples the desired properties, and one obtains the final answer by weighting these local properties in the proper way. A whole variety of problems can be tackled in this way, ranging from ground-state to finite-temperature properties and from continuum to lattice systems.

We will be concerned with techniques to obtain ground-state properties for lattice systems. In Sections 4.2 and 4.3, we will describe two standard Quantum Monte Carlo techniques for this purpose, which are commonly used. Section 4.4 treats ways of improving the statistical behavior of the sampling method. In Section 4.5 we explain why the standard technique cannot be applied to a system of fermions. Finally, Sections 4.6 and 4.7 describe how we adapt the standard technique to obtain a useful algorithm for dealing with fermions on a lattice.

In all of the following sections in this chapter we assume that we are working with a Hamiltonian $\mathcal{H}$, describing particles that reside on a lattice. This Hamiltonian operates in a space of configurations $\{R\}$, each of which is a complete description of the positions and spin states of all particles on the lattice (see Section 1.4).

4.2 Variational Monte Carlo

The simplest way of performing a Monte Carlo simulation is *Variational Monte Carlo* (VMC). In this method, an explicit form for the wave function is assumed, containing a set of parameters that can be varied. A simple example of such a *trial wave function* is a Slater determinant, resulting from a self-consistent mean-field calculation. One wants to optimize the wave function with respect to these parameters, to obtain the lowest possible energy

given this form. By means of a Monte Carlo simulation one can calculate the energy of the wave function as a function of the parameters, and thus find their optimized values.

Let us express the trial state $|\Psi_T\rangle$ in terms of the configurations of our lattice system:

$$|\Psi_T\rangle = \sum_R |R\rangle\langle R|\Psi_T\rangle, \tag{4.2.1}$$

where the coefficients $\langle R|\Psi_T\rangle$ depend on a set of parameters $\{q_i\}$. We want to minimize the energy of this state with respect to the q_i, which is expressed as

$$
\begin{align}
E_T &= \frac{\langle\Psi_T|\mathcal{H}|\Psi_T\rangle}{\langle\Psi_T|\Psi_T\rangle} \tag{4.2.2a} \\
&= \frac{\sum_R \langle\Psi_T|H|R\rangle\langle R|\Psi_T\rangle}{\sum_R \langle\Psi_T|R\rangle\langle R|\Psi_T\rangle} \tag{4.2.2b} \\
&= \frac{\sum_R E_T(R)\langle\Psi_T|R\rangle\langle R|\Psi_T\rangle}{\sum_R \langle\Psi_T|R\rangle\langle R|\Psi_T\rangle}, \tag{4.2.2c}
\end{align}
$$

where the *local energy* in R is defined by

$$E_T(R) = \frac{\langle\Psi_T|H|R\rangle}{\langle\Psi_T|R\rangle}. \tag{4.2.3}$$

If one would want to calculate this expression exactly, one would have to calculate $E_T(R)$ in all possible configurations R in configuration space. For the half-filled Hubbard model with an equal number of up and down spins on a modestly sized system of, say, 100 sites, that would mean approximately $2^{200} \approx 10^{70}$ configurations. In the limit of infinite U, still about 10^{35} configurations would remain. It may be clear that, even with modern computer power, this calculation is not feasible.

Instead, one considers the following expression:

$$E_T = \frac{\sum_R E_T(R)p(R)}{\sum_R p(R)}, \tag{4.2.4}$$

where one now interprets

$$p(R) = \langle\Psi_T|R\rangle\langle R|\Psi_T\rangle \tag{4.2.5}$$

as a probability distribution. One has to choose a set of, say, N, configurations, distributed according to the square of the trial wave function. Summing over those configurations only, the expression

$$E_{VMC} = \frac{\sum_R' E_T(R)}{N} \tag{4.2.6}$$

then is an estimate for E_T, and it will converge to E_T in the limit of large N.

Thus, in the VMC procedure, one chooses a limited number of configurations according to the trial wave function, and simply averages over all local values of the energy obtained. This average is subject to statistical fluctuations, as can be seen by choosing different sets of N configurations, and one can easily calculate the statistical error as

$$\delta E = \frac{1}{\sqrt{N}}\sqrt{\langle E^2\rangle - \langle E\rangle^2}, \tag{4.2.7}$$

where

$$\langle E \rangle = E_{\text{VMC}} \tag{4.2.8}$$

and

$$\langle E^2 \rangle = \frac{\sum_R' E_T^2(R)}{N}. \tag{4.2.9}$$

as usual.

There are several ways to choose the set of configurations. The most common one is the Metropolis scheme [23], which can be used to generate configurations by creating paths in configuration space. Starting from an arbitrary configuration, one selects a nearby configuration according to a predefined transition probability. One accepts the new configuration with a probability that depends on the energy difference ΔE between both configurations. In particular, if the new configuration has lower energy ($\Delta E < 0$), it is always accepted, otherwise it is accepted with probability $e^{-\beta \Delta E}$. In this way, one generates a Markov chain of configurations, which can then be proved to be distributed according to the trial wave function.

In VMC, it is important that the configurations be distributed properly, in order to obtain the correct result. For the method described in the following section, where the configurations are only used as a starting point for generating sequences in phase space, it is less important to establish the correct starting distribution. There, even a random set of configurations may suffice in some cases.

4.3   Green function Monte Carlo

A more sophisticated way to calculate ground-state properties is the *Green Function* Monte Carlo (GFMC) scheme. Here, one combines the random sampling of the configuration space with a simple method to filter out the ground state from the trial wave function.

We consider the following operator:

$$\mathcal{F} = 1 - \tau(\mathcal{H} - w). \tag{4.3.1}$$

If we take w to be the lowest eigenvalue of $\mathcal{H}$, the ground state of $\mathcal{H}$ is an eigenstate of $\mathcal{F}$ with eigenvalue 1 (in Section 5.2 we will discuss how this is implemented in practice). We may assume that the spectrum of eigenvalues of $\mathcal{H}$ is bounded, as is always the case in the finite systems we consider. If we then choose

$$\tau < \frac{2}{E_{\text{max}} - w}, \tag{4.3.2}$$

where E_{max} is the largest eigenvalue of $\mathcal{H}$, all other eigenvalues of $\mathcal{F}$ are between -1 and 1. We let $\mathcal{F}$ operate repeatedly on a trial state, of which we assume the wave function to be known in all configurations:

$$|\Psi^{(n)}\rangle = \mathcal{F}^n |\Psi_T\rangle. \tag{4.3.3}$$

If $|\Psi_T\rangle$ has any overlap with the ground state of $\mathcal{H}$, $|\Psi^{(n)}\rangle$ will closely resemble this ground state for large n, due to the fact that at every step the ground-state component of $|\Psi_T\rangle$ is

multiplied by a larger number than all other components. One can in fact use any operator that has this property; in continuum problems, one generally uses the exponential operator $e^{-\tau \mathcal{H}}$. For our purposes the simple linear form (4.3.1) is more convenient.

We approximate the energy of the ground state by the *mixed estimate*

$$E^{(n)} = \frac{\langle \Psi_T | \mathcal{H} | \Psi^{(n)} \rangle}{\langle \Psi_T | \Psi^{(n)} \rangle} \tag{4.3.4a}$$

$$= \frac{\sum_{\mathcal{R}} \langle \Psi_T | H | R_n \rangle \left[\prod_{i=1}^{n} \langle R_i | F | R_{i-1} \rangle \right] \langle R_0 | \Psi_T \rangle}{\sum_{\mathcal{R}} \langle \Psi_T | R_n \rangle \left[\prod_{i=1}^{n} \langle R_i | F | R_{i-1} \rangle \right] \langle R_0 | \Psi_T \rangle}. \tag{4.3.4b}$$

Here $\mathcal{R} = \{R_0, R_1, R_2, ..., R_n\}$ is a path in the configuration space. The reason for using this expression, instead of writing $\langle \Psi^{(n)} |$, is the fact that we need to know the value of this wave function in all end configurations R_n that are present in the summation. For the trial wave function, we assumed this to be the case, but $\Psi^{(n)}$ we do not know at all, as it is only a theoretical result after applying $\mathcal{F}$, n times. Note that this way of estimating the value of an observable works well for all operators that commute with the Hamiltonian. If that is not the case, one should adapt the estimate; this will be discussed in Section 5.2.

As in the VMC scheme, it is impossible to consider all of the terms in this summation. Again, we need to restrict ourselves to a very limited number of terms. We have to adapt (4.2.4) in the following way to obtain a suitable expression for this case:

$$E^{(n)} = \frac{\sum_{\mathcal{R}} E_T(R_n) m(\mathcal{R}) p(\mathcal{R})}{\sum_{\mathcal{R}} m(\mathcal{R}) p(\mathcal{R})}. \tag{4.3.5}$$

where $p(\mathcal{R})$ is now interpreted as the probability of choosing the complete path $\mathcal{R}$. Each path carries a *weight* or *multiplicity* $m(\mathcal{R})$, and the *contribution* of a path to the average energy is $m(\mathcal{R}) E_T(R_n)$. $E_T(R)$ is as defined in Eq. (4.2.3). One has to make sure that $p(\mathcal{R})$ and $m(\mathcal{R})$, up to a constant factor, satisfy

$$p(\mathcal{R}) m(\mathcal{R}) = \langle \Psi_T | R_n \rangle \left[\prod_{i=1}^{n} \langle R_i | F | R_{i-1} \rangle \right] \langle R_0 | \Psi_T \rangle. \tag{4.3.6}$$

Only then, the energy $E^{(n)}$ will be exactly given by

$$E_{\text{GFMC}}^{(n)} = \frac{\sum_{\mathcal{R}}' E_T(R_n) m(\mathcal{R})}{\sum_{\mathcal{R}}' m(\mathcal{R})} \tag{4.3.7}$$

in the limit of taking infinitely many paths distributed according to $p(\mathcal{R})$. If one chooses the probability function in a proper way, one can interpret these paths as *random walks in configuration space*. One starts with a set of configurations, called *walkers*, which may be chosen randomly or distributed in some specified way. Then, one lets each of these configurations perform a random walk, prescribed by the transition probabilities that are defined by $p(\mathcal{R})$. Finally, one collects the information from the paths, i.e., the weights and the local energies, to calculate the energy (4.3.7). For lattice problems, this process and its mathematical justification have been described in more detail by Trivedi and Ceperley [24].

In the following section, some possibilities to choose $p(\mathcal{R})$ and $m(\mathcal{R})$ will be treated. First, we will discuss how one must interpret the set of configurations that results when each walker has moved n steps. We take into account the multiplicity $m(\mathcal{R})$ of each path, according to which one may redistribute the configurations (by means of a procedure called *branching*, see next section). Therefore we define the probability of finding a configuration R after n steps as

$$p(R) = \sum_{\{\mathcal{R} | R_n = R\}} m(\mathcal{R}) p(\mathcal{R}). \tag{4.3.8a}$$

This can be easily calculated, since $p(\mathcal{R})$ and $m(\mathcal{R})$ satisfy (4.3.6):

$$\begin{aligned} p(R) &= \langle \Psi_T | R \rangle \sum_{\{R_0, R_1, \cdots, R_{n-1}\}} \left[\prod_{i=1}^{n} \langle R_i | F | R_{i-1} \rangle \right] \langle R_0 | \Psi_T \rangle \tag{4.3.8b} \\ &= \langle \Psi_T | R \rangle \langle R | \mathcal{F}^n | \Psi_T \rangle \tag{4.3.8c} \\ &= \langle \Psi_T | R \rangle \langle R | \Psi^{(n)} \rangle. \tag{4.3.8d} \end{aligned}$$

Thus, the resulting configurations with their attached weights can be viewed to be distributed according to the product of the trial wave function and the wave function onto which it is being projected. Note that this result is independent of how $p(\mathcal{R})$ and $m(\mathcal{R})$ are chosen, as long as they satisfy (4.3.6).

4.4 Importance sampling

In principle, the result of the algorithm described in the previous section does not depend on the choice of $p(\mathcal{R})$ and $m(\mathcal{R})$. It can be shown, though, that not all choices lead to a good result in practice. One is interested in finding a correct value with a reasonable error bar, within a limited amount of time, i.e., by using a limited number of samples and a limited path length. The error bar and convergence to the correct result do in general depend on $p(\mathcal{R})$ and $m(\mathcal{R})$. In fact, for bad choices, the variance may increase very rapidly,[1] such that the correct answer can never be obtained. One usually needs to perform *importance sampling* in order to be able to find the correct result with any certainty. In this section, we treat a few ways to define $p(\mathcal{R})$ and $m(\mathcal{R})$, and we will give a general prescription for using a *guiding function* to do importance sampling.

One apparently simple way to choose $p(\mathcal{R})$ and $m(\mathcal{R})$, which seems straightforward, would be to take the probability to find a path the same for all paths, and therefore

$$m(\mathcal{R}) = \langle \Psi_T | R_n \rangle \left[\prod_{i=1}^{n} \langle R_i | F | R_{i-1} \rangle \right] \langle R_0 | \Psi_T \rangle. \tag{4.4.1}$$

It is, however, in general not possible to do so. In GFMC, paths are never selected as a whole, in one stroke, but only step by step. In a practical situation,

---

[1] In any case, except when the trial wave function is exactly equal to an eigenfunction of $\mathcal{H}$, the variance increases exponentially with the path length. In this section some indications of the causes for this behavior will be given, and a method to suppress the variance (branching) will be discussed. For a more detailed analysis, see the review by De Raedt and Von der Linden [25].

1. the starting configuration can be chosen randomly, and

2. from any configuration, all possible continuations of the path can be given equal probability.

These conditions can be easily satisfied, but they are not sufficient to give all paths equal probability of appearing in the summation (4.3.7). For that, one would also need the condition that

3. the number of possible continuations is the same from each configuration.[2]

In an arbitrary system, where this is generally not the case, it is usually very difficult to choose the paths truly randomly, thus making this definition of $m(\mathcal{R})$ useless. However, one can redefine the multiplicity on the basis of conditions 1 and 2 being satisfied. Writing $C(R)$ for the number of possible continuations of a path in a configuration R, one easily finds that

$$p_{random}(\mathcal{R}) \sim \prod_{i=1}^{n} \frac{1}{C(R_{i-1})} \tag{4.4.2}$$

and

$$m_{random}(\mathcal{R}) = \langle \Psi_T | R_n \rangle \left[\prod_{i=1}^{n} C(R_{i-1}) \langle R_i | F | R_{i-1} \rangle \right] \langle R_0 | \Psi_T \rangle \tag{4.4.3}$$

satisfy (4.3.6). This is the simplest definition for $p(\mathcal{R})$ and $m(\mathcal{R})$, with which one can actually perform GFMC. It now becomes clear that the paths can be interpreted as random walks of configurations in the phase space. They are prescribed by a stochastic matrix, which gives the transition probabilities for a walker R to go to any configuration R'. Note once more that the set of walkers that is obtained after n steps represents the wave function $\Psi^{(n)}$, as described in the previous section by Eq. (4.3.8d).

Intuitively one assumes that the most important contributions to the calculation of a property come from regions where the absolute value of the wave function is large, as the local value of the property gets weighted by the square of the wave function. Thus, assuming that the trial wave function is a good approximation of the ground state, the first improvement on the completely random sampling scheme is to have the starting configurations of the paths distributed according to the trial wave function. We then get

$$p_{start}(\mathcal{R}) = \left[\prod_{i=1}^{n} \frac{1}{C(R_{i-1})} \right] \langle \Psi_T | R_0 \rangle \langle R_0 | \Psi_T \rangle \tag{4.4.4}$$

and

$$m_{start}(\mathcal{R}) = \frac{\langle \Psi_T | R_n \rangle}{\langle \Psi_T | R_0 \rangle} \left[\prod_{i=1}^{n} C(R_{i-1}) \langle R_i | F | R_{i-1} \rangle \right]. \tag{4.4.5}$$

---

[2]It is left as a simple exercise for the reader to check that, in general, equal probabilities for all paths cannot be obtained by considering the path step by step from the starting point. One can easily see this, e.g., for paths of two steps in a configuration space that looks like ▱ , where the bonds represent possible steps. Considering only hopping terms, and choosing the starting configurations randomly, one finds 20 paths which have probability $\frac{1}{24}$, and 6 with probability $\frac{1}{36}$ of being generated.

This is not sufficient to adequately suppress the statistical fluctuations in all cases. In fact, for large n, the effect of choosing the starting configurations in some special way is completely lost along the path. Thus, the next step is to vary the probabilities along the whole path, and perform importance sampling. We give a general prescription in terms of a *guiding function* Ψ_G. In many cases the trial wave function itself can be used for this purpose, but it can also be a different function. In any case, a good guiding function resembles the exact ground-state wave function as much as possible, and it does not restrict the paths to a smaller region of the phase space than is allowed by the Hamiltonian (it should, e.g., not be equal to zero in configurations where the exact wave function is non-zero). For the moment, we only require the guiding function to be positive in all configurations.

We define the probability of finding a path in the following way:

$$p_{\text{guiding}}(\mathcal{R}) = p_0(R_0) \prod_{i=1}^{n} p(R_i \leftarrow R_{i-1}), \tag{4.4.6}$$

where the probability to start in a configuration R is

$$p_0(R) = \langle \Psi_T | R \rangle \langle R | \Psi_T \rangle, \tag{4.4.7}$$

and the transition probability for going from R to R' is given by

$$p(R' \leftarrow R) = \frac{\langle \Psi_G | R' \rangle \langle R' | F | R \rangle}{\langle \Psi_G | R \rangle m(R)}. \tag{4.4.8}$$

Note that $p(R' \leftarrow R)$ is in fact the stochastic matrix prescribing the random walks in configuration space. The *weight factor* $m(R)$, serving as a normalization factor for $p(R' \leftarrow R)$, is defined as

$$m(R) = \sum_{R'} \frac{\langle \Psi_G | R' \rangle \langle R' | F | R \rangle}{\langle \Psi_G | R \rangle} = \frac{\langle \Psi_G | F | R \rangle}{\langle \Psi_G | R \rangle}, \tag{4.4.9}$$

and the total weight of the path is given by

$$m_{\text{guiding}}(\mathcal{R}) = \frac{\langle \Psi_G | R_0 \rangle}{\langle \Psi_T | R_0 \rangle} \left[\prod_{i=1}^{n} m(R_{i-1}) \right] \frac{\langle \Psi_T | R_n \rangle}{\langle \Psi_G | R_n \rangle}. \tag{4.4.10}$$

One can easily verify that the relation (4.3.6) holds for p_{guiding} and m_{guiding}.

