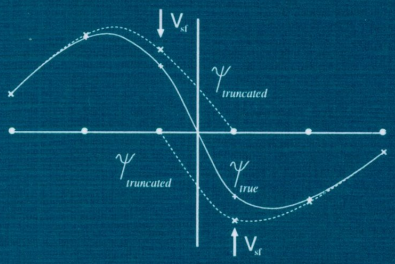
# REAL-SPACE ASPECTS OF CORRELATED FERMIONS




H.J.M. van Bemmel

22 SEP. 1995

RIJKSUNIVERSITEIT LEIDEN
0926 8389

RIJKSUNIVERSITEIT TE LEIDEN
BIBLIOTHEEK INSTITUUT-LORENTZ
Postbus 9506 - 2300 RA Leiden
Nederland

Kast dissertaties

# REAL-SPACE ASPECTS OF CORRELATED FERMIONS

# REAL-SPACE ASPECTS OF CORRELATED FERMIONS

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR
AAN DE RIJKSUNIVERSITEIT TE LEIDEN, OP GEZAG
VAN DE RECTOR MAGNIFICUS DR. L. LEERTOUWER,
HOOGLERAAR IN DE FACULTEIT DER GODGELEERDHEID,
VOLGENS BESLUIT VAN HET COLLEGE VAN DEKANEN
TE VERDEDIGEN OP WOENSDAG 27 SEPTEMBER 1995
TE KLOKKE 15.15 UUR

DOOR

Henricus Jozef Martinus van Bemmel

GEBOREN TE NIJMEGEN OP 9 JULI 1968

Promotiecommissie:

| | |
|---|---|
| Promotor: | Prof.dr.ir. W. van Saarloos |
| Referent: | Prof.dr. J.M.J. van Leeuwen |
| Overige leden: | Prof.dr. J. Amesz |
| | Prof.dr. C.W.J. Beenakker |
| | Prof.dr. H. de Raedt (RUG) |
| | Prof.dr. L.J. de Jongh |
| | Prof.dr. J.A. Mydosh |
| | Dr. J. Zaanen |

# Contents

# 1 Introduction

## 1.1 Perspective

Heavy-fermion [1] and high-$T_c$ [2] materials cannot be understood without considering the electrons as a many-body system. The advent of these strange metals brought the study of models of interacting fermions to the foreground. Compared to non-interacting fermions, qualitatively new features arise in two respects. First, electron-electron interactions cause tendencies towards magnetic order. Second, the nature of the charge carriers changes, in some cases. In this respect, the essential features known from free electrons are rather robust. If these are preserved by interactions, a system is said to show 'Fermi liquid' behaviour. In low dimensions, qualitatively different ('non-Fermi liquid') behaviour has been found experimentally [3].

On the theoretical side, the phase diagrams of systems of correlated fermions, such as the Hubbard, $t$-$J$, Periodic Anderson and Kondo Lattice models, have been studied using various techniques. Both the magnetic properties and the nature of the charge carriers (Fermi liquid or non-Fermi liquid) continue to be subject of debate, theoretically [4] as well as experimentally [3]. In this introductory chapter, a brief overview of methods and results will be given.

In this thesis, some aspects of correlated fermions will be treated in more detail. These are related to three fields of interest. Let us, for each field, formulate more specific questions we shall discuss. In the separate chapters these will be introduced more extensively, here they serve as motivation for presenting the background material in this chapter.

The first field is the comparison of Fermi liquid and non-Fermi liquid behaviour of the charge carriers. We shall not study the breakdown of Fermi liquid theory itself, rather, we extend an existing non-Fermi liquid phenomenology for the normal state of high-$T_c$ superconductors, the so-called marginal Fermi liquid.

Our first question is: "If the transport properties show marginal Fermi liquid behaviour, instead of Fermi liquid behaviour, will the response to a dynamic *local* disturbance also be different?". This is the subject of chapter 2. In section 1.2.2, we shall introduce the notion of a Fermi liquid and the issue of its stability. The marginal Fermi liquid hypothesis was originally formulated to describe the transport properties of the metallic phase, above $T_c$, of cuprate superconductors. We shall describe it in section 2.2.1. Local properties are very different from transport properties. We shall discuss the question whether an experimental test of our new theoretical findings for local properties would be able to discriminate between Fermi liquid and marginal Fermi liquid behaviour.

The second field of study is the possibility of obtaining ground state wave functions with spin and charge distributions which are inhomogeneous in space. We have

a proposal to experimentally scrutinize existing explanations of thermodynamic measurements in solid $^3He$, which are in terms of these structures.

In some approximation, a *ferromagnetic spin polaron* is obtained as the structure accompanying a hole, in the region of the phase diagram of the Hubbard model that is relevant for solid $^3He$. The question we address is: "If measurements are explained in terms of trapped vacancies in heavy spin polarons, could one perform a direct experimental test to verify this explanation?". Chapter 3 is devoted to this question. The fact that the existing explanation is in terms of *trapped* vacancies, and the state involved is, therefore, metastable, suggests a procedure of *magnetic annealing*, which will be the key feature of our proposed test. The more general issue of inhomogeneous solutions will reappear in later chapters.

The third field concerns methods to calculate properties of phase diagrams stochastically. For fermions, both the finite temperature and the ground state quantum Monte Carlo methods are hampered by the sign problem, i.e. the occurrence of cancellations because all sampled quantities have positive and negative contributions that are exponentially large compared to the final answer, causing statistical errors to be large.

Chapter 4 focuses on a new method for studying models of correlated fermions on a lattice. This fixed-node quantum Monte Carlo method involves random walks in configuration space. It is an attempt to supply an answer to the question: "Which approximation should be made in order to avoid the sign problem in stochastic simulations aiming at the ground state properties of lattice many-fermion systems?". The idea of forcing the zeroes of the wave function to be the same as those of a trial wave function, an approximation that limits the class of possible resulting wave functions but is variational in the continuum case, is extended to lattice models. The discreteness of lattice models makes the definition of the location of a node problematic. Our new algorithm deals with the non-locality in lattice problems by replacing kinetic terms by potentials, in such a way that the sign problem is avoided, without violating the requirement that a method should be variational.

Chapter 5 is an application of this method to the Kondo Lattice model, it addresses the question "How does the fixed-node quantum Monte Carlo method for lattice fermions perform in practice?". Magnetic properties, correlations that are in real-space language, are studied naturally by random walk Monte Carlo, in contrast to the issue of the dynamics close to the Fermi surface, i.e. the Fermi liquid or non-Fermi liquid properties of the charge carriers. For the one-dimensional Kondo Lattice model, we calculate spin-spin correlations. While studying the excitations of the system, we shall encounter an inhomogeneous structure in space, similar in many respects to the inhomogeneous structures in the Hubbard model mentioned before. The requantization of these structures, i.e. letting them propagate in space, is also studied using fixed-node Monte Carlo.

The new method for dealing with the sign problem was designed for lattice fermions. It also has the potential to become an important tool for studying different systems that exhibit a sign problem, most notably frustrated quantum spins. In chap-

ter 6, we discuss the following questions: "What is known about the sign structure of many-particle systems, how should one generalize theorems known for continuum systems to the lattice case, and which are the consequences for the applicability of the lattice fixed-node Monte Carlo method?".

Our new application of the marginal Fermi liquid ansatz, the localized structures we discuss and the stochastic method we introduce have one thing in common: they have a real-space character rather than a $\vec{k}$-space one. The title of this thesis is chosen to emphasize this unifying element. Different connections exist. As is discussed in ref. [5], the formation of a spin polaron around a hole is a change of the charge carrier. It becomes effectively very heavy and tends to be localized. This implies that the usual Fermi liquid picture is no longer applicable. Moreover, signs of the wave function are crucial in the argument in ref. [5] that in a one-hole doped antiferromagnet, the Fermi liquid picture breaks down, either due to this localization effect, or due to the disappearance of the quasi-particle weight (a notion we shall introduce in section 1.2.2). In this thesis, no work on these connections between the nature of quasi-particles, inhomogeneous structures and signs of wave functions will be presented. We shall discuss the separate questions mentioned before and our contributions to the marginal Fermi liquid phenomenology, possible experimental tests for the existence of inhomogeneous structures, and the development of a Monte Carlo method that involves signs of wave functions.

The appreciation of the subtleties of models of correlated fermions requires some insight in simpler cases. On the one hand, for most metals mean-field assumptions in theoretical models suffice to reproduce the experimental behaviour. In these cases, the elementary excitations are free-electron like. On the other hand, some insulators are quantum magnets, these consist of localized quantum spins. Either going beyond a band-structure calculation in the case of interacting electrons, or introducing doped charge carriers in a quantum magnet, yields a many-electron problem. The two simpler cases will be treated in section 1.2, the full problem in section 1.3.