This definition for $p(\mathcal{R})$ and $m(\mathcal{R})$ can be straightforwardly implemented in a GFMC program. The starting configurations are distributed according to the trial wave function, e.g., by means of the Metropolis algorithm. From each configuration, the next configuration is determined randomly, where the probability to step to a neighboring configuration is given by (4.4.8). At the same time, the multiplicity is adapted by a factor given by (4.4.9). If the trial wave function and the guiding function are not the same, the multiplicity has to be corrected by factors containing their values in the starting and end configurations of the path, as is shown in Eq. (4.4.10).

The variance can be reduced significantly by making a good guess for the guiding function. In the ideal case of the guiding function being equal to the ground-state wave function,

the variance is zero. However, if this is not the case, the variance still increases exponentially with the number of iterations n. In order to further improve the accuracy of the calculation, we use the concept of *branching* [24, 26, 27]. In this *birth-and-death* process, the set of configurations is re-expressed according to the accumulated multiplicities: configurations with a low weight are discarded, and multiple copies are made of configurations with a large weight. The multiplicity of each walker is turned into an integer number, by means of a random process to treat the decimal part. Subsequently, a number of copies of the concerning configuration is created according to this integer multiplicity. The total weight of all walkers remains approximately the same in this procedure; in the new set, each configuration starts with weight 1. Note that Eq. (4.3.8d) thus gives the distribution of the walkers *after* the branching process. In this way, large fluctuations in the weights are suppressed, and therefore the variance is further reduced. One should take care, though, not to apply the branching procedure too often, as it may introduce a bias of the set of walkers. This could cause the energy measured to deviate from the exact energy.

4.5 The sign problem

In the previous sections, one aspect of the problem has been completely ignored: the sign of the wave function. In a boson system, where the wave function must be symmetric under an exchange of identical particles, one may usually assume that the ground-state wave function is of the same sign everywhere.[3] In that case, if one uses a trial wave function which is also of the same sign everywhere, the algorithm described before may be perfectly suitable for calculating the properties of the ground state of the system. In a fermion system, the wave function must be antisymmetric, and thus can be of different sign in different configurations. This causes a significant problem, the *sign problem*, which is detrimental for the accuracy by which results may be obtained. In this section we will examine the source of this problem, and briefly indicate what may be done to circumvent this problem.

Let us take a closer look at the denominator in Eq. (4.3.4b), which denotes the mixed estimate for the energy:

$$\sum_{\mathcal{R}} \langle \Psi_T | R_n \rangle \left[\prod_{i=1}^{n} \langle R_i | F | R_{i-1} \rangle \right] \langle R_0 | \Psi_T \rangle. \tag{4.5.1}$$

If we make the following assumptions:

1. Ψ_T is of the same sign everywhere, and

2. All matrix elements of $\mathcal{F}$ are non-negative,

then trivially all terms in the summation (4.5.1) are non-negative. In a fermion system, it is clear that the first assumption does not hold. On a lattice, also the second assumption need

---

[3]Note that, on a lattice, this is not necessarily true. For instance, in case of next-nearest-neighbor hopping, sign changes may occur. See, e.g., Refs. 28 or 29, and references therein.

not be true, causing a possible sign problem also in bosonic systems. Note that *weight* is in fact a misnomer for $m(\mathcal{R})$ in such cases, as it can very well become negative.

For "paths" of length zero, i.e., in the VMC scheme, no sign problem exists. At each subsequent step, some percentage of the walkers goes to the region of opposite sign. For increasing path lengths, the percentage of walkers that has obtained a sign along their walks rapidly increases.[4] For large n, about half of the terms in (4.5.1) may be negative. The sum of all the terms becomes exponentially smaller than the sum of the positive or the negative terms separately. In the statistical process, this leads to an increasingly inaccurate result, due to the fluctuations which may become much larger than the actual value desired. Therefore the calculation becomes useless.

This problem cannot be solved. It has been tried in many different ways to find a solution for the sign problem [30–35], but the sign is an intrinsic feature of the system, and it cannot be removed without causing similar problems at a different level. De Raedt and Von der Linden [25] even state that 'it is likely that this minus-sign problem is detrimental for the GFMC method to be applicable to lattice fermion problems.' As we will see, this is too pessimistic a point of view, and we will describe a method by which lattice fermions can be tackled.

A seemingly unsophisticated approach is to improve the trial wave function as much as possible, and hope that convergence to the exact result can be reached before the sign problem arises. For most systems, this is very difficult, as modifying the trial wave function is a far from trivial thing to do, and one must take care that the prescription for the trial function does not become too complicated to handle in the simulations. This approach has been followed by Chen and Lee [36] for the simple *t-model*, which is equivalent to the Hubbard model with infinite U. They use what they call a variational Lanczos scheme, in order to improve on the trial wave function. Another approach, which has been successfully applied to continuum systems for quite some time [24, 26, 27, 37–40], is to discard all steps that generate a sign, when a walker would cross the nodal boundary of the trial wave function. In that case, one makes an approximation of the true result, but one can make sure that a variational principle is obeyed. In fact, having continuous coordinates, one can do this rather easily by decreasing the step size in the system, thus avoiding the unpredictable truncation error made when a step that causes a change of sign is discarded.

On a lattice, one cannot straightforwardly apply the latter idea, due to the discreteness of the phase space. If one would simply discard all wrong paths, one would effectively sample a different Hamiltonian which would not lead to a firm relation between the result found and the true result. It is possible, though, to adapt this approach for the lattice in such a way that it yields an upper bound for the energy of the true system. This *fixed-node* Monte Carlo (FNMC) method will be the topic of Section 4.6. In Section 4.7, the FNMC results are subsequently used as input for the straightforward approach, which we call the *power method*, in an attempt to find an estimate for the properties of the exact ground state.

---

[4]In a typical example for the Hubbard model on a small system (see Section 5.3), the percentage of steps that change the sign is of the order of 5%, and the sign problem arises after some 15 to 20 steps.

4.6   Fixed-node Monte Carlo

One possible way to deal with the sign problem is to avoid it. Let us, again, consider the denominator (4.5.1) in the expression for the mixed estimate of the energy. As was explained in the previous section, a sign problem can be introduced at any step ($R_i \leftarrow R_{i-1}$) along a path $\mathcal{R}$ if the trial wave function can be of opposite sign in R_i and R_{i-1}, *or* if the matrix element $\langle R_i|F|R_{i-1}\rangle$ can be negative. We can denote the *sign-flip condition* in the following way: if, for any pair of configurations R and R',

$$\langle\Psi_T|R'\rangle\langle R'|F|R\rangle\langle R|\Psi_T\rangle < 0, \tag{4.6.1a}$$

the GFMC calculation of the properties of the system suffers from the sign problem. One can rewrite this condition in terms of the Hamiltonian:

$$\langle\Psi_T|R'\rangle\langle R'|H|R\rangle\langle R|\Psi_T\rangle > 0. \tag{4.6.1b}$$

The apparent solution to the sign problem is to discard all of the steps in configuration space that satisfy (4.6.1b). This approach is called *Fixed-Node Monte Carlo* (FNMC).

The fixed-node method was developed for the case in which the electron coordinates are continuous variables [26, 41]. One then has to deal with kinetic terms of negative sign only, and the *nodal surface* of a trial wave function is uniquely defined by the set of configurations where it vanishes. The fixed-node constraint can be implemented by imposing the boundary condition that Ψ must vanish on the nodal surface of Ψ_T. In the limit of sufficiently small step sizes, one can make sure that Eq. (4.6.1b) is never violated since R and R' come closer together and Ψ_T will vanish. In this way one obtains the lowest energy under the condition that the wave function has the same nodal surface as the trial wave function. This energy yields an upper bound to the true ground-state energy; in practice, very accurate estimates for the ground-state energy of continuum problems can be obtained.

On a lattice, we have to deal with discrete steps, and we have to treat the steps that cause a change of sign in a different way. If we discard these steps, we in fact modify our Hamiltonian by removing a number of off-diagonal elements, and we do not have sufficient information on the Hamiltonian obtained in this way to be able to give a relation between the ground-state energies of both Hamiltonians. Indeed, the ground-state energy of the new Hamiltonian is generally not the same as the energy we are interested in. It may well be below the desired energy,[5] such that it does not give a bound on the exact ground-state energy of the original Hamiltonian.

This problem can be solved by defining an effective Hamiltonian that fulfills the condition of lacking sign-flipping steps, but that nevertheless yields an upper bound for the exact ground-state energy. In our implementation, we use a Hamiltonian $\mathcal{H}_{eff}$, of which the off-diagonal matrix elements are given by

$$\langle R'|H_{eff}|R\rangle = 0 \qquad \text{if Eq. (4.6.1b) holds,} \tag{4.6.2a}$$
$$= \langle R'|H|R\rangle \qquad \text{otherwise,} \tag{4.6.2b}$$

[5]This was found in an earlier attempt to perform FNMC on lattice fermions by An and Van Leeuwen [42].

and the diagonal elements by

$$\langle R|H_{\text{eff}}|R\rangle = \langle R|H|R\rangle + \langle R|V_{\text{sf}}|R\rangle. \tag{4.6.2c}$$

Thus, instead of only truncating the Hamiltonian by discarding all the sign-flipping terms, we replace those terms by non-negative diagonal terms, which can be viewed as local potentials that are present due to the *sign-flip boundary*. These potentials read:

$$\langle R|V_{\text{sf}}|R\rangle = \sum_{R'}^{\text{sf}} \langle R|H|R'\rangle \frac{\langle R'|\Psi_{\text{T}}\rangle}{\langle R|\Psi_{\text{T}}\rangle}. \tag{4.6.3}$$

The sum is over all R' for which (4.6.1b) holds, denoted by sf above the summation sign.

A somewhat similar procedure, called *model-locality*, has been used by Mitáš *et al.* [39] in continuum problems with a *non-local* potential that arises from replacing ionic cores with pseudo-potentials. As in a lattice system, they cannot solve the problem of crossing a node by making the step size of the walkers continuously smaller, because the non-local potential connects configurations at finite distances. In their approach the unwanted off-diagonal terms are truncated, and replaced by diagonal contributions as in Eq. (4.6.3), but with the sum over all R', not just over sign-flip configurations. A drawback of the model-locality procedure is that one does not obtain an upper bound for the ground-state energy.

Proof for the upper bound

We now show that the prescription given above for $\mathcal{H}_{\text{eff}}$ does indeed lead to an upper bound for the ground-state energy of $\mathcal{H}$. In order to do so, we define a truncated Hamiltonian $\mathcal{H}_{\text{tr}}$, and a *sign-flip* Hamiltonian $\mathcal{H}_{\text{sf}}$, by

$$\mathcal{H} = \mathcal{H}_{\text{tr}} + \mathcal{H}_{\text{sf}}, \tag{4.6.4}$$
$$\mathcal{H}_{\text{eff}} = \mathcal{H}_{\text{tr}} + \mathcal{V}_{\text{sf}}, \tag{4.6.5}$$

where the diagonal elements of H_{tr} are

$$\langle R|H_{\text{tr}}|R\rangle = \langle R|H|R\rangle, \tag{4.6.6}$$

and its off-diagonal elements are given by

$$\langle R|H_{\text{tr}}|R'\rangle = \langle R|H_{\text{eff}}|R'\rangle. \tag{4.6.7}$$

$\mathcal{V}_{\text{sf}}$ is the sign-flip potential, for which the matrix elements are given by (4.6.3), and H_{sf} contains only the off-diagonal elements of H which are put to zero in the effective Hamiltonian. We now take *any* state

$$|\Psi\rangle = \sum_{R} |R\rangle\langle R|\Psi\rangle, \tag{4.6.8}$$

and we compare its energy with respect to $\mathcal{H}$ and to $\mathcal{H}_{\text{eff}}$:

$$\Delta E = \langle\Psi|(\mathcal{H}_{\text{eff}} - \mathcal{H})|\Psi\rangle \tag{4.6.9a}$$
$$= \langle\Psi|(\mathcal{V}_{\text{sf}} - \mathcal{H}_{\text{sf}})|\Psi\rangle. \tag{4.6.9b}$$

ΔE can be written explicitly in terms of the matrix elements of V_{sf} and H_{sf}:

$$\Delta E = \sum_{R} \langle \Psi | R \rangle \left[\langle R | V_{sf} | R \rangle \langle R | \Psi \rangle - \sum_{R'} \langle R | H_{sf} | R' \rangle \langle R' | \Psi \rangle \right]. \tag{4.6.9c}$$

Using (4.6.3), we rewrite this expression in terms of the matrix elements of H:

$$\Delta E = \sum_{R} \langle \Psi | R \rangle \left[\sum_{R'}^{sf} \langle R | H | R' \rangle \frac{\langle R' | \Psi_T \rangle}{\langle R | \Psi_T \rangle} \langle R | \Psi \rangle - \sum_{R'}^{sf} \langle R | H | R' \rangle \langle R' | \Psi \rangle \right]. \tag{4.6.9d}$$

In this double summation each pair of configurations R and R' occurs twice. We combine these terms and rewrite (4.6.9d) as a summation over pairs:

$$\Delta E = \sum_{(R,R')}^{sf} \langle R | H | R' \rangle \left[\frac{\langle R' | \Psi_T \rangle}{\langle R | \Psi_T \rangle} \langle \Psi | R \rangle \langle R | \Psi \rangle + \frac{\langle R | \Psi_T \rangle}{\langle R' | \Psi_T \rangle} \langle \Psi | R' \rangle \langle R' | \Psi \rangle \right.$$
$$\left. - \langle \Psi | R \rangle \langle R' | \Psi \rangle - \langle \Psi | R' \rangle \langle R | \Psi \rangle \right]. \tag{4.6.9e}$$

Denoting by $s(R, R')$ the sign of the matrix element $\langle R | H | R' \rangle$, and using the fact that for all terms in this summation the sign-flip condition (4.6.1b) is satisfied, we can finally write ΔE as

$$\Delta E = \sum_{(R,R')}^{sf} |\langle R | H | R' \rangle| \left| \langle R | \Psi \rangle \sqrt{\left| \frac{\langle R' | \Psi_T \rangle}{\langle R | \Psi_T \rangle} \right|} - s(R, R') \langle R' | \Psi \rangle \sqrt{\left| \frac{\langle R | \Psi_T \rangle}{\langle R' | \Psi_T \rangle} \right|} \right|^2. \tag{4.6.9f}$$

Note that we do not have to worry about configurations R where $\langle R | \Psi_T \rangle = 0$: they do not occur in this summation. Obviously, ΔE is positive for any wave function Ψ. Thus the ground-state energy of $\mathcal{H}_{eff}$ is an upper bound for the ground-state energy of the original Hamiltonian $\mathcal{H}$.

Variation of the trial state

Let us consider the situation where we use the exact ground state $|\Psi_0\rangle$ of $\mathcal{H}$, with energy E_0, as trial state. Obviously, for the method to be useful, it is desirable that in that case the effective Hamiltonian has the same ground-state energy E_0, and the same ground state $|\Psi_0\rangle$, as that would make it possible to find the true ground state by varying the trial wave function in some way. In Eq. (4.6.9f) we substitute Ψ_0 for Ψ_T. In order to have ΔE equal to zero, each individual term in the summation (4.6.9f) has to vanish, thus leading to

$$\langle R | \Psi \rangle \sqrt{\left| \frac{\langle R' | \Psi_0 \rangle}{\langle R | \Psi_0 \rangle} \right|} - s(R, R') \langle R' | \Psi \rangle \sqrt{\left| \frac{\langle R | \Psi_0 \rangle}{\langle R' | \Psi_0 \rangle} \right|} = 0, \tag{4.6.10}$$

or,

$$\frac{\langle R | \Psi \rangle}{\langle R' | \Psi \rangle} = s(R, R') \left| \frac{\langle R | \Psi_0 \rangle}{\langle R' | \Psi_0 \rangle} \right| = \frac{\langle R | \Psi_0 \rangle}{\langle R' | \Psi_0 \rangle} \tag{4.6.11}$$

for all sign-flipping pairs (R, R'). This condition is trivially fulfilled for $\Psi = \Psi_0$. Thus, the true ground-state energy can be reached by variation of the trial wave function. One can further extend this result to show that as $\Psi_T \to \Psi_0$ the error in the fixed-node energy will be second order in the difference, $\Psi_T - \Psi_0$, with positive coefficient.

An example of the fixed-node procedure, applied to a small system which can be calculated exactly, is given in the appendix.

4.7 Beyond fixed-node: the power method

In the previous section, we have presented a method to perform GFMC on lattice fermions by using a modified Hamiltonian, such that the calculation does not suffer from the sign problem. The method yields an upper bound for the true ground-state energy. In this section we will show that it is possible to use the result of the FNMC procedure as input for a straightforward GFMC approach, the *power method*, which does suffer from the sign problem, but which may yield convergent results before the fluctuations destroy the accuracy of the measurements. It strongly resembles the *nodal relaxation method*, which has been described for a continuum problem by Ceperley and Alder [38]. It is also similar to the power method described by Chen and Lee [36] for a lattice system.