Once the models for correlated fermions are introduced, we shall give a brief overview of the methods used to study them, and of what is known about their phase diagrams, both about the magnetic properties and about the Fermi-liquid or non-Fermi liquid character. This sets the stage for the detailed discussion of our own contributions in later chapters.

## 1.2 Spins and electrons

A solid consists of many nuclei and electrons. In principle, Coulomb interactions exist among all particles, and the solution of the corresponding many-body Schrödinger equation would yield the properties of the system. This is not feasible in practice. Depending on the type of solid, different effective models are used. In all of these, a lattice of ions is assumed to exist. We shall not discuss the properties of these lattices, nor the complications that arise when defects cause disorder. We concentrate on the electrons.

### 1.2.1 Quantum spin systems

In one extreme situation, one electron is localized on each lattice site, only its spin can vary in time. The spin system that has the most relevance for later discussions, because it gives the low-energy behaviour of the half-filled Hubbard model, is the Heisenberg model. Its Hamiltonian is

$$\mathcal{H} = J \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j, \tag{1.2.1}$$

in which $\langle i, j \rangle$ denotes pairs of nearest neighbours, and in which the operators satisfy the usual commutation relations of spin operators

$$[S^x, S^y] = iS^z, \tag{1.2.2}$$

and cyclic.

At high temperatures, the quantum fluctuations introduced by the commutation relations Eq.(1.2.2) are unimportant compared to the thermal fluctuations. Neglecting the quantum fluctuations, one obtains the *classical* Heisenberg model, which has the same Hamiltonian Eq.(1.2.1), and interacting vectors instead of quantum spins obeying Eq.(1.2.2).

For the discussion in chapter 3 about Helium, and for the appreciation of the Monte Carlo method in later chapters, we need to know in which cases such a system shows long-range magnetic order. For $J > 0$, the system has a tendency to become antiferromagnetic. Whether it shows long-range order strongly depends on the dimension, and on the type of lattice. For the *classical* model, a general theorem (Hohenberg-Mermin-Wagner) [6] forbids long range order at finite temperatures in dimensions $d \leq 2$. In higher dimensions on a regular lattice, there is a classical phase transition at some finite temperature, separating an ordered low-temperature phase from a disordered high-temperature phase. The critical exponents characteristic of such a phase transition can be obtained using various techniques, most notably real-space [7] or $\vec{k}$-space [8] renormalization, and finite-size scaling combined with the notion of conformal invariance [9] in two dimensions.

At zero temperature, thermal fluctuations are absent. The tendency to order is counteracted only by the quantum fluctuations. Their role is illustrated by the observation that the classical ground state on a bipartite lattice, which has up spins on one sublattice and down spins on the other sublattice, is not an eigenstate of the Heisenberg Hamiltonian. The issue is whether there is a non-zero staggered magnetization (Néel order) in the, more complicated, ground state. Again, order is more easily destroyed in low dimensions. Often the dimension above which order is possible for a certain model, is one lower for the ground state than for finite temperatures. In the ground state, long-range order is possible in two dimensions [10].

For quantum spins, not only the dimension of the lattice determines whether there is long-range magnetic order, the quantum number $S$ ($S = n/2, S^z = -S, -S+1, ..., S$) of the localized spins is also important. In general, the tendency to order is

stronger for higher spin, quantum fluctuations are suppressed. In addition, there is a difference between integer and half-integer spin, especially in one dimension. This difference was uncovered by Haldane. It is of a topological nature [11, 12].

One-dimensional systems with a continuous symmetry and without long-range interactions do not order in the ground state. An important intermediate case between order (correlation functions that go to a constant value at long distances) and disorder (exponentially decaying correlation functions) exists: algebraic order (correlation functions decaying as a power-law) [13]. In classical statistical mechanics, this is the behaviour found at the phase transition point, and also in the low-temperature phase of the Kosterlitz-Thouless phase transition (two-dimensional XY-model; note that order is not forbidden by the Hohenberg-Mermin-Wagner theorem). The one-dimensional $S = 1/2$ antiferromagnetic Heisenberg model shows this power-law behaviour in the ground state. For this case, much is known from the Bethe-ansatz solution [14].

For $S = 1$, the correlation functions decay exponentially. There is a corresponding difference between integer and half-integer spin in the excitation spectrum. If a system with continuous symmetry orders, there are always gapless excitations (Goldstone modes). This is also true if the ordering is algebraic. Without order, with exponentially decaying correlation functions, there is a gap for excitations. Various numerical techniques have been used to provide evidence for the existence of the so-called Haldane gap in the one-dimensional, $S = 1$ Heisenberg model. Its value is $0.41050(2)J$ [15].

In two dimensions, order is not ruled out by a general theorem. For certain cases, proofs exist for the *presence* of order in the ground state. For the Heisenberg model on hypercubic lattices this is true for all $S$ in $d \geq 3$ and for $S \geq 1$ in $d = 2$. These theorems are due to Dyson, Lieb and Simon and to Neves and Peres. [16, 17, 18]. For $d = 2, S = 1/2$, there is no analytic proof of long-range order. A different theorem, due to Lieb, Schultz and Mattis [19] says that there cannot be a unique ground state with a gap (a Haldane phase) [20, 13]. Numerical evidence [21] shows that there is Néel order, with corresponding gapless Goldstone modes.

This latter fact is the most important observation about quantum spins if one wants to consider the two-dimensional Hubbard model as a $S = 1/2$ quantum magnet with charge fluctuations added. To supplement the picture, let us briefly consider aspects different than gaps and order in the ground state.

- Finite temperature properties of the Heisenberg model are often studied using the non-linear $\sigma$-model [22]. If there is long-range order at $T = 0$, the long-wavelength behaviour appears to be classical, with parameters renormalized by the quantum fluctuations.

- There is an interesting correspondence between the phase diagrams of crystal surfaces and quantum spin chains [23]. The Haldane phase corresponds to the disordered flat phase. Both have a non-local order parameter.

- This comparison extends to more complicated models, which in the spin case include squares of dot-products of spin operators on neighbouring sites [13]:

$$\mathcal{H} = \sum_i \left[ \vec{S}_i \cdot \vec{S}_{i+1} + \frac{1}{3} \left( \vec{S}_i \cdot \vec{S}_{i+1} \right)^2 \right], \tag{1.2.3}$$

or a next-nearest neighbour interaction, with the coupling to the next-nearest neighbour half as large as that to the neighbour:

$$\mathcal{H} = \sum_i \left( J_1 \vec{S}_i \cdot \vec{S}_{i+1} + J_2 \vec{S}_i \cdot \vec{S}_{i+2} \right), \tag{1.2.4}$$

with $J_2 = \frac{1}{2} J_1$. These models can be viewed as having been constructed to have a valence bond ground state. The simplest version of a valence bond state consists of a set of singlets, each singlet occupying two neighbouring sites. These states have gaps.

- The above cases have been chosen to obtain solvable models. In general, quantum fluctuations are enhanced by introducing frustration. This can be done by choosing a non-bipartite lattice or by introducing competing next-nearest neighbour interactions as in Eq.(1.2.4) with $J_1$ and $J_2$ of equal sign and arbitrary magnitude. Spin liquids, with gaps and without order, are favoured by frustration.

- Finally, let us mention a recent development in the study of the dimension-dependence of the phase-diagrams: two coupled spin-chains have been considered, as an intermediate case between one and two dimensions [24]. Interesting features arise, such as the analogy between a $S = 1/2$ double chain and a $S = 1$ single chain. Also double layers have become topical.

Having summarized what is known about order in quantum spin systems, let us go back to the questions we posed in section 1.1. For which parts of the thesis shall we need the above? The spin polaron, and in fact all inhomogeneous structures in proposed ground states of correlated fermion models, arise from *competition* between the properties of charge carriers (to be introduced later) and the magnetic order of the spin background. While for the discussion in chapter 3 we shall *assume* the existence of order, the fixed-node Monte Carlo method aims at calculating correlation functions, and thus at finding out whether the ground state of a model is magnetically ordered. As we shall discuss in chapter 6, the introduction of frustration like in Eq.(1.2.4) is intimately related to the sign structure of the ground state wave function, and we shall argue that our new method opens new possibilities for studying the nature of the emerging spin liquid states. An understanding of the relation between order and the existence of gaps is also needed to appreciate the calculation of the spin gap in the one-dimensional Kondo Lattice model, in chapter 5.

Pure quantum spin problems, especially those leading to spin liquid ground states, continue to be the subject of active research. Here, we have the ingredients we need

to compare spins with free electron-like systems, and to understand the difficulties of the combined problem of doped quantum magnets. We first turn to the opposite extreme: Fermi liquids.

### 1.2.2 The single electron picture

While the quantum spin systems of the previous section are naturally described in real-space, the essential properties of systems that are free electron-like are formulated in $\vec{k}$-space.

Free fermions are characterized by single-particle levels. In the ground state, these are filled up to the Fermi surface. At this surface in $\vec{k}$-space, the occupation jumps from $n(\vec{k}) = 1$ to $n(\vec{k}) = 0$.

The basic idea of Fermi liquid theory is that if interactions between the bare particles are included in the Hamiltonian, the elementary low energy quasi-particle excitations are in one to one correspondence with those of the non-interacting system. The phenomenology based on this assumption is due to Landau [25].

Starting from microscopic models and using diagrammatic perturbation theory [26], the phenomenology was given a sound basis, in higher dimensions and for weak coupling. Wilson renormalization group calculations, as performed e.g. by Shankar [27, 28] are also restricted to weak coupling; this method is able to decide on the relevance of certain interactions in the scaling sense, that is, the behaviour when the energy is scaled down. The unusual technical difficulty is that low energy phenomena are not at $\vec{k} = 0$, but at the Fermi surface.

Note that Fermi liquids are always unstable towards a BCS superconducting state, if (phonon-mediated) attractive interactions are present. The issue of the breakdown of the Fermi liquid picture is about different instabilities, which, in experimentally realized systems, occur at higher temperatures. How the BCS instability changes if it takes place in a non-Fermi liquid, is a separate question, which we shall not discuss.

Breakdown of Fermi liquid theory can occur for various reasons. The first reason is having strong coupling. Let us illustrate that on a well-known example. As we shall discuss later, the Hubbard model, which has nearest neighbour hopping ($t$) and on-site repulsion ($U$), reduces to the Heisenberg model at half filling (one electron per site on average) and large $U$ (which leads to having one electron per site and no double occupancies). The corresponding elementary excitations are spin excitations. Charge excitations are at energy scale $U$. This situation is very different from a Fermi liquid. In a Fermi liquid, the elementary excitations are quasi-particle excitations, both for the spins and for the charges. Taking $U = 0$ in the Hubbard model, of course *does* lead to a Fermi liquid. A fundamental question is whether one obtains non-Fermi liquid behaviour for every finite coupling, or whether one needs to exceed a threshold.

The occurrence of divergencies in responses with specific interactions in a specific dimension, calculated in perturbation theory, indicates that the starting point of perturbing around a Fermi sphere is incorrect, because the ground state is fundamentally different. The occurrence of relevant operators in renormalization group calculations

gives the same message: the assumed Fermi liquid fixed point is not stable, and the system will have the properties of a different fixed point, which then still needs to be identified. Such instabilities are the second, more fundamental reason for non-Fermi liquid phases.

The behaviour strongly depends on the dimension. In one dimension, phase space is such that the Fermi liquid is unstable [27]. For every finite interaction, the elementary excitations are different from electron-hole pairs, the analytic properties of response functions are different. The resulting Luttinger liquid will be described in more detail in chapter 2.

In two dimensions, taking exotic interactions leads to non-Fermi liquids. Anderson [29] advocates the Tomographic Luttinger liquid, related to singular interactions between quasi-particles of the form [30]

$$\Delta f_{pp'} \sim \frac{\vec{p}\,' \cdot (\vec{p} - \vec{p}\,')}{|\vec{p} - \vec{p}\,'|^2}. \tag{1.5}$$

It has been argued [30] that this leads to non-Fermi liquid behaviour, but this conclusion has been challenged [31]. The derivation of the singular effective interaction is beyond perturbation theory and is controversial [32]. Also super-long range interactions, as treated by Bares and Wen [33], lead to a breakdown of the Fermi liquid picture. Using non-exotic interactions in two dimensions and dividing the one-dimensional Fermi surface in patches as if one dealt with coupled Fermi points, leads to the conclusion that, treating the interactions among the patches correctly, one obtains a Fermi liquid [34]. This fixed point is also stable in higher dimensions, in other words, it is difficult to destroy the Fermi liquid in $d > 1$.

The fact that interactions are able to destroy the Fermi liquid makes it necessary to consider the possibility that making a certain approximation in calculating the properties of a model *assumes* Fermi liquid behaviour. If the true behaviour is different, that will never be the outcome of the calculation. A Hartree-Fock calculation tries to find the best Slater-determinant of single-particle levels, which is a Fermi liquid notion. Likewise, all band-structure calculations rely on the Fermi liquid picture.

We discussed the fact that electron-electron interactions can, depending on their range and on the dimension of the model, destroy the Fermi liquid. To complete the picture, let us mention two special situations in which there is non-Fermi liquid behaviour. The first is at a quantum phase transition, which is a second order phase transition at $T = 0$. Close to such a transition, interactions between electrons are mediated by the exchange of overdamped massless (magnetic) fluctuations. This non-standard effective electron-electron interaction yields non-Fermi liquid behaviour [35]. Impurity problems constitute the second special model that is known to be able to beat the stability of the Fermi liquid fixed point [36]. Scattering off localized impurities, in the two-channel or the two-impurity Kondo model, yields non-Fermi liquid behavionr for the conduction electrons. The Kondo model will reappear at different places in this thesis, and will be explained in more detail.

By discussing the stability of the Fermi liquid we mentioned an important issue in the field of correlated fermions, which we need to appreciate the discussion in chapter 2. There, however, the marginal Fermi liquid, which is a special type of non-Fermi liquid, will just be postulated to exist, and it will be used as a phenomenology. The one-dimensional Kondo Lattice model is, due to electron-electron interactions, insulating at half-filling. So we shall, more indirectly, encounter non-Fermi liquid features in chapter 5.

## 1.3 Models for correlated fermions

We have introduced spin systems and the issue of ordering in real-space. We also introduced the question whether a system has a jump in the momentum distribution function at the Fermi surface, if certain interactions are included in a system that otherwise consists of free electrons. This is a matter of ordering in $\vec{k}$-space.

In models of correlated fermions, both issues are important. In some regions of parameter space, one can view the problem as fermions moving in a spin background. On the one hand, one might wonder how this motion influences the spin ordering. On the other hand, the fermions effectively interact through the spins, and this possibly causes deviations from the Fermi liquid picture.

### 1.3.1 The Hamiltonians

Let us now introduce the models which are often used to study these issues. These are based on an (experimental) understanding of the relevant microscopic parameters.

It has been argued [37] that the appropriate model to describe high-$T_c$ materials is the two-dimensional one-band Hubbard model:

$$\mathcal{H} = -t \sum_{\langle i,j \rangle, \sigma} c^{\dagger}_{i\sigma} c_{j\sigma} + U \sum_{i} n_{i\uparrow} n_{i\downarrow}, \tag{1.3.1}$$

which has nearest neighbour hopping on a square lattice, and an on-site repulsion. The latter term involves four fermion operators. This model combines a kinetic term which is simple in $\vec{k}$-space, and an interaction term which is simple in real-space, while in combination this is a difficult problem. The formulation is convenient in real-space, because the kinetic term only couples nearest neighbours, while in a $\vec{k}$-space formulation the interaction term would couple all $\vec{k}$s to all other $\vec{k}$s.

Each site can be either empty, occupied by an up-spin electron, occupied by a down-spin electron, or doubly occupied. For $N$ sites, there are $4^N$ possible states in Fock space. The $S = 1/2$ Heisenberg model has $2^N$ possible states.

In the strong coupling limit, the Hubbard model reduces to the $t$-$J$ model, which has a Fock space of size $3^N$. This model is obtained as follows. Let us first consider half-filling, i.e. one electron per site on average. For large $U$, it is unfavourable to have doubly occupied positions. Taking $U = \infty$, the ground state is highly degenerate: each site has precisely one electron, with arbitrary spin. Taking $U$ large but finite, we

have a problem in degenerate perturbation theory, with $t/U$ as small parameter. If neighbouring sites have antiparallel spins, it is possible to have a doubly occupied site and an empty site next to each other *as an intermediate state*. The energy cost in the virtual state is $U$, the matrix element squared is $t^2$. Counting the possibilities, one concludes that the effective Hamiltonian is an antiferromagnetic Heisenberg model with $J = 2t^2/U$.

Doping holes into this magnet means that there is less than one electron per site, on average. For large $U$, doubly occupied sites remain forbidden. A site can be empty, or it can have one electron, of either spin. Motion of empty sites (holes) is a complication compared to the Heisenberg model. The complete effective Hamiltonian is

$$\mathcal{H} = -t \sum_{\langle i,j\rangle,\sigma} c^\dagger_{i\sigma} c_{j\sigma} + J \sum_{\langle i,j\rangle} \vec{S}_i \cdot \vec{S}_j, \tag{1.3.2}$$

with the local constraint

$$\sum_{\sigma=\uparrow,\downarrow} c^\dagger_{i\sigma} c_{i\sigma} = n_i = 0, 1. \tag{1.3.3}$$

This is the $t$-$J$ model, the presence of the constraint yields the reduction of the size of the Fock space to $3^N$.