Let us recall the expression (4.3.8d) for the distribution of the configurations in the GFMC procedure after n steps. For the FNMC procedure, as we sample a different Hamiltonian, we have to slightly adapt this expression:

$$p_{\text{FN}}^{(n)}(R) = \langle \Psi_T | R \rangle \langle R | \Psi_{\text{eff}}^{(n)} \rangle. \tag{4.7.1}$$

where $|\Psi_{\text{eff}}^{(n)}\rangle$ is supposed to be a good approximation of the ground state $|\Psi_{\text{eff}}\rangle$ of the effective Hamiltonian $\mathcal{H}_{\text{eff}}$. The idea is now that $|\Psi_{\text{eff}}\rangle$ is sufficiently close to the ground state $|\Psi_0\rangle$ of $\mathcal{H}$, such that in the ordinary GFMC procedure convergence can be obtained within a relatively small number of steps. Thus, we modify (4.3.3) in the following way:

$$|\Psi_P^{(n)}\rangle = \mathcal{F}^n |\Psi_{\text{eff}}\rangle. \tag{4.7.2}$$

Then, we also have to modify the expression (4.3.4b) for the mixed estimate of the energy:

$$E_P^{(n)} = \frac{\langle \Psi_T | \mathcal{H} | \Psi_P^{(n)} \rangle}{\langle \Psi_T | \Psi_P^{(n)} \rangle} \tag{4.7.3a}$$

$$= \frac{\sum_{\mathcal{R}} \langle \Psi_T | H | R_n \rangle \left[\prod_{i=1}^{n} \langle R_i | F | R_{i-1} \rangle \right] \langle R_0 | \Psi_{\text{eff}} \rangle}{\sum_{\mathcal{R}} \langle \Psi_T | R_n \rangle \left[\prod_{i=1}^{n} \langle R_i | F | R_{i-1} \rangle \right] \langle R_0 | \Psi_{\text{eff}} \rangle}. \tag{4.7.3b}$$

Note that we still cannot put anything but $\langle \Psi_T |$ in this expression, as that is the only available information we have in each configuration. However, in the starting configurations of the paths, the trial wave function has been replaced by the effective ground-state wave function, of which we have information through the distribution of the walkers. We have to pay

further attention to the fact that sign changes can now occur, which have to be embedded in the transition probabilities and multiplicities. We can use the following guiding procedure, adapted from the end of Section 4.4. The probability to start in a configuration R is given by the fixed-node resulting distribution:

$$p_0(R) = p_{\text{FN}}^{(n)}(R).$$ (4.7.4)

The transition probabilities are given by

$$p(R' \leftarrow R) = \frac{\langle \Psi_G | R' \rangle \, |\langle R' | F | R \rangle|}{\langle \Psi_G | R \rangle m(R)},$$ (4.7.5)

with the modified weight factor

$$m(R) = \sum_{R'} \frac{\langle \Psi_G | R' \rangle \, |\langle R' | F | R \rangle|}{\langle \Psi_G | R \rangle}.$$ (4.7.6)

The absolute value has to be taken to make sure that the transition probabilities remain positive, such that the procedure can still be interpreted as a stochastic walk. With these expressions replacing (4.4.7), (4.4.8), and (4.4.9), the equations (4.4.6) and (4.4.10) can again be used for the total probability and the total "weight" of a path, respectively. Note that the weights can now be negative, due to the possibly different signs of the wave function in the starting and end configurations of the path.

Appendix: Example of fixed-node calculation

In this appendix, we give an illustration of how the effective Hamiltonian is created for a very simple small system. All steps can be straightforwardly generalized to more complicated systems.

Consider the Hamiltonian for non-interacting spinless fermions

$$\mathcal{H} = \mathcal{H}_{\text{kin}} = -t \sum_{\langle i, j \rangle} c_i^\dagger c_j$$ (1)

on a loop of 4 sites with 2 particles. We define *configurations of labeled fermions* $[i_1 i_2]$, where particle j ($1 \leq j \leq 2$) sits on site i_j ($1 \leq i_j \leq 4$). We number the sites, as follows:

A valid (i.e., antisymmetric) fermion wave function Ψ must satisfy $\Psi([ij]) = -\Psi([ji])$. The configuration space of this system consists of 12 configurations, and can be depicted as follows:
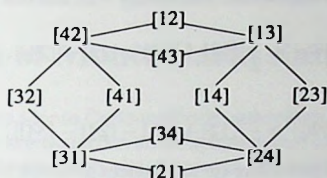
The lines (or bonds) represent valid steps in this space. The matrix elements of the Hamiltonian for this system $\langle[kl]|H|[mn]\rangle$ are $-t$ if there is a bond between $[kl]$ and $[mn]$ (in that case $k = m$ or $l = n$ must hold), or 0 otherwise. The ground state of this Hamiltonian is symmetric under exchange of the particles, and we have to restrict the wave function explicitly to be antisymmetric in order to find a valid fermion wave function. To obtain a Hamiltonian $\underline{H}$ which describes the fermion problem only, we define *antisymmetric states* $[\underline{kl}]$, which are antisymmetrized combinations of the configurations:

$$[\underline{kl}] = \frac{1}{\sqrt{2}} \left([kl] - [lk]\right). \tag{2}$$

In this way each pair of configurations $[kl]$ and $[lk]$ produces two states, $[\underline{kl}]$ and $[\underline{lk}]$, which only differ by their sign. One has the freedom to choose one of these states to obtain only one state per pair of configurations, and one can calculate the resulting Hamiltonian for the $[\underline{kl}]$:

$$\begin{aligned}
\langle[\underline{kl}]|H|[mn]\rangle &= \frac{1}{2} \sum_{\Pi_1} \sum_{\Pi_2} \mathrm{sg}(\Pi_1)\mathrm{sg}(\Pi_2)\, \langle\Pi_1[kl]|H|\Pi_2[mn]\rangle \\
&= \mathrm{sg}(\Pi)\langle[\underline{kl}]|H|\Pi[mn]\rangle,
\end{aligned} \tag{3}$$

where Π_1 and Π_2 denote permutations of the two particles, $\mathrm{sg}(\Pi)$ gives the sign of a permutation Π, and $\Pi[mn]$ is the permutation of $[mn]$ that can be reached by one step from $[kl]$, such that $\langle[kl]|H|\Pi[mn]\rangle = -t$. We can again denote the Hamiltonian in a picture, representing matrix elements $-t$ by thin lines and $+t$ by thick lines (we choose the $[\underline{kl}]$ with $k < l$; other choices give different pictures but the same results):
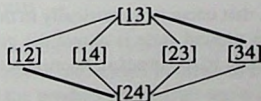
This structure fully contains the antisymmetry, and the corresponding Hamiltonian gives all the information there is on the fermion problem. Solving this 6×6 eigenvalue problem, we

find that the ground-state is degenerate, with energy $-2t$, and that possible ground states are

$$|\Psi_0\rangle = \frac{1}{2} \left|[13] + [14] + [23] + [24]\right\rangle, \tag{4}$$

and

$$|\Psi'_0\rangle = \frac{1}{2} \left|[12] + [13] - [24] - [34]\right\rangle. \tag{5}$$

It is easy to generalize this procedure for any system of lattice fermions.

Let us now consider a trial state, and calculate the effective Hamiltonian according to our fixed-node prescription. The trial wave function defines the *nodal regions* through its sign in all states, and, because we are working with negative hopping terms, the sign-flip constraint reduces to sign changes of the wave function only. We take a very simple trial state:

$$|\Psi_T\rangle = \frac{1}{\sqrt{6}} \left|[12] + [13] + [14] + [23] + [24] + [34]\right\rangle. \tag{6}$$

purposely chosen such that we only have to slightly adapt the previous picture to denote the effective Hamiltonian:
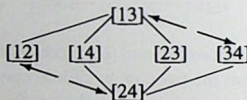
Here the thin lines are still matrix elements $-t$. The thick lines have been cut (we do not allow these steps in the effective Hamiltonian) and replaced by arrows, indicating diagonal matrix elements. In this simple case, they all become $+t$, because we have chosen equal weights for all the states in the trial wave function. The (non-degenerate) ground state of this effective Hamiltonian is

$$|\Psi_0^{\text{eff}}\rangle = 0.165|[12] + [34]\rangle + 0.448|[13] + [24]\rangle + 0.523|[14] + [23]\rangle \tag{7}$$

with energy $-1.709t$. Note that, e.g., the states [12] and [24] do *not* have the same wave function in this ground state, while they do in the trial state. As one could have expected from symmetry considerations, the wave function is the same in states that have an equivalent position in the picture, i.e., that occur symmetrically in the effective Hamiltonian. Note that the energy of the effective ground state is above the ground-state energy of the true problem, as it should be according to the proof in Section 4.6.

5 Application of Monte Carlo simulations

The implementation of the Monte Carlo methods, described in the previous chapter, is clarified. The methods are tested on small systems that can also be diagonalized exactly. Application of Monte Carlo simulations to the problem of ferromagnetism in the Hubbard model is discussed.

5.1 Introduction

Having described the theory of Monte Carlo simulations in the previous chapter, we now turn to the application of these methods to the Hubbard model. First, we will present some important features of the implementation of the Monte Carlo technique in computer programs, in Section 5.2. We discuss the issues of which set of starting configurations to use, and how to determine the parameters w and τ in the expression (4.3.1) for the projection operator $\mathcal{F}$. We describe a standard way to calculate the trial wave function during the computer runs, and we comment on the calculation of observables whose operators do not commute with the Hamiltonian. Then, in Section 5.3, we perform these calculations for small systems for which exact calculations can be performed. In this way, we can test the stability of both the fixed-node and power formalisms. We discuss the necessity to have a good trial wave function in Section 5.4. Finally, in Section 5.5, we discuss possibilities to apply the Monte Carlo methods to the problem of ferromagnetism in the Hubbard model.

5.2 Implementation

We have written a set of computer programs by which the Monte Carlo simulations can be performed. Important ingredients of a simulation are the trial wave function and the set of walkers (the *ensemble*), which we calculate in separate programs. In this section, we will briefly describe the structure of the main program, by which the actual Monte Carlo simulation is carried out. Then, we will make some additional remarks on the important details.

Description of the program

The fixed-node program mainly consists of two parts. In the first part (called *thermalization*), a number of projection steps is performed in order to obtain an ensemble of walkers that is distributed according to the product of the trial wave function and the ground-state wave function. If this process has become stable, the actual calculation of the energy and other quantities of interest takes place in the second part, as discussed below.

The program starts with an ensemble of walkers, each of which has weight 1. For each subsequent walker, a list is prepared of all possible steps in configuration space, and the

probability for each step to be selected from the list is calculated (according to the prescriptions given in Section 4.4, and using the effective Hamiltonian defined in Section 4.6). At the same time, the local values of several quantities are calculated, which are used to determine a weighted average over all configurations visited. By means of a random number, one of the steps is selected, and all available information (including the location of the particles, the multiplicity, and the value of the trial wave function) is updated to be consistent with the step chosen. In this way, each walker makes some fixed number of steps, corresponding to applying the projection operator as many times. Note that 'stepping' can also mean 'staying in the same configuration', which is due to the diagonal elements of the projection operator. In that case, the local values contribute to the weighted average once more.

After this procedure has been followed for each walker, branching (described in Section 4.4) is used to reduce the variance of the multiplicities. In this process, the sum of the multiplicities remains approximately the same, but each surviving walker has multiplicity 1 after the branching. Thus, the number of walkers may increase or decrease, depending on the parameters chosen. One of the aims of thermalization is to find the appropriate value for the parameter w, such that the population of walkers remains stable. After the branching process is completed, the walkers start stepping again and the whole procedure is repeated.

During thermalization, measurement of the energy only serves as an indication for the stability of the ensemble, together with the fluctuation in the number of walkers. In the second part of the program, a number of measurements of the energy is done, which are supposed to be independent of each other. Each measurement is an average over a large number of steps, which may or may not include branching. After each measurement, the number of walkers is brought back to its original value before starting the next measurement,[1] in order to make sure that the ensemble does not shrink or grow too much. The final result is the average over all measurements, and the statistical error is determined from the fluctuations in the measurements. Here, one has to be aware of possible correlations between subsequent measurements, which may cause an underestimate of the error. Typically, in a fixed-node run one would start with a set of 1000 walkers, one would collect 100 measurements based on 500 steps of each walker, and branching would take place every 100 steps.

The program for doing the power calculations is quite similar. Instead of the effective Hamiltonian, the original Hamiltonian is used, and the prescription for the transition probabilities has to be adapted according to Section 4.7. An important difference is in the averaging of the physical quantities: the values are not averaged along the paths in this case, but only over the set of walkers. Thus, an average is obtained after one step, after two steps, etc., such that one can see whether the values stabilize as a function of the number of steps taken. In order to obtain sufficient accuracy, one would use of the order of a million walkers, to compensate for the fact that one cannot average along the paths. Depending on the values of the parameters used, a walker can only perform some 50 steps at the most before the sign problem arises. Branching cannot be applied in the power scheme.

---

[1]This usually coincides with the last branching procedure in the measurement, in which all multiplicities are multiplied by the same factor in order to obtain the total weight desired.

Choice of parameters

One of the important issues that has to be considered is the choice of the parameters w and τ. The other parameters (the number of walkers, the number of steps each walker takes in one measurement, the number of measurements, etc.) are mainly related to the accuracy by which one wants to obtain the results, and should be fixed accordingly. For w and τ some clear guidelines can be given, which should be followed in order to obtain a stable Monte Carlo run.

The parameter τ can be used to speed up the projection process. In the implementation, this can be seen from the fact that the probability for a single walker to stay in the same configuration in a specific step becomes smaller with increasing τ. In Eq. (4.3.2) we defined limits for τ such that the GFMC method would formally yield the ground state. However, in FNMC, there is an additional problem due to the sign-flip potential. As one can see from Eq. (4.6.3), there is no bound on the value that the sign-flip potential can take. Incidentally, this potential becomes so large, that the probability to stay as prescribed by Eq. (4.4.8) [with $R' = R$] becomes negative. If this does not happen too often, one can safely truncate these negative values to be zero, but one must beware of introducing a systematic error in this way. One can avoid this by checking that the results do not depend on τ in some reasonable range of this parameter. The influence of the potential can be reduced by choosing a smaller value for τ. Apart from this problem, one usually wants to choose τ as large as possible, within the limits dictated by the energy spectrum.

The influence of the parameter w can be easily understood by looking at the form of the operator $\mathcal{F}$, given by Eq. (4.3.1). Assuming that we are dealing with the ground state, we see that at each projection step the ground-state wave function is multiplied by 1 if w is exactly equal to the ground-state energy. If w is below the ground-state energy, this multiplicative factor is smaller than 1, otherwise it is larger than 1. In the implementation, this is directly related to the total weight carried by the walkers. If w equals the ground-state energy, the sum of the multiplicities tends to remain unchanged (apart from fluctuations), and therefore the number of walkers remains approximately stable in the branching process. If w is below this energy, the number of walkers decreases, whereas it increases if w is above the ground-state energy. In the fixed-node program, one can let w obtain a proper value in the thermalization process. Starting with a trial value, one can adapt its value by taking some average of the energy measured and the old value. After some time, one will find that the value of w remains almost constant, as well as the number of walkers. In this way, one in fact obtains a first approximation of the ground-state energy, which is then calculated more accurately in the second part of the program. In the power procedure, one can safely use a value for w that is slightly below the fixed-node result. As the number of steps in a power run is rather limited, this method is not very sensitive to the exact value of w chosen, because the excursions of the walkers are too short to build up a large variance in the multiplicities.

Calculation of observables

In Section 4.3 we explained the use of a mixed estimate for calculating the energy and other observables. As we pointed out, this is a good way of estimating the value of an observable only if the operator O concerned commutes with the Hamiltonian. In that case we have, for n large enough,

$$\frac{\langle \Psi_0|O|\Psi_0\rangle}{\langle \Psi_0|\Psi_0\rangle} = \frac{\langle \Psi_T|\mathcal{F}^n O|\Psi_0\rangle}{\langle \Psi_T|\mathcal{F}^n|\Psi_0\rangle} = \frac{\langle \Psi_T|O\,\mathcal{F}^n|\Psi_0\rangle}{\langle \Psi_T|\mathcal{F}^n|\Psi_0\rangle} = \frac{\langle \Psi_T|O|\Psi_0\rangle}{\langle \Psi_T|\Psi_0\rangle}, \tag{5.2.1a}$$

as making $\mathcal{F}^n$ work to the right yields the same factor in the numerator and the denominator. However, this obviously does not hold in the case that O does not commute with the Hamiltonian. In that case, one should use a different approximation for the expectation value of O. The *extrapolated estimate* [25,43]

$$\frac{2\langle \Psi_T|O|\Psi_0\rangle}{\langle \Psi_T|\Psi_0\rangle} - \frac{\langle \Psi_T|O|\Psi_T\rangle}{\langle \Psi_T|\Psi_T\rangle} \tag{5.2.2}$$

is a much better approximation for this expectation value.

The trial wave function

The single-most important ingredient of the Monte Carlo methods described here is the trial wave function. In the fixed-node scheme, the energy that one is able to find is determined implicitly through the nodal structure of the trial function. In the power method, the sign problem is prohibitive when the trial wave function does not resemble the true ground-state wave function close enough, since the ground state has to be reached within a limited number of steps. Therefore, it is worthwhile to invest a fair amount of time in finding the best trial state possible.

There are two requirements that the trial wave function must fulfill in any case. The first is that it must contain the proper symmetry for the particles concerned. In case of the Hubbard model, one has to make sure that the trial wave function is antisymmetric under exchange of two electrons of the same spin, in order to obtain results that hold for the fermion problem. As the Hamiltonian does not change this symmetry, it will be conserved in the projection process, and thus properties are obtained for a state of the correct symmetry. The second requirement is of more practical nature: the trial function must be easily calculable in any given configuration. Most of the effort in the Monte Carlo program is put in the calculation of the trial wave function in every configuration that is visited, and in all of its neighbors. Any extra work that has to be done here causes a serious slow down of the program.

In our implementation we make use of Slater determinants to define the trial wave function. They automatically fulfill the antisymmetry condition for the electrons, and, as we will see, they enable an efficient way for calculating the wave function in neighboring configurations and for updating the information. There are several ways to define the basis functions for such determinants. We use a self-consistent Hartree-Fock (or mean-field) approach that

provides an approximation of the ground state. As we work with a fixed number of parti-
cles and fixed total spin in the z-direction S^z, it is straightforward to use the average oc-
cupation numbers $\langle n_{i\sigma} \rangle$ as the parameters that have to be solved self-consistently from the
Schrödinger equation, with the Hartree-Fock Hamiltonian

$$\mathcal{H}_{HF} = \mathcal{H}_{kin} + U \sum_i \left(n_{i\uparrow}\langle n_{i\downarrow}\rangle + n_{i\downarrow}\langle n_{i\uparrow}\rangle - \langle n_{i\uparrow}\rangle\langle n_{i\downarrow}\rangle \right). \tag{5.2.3}$$

This can be done by iteratively diagonalizing the one-particle Hamiltonians for the up and
the down spins separately:

$$\mathcal{H}^1_\uparrow = \mathcal{H}_{kin} + U \sum_i n_{i\uparrow}\langle n_{i\downarrow}\rangle, \tag{5.2.4a}$$

and

$$\mathcal{H}^1_\downarrow = \mathcal{H}_{kin} + U \sum_i n_{i\downarrow}\langle n_{i\uparrow}\rangle. \tag{5.2.4b}$$

The fields $\langle n_{i\sigma} \rangle$ must fulfill the condition

$$\sum_i \langle n_{i\sigma} \rangle = N_\sigma, \tag{5.2.5}$$

where N_σ is the number of electrons with spin σ. Starting with some well chosen field for,
e.g., the down spins, one can diagonalize $\mathcal{H}^1_\uparrow$ in order to obtain the one-particle levels for
the up spins. As one is interested in the state with lowest energy, one has to fill the $N_\uparrow$ low-
est levels, and one can then calculate the resulting field for the up spins. Using this as in-
put in $\mathcal{H}^1_\downarrow$, one finds in a similar way a new field for the down spins. Iteratively repeating
this procedure, one finally finds a set of self-consistent one-particle wave functions $\{\varphi^\sigma_j\}$,
with $j = 1, 2, \ldots, N_\sigma$. The Hartree-Fock wave function is then constructed from these one-
particle functions as a product of two Slater determinants:

$$\langle R|\Psi_{HF}\rangle = \big|D_\uparrow(R)\big|\,\big|D_\downarrow(R)\big|. \tag{5.2.6}$$

Denoting the position of the kth electron of spin σ by i^σ_k, we can write

$$|D_\sigma(R)| = \begin{vmatrix} \varphi^\sigma_1(i^\sigma_1) & \varphi^\sigma_1(i^\sigma_2) & \cdots & \varphi^\sigma_1(i^\sigma_{N_\sigma}) \\ \varphi^\sigma_2(i^\sigma_1) & \varphi^\sigma_2(i^\sigma_2) & \cdots & \varphi^\sigma_2(i^\sigma_{N_\sigma}) \\ \cdots & \cdots & \cdots & \cdots \\ \varphi^\sigma_{N_\sigma}(i^\sigma_1) & \varphi^\sigma_{N_\sigma}(i^\sigma_2) & \cdots & \varphi^\sigma_{N_\sigma}(i^\sigma_{N_\sigma}) \end{vmatrix}. \tag{5.2.7}$$

One can always use such a numerical approach, but in many cases it is also possible to solve
the Hartree-Fock problem analytically. An example is the case of half filling, where the so-
lution can be found easily by making a transformation to reciprocal space.

The Hartree-Fock function itself is quite suitable as a starting point for a fixed-node
Monte Carlo calculation, but one can also build in correlations that one expects to be present,
like a Gutzwiller factor (by which double occupancy can be reduced) or a Jastrow factor (by
which, e.g., spatial correlations between the electrons spins can be enhanced). Both of these

examples[2] are local[3] modifications of the wave function, which means that they can be easily calculated. They provide extra parameters whose values can be optimized, to lead to the best result in the Monte Carlo method. Note that, in general, the best variational wave function (i.e., the wave function with lowest energy that can be obtained by varying all available parameters) does not necessarily lead to the best fixed-node result. Much more than the exact energy, the sign-flip structure of the trial wave function determines how good the fixed-node solution is. As has been pointed out by Ceperley [26], 'the knowledge of the nodes of the many-fermion wave function would enable exact calculation of the properties of fermion systems by Monte Carlo methods.' This specifically holds for the fixed-node method.

An important advantage of the Slater determinants is the efficient way in which the wave function can be calculated. In our methods, for a walker that is sitting in some configuration R, one needs to calculate the wave function in all neighboring configurations R', as that is needed to perform importance sampling as well as to calculate the energy and other observables. If one would have to calculate the determinants for each of these neighbors separately, each of which takes of the order $O(N_\sigma^3)$ operations, the program would be rather slow. However, as we only need to scan neighboring configurations, we can make use of the fact that one of the determinants representing the wave function in a neighbor R' is the same as the corresponding determinant in R, and the other one is different in one column only. Ceperley et al. [44] have derived formulas which enable updating the determinant in going from R to R', without the need to invert a matrix in R'.[4] Let us consider the determinants for R and its neighbor R', that differs from R by one electron, with number k and with spin up, sitting on site j in R'. As the spin-down electrons remain at the same positions, the determinant $|D_\downarrow(R')|$ equals $|D_\downarrow(R)|$. We only have to consider the determinants for the spin-up electrons. We can express the ratio q between these determinants for R and R' in the following way (dropping the spin index):

$$q \equiv \frac{|D(R')|}{|D(R)|} = \sum_l \varphi_l(j)[D(R)^{-1}]_{lk}^{\mathsf{T}}. \tag{5.2.8}$$

This means that we only have to calculate the dot product between the j-th column of the matrix D at R' and the k-th column of the transposed inverse matrix D^{-1} at R, to derive the value of the wave function in R' from that in R. This reduces the number of operations to obtain the value of the determinant in a neighboring configuration to $O(N_\sigma)$. Furthermore, one can construct the new transposed inverse matrix from the old one when one has decided to which neighbor the walker must step, in the following way:

$$[D(R')^{-1}]_{lk}^{\mathsf{T}} = \frac{[D(R)^{-1}]_{lk}^{\mathsf{T}}}{q} \tag{5.2.9a}$$

for all l, and

$$[D(R')^{-1}]_{li}^{\mathsf{T}} = [D(R)^{-1}]_{li}^{\mathsf{T}} - [D(R')^{-1}]_{lk}^{\mathsf{T}} \sum_m \varphi_m(j)[D(R)^{-1}]_{mi}^{\mathsf{T}} \tag{5.2.9b}$$

---

[2]They will be treated in more detail in Section 5.4.
[3]That is, local in configuration space.
[4]These formulas can also be found in a report by Von der Linden [45].

for all l and $i \neq k$. Note that in the second formula, the result of the first one has been used. Calculating the new matrix D^{-1} in this way takes $O(N_\sigma^2)$ operations.

Full calculation of the transposed inverse of a matrix involves a number of operations which is of the same order $[O(N_\sigma^3)]$ as calculation of a determinant. Clearly, as it has to be calculated only for the configuration from which the walker starts hopping, the number of operations required for this procedure is very small compared to the full calculation of all determinants for the configurations that have to be considered. Once the transposed inverse matrix has been calculated for the first configuration of the path, it can be used for calculating the wave function in all neighbors and updating the inverse for subsequent configurations with relatively little effort.

Distribution of the starting configurations

Our last remark concerns the ensemble of walkers that one can use to start the Monte Carlo program. One of the advantages of the fixed-node scheme is that it does not rely very much on the starting configurations. In principle, one can use a more or less random ensemble, since during the thermalization process it should automatically become distributed according to Eq. (4.3.8d). This is caused by the fact that the vector defined by this distribution is a stable fixed point[5] of the stochastic matrix $p(R' \leftarrow R)$, such that any vector representing a physical distribution converges to it when repeatedly applying $p(R' \leftarrow R)$. On large systems, however, it is wise to use some knowledge on what the ensemble should be. Otherwise, due to large fluctuations in the multiplicities and the trial value of w, one may find that the number of walkers increases or decreases very rapidly, and that it is impossible to obtain a stable set.

Again, several ways exist to produce a set of walkers that is suitable as a starting ensemble. One possibility, that we mentioned in Section 4.2, is making use of the Metropolis scheme. In this scheme one generates sequences of configurations by creating paths in configuration space, in such a way that the configurations are distributed according to the trial wave function. However, as the number of walkers to be generated for the ensemble is usually rather restricted, the sequences may be too short to make the configurations be properly distributed. In that case, one can use a similar but more straightforward method that may yield better results with less effort. Generating a random configuration, one can compare its wave function to a random number scaled to some threshold, and accept it if the wave function is larger than the random number. This means that the configuration is always accepted if its wave function is larger than the threshold, and that below the threshold the probability that it is accepted decreases with decreasing wave function. Usually, this is sufficient to obtain a set of walkers by which a stable Monte Carlo run can be performed. In order to save a lot of time, when doing several runs for the same system, one can store the ensemble that is obtained after the thermalization in the first run, and use it as input for the other runs.

For the power method, it is crucial that the starting walkers be distributed according to the product of the trial wave function and the ground-state wave function of the effective

---

[5] See, e.g., the paper by Hetherington [46] on the statistical iteration of matrices.

Hamiltonian, as required by Eq. (4.7.4). The paths in this procedure are too short for the ensemble to converge to the correct distribution. Therefore, one has to use an ensemble resulting from the fixed-node program. It can be taken from a fixed-node run, at any stage after the thermalization process has been completed.

5.3   Testing the method

As with any method, it is important to check whether the Monte Carlo procedures we described for fermions are capable of producing correct results in cases where exact results are available. On the other hand, it is useful to have a small and simple model system that can serve as a playground for testing new ideas concerning efficiency or improvements of the wave function. Doing this, one should of course take into account the fact that finite-size effects possibly dominate the physics of the system, but such small systems usually still give a fairly good impression of what tendencies exist.

The $2 \times 2 \times 2$ cube

We choose the $2 \times 2 \times 2$ cube for testing purposes. This system has its largest diagonalization problem at half filling with zero total S^z (i.e., with 4 electrons carrying spin up, and 4 with spin down), in a space containing 4900 states. Using the symmetries of the cube,[6] this problem can be reduced to diagonalizing 10 smaller matrices, of which the largest has dimension 960. Such a matrix can be diagonalized on a simple workstation in a few minutes. In case one considers an antiferromagnetic arrangement of spins on the cube, one can use only half of the symmetry operations, leading to a largest matrix of dimension 1896. This can still be tackled within a reasonable amount of time. Thus this system is small enough to be able to do exact calculations, and yet it is complicated enough to expect that it provides a good test case for checking whether the Monte Carlo algorithms work properly.

In Table 5-1 we give an impression of how good the fixed-node *theory* works for this small system, by exact calculations of the energies. We consider the case of half filling with $S^z = 0$, and of 4 electrons with spin up and 1 with spin down.[7] We use different types of self-consistent mean-field solutions as trial wave functions in the fixed-node calculation, and we compare the mean-field energy (MF), the ground-state energy of the effective Hamiltonian (FN), and the true ground-state energy of the problem, for different values of the parameter

---

[6]In order to do so, one has to make use of group theory in order to find the irreducible representations of the symmetry group of the cube, create new basis vectors in state space related to these representations, and re-express the Hamiltonian in this basis. As vectors in different representations do not communicate with each other in the Hamiltonian, it is reduced in the new basis to separate blocks related to the representations. For a detailed description of this procedure see, e.g., Ref. 47. The symmetry group of the cube can be found in, e.g., Ref. 48.

[7]There is a practical reason for choosing these fillings. In order to be able to use the group theoretical approach also for calculating the ground-state energy of the fixed-node effective Hamiltonian, the trial wave function must fulfill the requirements of the symmetry group of the cube. In most cases, mean field breaks the symmetry, due to a degeneration of the one-particle levels. With 4 or 1 particles of the same spin, the mean-field solution turns out to be non-degenerate and of the correct symmetry.

Table 5-1. Comparison of the true ground-state energy with the energy of different types of self-consistent mean-field (MF) solutions, and with the ground-state energy of the fixed-node (FN) effective Hamiltonian, for the $2 \times 2 \times 2$ cube with different fillings and at various finite U, by exact calculations only. For a clarification of the type of mean-field wave function used, see text.

	U/t	MF type	energies (in units of t)		
			MF	FN	true
4↑	0		-9	-9	-9
1↓	1	H	-8.5	-8.5419	-8.5420
	6	H	-6	-7.2508	-7.2533
	6		-6.0701	-7.2424	-7.2533
	10	H	-4	-6.8400	-6.8442
	10	AF	-4.2551	-6.7476	-6.8442
	10		-5.3271	-6.7637	-6.8442
4↑	0		-12	-12	-12
4↓	1	H	-10	-10.1148	-10.1188
	2.5	H	-7	-7.7257	-7.7510
	2.5	AF	-7.0061	-7.6942	-7.7510
	10	H	8	-2.6597	-2.8652
	10	AF	-2.3113	-2.6382	-2.8652

U. The mean-field wave functions are obtained as described in the previous section, but with different restrictions on the average number of electrons with spin up and down per site. Writing $\langle n_{i\sigma} \rangle = \langle n_\sigma \rangle + (-1)^\sigma q_{i\sigma}$ for the average number of spin-σ particles on site i, we denote $q_{i\sigma} = 0$ by H (homogeneous), and $q_{i\sigma} = (-1)^i q_\sigma$, with q_σ a constant, by AF (antiferromagnetic, or Néel order favored). In the case of 4 up and 1 down spins (i.e., off half filling), the self-consistent mean-field solution with lowest energy turns out to have a spatial symmetry different from both H and AF for larger values of U. As the ground states of both the original and the effective Hamiltonians are not degenerate in these cases, their symmetry cannot be broken. Note that, for $U = 0$, the mean-field approximation yields the true ground state. We find that also the fixed-node result equals the true ground-state energy in that case, as predicted.

One can see that the fixed-node approach (i.e., the use of an effective Hamiltonian) on this small system yields a significant improvement on the upper bound for the ground-state energy, compared to the mean-field approximations. As we pointed out before, the mean-field wave function with lowest energy does in general not give the best fixed-node result. In all cases, the trial wave function marked H yields the highest variational energy and the lowest fixed-node energy. This is most striking in the case of half filling with $U/t = 10$. We want to stress again that, clearly, the sign-flip structure of the trial wave function determines how good the fixed-node approximation is.

The next step is to check that the computer program is indeed capable of producing the results that are expected from the exact calculations. In Table 5-2 we show the results of actual Monte Carlo simulations for four of the cases given in Table 5-1. We have calculated

Table 5-2. Comparison of FNMC simulations with exact calculations of the fixed-node energy E_{FN} and structure factor $S(\pi, \pi, \pi)$, for the 2 × 2 × 2 cube, using four different trial wave functions. The statistical error in the last digit is indicated between brackets.

| | | 4 ↑, 1 ↓, $U/t = 6$ | | 4 ↑, 4 ↓, $U/t = 2.5$ | |
		1. MF H	2.	3. MF H	4. MF AF
E_{FN}/t	exact	-7.2508	-7.2424	-7.7257	-7.6942
	MC	-7.248(2)	-7.230(6)	-7.725(2)	-7.688(5)
$4S(\pi, \pi, \pi)$	exact	0.6753	0.6448	1.2711	1.4841
(FN mixed estimate)	MC	0.6760(6)	0.644(1)	1.271(3)	1.488(5)

Table 5-3. Comparison of the true ground-state energy with the energy of Gutzwiller-projected mean-field (GMF) solutions, and with the ground-state energy of the fixed-node (FN) effective Hamiltonian using the GMF wave function as trial function. Exact calculations performed for the 2 × 2 × 2 cube at infinite U, with 1 to 3 holes in all possible spin backgrounds (for reasons of symmetry, only the cases where the number of up spins is larger than the number of down spins need to be given). Energies are given in units of t.

	1 hole				2 holes				3 holes		
$N_\uparrow$	7	6	5	4	6	5	4	3	5	4	3
$N_\downarrow$	0	1	2	3	0	1	2	3	0	1	2
GMF	-3	-2.25	-2.143	-2.20	-4	-4	-4	-4.120	-5	-5.625	-5.294
FN	-3	-2.751	-2.601	-2.505	-4	-4.391	-4.469	-4.531	-5	-5.975	-5.635
true	-3	-3	-3	-3	-4	-4.510	-4.673	-4.789	-5	-5.984	-5.984

the energy and, in addition, the structure factor (2.6.9) for $\mathbf{q} = (\pi, \pi, \pi)$ for these cases, using the absolute value of the trial wave function as guiding function to perform importance sampling. As can be seen, the results agree fairly well with the exact values. The Monte Carlo runs were done with 2000 walkers, each doing 250 × 500 steps, except for run no. 2. (using a mean-field trial wave function with a broken symmetry, but not AF), in which 4000 walkers were used. This run caused most problems in finding the best parameters to lead to a stable population of walkers, and fluctuations remained relatively large although more walkers were used than in the other runs. Also in run no. 4., the error bars remain somewhat larger than in the runs where a symmetric H wave function was used. Apparently a wave function with less symmetry not only causes the final energy to be higher, but also leads to a larger variance. We assume that this is due to the fact that measurements in different but symmetric regions of the configuration space yield different results for such a wave function, whereas in the case of a wave function of high symmetry they would be the same. This causes larger fluctuations in the measurements.

In order to give an impression of the Monte Carlo runs, we show the thermalization for one of the above simulations in Figure 5-1. After some 200 steps, the energy measured starts fluctuating around the expected value. After averaging over several thousands of steps (not shown in this figure), one obtains the value indicated by the square at the right.
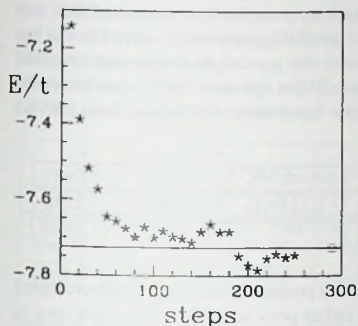
Figure 5-1. Thermalization in a FNMC simulation, for the $2 \times 2 \times 2$ cube at half filling, with zero total S^z, using a homogeneous mean-field trial wave function. Each star indicates the average of the energy sampled in 10 Monte Carlo steps. The drawn line indicates the exact ground-state energy of the fixed-node effective Hamiltonian. The square at the right indicates the resulting energy, obtained after averaging over several thousands of steps, with a statistical error which is smaller than the symbol size.
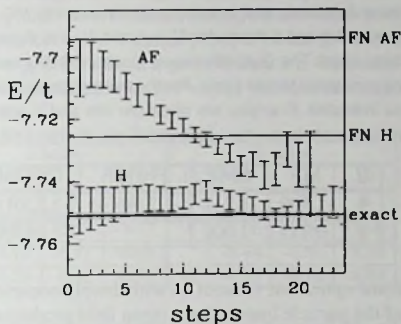
Figure 5-2. Power method for the $2 \times 2 \times 2$ cube at half filling, with zero total S^z, using the H and AF mean-field trial wave functions. An error bar indicates the estimated value of the ground-state energy after each power step.

For the same system (the cube at half filling, with zero total S^z) we have also performed a test of the power method. In Figure 5-2 we show two power Monte Carlo calculations, using the AF and H mean-field wave functions, respectively. For each calculation, the proper starting ensemble of walkers, resulting from a FNMC calculation, is used. As can be seen, the average energy in the AF calculation does not converge to the true ground-state value before the sign problem arises, after some 20 steps. In the calculation using the H wave function, the energy starts at a lower value than the FN result, due to fluctuations in the ensemble. It does converge to the correct value, however, giving a fairly accurate estimate of the ground-state energy. From these examples it is clear that, indeed, the power method cannot be used to find a reliable estimate of the true ground-state properties, if the fixed-node method does not yield a good approximation. Therefore, we will concentrate on preparing applications of FNMC, and show only few examples of application of the power method.