The $t$-$J$ model is the strong coupling limit of the Hubbard model. The weak coupling limit assumes a close correspondence with free electrons, and therefore yields a Fermi liquid picture [12].

For heavy-fermion materials, two bands are important. The electrons not only have interactions within the bands, there is also hybridization between the bands. Let us consider the one-dimensional case. The Periodic Anderson Hamiltonian reads

$$\mathcal{H} = -t \sum_{i\sigma} (c^\dagger_{i\sigma} c_{i+1\sigma} + c^\dagger_{i+1\sigma} c_{i\sigma}) + \epsilon_f \sum_{i\sigma} n^f_{i\sigma} + U \sum_i n^f_{i\uparrow} n^f_{i\downarrow} + V \sum_{i\sigma} (c^\dagger_{i\sigma} f_{i\sigma} + f^\dagger_{i\sigma} c_{i\sigma}), \tag{1.3.4}$$

the $f$-operators are fermion operators for localized $f$-orbitals, the $c$-operators are for the conduction band. $V$ is the on-site hybridization, $\epsilon_f$ is the energy of the $f$-orbital, $U$ is the energy cost for having two $f$-electrons on one site and $t$ is the hopping matrix element for the conduction electrons. For $N$ sites, the Fock space has size $16^N$. The original Anderson model is for *one* $f$-impurity in a metal. The periodic version has one such level per unit cell and is relevant for compounds that have, for example, one uranium atom per unit cell.

Again, like in the case of the Hubbard model, an effective model with a smaller Fock space exists for strong coupling, i.e. for large $U$ and small $V$, more precisely $\pi V^2/2t(\epsilon_f + U) \ll 1$ and $\pi V^2/2t\epsilon_f \ll 1$. The constraint is that each site has one $f$-electron. Since eight configurations are possible per site, Fock space has size $8^N$, for $N$ sites. The effective Hamiltonian becomes

$$\mathcal{H} = -t \sum_{i\sigma} (c^\dagger_{i\sigma} c_{i+1\sigma} + c^\dagger_{i+1\sigma} c_{i\sigma}) + J_{eff} \sum_i \vec{S}^f_i \cdot \vec{s}^c_i, \tag{1.3.5}$$

with

$$J_{eff} = -\frac{2|V|^2 U}{\epsilon_f(\epsilon_f + U)}. \tag{1.3.6}$$

This is the Kondo Lattice model. It can be viewed as a lattice version of the familiar Kondo impurity model. So, both impurity models have lattice versions, and both Kondo models are limiting cases of Anderson models.

With the Hubbard model and the Periodic Anderson model and their strong coupling effective Hamiltonians, the $t$-$J$ model and the Kondo Lattice model, we have introduced the most important models for interactions within one band, and for interactions and hybridization in two bands. In some cases, microscopic parameters suggest the necessity of considering a three band Hubbard model. Spectroscopic data can be related to the parameters in such a model, as has been done by Zaanen, Sawatzky and Allen [38]. For many purposes, however, it has been argued that the one-band Hubbard model already gives the essential correlation effects.

### 1.3.2 Calculation methods and phase diagrams

In this section, we focus on the various techniques for obtaining information about the phase diagrams of models of correlated fermions. It also summarizes the salient features of those diagrams. Since from a theoretical point of view, models of interacting fermions are a natural extension of models of interacting spins (charge fluctuations are added), some methods known to work for spins are also used for correlated fermions.

In a few cases, rigorous statements have been proven using analytic methods. The one-dimensional Hubbard model was solved using the Bethe-ansatz [39], a method introduced for solving the one-dimensional $S = 1/2$ Heisenberg model [40]. The technical treatment is rather involved, for an elaborate discussion see e.g. ref. [14]. Some important ingredients and insights are the following [41]. One starts from an ansatz wave function which in the large $U$ limit reduces to a product of a Slater-determinant for the charge degrees of freedom, and the exact solution for the Heisenberg spin chain in 'pseudo coordinates', meaning that empty sites are omitted. Sites that are a certain distance apart in usual coordinates are neighbours in 'pseudo coordinates' if the sites in between are empty. The factorization of the wave function is an indication of spin-charge separation, i.e. the fact that the excitations come in two classes, one involving charge, the other involving spin.

This method yields energies. From the differences between energies one concludes that the half-filled case is insulating for all $U > 0$. Spin correlation functions are more difficult to obtain, but have been shown to decay as a power-law. As far as the issue whether the one-dimensional Hubbard model is a Fermi liquid or not is concerned, we already indicated that in one dimension, the Fermi liquid is unstable. The way this is seen in the analytic properties of correlation functions in the exact solution (the quasi-particle peak disappearing in a branch-cut), will be discussed briefly in chapter 2. The Bethe-ansatz is restricted to one dimension.

The case for which Nagaoka [42] was able to obtain exact results is for one hole in an otherwise half-filled Hubbard model with $U = \infty$ and $d > 1$. While the half-filled case is antiferromagnetic for large $U$, the fact that a single hole has more freedom to move in a ferromagnetic surrounding is, in the limit mentioned, enough to cause the

ground state to be ferromagnetic. The size of the region of the phase diagram (with the hole density and $t/U$ as parameters) to which this result extends, is the subject of much numerical effort. We shall come back to the issue when discussing ferromagnetic polarons in solid $^3He$ in chapter 3. Let us add here that a similar result exists for the Kondo Lattice model (Eq.(1.3.5)). In one dimension, one conduction electron is able to cause the $f$-electrons to become ferromagnetic in the ground state [43]. From these observations it is clear that the motion of charge carriers has a large effect on the spin distribution.

Besides the two abovementioned cases, all insight in the thermodynamic phase diagrams in the relevant dimensions (mainly the two-dimensional Hubbard model for high-$T_c$ and the three-dimensional Periodic Anderson model for heavy-fermions) is gained from approximate calculations.

One approach is to infer results for the thermodynamic limit from calculations on finite clusters [44]. This can be done by performing exact diagonalization calculations, accompanied by a finite-size scaling analysis. With Monte Carlo calculations one is able to consider larger, but still far from thermodynamic, systems, with a stochastic uncertainty as additional limitation. Both the finite temperature and the ground state algorithms are plagued by the sign problem, an issue we shall discuss extensively in chapter 4.

To understand the low doping regime, much effort has been devoted to the study of a single hole in an antiferromagnet. It should be noted that work by Stephan and Horsch [45] indicates that considering two holes at the same time changes the geometry of the Fermi surface rigorously (from hole pockets in the single hole case to a large Fermi surface in the two hole case). The picture of calculating the single hole dispersion and assuming that the many-hole case is a matter of rigid-band filling is, for this reason, subject to criticism. The numerical work [45] is, however, on small clusters, so having two holes means a substantial increase in doping. The discussion about the aptness of the single hole picture and about the geometry of the Fermi surface for small doping is still an active one.

Correlations cause the single hole dispersion to be flat near the Fermi surface. Many authors (see e.g. the review by Dagotto [44]) claim that this observation not only explains photo-emission data, but is also of importance for the mechanism of superconductivity. The Van Hove scenario is that there is BCS coupling, the effect of which is enhanced by a peak in the density of states near the Fermi surface. This scenario is rendered plausibility by the flatness of the dispersion, since then it does not need fine-tuning of the parameters.

For the one-dimensional Kondo Lattice model, an important result from exact-diagonalization studies is that at half-filling, there is a gap both for the spin and for the charge excitations, and that the charge gap is larger than the spin gap [46].

As far as the discussion about whether or not higher-dimensional systems show Fermi liquid behaviour is concerned, some numerical work on the $t$-$J$ model and the Hubbard model favours the existence of quasi-particles, other work contradicts this conclusion [44]. This issue is far from settled.

As far as the magnetic properties of the Hubbard model are concerned, two results about magnetic order have emerged from numerical studies [44]. One is that at half filling, there is long-range antiferromagnetic order, not only in the limit of large $U$ in which the system is effectively a Heisenberg model, but at every finite $U$. The other result that stands out is that every finite doping yields a finite spin correlation length, true long range order is destroyed.

Reliable renormalization group results for correlated fermions have been obtained only recently, after the introduction of the Density Matrix Renormalization Group by White [47]. The method is only operational in one-dimensional and ladder (double chain) systems, up to this moment. Some results for the one-dimensional Kondo lattice model will be discussed in chapter 5.

After the approach via numerical work on small clusters, we discuss work in which effective Hamiltonians are used.

While it is clear that the motion of holes disturbs the spin background, the way to describe this and the precise mechanism of destroying long range order, are subject to debate. One attempt has been to incorporate the effect of doping on a spin model by adding frustrating spin-spin interactions [48, 49]. While frustration and the spin-liquid phases it leads to are interesting in themselves, it has become clear that doping works differently [44]. One should treat the full problem of correlated fermions. In chapter 6, both the effect of frustration and the effect of doping on the *sign structure* of the wave function of a spin system will be discussed.

An important class of methods has a mean-field character. Various formulations exist. Making a saddle point approximation in a functional integral for the partition function [50], taking only a certain class of diagrams (Hartree-Fock) along in a perturbation expansion, or considering the system in $d = \infty$ [51] have yielded insight in the properties of correlated fermions. It should be clear that all of these approaches underestimate the effects of quantum fluctuations, like mean-field approximations underestimate the effects of thermal fluctuations in classical statistical mechanics.

In later chapters on Monte Carlo methods, we shall use mean-field results as starting point for further calculations that improve on the input wave function. The way we proceed for obtaining the mean-field result is decoupling the interaction term in the Hamiltonian and solving the problem self-consistently, which is possible because the Hamiltonian is then bilinear in the fermion operators. This can also be formulated as finding the best Slater-determinant wave function, which is an example of taking classes of wave functions (variational wave functions) and optimizing parameters.

The main virtue of a *slave-boson* formulation of the Hamiltonians [41, 51], is that it admits additional ways to decouple. How to deal with the local constraints has remained a difficult problem.

From the above, a picture emerged of competition between a spin background that favours a certain magnetic order, and charge carriers that prefer a different surrounding. In the case of the Hubbard model, a hole prefers the spins to be ferromagnetic, the spins prefer antiferromagnetic order. Mean-field results indicate that doping does not simply reduce the antiferromagnetic order uniformly, but possibly

leads to inhomogeneous solutions. These can be viewed as compromises between the two types of order that yield the best mean-field energy. The inhomogeneous ground states include phase separation, domain walls of charge separating antiferromagnetic domains, and spin polarons in the Hubbard model. A more elaborate discussion will be given in chapter 3. In chapter 5 we shall encounter a similar structure in the Kondo Lattice model. An important general question is how quantum fluctuations affect these mean-field solutions. This question can be answered in part by our lattice fixed-node quantum Monte Carlo method.

In this section, we mainly discussed magnetic properties and the possible destruction by correlations of the Fermi liquid picture. We omitted the issue of superconductivity. While we shall not study this issue in the rest of this thesis, it should be mentioned that electronic mechanisms of superconductivity have been proposed [44], and that correlations could also alter more conventional (BCS with electron-phonon coupling as origin of the electron-electron attraction) mechanisms, such that these can be reconciled with the non-standard superconducting properties of the cuprates. One possibility is that the single-particle properties of the electrons are peculiar because the density of states has singularities, leading to a non-standard transition to the superconducting state (the Van Hove scenario [52]). A different non-standard transition occurs if it does not take place in a Fermi liquid, but in a marginal Fermi liquid [53]. The instability towards a marginal Fermi liquid is, phenomenologically, at a higher temperature than that towards superconductivity.

Having given an overview of the field of correlated fermions, let us now indicate which methods and which features of phase diagrams will be of importance to the rest of this thesis. As far as methods are concerned, our lattice fixed-node Monte Carlo is a contribution to the numerical approach. The absence of a sign problem allows to go to larger clusters. As input, we need mean-field results, or optimized variational wave functions. The models on which we performed calculations are the Hubbard model and, mainly, the Kondo Lattice model, Eq.(1.3.5). Mean-field and fixed-node Monte Carlo calculations on this model will be presented in chapter 5. The stability against quantum fluctuations of the inhomogeneous solutions in space, which we mentioned in this section, will be our most important application of the Monte Carlo method we developed. As mentioned before, the spin polaron of chapter 3 is a consequence of the competition between the tendency towards delocalization of fermion wave functions (the argument of Nagaoka says that this leads to ferromagnetism), and the way this destroys the antiferromagnetic spin ordering of the background. Finally, the fact that in the one-dimensional Hubbard model the Fermi liquid picture breaks down will be the inspiration for discussing new consequences of non-Fermi liquid behaviour in the next chapter.

# References

[1] For an introduction to the experimental properties of heavy-fermion materials see e.g. P. Fulde, *Electron correlations in molecules and solids*, Springer, Berlin

(1991), chapter 13. References to more extensive reviews are given there.

[2] For an overview of the experimental properties of high-Tc materials see e.g. the articles by B. Batlogg and by P.A. Lee in *High temperature superconductivity* The 1989 Los Alamos Symposium, K.S. Bedell, D. Coffey, D.E. Meltzer, D. Pines, J.R. Schrieffer eds., Addison-Wesley, Redwood City, (1990)

[3] See e.g. B. Batlogg, Physics Today **44**, 44 (1991)

[4] See e.g. Y.C. Chen, A. Moreo, F. Ortolani, E. Dagotto, T.K. Lee, Phys. Rev. B [illegible]. 655 (1994)

[5] Z.Y. Weng, Y.C. Chen, and D.N. Sheng, cond-mat preprint 9505124

[6] N.D. Mermin and H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966)

[7] For a review of real-space renormalization see e.g. Th. Niemeijer and J.M.J. van Leeuwen, in *Phase transitions and critical phenomena* vol. 6, C. Domb and M.S. Green eds., Academic Press, London (1976)

[8] For a review of $\vec{k}$-space renormalization see e.g. F.J. Wegner, in *Phase transitions and critical phenomena* vol. 6, C. Domb and M.S. Green eds., Academic Press, London (1976)

[9] For a review see e.g. J.L. Cardy, in *Phase transitions and critical phenomena* vol. 11, C. Domb and J.L. Lebowitz eds., Academic Press, London (1987)

[10] For discussions of symmetry breaking see G. Parisi, *Statistical field theory*, Addison-Wesley, Redwood City, (1988) and P.W. Anderson, *Basic notions of condensed matter physics*, Addison-Wesley, Redwood City, (1984) and also L. Pitaevskii and S. Stringari, J. Low Temp. Phys. **85**, 377 (1991) and L. Pitaevskii and S. Stringari, Phys. Rev. B **47**, 10915 (1993)

[11] F.D.M. Haldane, Phys. Lett. **93A** 464 (1983), Phys. Rev. Lett. **50** 1153 (1983), and J. Appl. Phys. **57**, 3359 (1985)

[12] E. Fradkin, *Field theories of condensed matter systems*, Addison-Wesley, Redwood City, (1991), chapter 5

[13] I. Affleck, J. Phys.: Condens. Matt. **1**, 3047 (1989)

[14] For a review of the Bethe-Ansatz see e.g. H.J. Schulz, lecture notes Les Houches summer school 1994, cond-mat preprint 9503150

[15] S.R. White and D.A. Huse, Phys. Rev. B **48**, 3844 (1993)

[16] F. Dyson, E.H. Lieb and B. Simon, J. Stat. Phys. **18**, 335 (1978)

[17] E.J. Neves and J.F. Peres, Phys. Lett. **114A**, 331 (1986)

[18] K. Kubo and T. Kishi, Phys. Rev. Lett. **61**, 2585 (1988)

[19] E.H. Lieb, T. Schultz, and D.J. Mattis, Ann. Phys. NY **16**, 407 (1961)

[20] I. Affleck, Phys. Rev. B **37**, 5186 (1988)

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

[22] S. Chakravarty, B.I. Halperin and D.R. Nelson, Phys. Rev. B **39**, 2344 (1989)

[23] M. den Nijs and K. Rommelse, Phys. Rev. B **40**, 4709 (1989)

[24] S.R. White, cond-mat preprint 9503104

[25] For elaborate treatments of Landau Fermi liquid theory see e.g. G. [illegible] and C.J. Pethick, *Landau Fermi liquid theory, concepts and applications*, John Wiley and Sons, New York (1991) or D. Pines and P. Nozières, *The theory of quantum liquids* Vol I, Benjamin, New York (1966)

[26] For a thorough discussion of the diagrammatic perturbation theory see A.A. Abrikosov, L.P. Gorkov and I.E. Dzyaloshinski *Methods of quantum field theory in statistical physics*, Prentice-Hall, Englewood cliffs, N.J. (1963)

[27] R. Shankar, Rev. Mod. Phys. **66**, 129 (1994)

[28] A.C. Hewson, cond-mat preprint 9410013

[29] P.W. Anderson, Phys. Rev. Lett. **64**, 1839 (1990)