In view of our interest in the large-U Hubbard model, it is also useful to test the fixed-node method on the cube at infinite U. In Table 5-3 we compare true ground-state energies for this system with up to three holes to values obtained by diagonalizing the fixed-node effective Hamiltonian. As trial wave functions we use restricted (H) mean-field solutions for

Table 5-4. Comparison of the exact ground-state energy to the energy of mean-field and Gutzwiller wave functions, and to the energy obtained by MC simulations, for the 4 × 4 square (PBC), with 5 spins up and 5 down, for $U = 4$ and $U = 8$. A mean-field (MF) restricted (H) wave function has been used. The Gutzwiller wave function (GMF) was used with $g = 0.6$; its energy was calculated by variational Monte Carlo. For the fixed-node simulations (FN), a trial wave function has been used as indicated. Energies are given per site and in units of t. Constrained Path Monte Carlo (CPMC) results were taken from [35], exact results from [49].

U	MF	GMF	FN/MF	FN/GMF	power MC	CPMC	exact
4	-1.109	-1.212	-1.2186(4)	-1.2201(4)	-1.2234(6)	-1.2238(6)	-1.2238
8	-0.719	-1.007		-1.0858(2)		-1.0925(7)	-1.0944

some optimized value of U, with doubly occupied sites projected out.[8] Due to the averaging of the particle interactions, mean field produces a rather poor wave function in the case of large U. The finite-U wave function with double occupation projected out yields considerably better results, but still gives rather bad approximations of the ground-state energy. In case of all spins up, both mean field and fixed node produce the exact ground state correctly, but with increasing number of spin flips the approximation becomes poorer.

The 4 × 4 square

Another interesting test system is the 4 × 4 square lattice. Various authors [35, 47, 49–51] have performed calculations for the Hubbard model on this system, and a number of exact and approximate results has been obtained. For comparison, we have also performed some Monte Carlo calculations for the 4 × 4 square, in order to check whether our method still gives good results on a somewhat larger system. In Table 5-4 we show two series of calculations for this system, applying periodic boundary conditions (PBC), with 5 particles carrying spin up and 5 spin down, for two values of the interaction U. As can be seen, FNMC produces rather good approximations, and in the case of $U = 4$ we succeeded in approaching to the exact ground-state energy by means of the power method. For $U = 8$ we performed only one FNMC simulation, and the FN result is not good enough to be able to use the power method.

We also show a few calculations for the 4 × 4 square at infinite U, in Table 5-5. As trial wave functions we use simple mean-field solutions for $U = 0$, with Gutzwiller factor $g \to 0$ to project out the configurations with doubly occupied sites. As in the case of the cube at infinite U, we see that the FNMC results are a substantial improvement of the Gutzwiller mean-field approximation, but still rather far from the exact value. They become poorer with increasing number of spin flips.

---

[8]This is equivalent to using a Gutzwiller wave function with $g = 0$; see Eq. (5.4.1) for the definition of the Gutzwiller factor g.

Table 5-5. Comparison of the exact ground-state energy to the energy of Gutzwiller-projected mean-field (GMF) solutions, and to the energy resulting from FNMC simulations using the GMF wave function as trial wave function. Calculations performed for the 4 × 4 square (PBC) at infinite U, with 1 to 3 holes in some spin backgrounds. The total energy of the system is given, in units of t. The statistical error is of the order of the digit after the last one indicated. The GMF energy has been calculated by VMC, exact results have been taken from [50].

	1 hole				2 holes				3 holes			
$N_\uparrow$	15	14	13	12	14	13	12	11	13	12	11	10
$N_\downarrow$	0	1	2	3	0	1	2	3	0	1	2	3
GMF	-4	-3	-2.75	-2.84	-6	-5.67	-5.53	-5.57	-8	-8.06	-8.04	-7.46
FN		-3.77	-3.66	-3.46		-6.10	-6.11	-6.08		-8.44	-8.59	-8.34
exact	-4	-4	-4	-4	-6	-6.28	-6.41	-6.52	-8	-8.55	-8.76	-8.82

Table 5-6. Various mean-field and quantum Monte Carlo calculations (QMC) of the exact ground-state energy of a 10 × 10 square (PBC), at half filling and with zero total S^z, for $U = 4$. Slave Boson MF result has been taken from [52], QMC results from [53] and [54].

method	energy per site
Mean Field (AF)	-0.797
Slave Boson MF	-0.839
Gutzwiller VMC	-0.842
QMC (Hirsch '85)	-0.88(3)
QMC (White '89)	-0.864(1)
FNMC	-0.852(2)
Power MC	-0.860(5)

Larger systems

Many results exist in literature also for the 8 × 8, 10 × 10, and larger systems. Although we have not performed many calculations for these systems, in order to demonstrate that our methods can be used for larger systems we give one last example of calculations that we have done for the 10 × 10 system. In Table 5-6 we compare several mean-field and quantum Monte Carlo calculations of the ground-state energy of this system at half filling, with zero z-component of the total spin and with periodic boundary conditions. Our results are in very good agreement with previous QMC calculations by Hirsch [53] and White *et al.* [54]. Therefore we conclude this section by stating that the application of FNMC and the power method looks very promising, but that an investigation of possible trial wave functions is necessary before these larger systems can be handled well. This especially holds for the study of systems at large or infinite U, where mean-field theory fails to give a reasonable description.

5.4   Improving the trial wave function

As we have seen in the previous section, mean-field wave functions sometimes lead to rather
good fixed-node results, but in other cases they seem to lack the proper structure. As we
pointed out in Section 5.2, a good trial wave function is vital for the fixed-node method (and
even more for the power method) to produce useful information. In this section we will list
a few simple options for constructing wave functions, which we have investigated. None
of them seems to provide a significant improvement on mean-field functions, however, and
further studies will be necessary.

Slater determinants

Usually we exploit wave functions that are based on Slater determinants, obtained from
mean-field calculations. The advantage of these functions is that they are explicitly antisym-
metric, as required, and that they provide a reasonable approximation of the ground-state
wave function. In some situations, especially if U is large, they are not capable of produc-
ing the proper sign structure. Another disadvantage, which becomes less significant with
increasing system size, is that the Slater determinants tend to vanish in a number of con-
figurations, while the exact ground-state wave function need not be zero in these config-
urations. In the fixed-node method, a configuration with vanishing trial function remains
a zero in the fixed-node wave function. Even if the signs in the other configurations were
completely correct, FNMC would not be able to produce the exact ground-state properties,
as the contributions of these zero configurations are omitted. Therefore, one should as much
as possible try to avoid having a large number of zeroes in the trial function.

Variational parameters

A simple means of varying a given trial wave function is the introduction of variational
parameters. A well known example is the *Gutzwiller factor g*, by a power of which the wave
function is multiplied, depending on the number of doubly occupied sites $n_d(R)$ in each
configuration R:

$$\langle R|\Psi_G\rangle = g^{n_d(R)}\langle R|\Psi\rangle. \tag{5.4.1}$$

Note that this factor is of no use in the case of infinite U, where doubly occupied sites are
excluded. Another example is the *Jastrow factor*, which can be used to enhance certain cor-
relations in the trial wave function. Taking $C(R)$ to be some correlation calculated locally
in R, we can write

$$\langle R|\Psi_J\rangle = e^{C(R)}\langle R|\Psi\rangle. \tag{5.4.2}$$

Usually, these extra factors can easily be calculated locally, such that they do not involve a
significant increase of the effort to obtain the wave function. In general, one would try to ob-
tain the best variational wave function of such kind, minimizing the energy by varying the
parameter(s) involved. However, as we have stated before, it is not true that the wave func-
tion so obtained leads to the best fixed-node result. It would therefore be necessary to use
different values for these parameters also in the fixed-node runs, which makes this approach

rather laborious. Another disadvantage is that these factors usually do not change the sign of the wave function. As we suspect that the sign structure of the wave function is far more important than its actual value, this is an important restriction which causes the variational parameters to be less useful for the fixed-node program. An example of the application of the Gutzwiller factor, in which it provides a significant improvement on the ground-state energy of the trial wave function but only slightly changes the fixed-node results, was presented in Table 5-4.

Variational Lanczos approach

Another variational method which does not suffer from this restriction, but is rather laborious, is the so-called *variational Lanczos* approach as proposed by Chen and Lee [36]. In the Lanczos diagonalization scheme, by applying one projection step one adapts the wave function $\Psi = \Psi_L^{(0)}$ in the following way:

$$\Psi_L^{(1)} = \Psi_L^{(0)} + \lambda \mathcal{H} \Psi_L^{(0)}. \tag{5.4.3}$$

where the value of the parameter λ is to be calculated. Instead of calculating the exact value of this parameter according to the Lanczos method, one can also determine it variationally, such that $\Psi_L^{(1)}$ is in fact a variational wave function. As we have seen, for a walker located in a configuration R, one needs to know the value of the wave function in all neighboring configurations. In the (one-step) variational Lanczos scheme, calculation of the wave function in one of these neighbors R' involves the evaluation of Ψ in R' and in all of its neighbors R'', due to the application of the Hamiltonian. Thus, in order to obtain the wave function $\Psi_L^{(1)}$ in the neighbors of R, one needs to calculate the wave function Ψ in all configurations that can be reached from R within two steps. This implies an increase of the number of operations needed to calculate the trial wave function, by a multiplicative factor which is of the order of the average number of neighbors. Using Slater determinants for Ψ, this may well be feasible in a system with low connectivity, such as in case of infinite U and low hole doping. In case of finite U, however, the average number of neighbors is of the order of the dimension times the number of particles in the system, and calculation of all these values of the wave function would take too much time. One needs a much simpler prescription for Ψ then, such that the extra effort due to the need to calculate Ψ in more configurations is compensated by a much faster evaluation of Ψ in each configuration. If one could find a sufficiently simple Ψ, one could even think of applying more Lanczos steps on the trial wave function. In the examples we considered, however, the nodes are moved only rather slowly, such that one obtains relatively small improvements with each further Lanczos step.

Hole wave functions

We have examined several possibilities for defining Ψ in such a way that it can be calculated very easily. One idea, which would be especially suited for infinite U and few holes, would be to express the wave function in terms of Slater determinants for the *holes*. If the number of holes is very small, this drastically reduces the effort for calculating the determinants. The

problem is, however, that hole wave functions are destined to produce a ferromagnetic state, as the spin background in which the holes move around is completely ignored. Neither the Lanczos projection nor the fixed-node procedure can change this, such that it is not possible to use such a wave function for checking whether the ferromagnet is the ground state of the system.

Simple sign rules

Having in mind our impression that the value of the wave function does not matter so much as its sign, one could also think of taking the absolute value of the (zeroth-order) wave function to be constant, and giving a prescription for its sign only. Of course, the wave function still has to be antisymmetric. The simplest way to establish this is to define a standard order of the numbered particles on the lattice, by numbering the sites in some arbitrary way and ordering the particles according to the site numbers. One can then define the sign by taking the permuted configuration in which the particles have been rearranged with increasing particle number to have positive sign. The sign for neighboring configurations can easily be determined from the (local) permutations of the standard order of the particles. We have tried this for small systems, where the results are not very encouraging. Even with a relatively large number of Lanczos steps, it is hard to improve on the mean-field wave function. The sign convention itself yields a very bad wave function, and the Lanczos projection makes it only slowly converge to the ground-state wave function. We fear that this can only get worse for larger systems, while mean-field solutions tend to become better. An extra complication is in the fact that, depending a little bit on the actual sign convention used, the wave function hardly depends on the particle spins. This means that, again, a (ferromagnetic) hole wave function is produced.

Understanding of the nodal structure

So far, little information is available on how to tackle these problems. In his study on fermion nodes, which basically treats continuum problems but which is also very well applicable to lattice systems, Ceperley [55] states that he has '... made little progress in stating exact conditions that nodes must obey.' Most studies of wave functions aim at improving the energy of the wave function itself, while for our purposes it would be necessary to explore its sign-flip structure, in order to improve the energy that results in a fixed-node Monte Carlo simulation. We do not have a clear understanding yet of the requirements that a wave function must fulfill to make a good fixed-node trial function. Before doing the calculations described in the next section, that may lead to a more conclusive answer with respect to the problem of magnetism in the large-U Hubbard model, it is necessary to obtain better insight in the features of trial wave functions that define whether they are suited for the fixed-node method or not.

5.5 How to find ferromagnetism?

In the last section of this chapter, we will outline how the QMC methods may be used to treat the problem of ferromagnetism in the Hubbard model. Several authors [50,56,57] have followed this line of research, and obtained some interesting results on rather small systems already.

In order to study the stability of the ferromagnetic ground state, it is sufficient to investigate whether it is possible to find a non-ferromagnetic state that has lower energy than the ferromagnet. A ferromagnetic state is easy to find, by considering a system (at any level of hole doping) with all particles carrying spin up (i.e., a fully polarized ferromagnet, FPFM). Mean-field theory yields the exact ground state in that case, as the particles have no interaction other than the exclusion principle. The question is now whether one can find states with a lower energy by introducing spin flips in the fully polarized system. Let us consider a system with a fixed number of particles N_s and with total spin in the z-direction $S^z > 0$, and an eigenstate $|\Psi\rangle$ of this system. As the spin lowering operator

$$S^- = \sum_j c_{j\downarrow}^\dagger c_{j\uparrow} \tag{5.5.1}$$

commutes with our Hamiltonian, the state $S^-|\Psi\rangle$ is an eigenstate of the system with $S^z - 1$, with the same energy as $|\Psi\rangle$. In particular, this means that the fully polarized ferromagnet is an eigenstate of the system with any number of spin flips. In Figure 5-3 we visualize the main possibilities that one can find for the behavior of the minimum energy as a function
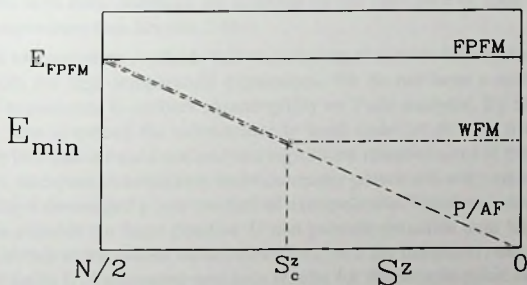
Figure 5-3. Possible behavior of the minimum energy as a function of the total spin in the z-direction, in a system with N particles. If the lowest minimum energy is found to be the same for all S^z, the ground state is a fully polarized ferromagnet (FPFM, all spins aligned). If the lowest energy is found for $S^z < S_c^z$, the ground state is weakly ferromagnetic (WFM). If the minimum is at $S^z = 0$, no ferromagnetic tendencies are present, and the ground state can be paramagnetic (P) or antiferromagnetic (AF), or it can contain more complex symmetries.

of the number of spin flips. If one finds the energy of the ferromagnet to be the lowest energy possible, for all possible spin backgrounds, the fully polarized ferromagnet is indeed the ground state of the system. If the lowest possible energy decreases as a function of the number of spin flips,[9] having a minimum only at $S^z = 0$,[10] one should conclude that the ground state is certainly not ferromagnetic, and one should investigate the possibilities that it is paramagnetic, antiferromagnetic, or contains still a different kind of correlation. The last possibility is that the lowest energy reaches its minimum at $S^z = S_c \neq 0$ (and then the same lowest energy is found for $0 \leq S^z < S_c$). In that case, the ground state exhibits weakly ferromagnetic tendencies, but the spins are not fully aligned.

One can now consider systems with various numbers of holes (as illustrated in Tables 5-3 and 5-5), to check how the ground-state properties behave as a function of the system size. Furthermore, considering systems of increasing size and extrapolating the tendencies found, it should be possible to derive useful information on the ground-state behavior of the system in the thermodynamic limit. For instance, one can look at the total spin per electron (or magnetization), which should obtain a finite value in the case of ferromagnetism. Other properties of interest are, e.g., the staggered magnetization and the (nearest-neighbor) spin-correlation functions. This program has been carried out by Riera and Young [50] by means of exact diagonalization, for systems up to 4×4 sites. In order to find consistent results, however, they have to average over various boundary conditions, as otherwise finite-size effects make the results erratic. It would be very interesting, and it certainly appears to be feasible, to use FNMC to obtain similar comparisons for larger systems.

---

[9]Note that it cannot increase as long as $S^z > 0$.

[10]In case of an even number of spins, otherwise $S^z = \pm 1/2$.

Summary, discussion, and conclusions

By using a simple perturbation theory and a cluster expansion technique, we have calculated high-temperature series expansions for the Hubbard model. These were the topics of the first part of this thesis. We have obtained expansions up to eighth order in the parameter βt, for the grand potential and several of its derivatives and for the spin-correlation functions, on the square and simple cubic lattices. Our most interesting observation, following directly from these expansions, is that for sufficiently large interaction U the nearest-neighbor correlations are ferromagnetic in a range of densities. The stronger the interaction, the higher the temperature below which the correlations are ferromagnetic. We see this as an indication for the existence of a ferromagnetic phase at low temperatures. Apparently the predicted Nagaoka phase at zero temperature leaves its imprints at the relatively high temperatures that we analyze.

This work is most closely related to the work of Hone *et al.* [7], Kubo and Tada [8,9], Pan and Wang [10], and Henderson *et al.* [12]. Hone *et al.* and Kubo have been the first to present the full set of coefficients up to fourth order for the grand potential and the uniform and staggered magnetic susceptibility of the simple cubic lattice, for arbitrary value of U. Pan and Wang have provided a correction on their work and extended it to sixth order. Our work, in turn, is an extension of these results, by including the eighth order and considering also correlation functions. Henderson *et al.* have addressed the same problem, extending the expansions to eighth order also for non-bipartite lattices, and their overall conclusions are consistent with ours, although we disagree on the relevance of their estimates for the transition temperature (see Section 2.8).

Standard extrapolation methods fail when trying to extract information for lower temperatures from the high-temperature expansions. We do not have a sufficient number of terms in the expansions to embark meaningfully on Padé analysis. By the cluster method one may be able to extend the calculations to tenth order in βt with use of a supercomputer, but we feel that it would not lead to a significant improvement of the results for lower temperatures, and even considerably more computer power will not extend the method any further. We have developed a new method of extrapolation, based on a density of states for holes, that is suitable for large positive U and particle densities near half filling. By this method, we obtain extrapolated values down to $T = 0$ for infinite U, and to very low temperatures for finite U. Comparing previous results for the simple cubic lattice (depicted in Figure 2-6) to our current results (Figure 3-5), we see that we now find Curie temperatures that are an order of magnitude smaller than before. Also, subsequent approximations in this method give consistent results, unlike in the old situation where alternately Curie temperatures were found or not. These convergence problems in the primitive series expansions are likely due to the Fermi degeneracy of the electron gas. At $\beta t \approx 1$, the wavelength of the electrons becomes of the order of the lattice distance. Above this temperature, the electrons behave classically, but below it degeneracies appear which hardly show up in a high-

temperature expansion, and there appear to be divergences in the series expansions which are not present in the physical quantities. This problem exists already when trying to describe a non-interacting gas by means of a series expansion, and it masks the physics of the correlated system. When applying a straightforward extrapolation technique, one does not sufficiently account for the degeneracy, leading to results that are erroneous for $\beta t \gtrsim 1$. In our new approximation, using a density of states for holes, we take the Fermi degeneracy into account, and therefore we are able to proceed to lower temperatures.