[30] P.C.E. Stamp, Phys. Rev. Lett. **68**, 2180 (1992)

[31] A.E. Ruckenstein and W. van Saarloos, private communication

[32] J. Engelbrecht and M. Randeria, Phys. Rev. Lett. **65**, 1032 (1990)

[33] P.-A. Bares and X.-G. Wen, Phys. Rev. B **48**, 8636 (1993)

[34] For a review of the extension of one-dimensional techniques to higher dimensions see e.g. the article by F.D.M. Haldane in *Perspectives in Many-Particle Physics*, proceedings of the international school of physics Enrico Fermi, course CXXI, North-Holland, Amsterdam (1994)

[35] P. Coleman, cond-mat preprint 9410024

[36] B.A. Jones, C.M. Varma, J.W. Wilkins, Phys. Rev. Lett. **61**, 125 (1988)

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

[38] J. Zaanen, G.A. Sawatzky, and J.W. Allen, Phys. Rev. Lett. **55**, 418 (1985)

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

[40] H.A. Bethe, Z. Phys. **71**, 205 (1931)

[41] P. Fulde, *Electron correlations in molecules and solids*, Springer, Berlin (1991)

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

[43] M. Sigrist, H. Tsunetsugu, and K. Ueda, Phys. Rev. Lett. **67**, 2211 (1991)

[44] See e.g. E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994)

[45] W. Stephan and P. Horsch, Phys. Rev. Lett. **66**, 2258 (1991)

[46] H. Tsunetsugu, Y Hatsugai, K. Ueda and M. Sigrist, Phys. Rev. B **46**, 3175 (1992)

[47] S.R. White and R.M. Noack, Phys. Rev. Lett. **68**, 3487 (1992); S.R. White, Phys. Rev. Lett. **69**, 2863 (1992); S.R. White, Phys. Rev. B **48**, 10345 (1993)

[48] M. Imui, S. Doniach and M. Gabay, Phys. Rev. B **38**, 6631 (1988)

[49] A. Aharony, R.J. Birgeneau, A. Coniglio, M.A. Kastner and H.E. Stanley, Phys. Rev. Lett. **60**, 1330 (1988)

[50] See e.g. J.W. Negele and H. Orland, *Quantum many-particle systems*, Addison-Wesley, Redwood City (1992)

[51] For a review see e.g. the article by D. Vollhardt in *Perspectives in Many-Particle Physics*, proceedings of the international school of physics Enrico Fermi, course CXXI, North-Holland, Amsterdam (1994)

[52] D.M. Newns, C.C. Tsuei, P.C. Pattnaik and C.L. Kane, Comments Condens. Matt. Phys. **15**, 273 (1992)

[53] M.L. Horbach, F.L.J. Vos and W. van Saarloos, Phys. Rev. B **48**, 4061 (1993)

# 2 Local dynamic disturbances of marginal Fermi liquids

## 2.1 Motivation

In this chapter, we discuss our findings on a new application of an existing phenomenology [1] for the normal state of high-$T_c$ superconductors. Our aim is to answer the question: "If the transport properties of a material show marginal Fermi liquid behaviour instead of Fermi liquid behaviour, will the response to a dynamic *local* disturbance also be different?".

The motivation to pursue the calculation of these responses comes from considerations at different levels. As already indicated in chapter 1, the issue of Fermi liquid versus non-Fermi liquid behaviour is an important one in the field of correlated fermions. The study of the possible existence and the experimental consequences of a phase with new properties of the charge carriers is, therefore, of fundamental importance. Moreover, the marginal Fermi liquid was put forward [1] as a phenomenological description of high-temperature superconductors. In order to understand these materials, it is useful to try to assess the range of validity of this phenomenology as much as possible, by comparing its predictions with experiments.

Our motivation to study local disturbances is the observation by Zhang *et al.* [2] that one particular property of this type, the hopping rate of a heavy charged particle, is, theoretically, very different in a marginal Fermi liquid compared to the same quantity in a normal Fermi liquid. We study a wider range of properties of this type, which includes this hopping rate as a limiting case. We elucidate the origin of the anomalous behaviour in the limiting case by considering the behaviour close to this limit. The way the anomaly shows away from this limit is relevant for X-ray photoemission experiments.

While the marginal Fermi liquid phenomenology was topical at the time we started this line of research, it should be noted that it is not often used anymore for interpreting data, at the time of writing this thesis [3]. The reason is that the theory remained largely at a phenomenological level. Despite the usefulness of the description for correlating a number of experimental data, its consistency as well as its microscopic basis are unclear. The difficult task of extracting the properties that can be compared with experimental data from microscopic models, as well as the questions concerning the existence of microscopic models with non-Fermi liquid behaviour (see chapter 1), has been the main interest in recent years. In this chapter, we shall remain within the phenomenological framework.

This chapter is organized as follows. First, we shall introduce the two ingredients: the marginal Fermi liquid phenomenology and the local dynamic disturbance of a

system of conduction electrons. Then, in section 2.3, the results of our calculations and the assessment of experimental testability will be presented[1]. At the end of the chapter, we include a discussion of 'Fermi surface effects' in different anomalous metals, like one dimensional Luttinger liquids, two dimensional metals exhibiting a Van Hove singularity in the density of states, and half-metallic materials, which have a Fermi surface for one spin species only.

## 2.2 Ingredients

### 2.2.1 Marginal Fermi liquids

Before turning to the marginal Fermi liquid phenomenology, let us first [illegible] the Fermi liquid and known ways in which it can break down. In a normal Fer[illegible] [5], the one-particle Green function has the functional form

$$G(\vec{k},\omega) = \frac{1}{\omega - \epsilon_k - \Sigma(\vec{k},\omega)} = \frac{Z_k}{\omega - E_k + i\Gamma_k} + G_{\text{incoh}}, \qquad (2.2.1)$$

which has basically the same analytic structure as the free electron Green function

$$G_{\text{free}}(\vec{k},\omega) = \frac{1}{\omega - \epsilon_k}. \qquad (2.2.2)$$

The pole that is still present in Eq.(2.2.1) determines the response functions to a large extent. The differences between the two Green functions, consisting of the incoherent part in Eq.(2.2.1) and the imaginary part of the self-energy that places the pole off-axis, are less important. As can be seen from the notation in Eq.(2.2.1), the real part of $\Sigma$ corresponds to a change in dispersion (and can thus be viewed as a change of effective mass), while a non-zero imaginary part corresponds to the inverse of the quasi-particle life time. In the momentum distribution function $n(\vec{k})$, Eq.(2.2.1) corresponds to a jump of magnitude $Z_k$ at the Fermi surface, which remains at the same position as in the absence of interactions.

In higher dimensions, this is a stable phase, since the perturbation theory in which one calculates the corrections to the free electron picture shows no divergencies. In the space of interaction parameters, a finite basin of attraction that yields the behaviour corresponding to Eq.(2.2.1) is found, in renormalization group calculations.

In one dimension, by contrast, one does encounter divergencies in doing perturbation theory [6], indicating the breakdown of Fermi liquid theory. We shall describe the main features of the the exact solution instead of discussing the divergencies.

As can be found in standard textbooks, e.g. ref. [7], one starts with a Fermi surface, consisting of two points, $k_F$ and $-k_F$. In the free problem, there is a jump from one to zero in the momentum distribution function, at those points. Including

[1]This part is an adapted version of the paper we published in *Journal of Physics: condensed matter* [4].

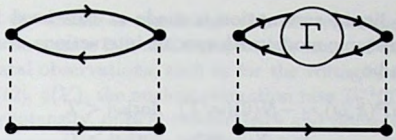



Figure 2-1. Contribution to the self-energy: a) virtual excitation of a particle-hole pair b) virtual excitation of the quantity replacing the particle-hole pair in an MFL

which cause scattering from one Fermi point to the other yields the so-called Luttinger model [7], which is exactly solvable. For later comparison, we need to know the analytic structure of the Green functions.

For the one particle Green function, one finds in space and time coordinates [8]

$$G^R(x,t) = -\frac{\Theta(t)}{\pi} e^{ik_F x} \mathrm{Re}\left\{ \frac{1}{\sqrt{(u_\rho t - x)(u_\sigma t - x)}} \left[ \frac{\alpha^2}{(\alpha + i u_\rho t)^2 + x^2} \right]^{\delta/2} \right\}, \quad (2.2.3)$$

indicating the separation of spin and charge: these move at different velocities. Fourier transforming this is somewhat cumbersome, the results are [9, 10] that, even for the weakest interaction strength, the pole in the one-particle Green function disappears. At $k_F$, $G(\vec{k},\omega)$ diverges as a *power law*. Correspondingly, the momentum distribution becomes [8]

$$n(k) \approx n(k_F) - \text{const} \cdot \text{sign}(k - k_F)|k - k_F|^\delta. \quad (2.2.4)$$

This still is a non-analytic function, but there is no jump at $k_F$.