We are confident that our present results do not suffer from the above-mentioned convergence problems. As we have shown (Figure 3-3), the difference between approximations to different orders in the parameter βt is rather small, and adding the interaction also does not change the result considerably. Thus we believe the eighth-order non-interacting hole approximation to be sufficient to describe the qualitative behavior, and to obtain a good indication for numerical values. We may add that, as a check, we have compared the free energy from calculations by this method to results following directly from the series expansions, at $\beta t \lesssim 0.5$, where the expansions are almost exact, and that these results agree very well. It is clear, however, that also this new method is not valid for very low temperatures. From Figures 3-4 to 3-6 one can see that it is unreliable for $\frac{t}{U} \gtrsim \frac{kT}{4t}$, and it is not evident at which temperature exactly the approximation breaks down. We believe, however, that our method gives an accurate description of the tendencies in the half-filled system at infinite U, and of the qualitative behavior up to $n_h \approx 0.2$.

We observe strong indications for a region of ferromagnetic behavior in the magnetic phase diagram, near half filling. From these observations, based on the work on series expansions, we cannot definitely conclude of which type possibly existing magnetic phases may be. In a way similar to the theorem of Mermin and Wagner for the Heisenberg model, Walker and Ruijgrok [58] and Ghosh [59] have proved that the Hubbard model in one and two dimensions can not have long-range magnetic ordering, and Uhrig [60] has extended this result to also exclude planar magnetic ordering in the 3-dimensional case. These arguments do not rule out other kinds (like the Kosterlitz-Thouless type) of phase transitions at finite temperature, nor the possibility of long-range magnetic order at $T = 0$. A complication is in the fact that our method relies heavily on the short-range information from the high-temperature expansion (which is obtained via calculations on small systems). It gives similar results for the square and the simple cubic lattices, although one may expect different kinds of phases to occur. Furthermore, due to the thermodynamic approach in which all possible states are taken into account, our method cannot distinguish special states that may start to dominate the system at low temperatures. Such states, if any, are not recognized by the high-temperature expansion.

Our method currently fails to describe the case of a paramagnetic-antiferromagnetic transition, due to the fact that the spin background, which gives rise to divergences, is not treated correctly off half filling. We can therefore calculate only possible second-order phase transitions between a paramagnetic and a ferromagnetic phase, for the case of finite hole density. At half filling, we do find a finite Néel temperature for any finite value of the parameter βU (see Figure 3-7). This implies that, near or at half filling, there is a transition from a paramagnetic to an antiferromagnetic state at a higher temperature than the calculated

paramagnetic-ferromagnetic transition. Thus, at such low density, the paramagnetic-ferromagnetic transition is preempted by the paramagnetic-antiferromagnetic transition, and one must study the antiferromagnetic-ferromagnetic transition to determine the ground-state behavior.

In the second part of this thesis we describe a way to apply Green function Monte Carlo to the problem of fermions on a lattice. In principle, by quantum Monte Carlo methods, exact information on the ground state can be obtained directly. Due to the fermionic character of the particles the sign problem arises, making it difficult to control the statistical fluctuations in these methods. We have developed a method to avoid this problem. In first instance, it leads to an upper bound for the ground-state energy by means of fixed-node Monte Carlo. After that, it gives a possibility to obtain exact values by nodal relaxation (the power method) if the fixed-node approximation is good enough. As we have demonstrated by a number of examples in Section 5.3, the fixed-node Monte Carlo method works well in some cases, but in other cases better understanding of the representation of physical properties in the structure of the trial wave function is needed. We have not yet been able to obtain information on the magnetic behavior of the Hubbard model by means of these Monte Carlo simulations, so we will have to rely on exact and numerical calculations and their implications as described by other authors.

Recently, Zhang *et al.* [35] presented a method that is very similar to FNMC. They combine diffusion in a space of Slater determinants, as discussed by Fahy and Hamann [61], with the fixed-node idea to keep a positive overlap with the trial wave function. They claim that this *constrained-path* Monte Carlo method (CPMC) has important advantages over FNMC, amongst which a much better way to calculate the expectation value of non-commuting operators. However, as Ceperley [62] pointed out, their method is restricted to Slater determinants as trial functions, whereas FNMC can handle many different kinds of trial wave functions. Therefore, CPMC is likely to work well only if the ground state is close to a Hartree-Fock solution.

Let us return to the question of magnetism in the Hubbard model. One issue of interest in the literature is for what range of hole densities ferromagnetism may exist. If a ferromagnetic phase exists, it is very likely to occur at hole densities below $n_{\rm h} \approx 1/3$. Many authors find indications that support this statement. Yedidia [63] and Yang and Thompson [64] have pointed out that the particles in the strongly correlated Hubbard model behave like free fermions at about this density. Shastry *et al.* [65] have shown that the Nagaoka state becomes unstable above a certain hole concentration. Von der Linden and Edwards [66] use a variational approach to find a ferromagnetic region in the $T = 0$ phase diagram of the square-lattice Hubbard model. They rigorously conclude that the state of complete spin alignment is unstable when $n_{\rm h} > 0.29$, for all U, and when $U/t < 42$, for all $n_{\rm h}$. The latter value is significantly higher than the value $U/t \approx 15$ above which we find ferromagnetism, but we assume that that is due to the fact that they consider only strong ferromagnetism (full alignment of the spins), whereas our method may also include weak ferromagnetism. Putikka *et al.* [67] have performed a high-temperature expansion like ours, for the related

t–J-model. For small positive J and near half filling, this model becomes equivalent to the Hubbard model for large U. As in the case of the infinite-U Hubbard model, its configuration space is of dimension 3^N instead of 4^N, since doubly occupied sites are excluded. They have succeeded in pushing the series to higher orders for this model (up to 10 in case of finite J and up to 14 for $J = 0$). They find a region of weak ferromagnetism (or ferrimagnetism, i.e., the spins are not fully aligned) for small positive J, at hole density $n_h < 0.28 \pm 0.05$. From our own extrapolations, we find similar results ($n_h \approx 0.28$, see Figure 3-3). Using an approximation of an entirely different nature, the slave-boson mean-field approach (at $T = 0$), Denteneer and Blaauboer [52] find a critical hole density $n_h = 1/3$ for ferromagnetism to occur at $U = \infty$. It is encouraging to note that some of our other results are in reasonable quantitative agreement with theirs: they find that the value of U/t above which ferromagnetism can occur, is $U/t = 20$ (at $n_h = 0.17$), whereas one may extrapolate the results of Figure 3-4(a) to $T = 0$ to find $U/t = 15$ (at $n_h = 0.15$). Another comparison that can be made is for the relation between the Néel temperature and U/t in the half-filled system. From Figure 3-7 one can calculate that the paramagnetic-antiferromagnetic transition occurs for $kT_N \approx 3.85t/U$. The large-U Hubbard model at half filling is known to be equivalent to an antiferromagnetic Heisenberg spin model, for which estimates of the values of the critical temperature are given in [68]. According to the results mentioned there, the relation would be $kT_N \approx 3.80t/U$, which is in very good agreement.

Another question concerns the type of ferromagnetism that may occur. In the Nagaoka state, all spins are aligned to point in the same direction, but this is not necessarily true for ferromagnetic behavior. In fact, many authors point out that the fully polarized ferromagnet cannot be the ground-state of the system at any finite hole concentration. Wan [69] claims that he has proved this for the $U = \infty$ problem, by using a transformation to map this problem onto a bosonic representation, and showing that the ground state in this representation is below the lowest ferromagnetic solution. Although the ferromagnetic state appears to be stable against a single spin reversal, as claimed by Von der Linden and Edwards [66], Wan's findings are supported by Fang *et al.* [56], Douçot and Wen [57], Anglès d'Auriac *et al.* [70], and Sütő [71]. They all show by exact calculations for finite systems with two or more holes that the fully polarized state is not stable against multiple spin flips in this case. On the other hand, several indications have been found that ferromagnetic behavior does survive the thermodynamic limit. Riera and Young [50] consider small systems of increasing size to find the total spin for which the lowest energy is obtained (as described in Section 5.5), again in the case of infinite U. They note that the spin per electron in the ground-state, when averaging over different boundary conditions, *increases* with increasing system size. This is believed to be an indication that the spin per electron retains a finite value in the thermodynamic system, implying weak ferromagnetic behavior. This statement is supported by Barbieri *et al.* [72] and by Chiappe *et al.* [51], who claim to find indications for a stable magnetic phase in finite systems. It is not clear, however, in what respect finite-size effects are of influence in these calculations. We think that further investigations on larger systems will be needed to settle the answer to this question, and that fixed-node Monte Carlo simulations will be an important tool for this purpose.

For finite temperature the indications are less convincing. From series expansion results, tendencies towards ferromagnetic behavior have been found, but an actual phase transition is hard to predict. Also, little is known about the antiferromagnetic tendencies, that dominate the system at half filling and are likely to persist at small hole densities for finite temperature. A noteworthy point is that with increasing temperature ferromagnetic nearest-neighbor correlations reverse into antiferromagnetic correlations (see Section 2.7). This behavior is hard to predict from an effective-spin Hamiltonian with temperature-independent coupling. It means that the notion of an effective-spin Hamiltonian is less useful for the Hubbard model. This approach has in fact been followed by Takahashi [73] and MacDonald *et al.* [74], who have derived an effective-spin Hamiltonian for the case of half filling, where we find that the correlations are antiferromagnetic for all temperatures. It may however not be possible to meaningfully extend this off half filling. As was pointed out before, no long-range magnetic order can exist in the two-dimensional model, and no planar long-range order in three dimensions. Thompson *et al.* [64, 75] suggest from their results in an infinite number of spatial dimensions that the strongly correlated Hubbard model does not undergo a magnetic transition at a finite temperature, in any dimension and for any density. However, they do not exclude transitions of the Kosterlitz-Thouless type, where the magnetic susceptibility is finite at non-zero temperature, but where correlations have a power-law decay below a certain critical temperature.

Finally, we mention the possibility of inhomogeneous phases appearing. In particular, antiferromagnetically structured domain wall phases have been reported to be stable against the incommensurate antiferromagnetic state at relatively small values of U, by Hartree-Fock [76], variational [77], and Monte Carlo [42, 78] methods. Other interesting peculiarities include the magnetic textures and vortex-like states reported by Vergès *et al.* [79] for intermediate values of U and high hole-doping, which Yonemitsu and Bishop [80] claim to be unstable against transverse spin fluctuations. These statements are all based on unrestricted inhomogeneous Hartree-Fock approximations. The Monte Carlo methods presented in this thesis provide suitable tools to investigate these and other kinds of inhomogeneous structures.

References

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

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

[3] For a summary, see, e.g., E. Müller-Hartmann, Th. Hanish, and R. Hirsch, Physica B **186–188**, 834 (1993), and references therein.

[4] R. Micnas, J. Ranninger, and S. Robaszkiewicz, Rev. Mod. Phys. **62**, 113 (1990).

[5] M.P. Gelfand, R.R.P. Singh, and D.A. Huse, J. Stat. Phys. **59**, 1093 (1990).

[6] C. Domb, Adv. Phys. **9**, 149 (1960).

[7] D.W. Hone and P. Pincus, Phys. Rev. B **7**, 4889 (1973); G. Beni, T. Holstein, and P. Pincus, *ibid.* **8**, 312 (1973); G. Beni, P. Pincus, and D. Hone, *ibid.* 3389 (1973).

[8] K. Kubo, Prog. Theor. Phys. **64**, 758 (1980).

[9] K. Kubo and M. Tada, Prog. Theor. Phys. **69**, 1345 (1983); **71**, 479 (1984); **72**, 15 (1984).

[10] K.K. Pan and Y.L. Wang, Phys. Rev. B **43**, 3706 (1991); J. Appl. Phys. **69**, 4656 (1991).

[11] W.O. Putikka, *private communication*.

[12] J.A. Henderson, J. Oitmaa, and M.C.B. Ashley, Phys. Rev. B **46**, 6328 (1992).

[13] J.A. Henderson, *private communication*.

[14] A. Moreo, D. Scalapino, and E. Dagotto, Phys. Rev. B **43**, 11442 (1991).

[15] E. Manousakis, Rev. Mod. Phys. **63**, 1 (1991).

[16] W.F. Brinkman and T.M. Rice, Phys. Rev. B **2**, 1324 (1970).

[17] Y.S. Yang, C.J. Thompson, and A.J. Guttmann, Phys. Rev. B **42**, 8431 (1990).

[18] E.N. Economou, *Green's Functions in Quantum Physics*, in *Springer Series in Solid-State Sciences*, Vol. 7, edited by P. Fulde (Springer-Verlag, Berlin, 1979).

[19] S.V. Meshkov and D.V. Berkov, Phys. Lett. A **170**, 448 (1992).

[20] See, e.g., L.E. Reichl, *A Modern Course in Statistical Physics* (University of Texas Press, Austin, 1980).

[21] See, e.g., N.N. Bogolyubov, Jr., *A Method for Studying Model Hamiltonians* (Pergamon Press, Oxford, 1972).

[22] P.W. Brouwer, graduation thesis, University of Leiden (1993).

[23] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[24] N. Trivedi and D.M. Ceperley, Phys. Rev. B **41**, 4552 (1990).

[25] H. de Raedt and W. von der Linden, *Quantum Lattice Problems*, in *Topics in Applied Physics*, Vol. 71: *The Monte Carlo Method in Condensed Matter Physics*, edited by K. Binder (Springer-Verlag, Berlin, 1992).

[26] D.M. Ceperley and B.J. Alder, Phys. Rev. Lett. **45**, 566 (1980); Science **231**, 555 (1986).

[27] N. Trivedi and D.M. Ceperley, Phys. Rev. B **40**, 2737 (1989).

[28] H.J.M van Bemmel and W. van Saarloos, unpublished.

[29] A.V. Dotsenko and O.P. Sushkov, Phys. Rev. B **50**, 13821 (1994).

[30] S. Sorella, S. Baroni, R. Car, and M. Parrinello, Europhys. Lett. **8**, 663 (1986).

[31] E.Y. Loh, Jr., J.E. Gubernatis, R.T. Scalettar, S.R. White, D.J. Scalapino, and R.L. Sugar, Phys. Rev. B **41**, 9301 (1990).

[32] S. Fahy and D.R. Hamann, Phys. Rev. B **43**, 765 (1991).

[33] M. Boninsegni and E. Manousakis, Phys. Rev. B **46**, 560 (1992).

[34] M. Suzuki, Physica A **194**, 432 (1993).

[35] S. Zhang, J. Carlson, and J.E. Gubernatis, Phys. Rev. Lett. **74**, 3652 (1995).

[36] Y.C. Chen and T.K. Lee, Phys. Rev. B **47**, 11548 (1993); **51**, 6723 (1995).

[37] P.J. Reynolds, D.M. Ceperley, B.J. Alder, and W.A. Lester, Jr., J. Chem. Phys. **77**, 5593 (1982).

[38] D.M. Ceperley and B.J. Alder, J. Chem. Phys. **81**, 5833 (1984).

[39] L. Mitáš, E.L. Shirley, and D.M. Ceperley, J. Chem. Phys. **95**, 3467 (1991).

[40] G. Ortiz, D.M. Ceperley, and R.M. Martin, Phys. Rev. Lett. **71**, 2777 (1993).

[41] J.B. Anderson, J. Chem. Phys. **63**, 1499 (1975); **65**, 4121 (1976).

[42] Guozhong An and J.M.J. van Leeuwen, Phys. Rev. B **44**, 9410 (1991).

[43] D.M. Ceperley and M.H. Kalos, in *Topics in Current Physics*, Vol. 7: *Monte Carlo Methods in Statistical Physics*, edited by K. Binder (Springer-Verlag, Berlin, 1979); K.E. Schmitt and M.H. Kalos, in *Topics in Current Physics*, Vol. 36: *Applications of the Monte Carlo Method in Statistical Physics*, edited by K. Binder (Springer-Verlag, Berlin, 1984).

[44] D.M. Ceperley, G.V. Chester, and M.H. Kalos, Phys. Rev. B **16**, 3081 (1977).

[45] W. von der Linden, Phys. Rep. **220**, 53 (1992).

[46] J.H. Hetherington, Phys. Rev. A **30**, 2713 (1984).

[47] G. Fano, F. Ortolani, and A. Parola, Phys. Rev. B **46**, 1048 (1992).

[48] J.C. Slater, *Quantum Theory of Molecules and Solids*, Vol. 2: *Symmetry and Energy Bands in Crystals*, p. 375 (McGraw-Hill, New York, 1965).

[49] A. Parola, S. Sorella, S. Baroni, R. Car, M. Parrinello, and E. Tosatti, Physica C **162-164**, 771 (1989).

[50] J.A. Riera and A.P. Young, Phys. Rev. B **40**, 5285 (1989).

[51] G. Chiappe, E. Louis, J. Galán, F. Guinea, and J.A. Vergés, Phys. Rev. B **48**, 16539 (1993).

[52] P.J.H. Denteneer and M. Blaauboer, J. Phys.: Condens. Matter **7**, 151 (1995).

[53] J.E. Hirsch, Phys. Rev. B **31**, 4403 (1985).

[54] S.R. White, D.J. Scalapino, R.L. Sugar, E.Y. Loh, J.E. Gubernatis, and R.T. Scalettar, Phys. Rev. B **40**, 506 (1989).

[55] D.M. Ceperley, J. Stat. Phys. **63**, 1237 (1991).

[56] Y. Fang, A.E. Ruckenstein, E. Dagotto, and S. Schmitt-Rink, Phys. Rev. B **40**, 7406 (1989).

[57] B. Douçot and X.G. Wen, Phys. Rev. B **40**, 2719 (1989).

[58] M.B. Walker and Th.W. Ruijgrok, Phys. Rev. **171**, 513 (1968).

[59] D.K. Ghosh, Phys. Rev. Lett. **27**, 1584 (1971).

[60] G.S. Uhrig, Phys. Rev. B **45**, 4738 (1992).

[61] S.B. Fahy and D.R. Hamann, Phys. Rev. Lett. **65**, 3437 (1990).

[62] D.M. Ceperley, *private communication*.

[63] J.S. Yedidia, Phys. Rev. B **41**, 9397 (1990).

[64] Y.S. Yang and C.J. Thompson, J. Phys. A: Math. Gen. **24**, L279 (1991).

[65] B.S. Shastry, H.R. Krishnamurty, and P.W. Anderson, Phys. Rev. B **41**, 2375 (1990).

[66] W. von der Linden and D.M. Edwards, J. Phys.: Condens. Matter **3**, 4917 (1991).

[67] W.O. Putikka, M.U. Luchini, and T.M. Rice, Phys. Rev. Lett. **68**, 538 (1992); W.O. Putikka, M.U. Luchini, and M. Ogata, *ibid.* **69**, 2288 (1992).

[68] G.S. Rushbrooke, G.A. Baker, Jr., and P.J. Wood, in *Phase Transitions and Critical Phenomena*, Vol. 3, edited by C. Domb and M.S. Green (Academic, London, 1974).

[69] C.C. Wan, Commun. Theor. Phys. **22**, 233 (1994).

[70] J.C. Anglès d'Auriac, B. Douçot, and R. Rammal, J. Phys. Condens. Matter **3**, 3973 (1991).

[71] A. Sütő, Commun. Math. Phys. **140**, 43 (1991).

[72] A. Barbieri, J.A. Riera, and A.P. Young, Phys. Rev. B **41**, 11697 (1990).

[73] M. Takahashi, J. Phys. C: Solid State Phys. **10**, 1289 (1977).

[74] A.H. MacDonald, S.M. Girvin, and D. Yoshioka, Phys. Rev. B **37**, 9753 (1988); **41**, 2565 (1990).

[75] C.J. Thompson, Y.S. Yang, A.J. Guttmann, and M.F. Sykes, J. Phys. A: Math. Gen. **24**, 1261 (1991).

[76] D. Poilblanc and T.M. Rice, Phys. Rev. B **39**, 9749 (1989); H.J. Schulz, J. Phys. France **50**, 2833 (1989); K. Machida, Physica C **158**, 192 (1989); J. Zaanen and O. Gunnarsson, Phys. Rev. B **40**, 7391 (1989); M. Inui and P.B. Littlewood, Phys. Rev. B **44**, 4415 (1991).

[77] T. Giamarchi and C. Lhuillier, Phys. Rev. B **42**, 10641 (1990).

[78] H.J.M. van Bemmel, D.F.B. ten Haaf, W. van Saarloos, J.M.J. van Leeuwen, and Guozhong An, Phys. Rev. Lett. **72**, 2442 (1994).

[79] J.A. Vergés, E. Louis, P.S. Lomdahl, F. Guinea, and A.R. Bishop, Phys. Rev. B **43**, 6099 (1991); J. Galán, F. Guinea, J.A. Vergés, G. Chiappe, and E. Louis, *ibid.* **46**, 3163 (1992).

[80] K. Yonemitsu and A.R. Bishop, Phys. Rev. B **48**, 6680 (1993).

List of publications

D.F.B. ten Haaf and J.M.J. van Leeuwen, *High-temperature series expansions for the Hubbard model*, Phys. Rev. B **46**, 6313 (1992).
Chapter 2 of this thesis.

H.J.M. van Bemmel, D.F.B. ten Haaf, W. van Saarloos, J.M.J. van Leeuwen, and Guozhong An, *Fixed-node quantum Monte Carlo method for lattice fermions*, Phys. Rev. Lett. **72**, 2442 (1994).
Chapter 4 of this thesis.

D.F.B. ten Haaf, P.W. Brouwer, P.J.H. Denteneer, and J.M.J. van Leeuwen, *Low-temperature behavior of the large-U Hubbard model from high-temperature expansions*, Phys. Rev. B **51**, 353 (1995).
Chapter 3 of this thesis.

D.F.B. ten Haaf, H.J.M. van Bemmel, J.M.J. van Leeuwen, W. van Saarloos, and D.M. Ceperley, *Proof for an upper bound in fixed-node Monte Carlo for lattice fermions*, Phys. Rev. B **51**, 13039 (1995).
Chapters 4 and 5 of this thesis.

J.M.J. van Leeuwen and D.F.B. ten Haaf, *On the phase diagram of the Hubbard model*, to appear in Physica A (October 1995).

H.J.M. van Bemmel, W. van Saarloos, and D.F.B. ten Haaf, *Fixed-node Monte Carlo calculations for the 1D Kondo lattice model*, in preparation.

Nederlandse samenvatting

Over het magnetisch fasediagram van het Hubbard-model

Inzicht in het gedrag van wisselwerkende elektronen is van het grootste belang om de eigenschappen van zeer verschillende klassen van materialen te begrijpen, zoals de energetisch gunstigste vorm van kleine moleculen of magnetische en supergeleidende instabiliteiten in roostersystemen. Het eenvoudigste model waarmee gecorreleerde elektronen in een vaste stof kunnen worden beschreven is het Hubbard-model. Oorspronkelijk werd dit model toegepast om magnetisme in overgangsmetalen te beschrijven. Recentere voorbeelden waarbij het model een rol speelt zijn pogingen om theorieën op te stellen met betrekking tot supergeleiding bij hoge temperaturen. Waarschijnlijk zijn voor de beschrijving van dat fenomeen eigenlijk ingewikkelder modellen nodig, maar zelfs dit simpele model is nog niet goed begrepen. Voor één-dimensionale systemen zijn exacte resultaten bekend, maar verder zijn exacte berekeningen alleen gedaan voor kleine systemen, of zijn Monte Carlo-simulaties gebruikt.

In het Hubbard-model worden elektronen voorgesteld als deeltjes die gekarakteriseerd worden door een positie en door een spin toestand. De deeltjes kunnen zich alleen op roosterplaatsen bevinden (die de ionen in een metaal voorstellen), en ze kunnen van de ene plaats naar een andere, naastgelegen plaats springen. Dit springen gebeurt met een frequentie t/h, waarbij t de overlap aangeeft tussen twee toestanden die van elkaar verschillen doordat één elektron naar een naastgelegen positie is gesprongen, en waarbij h de constante van Planck is. De spin van een elektron heeft slechts twee mogelijke waarden, meestal aangeduid met 'op' ($+1/2$ of $\uparrow$) en 'neer' ($-1/2$ of $\downarrow$). Op grond van de halftallige spin moeten de elektronen voldoen aan het Pauli- (ook wel *uitsluitings-*)principe voor fermionen. Dit betekent dat de elektronen beperkt zijn in hun bewegingsvrijheid: twee elektronen met gelijke spin kunnen zich niet op dezelfde positie bevinden. Elektronen met tegengestelde spin kunnen dit wel, maar dit gaat gepaard met een toename van de energie met een bedrag U, wat een versimpelde vorm is van de Coulomb-afstoting tussen gelijknamige ladingen. Door de wederzijdse beïnvloeding van de spins van de elektronen bij het springen over het rooster, vanwege het Pauli-principe en de prijs voor het bij elkaar zetten van twee elektronen met tegengestelde spin, kunnen de elektronen magnetisch gedrag vertonen (itinerant magnetisme).

Hoewel er slechts twee energieparameters zijn (U en t), bezit het model een rijk geschakeerd fasediagram, waarvan de lage-temperatuur fase nog steeds vele geheimen herbergt. Uit een groot aantal al dan niet exacte berekeningen is reeds min of meer betrouwbare informatie verkregen over de mogelijke fasen. Er kunnen zowel antiferromagnetische als ferromagnetische tendensen worden aangetoond, en naar het zich laat aanzien een hele reeks van variaties hierop. Er is een algemene tendens naar antiferromagnetische ordeningen, welke overtuigend zijn aangetoond in het één-dimensionale geval. In dat geval is voor iedere eindi-

ge waarde van U de grondtoestand van het halfgevulde systeem[1] een antiferromagnetische isolator. Voor hoger-dimensionale systemen zijn de mogelijk voorkomende fasen minder goed begrepen. Men neemt aan dat ook hier het halfgevulde systeem antiferromagnetisch is. Verder kan men laten zien dat het model supergeleidend is bij negatieve U, onder omstandigheden die te relateren zijn aan die waar het antiferromagnetisch is bij positieve U. Een verrassend resultaat is bekend voor bipartiete roosters[2] met één gat,[3] die een ferromagnetische grondtoestand blijken te hebben als de afstoting tussen de elektronen oneindig sterk is.

Vele onderzoekers hebben het gedrag van het model bij grote waarde van de interactie U en lage gatendichtheid bestudeerd, om te bepalen of de ferromagnetische grondtoestand beperkt is tot één punt in het fase-diagram, of deel uitmaakt van een groter gebied met ferromagnetisch gedrag. Diverse methoden zijn en worden daarvoor ingezet, waaronder exacte berekeningen aan kleine systemen, gemiddelde-veld benaderingen, Monte Carlo-simulaties en reeksontwikkelingsmethoden. Vele interessante resultaten geven enig inzicht in de kwestie van het magnetisme. Een éénduidig antwoord op deze vraag is echter nog niet gegeven.

In dit proefschrift pak ik dit probleem op twee zeer verschillende manieren aan. Na een beknopte beschrijving van het Hubbard-model (in hoofdstuk 1) geef ik in het eerste deel een beschrijving van het genereren van hoge-temperatuur reeksontwikkelingen voor het Hubbard-model en het interpreteren van de resultaten daarvan (hoofdstukken 2 en 3). De reeksen worden verkregen door de kinetische term ($\sim t$) te beschouwen als een storing op de lokale (interactie) term ($\sim U$). Door een handige herschikking van de te berekenen termen kan de reeksontwikkeling herschreven worden als een sommatie over kleine systemen (de *cluster expansie*), waarvan de maximale grootte bepaald wordt door de orde in t die verkregen moet worden. De berekeningen kunnen dan eerst voor de kleinste clusters gedaan worden, en daarna voor steeds grotere systemen, tot de beschikbare computercapaciteit niet meer toereikend is om nog grotere clusters te behandelen. Op deze wijze bereken ik reeksontwikkelingen tot op achtste orde in t voor de groot-kanonieke potentiaal en diverse grootheden die daarvan met behulp van thermodynamische vergelijkingen kunnen worden afgeleid, en verder voor correlatiefuncties tussen spins. Hieruit kunnen bij relatief hoge temperaturen enige aanwijzingen voor ferromagnetisch gedrag worden gevonden, met name in de vorm van ferromagnetische naaste-buur correlaties. Het is echter niet eenvoudig om de resultaten voor hoge temperaturen naar lagere temperaturen te extrapoleren. Omdat op de bipartiete roosters waarmee ik werk oneven ordes in t niet voorkomen, staan er feitelijk maar 5 termen ter beschikking. Dit blijkt niet voldoende te zijn om met standaardmethoden betrouwbare informatie over het lage-temperaturen gedrag af te leiden.

In hoofdstuk 3 behandel ik daarom een nieuwe methode om een extrapolatie van de resultaten van de reeksontwikkelingen te maken. Deze methode is toegespitst op het ge-

---

[1] Dit is een systeem waarin precies evenveel elektronen aanwezig zijn als er roosterplaatsen zijn.

[2] Dit zijn roosters die men kan beschouwen als opgebouwd uit twee equivalente subroosters, waarbij alle buren van een willekeurig roosterpunt op het andere subrooster liggen.

[3] Indien op een roosterplaats geen elektronen aanwezig zijn spreekt men over een gat. Op een rooster met één gat is derhalve één elektron minder aanwezig dan er roosterplaatsen zijn.

bied van het fasediagram waarin ferromagnetisme wordt verwacht: lage gatendichtheid en grote interactiesterkte. De thermodynamische grootheden worden uitgedrukt in termen van een toestandsdichtheid voor de *gaten*. Het magnetisch karakter van de spins wordt hierbij opgenomen in een beschrijving van het halfgevulde systeem, waarop vervolgens correcties worden aangebracht door al dan niet wisselwerkende gaten te beschouwen. De toestandsdichtheid kan worden uitgerekend door momenten te bepalen op grond van de hogetemperatuur reeksontwikkelingen. Bij oneindige U kan hiermee voor de uniforme susceptibiliteit een extrapolatie tot temperatuur nul verkregen worden. Voor eindige U moet de spinachtergrond nog steeds door een reeksontwikkeling beschreven worden, maar kan wel tot zeer lage temperatuur geëxtrapoleerd worden. Uit analyse van de uniforme susceptibiliteit vind ik aanwijzingen voor het bestaan van een gebied in het fasediagram waar het systeem ferromagnetisch geordend kan zijn. Om echter met zekerheid te kunnen concluderen dat dat ook werkelijk de toestand van het systeem is, zou ook met antiferromagnetische of nog andere ordeningen vergeleken moeten worden. Daarvoor blijkt deze methode echter niet geschikt te zijn.

In het tweede deel van het proefschrift (hoofdstukken 4 en 5) beschrijf ik een methode waarmee het probleem vanuit een heel andere invalshoek bekeken kan worden. Met Green-Functie Monte Carlo-simulaties (GFMC) kan direct informatie over de grondtoestand van het systeem verkregen worden. Het idee van Monte Carlo-simulaties is dat een sommatie (of integraal) in een grote ruimte die vanwege de grote dimensie van die ruimte niet volledig kan worden uitgevoerd, toch zeer nauwkeurig kan worden uitgerekend door middel van stochastische wandelingen door die ruimte. De sommand wordt daarbij steeds lokaal berekend en gewogen, en na een voldoende aantal wandelingen van voldoende lengte wordt de exacte uitkomst verkregen met een maximaal gewenste statistische foutenmarge. Het is hierbij van belang dat de sommand geschreven kan worden als het produkt van een waarschijnlijkheid en een gewicht, zodanig dat de wandelingen gegenereerd kunnen worden volgens die waarschijnlijkheid en gewogen met dat gewicht. In GFMC wordt hiervan gebruik gemaakt door de wandelingen te genereren met behulp van een 'projectie' operator die uitgedrukt wordt in termen van de Hamiltoniaan, waarmee (formeel[4]) de grondtoestand uit een zo goed mogelijk gekozen begintoestand wordt gefilterd.

Een belangrijk probleem in het geval van fermionen, zoals in het Hubbard-model, komt voort uit het feit dat de golffunctie voor een dergelijk systeem antisymmetrisch[5] moet zijn op grond van het Pauli-principe. Het gevolg hiervan is dat, bij toepassing van GFMC zonder meer, de gewichten die aan de wandelingen moeten worden toegekend negatief kunnen worden. Dit *tekenprobleem* kan zich op verschillende manieren manifesteren, maar het uiteindelijke resultaat is vrijwel altijd dat de onnauwkeurigheid in de berekening veel te groot wordt. Er zijn verschillende methoden bedacht om hier een mouw aan te passen. Ik beschrijf een methode (*Fixed-Node Monte Carlo*, FNMC), waarmee het probleem kan worden omzeild. Men kan de mogelijke stappen tijdens een wandeling onderscheiden in stap-

---

[4] In een praktische situatie kan men geen informatie krijgen over de waarde van de golffunctie in de grondtoestand, maar alleen over de eigenschappen van de grondtoestand.

[5] Dit betekent dat de golffunctie een tegengestelde waarde moet aannemen als twee deeltjes met gelijke spin worden verwisseld.

pen die wel en stappen die niet tot het tekenprobleem leiden. Indien men nu de stappen die tot het tekenprobleem aanleiding geven eenvoudigweg verwijdert, wordt het probleem vermeden. De prijs die daarbij betaald moet worden is dat feitelijk de eigenschappen van een andere Hamiltoniaan worden berekend, die niet dezelfde zijn als die van de Hamiltoniaan waarin men eigenlijk geïnteresseerd is. Voor systemen waarbij de wandelingen in een continue ruimte worden uitgevoerd, wordt FNMC al enige tientallen jaren met succes toegepast. Men kan laten zien dat in dat geval de waarde van de energie die zo verkregen wordt altijd boven de ware waarde ligt, en in de praktijk meestal vrij dicht erbij. In het geval van het Hubbard-model hebben we echter te maken met een discrete configuratieruimte. Zonder meer weglaten van de ongewenste stappen leidt hier tot een resultaat voor de energie waarvoor geen enkel verband meer kan worden gegeven met de ware energie. In hoofdstuk 4 bewijs ik dat met een aantal extra aanpassingen van de Hamiltoniaan op basis van de begintoestand toch een bovengrens voor de energie gevonden kan worden, zodanig dat FNMC ook toepasbaar is op fermionen op een rooster.

Het laatste obstakel om deze methode met succes te kunnen toepassen is het vinden van een geschikte begintoestand. In hoofdstuk 5 laat ik zien dat in specifieke voorbeelden op relatief kleine systemen redelijke tot goede resultaten gevonden kunnen worden, maar ook dat de benadering staat of valt met de kwaliteit van de tekenstructuur van de begintoestand. Er zijn nog nauwelijks criteria bekend om te kunnen bepalen wat een goede tekenstructuur is, zodat er geen systematische manier bestaat om een goede begintoestand te construeren. Voordat met FNMC zinnige en betrouwbare informatie over het probleem van ferromagnetisme in het Hubbard-model verkregen kan worden, is nader inzicht hieromtrent nodig.

In de afsluitende paragraaf bespreek ik de resultaten die in dit proefschrift zijn afgeleid, in vergelijking met diverse bronnen uit de literatuur. De conclusie hiervan luidt dat er zeker aanwijzingen zijn voor het bestaan van een ferromagnetische fase in het twee- en drie-dimensionale Hubbard-model, maar dat die nog uitgebreid vergeleken dienen te worden met andere mogelijke ordeningen, voor een definitief antwoord op de gestelde vragen gegeven kan worden.