The stability of the Fermi liquid is intimately related to phase space arguments. In three dimensions, when considering the quasi-particle life time, one has to calculate the contribution of self-energy diagrams like the one depicted in Fig.(2-1 a), the free electron propagator becomes 'dressed' by the interactions. Taking into account conservation laws and the occupancy of the various single-particle levels, one concludes that $\mathrm{Im}\Sigma \sim \omega^2$. At the Fermi surface, this goes to zero quickly and the inverse, the quasi-particle life time, goes to infinity. By contrast, those terms in the perturbation expansion do not go to zero as quickly as $\omega^2$ in one dimension. The self-energy $\Sigma$ even becomes large compared to the quasi-particle energy itself, which is $\omega$. The reason is that phase space is very different. The life time of quasi-particles is small and quasi-particles are ill-defined. In the exact solution, these appear to be replaced by collective spin and charge excitations.

What happens to the one-particle Green function depends on how easily particle-hole pairs are created, as is indicated by the presence of the particle-hole bubble (or 'polarizability bubble') in Fig.(2-1 a). Later, for different purposes, we shall need the behaviour of this bubble in a Fermi liquid, it is linear in $\omega$ and the proportionality constant includes the density of states at the Fermi surface (see Eq.(2.3.5)).

The marginal Fermi liquid assumption is made at this level [1]. Without any justification from a microscopic model with specific interactions, the charge and spin polarizability is taken to be

$$\begin{aligned} \mathrm{Im}P(\vec{k},\omega) &\sim -N(0)(\omega/T) \quad \text{for} |\omega| < T \\ &\sim -N(0)\mathrm{sign}\omega \quad \text{for} |\omega| > T. \end{aligned} \tag{2.2.5}$$

The slope of this linear function of $\omega$, as well as the point at which it turns over to become a constant function, are both determined by the temperature, not by $\epsilon_F$ or some antiferromagnetic coupling constant $J$. Note, moreover, the absence of any $\vec{k}$-dependence. We will first consider this assumption in terms of the analytic structure of Green functions, later we shall discuss the experimental reason why it was put forward to describe the normal state of high-$T_c$ materials.

If one calculates the self-energy by inserting Eq.(2.2.5) as the behaviour of the bubble, this means that the excitations that dress the single-particle propagator are not normal particle-hole pairs, but excitations of the postulated type. The corresponding diagram is depicted in Fig.(2-1 b).

The result is [1, 11] that the single-particle Green function can be written in the form of Eq.(2.2.1), but that

$$Z_k^{-1} = [1 - \partial \mathrm{Re}\Sigma/\partial\omega]_{\omega=E_k} \sim \ln|\omega_c/E_k|, \tag{2.2.6}$$

$\omega_c$ is an ultraviolet cutoff. Since at the Fermi surface $E_k = 0$, the factor $Z_k$ goes to zero, there.

Comparing a Luttinger liquid and a marginal Fermi liquid, the corresponding quasi-particle weights are, respectively, a power law (cf. Eq.(2.2.4)), and one over a logarithm, as a function of energy. The latter way of vanishing is slower than a power law, it is the weakest way $Z_k^{-1}$ can diverge. Hence the name 'marginal' Fermi liquid.

A different way of explaining this term is according to the behaviour of $\mathrm{Im}\Sigma$. If we write $\mathrm{Im}\Sigma(\omega) \sim \omega^\alpha$, this becomes small compared to the quasi-particle energy itself, $\omega$, if $\alpha > 1$. This means that the closer to the Fermi surface one comes, the better defined the quasi-particles are. This yields Fermi liquid behaviour. If $\alpha < 1$, the behaviour is non-Fermi liquid like, because the notion of quasi-particles becomes less well defined closer to the Fermi surface, which is where the important low energy physics is. The *marginal* case is $\alpha = 1$, the ratio of the quasi-particle energy and its inverse life-time is constant as a function of energy. The ansatz of Eq.(2.2.5) yields $\alpha = 1$ in $\mathrm{Im}\Sigma(\omega) \sim \omega^\alpha$.

Let us now describe the merits of the marginal Fermi liquid phenomenology (we shall use the abbreviation MFL, and FL for a conventional Fermi liquid) for the description of high-$T_c$ materials. It has been introduced [1] because it allows to obtain several anomalous experimental results for transport and thermodynamic properties, from one single assumption. The most well-known anomalous result is that the temperature dependent resistivity $\rho(T)$ is linear in $T$, while in a normal three dimensional Fermi liquid it is proportional to $T^2$. Since the resistivity is proportional to the inverse of the quasi-particle life time (and thus proportional to $\mathrm{Im}\Sigma$), the behaviour

that follows from the ansatz is consistent with the experimental observation. Along similar lines, such an ansatz is found to reproduce some of the other main features in various experimental observations, such as for the voltage dependence of the tunneling conductance [12], $g(V)$, the nuclear relaxation rate $T_1^{-1}(T)$, thermal conductivity, Raman scattering intensity, optical conductivity and photo-emission [1].

As far as the superconducting properties are concerned, the same mode Eq.(2.2.5) that causes the anomalous scattering which destroys the Fermi liquid, has been considered [13, 14] as possible 'glue' for superconductivity, replacing the usual BCS interaction, which is phonon-mediated. A different possibility is that the pairing is caused by a usual BCS coupling, and that *pair breaking* due to scattering off the MFL-mode explains the anomalous superconducting properties of cuprate materials [15].

The microscopic origin of the MFL behaviour is subject of debate. One scenario is related to a different part of this thesis, in which inhomogeneous structures are discussed, chapter 3. According to this scenario, advocated in Leiden by Zaanen, Horbach and Van Saarloos [16], the existence of a liquid consisting of fluctuating domain walls of holes explains the occurrence of the temperature as relevant scale in the polarizability Eq.(2.2.5). Different possible explanations were treated in chapter 1.

We only need the phenomenological level of description. Like in the philosophy of the MFL phenomenology itself, we shall assume Eq.(2.2.5) to give the relevant charge fluctuations, replacing the usual virtual electron-hole pairs. We shall dress the propagator of an *external* particle with those fluctuations, as was done for the particles in the MFL itself.

One further remark on the nature of the marginal Fermi liquid assumption we wish to make, is that the fact that spin and charge behave similarly contrasts with the Luttinger liquid, which shows spin-charge separation.

### 2.2.2 Local dynamic disturbances

The singularities in responses to local dynamic disturbances, which occur due to the presence of a standard filled Fermi sea, are treated in several textbooks [17, 7]. Let us discuss the essentials of the problem using the Hamiltonian for the X-ray absorption process as given in Mahan [7]

$$\mathcal{H} = \epsilon_h d^\dagger d + \sum_{\vec{k}} \epsilon\,(k) c^\dagger_{\vec{k}} c_{\vec{k}} + \frac{1}{\nu} \sum_{\vec{k}\vec{k}'} V(\vec{k},\vec{k}') c^\dagger_{\vec{k}} c'_{\vec{k}} d^\dagger d. \tag{2.2.7}$$

The $d$-operators relate to a localized level, e.g. a deep lying core level, which does not belong to the system of conduction electrons.

The basic idea is that when the occupancy of the $d$-level changes, the conduction electrons experience the sudden appearance of a local scattering potential. This happens e.g. if an X-ray photon is absorbed by a core electron and the core electron is either promoted to the conduction band (X-ray absorption, this is the more complicated case because one needs to consider the interaction of this new conduction

electron with the other conduction electrons), or the core electron is shot out of the material to a plane wave state (X-ray photo-emission).

Although in the ground state with a core hole present, the wave function of each single-particle level is changed only slightly, the fact that *all* single-particle levels have an overlap with the corresponding wave function without core hole that is a bit smaller than 1, causes the many-body ground states to be orthogonal [18]. This is often referred to as 'Anderson's orthogonality catastrophe'. The transient behaviour of the conduction electrons, during their adaptation to the new situation after the core level is emptied by the X-ray , influences the Green function of the core level and thus the observed spectroscopic behaviour. The logarithmic divergences that occur cause the core level spectral function to have a power law divergence, not a pole, as the free core level would have. The general name for the consequences of this observation is 'X-ray edge singularity'. How these divergencies in a Fermi liquid depend on temperature will be discussed in section 2.3, together with the behaviour in a marginal Fermi liquid.