Resumo en Esperanto

Pri la magneta fazodiagramo de la Hubbard-modelo

Pli profunda scio pri la konduto de interagantaj elektronoj tre gravas por kompreni la ecojn de tre malsamaj klasoj da materialoj, kiel la energie plej konvenan formon de etaj molekuloj, aŭ magnetajn kaj superkonduktivajn malstabilecojn en latiso-sistemoj. La plej simpla modelo, per kiu oni povas priskribi korelativigitajn elektronojn en solido, estas la Hubbard-modelo. Origine tiu ĉi modelo estis uzata por priskribi magnetismon en transiraj metaloj. Pli freŝdataj ekzemploj, en kiuj la modelo ludas rolon, koncernas klopodojn starigi teoriojn pri superkonduktiveco ĉe altaj temperaturoj. Por priskribi tiun lastan fenomenon oni supozeble bezonas pli komplikajn modelojn, sed eĉ tiu ĉi simpla modelo ne jam estas bone komprenata. Por unudimensiaj sistemoj ekzaktaj rezultoj estas konataj, sed krome efektiviĝis ekzaktaj kalkuloj nur por malgrandaj sistemoj, aŭ uziĝis Montekarlo-simuladoj.

En la Hubbard-modelo elektronoj prezentiĝas kiel partikloj karakterizitaj de pozicio kaj de spinmomanto. La partikloj povas troviĝi nur en la latislokoj (kiuj reprezentas la jonojn en metalo), kaj ili povas salti de unu loko al alia, najbara loko. Tiu saltado okazas kun frekvenco t/h, kie t estas la interkovriĝo de du statoj, kiuj diferencas unu disde la alia per tio ke unu elektrono saltis al la najbara pozicio, kaj kie h estas la konstanto de Planck. La elektronspino havas unu el du eblaj valoroj, indikitaj per 'supren' ($+1/2$ aŭ ↑) kaj 'malsupren' ($-1/2$ aŭ ↓). Pro la duonvalora spinmomanto de la elektronoj validas la (*ekskludo-*)principon de Pauli por fermionoj. Tio limigas la moviĝivon de la elektronoj: du elektronoj kun samaj spinoj ne povas troviĝi en la sama loko. Elektronoj kun malsamaj spinoj ja povas troviĝi en la sama loko, sed tion akompanas kresko de la energio per kvanto U. Tio estas simpligita formo de la forpuŝa kulombforto inter samspecaj ŝargoj. Pro la reciproka influado de la spinoj dum la elektronsaltoj sur la latiso, rezulte de la principo de Pauli kaj de la kosto (t.e. la energikresko) por kunmeti du elektronojn kun kontraŭaj spinoj, la elektronoj povas elmontri magnetan konduton (itineranta magnetismo).

Kvankam ĝi havas nur du energiparametrojn t kaj U, la modelo posedas riĉe nuancitan fazodiagramon, kies malaltatemperatura fazo daŭre tenas multajn sekretojn. El multaj foje jes, foje ne ekzaktaj kalkuloj jam akiriĝis pli aŭ malpli fidindaj informoj pri la eblaj fazoj. Pruveblas kaj antiferomagnetaj kaj feromagnetaj tendencoj, kaj atendindas multegaj variaĵoj inter tiuj du. La modelo ĝenerale emas al antiferomagnetaj ordiĝoj, kiujn oni konvinke pruvis por la unudimensia sistemo. Tiukaze ĉe ĉu finia valoro de U la fundostato de la duonokupita sistemo[1] estas antiferomagneta izolaĵo. La eblajn fazojn de sistemoj en pli altaj dimensioj oni malpli bone komprenas. Oni supozas ke ankaŭ tie la duonokupita sistemo estas antiferomagneta. Krome oni povas montri ke la modelo estas superkonduktiva je negativa U sub cirkonstancoj kiuj rekte rilatas al tiuj sub kiuj ĝi estas antiferomagneta

---

[1] Duonokupita sistemo enhavas tiom da elektronoj kiom da latislokoj.

je pozitiva U. Surpriza rezulto estas konata por dupartiaj latisoj[2] kun unu truo,[3] en kiuj oni trovas feromagnetan fundostaton, se la forpuŝa forto inter la elektronoj estas infinita.

Multaj esploristoj pristudis la konduton de la modelo en la kazo de granda interagado U kaj malgranda truodenseco, por precizigi ĉu la feromagneta fundostato estas limita al unu punkto en la fazodiagramo, aŭ ĉu ĝi estas parto de pli granda regiono de feromagneta konduto. Por tio oni uzas diversajn metodojn, inter kiuj ekzaktaj kalkuladoj de malgrandaj sistemoj, mezkampaj aproksimadoj, Montekarlo-simuladoj kaj serioekspansioj. Multaj interesaj rezultoj iom prilumas la demandon pri magnetismo. Tamen, klara respondo al tiu demando ankoraŭ ne ekzistas.

En tiu ĉi disertaĵo mi traktas tiun problemon laŭ du tre malsamaj manieroj. En la unua parto, koncize priskribinte la Hubbard-modelon en la unua ĉapitro, mi klarigas la starigon de altatemperaturaj serioekspansioj por la Hubbard-modelo, kaj la manieron interpreti ties rezultojn (ĉapitroj 2 kaj 3). Mi akiras la seriojn, konsiderante la kinetan termon ($\sim t$) perturbo de la loka (interaga) termo ($\sim U$). Per lerta reordigo de la kalkulotaj termoj la ekspansioj povas esti reesprimataj kiel sumo de kontribuoj el etaj sistemoj (faskoekspansio), kies maksimuman grandecon difinas la ordo en t, kiun oni volas kalkuli. La kalkuladoj estas farataj, unue por la plej malgrandaj faskoj, poste por pli kaj pli grandaj, ĝis la disponebla komputila kapacito ne plu sufiĉas por trakti pli grandajn sistemojn. Tiel mi kalkulas seriojn ĝis la oka ordo en t por la grandkanona potencialo kaj diversaj grandoj kiuj deriviĝas de ĝi pere de la termodinamikaj ekvacioj, kaj ankaŭ por korelaciofunkcioj inter la spinoj. Ĉe relative altaj temperaturoj troveblas indikoj el tio por feromagneta konduto, en formo de feromagnetaj korelacioj inter najbaraj lokoj. Ne estas facile, tamen, eksterpoli la rezultojn de altaj al malaltaj temperaturoj. Ĉar en dupartiaj latisoj, kiujn mi traktas, ne aperas malparaj ordoj en t, fakte nur 5 termoj disponeblas. Tio montriĝas nesufiĉa por derivi, pere de normaj metodoj, fidindajn informojn pri la konduto ĉe malaltaj temperaturoj.

Pro tio mi traktas en ĉapitro 3 novan metodon por eksterpoli la rezultojn de la serioekspansioj. Ĝi taŭgas precipe por la parto de la fazodiagramo kie oni atendas feromagnetismon, t.e. je malgranda truodenseco kaj granda interagado. La termodinamikaj grandoj esprimiĝas en termoj de statodenso de la *truoj*. La magneta karaktero de la spinoj estas enplektita en priskribo de la duonokupita sistemo, al kiu priskribo konsideroj pri truoj sen aŭ kun interago alportas korektojn. La statodenson oni povas kalkuli, fiksante ĝiajn momantojn surbaze de la serioekspansioj. Per tio oni je infinita U povas eksterpoli la unuforman susceptiblon ĝis temperaturo nula. Je finia U la spinfonon tamen priskribas serioekspansio, sed eblas ĝin eksterpoli ĝis tre malaltaj temperaturoj. Analizante la unuforman susceptiblon, mi trovas indikojn, ke ekzistas regiono en la fazodiagramo kie la sistemo povas ordiĝi feromagneta. Por konkludi kun certeco, ke tio estas la efektiva stato de la sistemo, necesas komparo kun antiferomagneta aŭ aliaj ordiĝoj. Tiu ĉi metodo por tio tamen ne taŭgas.

En la dua parto de la disertaĵo (ĉapitroj 4 kaj 5) mi okupiĝas pri metodo traktanta la problemon de tute alia elirpunkto. Per Green-Funkcio Montekarlo-simuladoj (*Green Func-*

---

[2]Dupartia latiso estas latiso kiun oni povas konsideri kiel kunmeton de du ekvivalentaj sublatisoj, tiel ke ĉiuj najbaroj de ajna latisloko troviĝas en la alia sublatiso.

[3]Se en latisloko ne estas elektrono, oni parolas pri truo. Sekve, latiso kun unu truo enhavas unu elektronon malpli ol latislokojn.

tion Monte Carlo, GFMC) akireblas informoj rekte pri la fundostato de la sistemo. En Montekarlo-simulado, sumigo (aŭ integralo) en granda spaco, kiun oni pro la granda dimensio de la spaco ne povas fari, tamen povas esti kalkulata tre precize per stokastaj promenoj en tiu spaco. La sumigato dum la promenoj ĉiam loke kalkuliĝas kaj pesiĝas, kaj post sufiĉe multe da promenoj kun sufiĉa longeco, oni akiras precizan respondon kun dezirita maksimuma deviomarĝeno. Gravas en tio ke la sumigato povas esti interpretata kiel la produkto de probablo kaj pezo, tiel ke la promenoj povas esti elektataj laŭ tiu probablo, kaj pesataj laŭ tiu pezo. En GFMC oni tion eluzas, kreante la promenojn per 'projekcia' operatoro esprimita en termoj de la Hamiltoniano, per kiu operatoro oni (formale[4]) filtras la fundostaton el kiel eble plej bone elektita komenca stato.

Grava problemo en la kazo de fermionoj, do ankaŭ en la Hubbard-modelo, devenas de la fakto ke la ondofunkcio por tia sistemo devas esti antisimetria[5] pro la principo de Pauli. Sekve, se oni aplikas GFMC senkonsidere, la pezoj atribuendaj al la promenoj povas iĝi negativaj. Tiu ĉi *signoproblemo* povas manifestiĝi en diversaj manieroj, sed preskaŭ ĉiam la malprecizo iĝas multe tro granda. Ekzistas pluraj metodoj por elturniĝi el tiu problemo. Mi priskribas metodon, Fiks-Nodan Montekarlon (*Fixed-Node Monte Carlo*, FNMC), per kiu oni ĉirkaŭiras ĝin. Oni povas distingi la eblajn paŝojn dum promenado kiuj kaŭzas la signoproblemon, disde tiuj kiuj ne kaŭzas ĝin. Se oni simple ellasas la paŝojn kiuj kaŭzas la problemon, oni evitas ĝin. La malavantaĝo kiu tion akompanas estas ke oni fakte kalkulas la ecojn de alia Hamiltoniano, malsamajn al tiuj de la Hamiltoniano pri kiu oni interesiĝas. Jam dum kelkaj jardekoj oni uzas FNMC en sistemoj ĉe kiuj la promenoj okazas en kontinuaj spacoj. Tie oni povas tiel aranĝi la aferon, ke la valoro de la energio trovita ĉiam superas la veran valoron, kaj en kutima situacio sufiĉe bone proksimas ĝin. En la Hubbard-modelo ni havas alian situacion, ĉar tie temas pri diskreta spaco. Ellasi la nedeziritajn paŝojn sen pliaj adaptoj tiukaze rezultigas valoron de la energio sen klara rilato kun la vera energio. En ĉapitro 4 mi pruvas ke oni povas adapti la Hamiltonianon surbaze de la komenca stato, tiel ke oni tamen akiras supran limon por la energio, kio igas FNMC aplikebla al la fermionoj sur latiso.

La lasta obstaklo por sukcesa uzado de tiu ĉi metodo estas la neceso trovi taŭgan komencan staton. En ĉapitro 5 mi montras per kelkaj specifaj ekzemploj por malgrandaj sistemoj ke akireblas sufiĉe bonaj rezultoj, sed ankaŭ ke la aproksimo tre dependas de la kvalito de la signostrukturo de la komenca stato. Apenaŭ konatas kriterioj por scii kio estas bona signostrukturo, tiel ke ne ekzistas sistema maniero por konstrui bonan komencan staton. Necesas pli adekvata kompreno pri tio, antaŭ ol oni povas pere de FNMC ekhavi utilajn kaj fidindajn informojn pri la problemo de feromagnetismo en la Hubbard-modelo.

En la fina paragrafo mi diskutas la rezultojn de la laboro priskribita en la prezentita disertaĵo, komparante ilin kun diversaj fontoj en la fakĵurnalaro. Konklude, certe troveblas indikoj pri la ekzisto de feromagneta fazo en la du- kaj tridimensia Hubbard-modelo, sed necesas ampleksa komparo kun aliaj eblaj ordiĝoj antaŭ ol oni povas doni definitivan respondon al la demandoj starigitaj.

---

[4]Praktike ne akireblas informoj pri la valoro de la fundostata ondofunkcio, nur pri la ecoj de la fundostato.
[5]Tio signifas ke la ondofunkcio devas preni kontraŭan valoron se du samspinaj partikloj interŝanĝiĝas.

Curriculum vitae

Op 11 september 1968 werd ik geboren te Mook, alwaar ik de lagere school St. Adalbert bezocht. Van 1980 tot 1986 genoot ik voorbereidend wetenschappelijk onderwijs aan het Dominicus College te Nijmegen, en ik behaalde het diploma gymnasium β. Daarna begon ik met de studie natuurkunde aan de Katholieke Universiteit te Nijmegen, waar ik in augustus 1987 het propaedeutisch examen behaalde. Na ook het tweede jaar van de studie in Nijmegen te hebben afgerond, zette ik de studie in 1988 aan de Rijksuniversiteit te Leiden voort. In 1989 werkte ik tijdens de introductieperiode laboratorium (IPL) in de werkgroep metalen (MtIV) aan *Numerieke methoden in de supergeleiding*, onder begeleiding van dr. J. Aarts. Ik zocht hierbij numerieke oplossingen van de gelineariseerde Ginzburg-Landau vergelijking, met als doel het kritiek magnetisch veld te berekenen waarbeneden supergeleiding optreedt in de oppervlaktelaag van een magnetisch materiaal, als functie van de dikte van het materiaal en van de hoek waaronder het veld wordt aangelegd. Van augustus 1989 tot januari 1990 volgde een (extracurriculaire) stage in het Philips Natuurkundig Laboratorium te Eindhoven, waar ik onder supervisie van dr.ir. J.J.M. Ruigrok werkte aan *Theorie en berekeningen met betrekking tot magnetische-kracht microscopie*. Op grond van een simpel model voor magnetische wisselwerking berekende ik de kracht die door een in ideale streep-patronen gemagnetiseerde dunne film wordt uitgeoefend op een gemagnetiseerd object. Doel was een theoretische onderbouwing op te zetten voor de metingen die in het Nat.Lab. met een magnetische-kracht microscoop werden verricht. Terug in Leiden maakte ik in februari 1990 een aanvang met de theoretische afstudeeropdracht *Reeksontwikkelingen voor het Hubbard model* onder begeleiding van prof.dr. J.M.J. van Leeuwen, in de vakgroep theoretische natuurkunde. Na mijn doctoraal examen in februari 1991 werkte ik verder aan dit onderwerp, eerst als Assistent in Opleiding (AIO), en vanaf juli 1991 als Onderzoeker in Opleiding (OIO). Tijdens de promotiewerkzaamheden heb ik de 18e IUPAP conferentie over statistische fysica in Berlijn (augustus 1992) bezocht, aan zomerscholen in Duitsland (Altenberg, juli 1993) en Italië (Trieste, mei 1994) deelgenomen, en een werkbezoek in Zwitserland (Zürich, november 1991) afgelegd. Ik volgde AIO-cursussen in Enschede (maart 1991 en mei 1992) en Groningen (december 1993). Voorts heb ik deelgenomen aan de FOM managementtraining op Nijenrode (september 1994).

And now, for something completely different ...
Monty Python's Flying Circus

STELLINGEN

1. De nauwkeurigheid die veel kristallografen aan Rietveld-analyse toekennen berust op een grove overschatting van de toepasbaarheid van de methode der kleinste kwadraten.

   H.M. Rietveld, J. Appl. Cryst. 2, 65 (1969).

2. Zij gegeven de Fibonacci-getallen F_k:

   $$F_1 = 1 \; ; \quad F_2 = 1 \; ; \quad F_k = F_{k-1} + F_{k-2} \quad \text{voor gehele } k \geq 3,$$

   dan geldt voor gehele en positieve k en m:

   $$k \bmod m = 0 \Longrightarrow F_k \bmod F_m = 0.$$

3. Het is niet zinvol te proberen een golffunctie door middel van Jastrow-achtige factoren geschikter te maken als startfunctie in een Fixed-Node Monte Carlo-simulatie, tenzij de golffunctie in (bijna) de gehele faseruimte het juiste teken heeft.

   Dit proefschrift, hoofdstuk 5.

4. Het indiceren van poederdiagrammen kan zeer veel efficiënter dan tot op heden gedaan worden door gebruik te maken van recente ontwikkelingen op getaltheoretisch gebied.

   A.K. Lenstra, H.W. Lenstra Jr. en L. Lovász, Math. Ann. 261, 515 (1982).

5. Voor een volledige analyse van de diep-inelastische verstrooiingsexperimenten om de structuurfunctie F_2 te bepalen bij kleine schalingsvariabele x, is het nodig om de 3e orde QCD correcties op de anomale dimensie uit te rekenen.

   E.B. Zijlstra, proefschrift, Rijksuniversiteit te Leiden (1993).

6. Voor een willekeurige set eigenwaarden $\{\varepsilon\}$ geldt

   $$\sum_{\varepsilon_j} \frac{-2}{\varepsilon_1 - \varepsilon_2 + \varepsilon_3 - \varepsilon_4} \left[\prod_{j=1}^{4} \frac{e^{(-1)^j \varepsilon_j}}{1 + e^{(-1)^j \varepsilon_j}} \right] = \sum_{\varepsilon_i, l_j} \delta_{l_1 + l_3, l_2 + l_4} \left[\prod_{j=1}^{4} \frac{1}{\pi i (2l_j + 1) - \varepsilon_j} \right],$$

   waarbij de ε_j ($j = 1, 2, 3, 4$) gesommeerd worden over de set $\{\varepsilon\}$ en de l_j ($j = 1, 2, 3, 4$) over de gehele getallen van $-\infty$ tot ∞.

   D.F.B. ten Haaf, afstudeerscriptie, Rijksuniversiteit te Leiden (1991).

7. Een gesprek tussen twee personen met verschillende moedertalen, dat in één van beide talen wordt gevoerd, is zelden diepgaand.

8. Het is verwonderlijk dat de meeste Nederlanders, ondernemers uitgezonderd, directe belastingen als meer belastend ervaren dan indirecte belastingen.

9. Om onderzoekers in opleiding te stimuleren snel hun proefschrift af te ronden, zou het effectiever zijn de korting op hun salaris gedurende de contractperiode niet af maar toe te laten nemen.

10. Niet de taal is seksistisch, maar degene die haar gebruikt.

<div align="right">

Danny ten Haaf
6 september 1995

</div>