Instead of the term X-ray photo-emission, we sometimes also use the common abbreviation XPS, for X-ray Photoelectron Spectroscopy. In XPS, the electrons come out of the material with different kinetic energies, large enough for there to be no no final state interactions with the material. The X-ray energy is used to overcome the binding energy of the core electron and for the creation of different amounts of electron-hole pairs needed to overcome the orthogonality of the many-electron wave functions. The fact that the way electron-hole pairs are created is relevant here, indicates that interesting new behaviour is to be expected if the MFL-bubble of Fig.(2-1 b) is used instead of the usual particle-hole bubble of an FL. This bubble dresses the propagator of a particle that belongs to the system of conduction electrons in such a way that it shows anomalous behaviour, in this chapter we wish to investigate how it affects the propagator of an external particle, i.e. one that does not belong to the system of conduction electrons.

As we shall use later, the effect of the conduction electrons on core hole spectra is related to the physics of a heavy particle hopping in a metal in the presence of interactions with conduction electrons [19]. In X-ray photo-emission, one measures the distribution of outcoming kinetic energy, which is determined by the spectral function of the core hole, $A_h(\omega)$. The *hopping problem*, on the other hand, is the following. The probability of a heavy charged particle to hop to a neighbouring site is considered. Such a hop constitutes a change of the local potential experienced by the conduction electrons. The corresponding singular effects lead to a power law divergence of the hopping rate as a function of decreasing temperature, in an FL. The energy of the hopping particle has to match the energy of the energy level at the neighbouring site. If the energies at different sites are the same, in the simplest models, the hopping rate $\nu$ is proportional to the core hole spectral function at zero frequency, $\nu \propto A_h(0)$ (see Eqs(2.3.1) and (2.3.2)). If there is a distribution of levels, however, we have to average $A_h(\omega)$ over a certain energy (or frequency) interval to obtain the hopping rate.

Having described the essence of the X-ray problem as the transient adaptation of many electrons to the sudden switch of a local potential, let us now consider the Kondo Hamiltonian. It is given by

$$\mathcal{H} = \sum_{\vec{k}} \epsilon\,(k) c^{\dagger}_{\vec{k}} c_{\vec{k}} + J\vec{S}^{f}_{0} \cdot \vec{s}_{0}, \tag{2.2.8}$$

there is a localized spin that interacts with otherwise free conduction electrons. Spin flips of the $f$-impurity act as sudden switches of a local potential for the conduction electrons. It has been argued [20] that the similarity is strong enough for the Kondo problem to be viewed as a *succession* of X-ray problems.

The question what happens to the singular responses in X-ray edge problems if interactions among the conduction electrons are introduced in the Hamiltonian has been considered in the literature [21]; it has been summarized by Nozières [22] as: "nothing changes, qualitatively, as long as the conduction electrons remain a Fermi liquid".

In the next section, we take the conduction electrons to be a marginal Fermi liquid, i.e., we drastically change the dynamics around the Fermi surface, and discuss how that changes the singular responses. The description of the X-ray photo-emission problem is in terms of $\mathrm{Im}P$, or, in terms of diagrams, of particle-hole bubbles. Cases in which the core electron remains in the material and interacts with the other conduction electrons, like in X-ray absorption, would require knowledge about the effective electron-electron interaction, which we do not have for the MFL case. An analytic treatment of the Kondo effect in an MFL is, for the same reason, not feasible. The only singular effects that can be treated are a heavy charged particle hopping in the material and the X-ray photo-emission line shape, because these effects are governed by orthogonality alone, since the number of electrons in the conduction band remains the same.

A comparison of these effects in Fermi liquids and in marginal Fermi liquids will be made in the next section. We were inspired by a paper by Zhang, Toyoda and Takahashi [2, 23], who pointed out that in the particular case of the hopping rate of a charged particle, the singular behaviour is modified in an MFL compared to an FL. We worked this out in more detail.

## 2.3 X-ray photo-emission and hopping in an MFL

We study the spectral properties of a localized particle (a deep core level, or a heavy particle hopping in a solid), coupled to conduction electrons that are described by the marginal Fermi liquid hypothesis. Our main result is that, in this model, the core level line shape in an X-ray photo-emission experiment shifts with temperature. The spectrum we find is consistent with the decrease of the hopping rate with decreasing temperature obtained by Zhang *et al.* [2], which is in sharp contrast to the increase of the hopping rate as a function of decreasing temperature in a normal Fermi liquid. In

the presence of a distribution of energy levels for the hopping particle, the temperature dependence of the hopping rate is found to be less pronounced.

As we mentioned before, the hopping rate corresponds to the zero-frequency behaviour of the spectrum seen in an X-ray photo-emission experiment. Whether the decrease of the hopping rate, found by Zhang *et al.*, corresponded to a broadening and decreasing of the whole spectrum, or to a shift of the position of the maximum away from zero frequency, was not known beforehand. The calculation we shall present shows that the latter is the correct explanation of the decrease of the zero frequency value. This is also relevant for the hopping problem itself. Averaging over an energy interval would yield an unchanged drop if the whole spectrum would broaden. The fact that the peak remains of the same magnitude and stays close to zero frequency, means that the decrease will be less pronounced after averaging.

The observability of both the shift in the X-ray photo-emission spectrum and of the decrease of the hopping rate depends on the strength of the interaction. Our rough estimate of the value of the interaction parameter indicates that the effect might be observable in X-ray photo-emission experiments.

## 2.3.1 Formulation of the problem

Various formulations exist of both the hopping problem [19, 25, 26, 27] and of many-body effects in X-ray spectroscopy [28, 29, 30]. We choose a formulation which involves the electronic polarizability in the frequency-domain.

The model that we consider is that of a particle (the hopping particle or a core-level electron) in a localized state coupled to the conduction electrons through an interaction potential $V$. The properties of the conduction electrons are assumed to be described by the MFL ansatz for the electronic polarizability $P(\vec{q},\omega)$. For both problems, the interaction is of the form $\sum_{\vec{q}\vec{k}} V(\vec{q}) c^{\dagger}_{\vec{k}+\vec{q}} c_{\vec{k}}$, where $c_{\vec{k}}$ and $c^{\dagger}_{\vec{k}}$ are creation and annihilation operators for the conduction electrons. For the hopping problem, $V(\vec{q})$ is the Fourier transform of the difference in the potential that the conduction electrons feel when the particle hops to a neighbouring site with the same energy, while for the X-ray problem, $V(\vec{q})$ is the effective potential that the conduction electrons experience when the core level is unoccupied [7].

For the hopping problem, the hopping rate $\nu$ between two levels of the same energy is given by [2]

$$\nu = 4\Delta^2 \mathrm{Re} \int_0^\infty dt \rho(t), \tag{2.3.1}$$

in which $\Delta$ is the tunneling matrix element without any interaction and $\rho(t) = \langle e^{i\mathcal{H}(\vec{R}_1)t} e^{-i\mathcal{H}(\vec{R}_2)t} \rangle$, where $\mathcal{H}(\vec{R}_1)$ is the interaction Hamiltonian with the hopping particle at site $\vec{R}_1$ and $\mathcal{H}(\vec{R}_2)$ the Hamiltonian when the particle is at site $\vec{R}_2$. As usual, we take $\hbar = 1$.

In the case of photo-emission, the relevant quantity is $A_h(\omega)$, the spectral function of the core hole. This quantity directly determines the line shape, and is given by

$$A_h(\omega) = 2\mathrm{Re}\int_0^\infty dt e^{i\omega t}\rho(t), \tag{2.3.2}$$

where in this case $\rho(t) = \langle e^{i\mathcal{H}_g t}e^{-i\mathcal{H}_f t}\rangle$, with $\mathcal{H}_g$ the ground state Hamiltonian (with the core level occupied) and $\mathcal{H}_f$ the Hamiltonian with the core hole present [7]. The core hole Green function equals $-i\Theta(t)\rho(t)$. In Eq.(2.3.2), we measure the energy $\omega$ relative to the energy in the absence of interaction with the conduction electrons; without these interactions, the spectral density is just a delta function at $\omega = 0$. While Zhang *et al.* analyze Eq.(2.3.1) for an MFL, we take Eq.(2.3.2), i.e. we consider the frequency dependence, Zhang *et al.* take a problem for which $\omega = 0$ is relevant. Otherwise, the approach of taking the MFL ansatz instead of the usual FL is the same in ref. [2] and this work. We wish to extend and elucidate this.

$\rho(t)$ is normally analyzed by writing $\rho(t) = e^{F(t)}$ and calculating $F(t)$ with a linked cluster expansion, $F(t) = \sum_l F_l(t)$. The general expression for $F_2$ [25]

$$F_2(t) = \frac{1}{\Omega\pi}\int_{-\infty}^{\infty}\frac{d\omega}{\omega^2}\,\frac{1-e^{-i\omega t}}{1-e^{-\beta\omega}}\sum_{\vec{q}}\left(\frac{V(\vec{q})}{\epsilon(\vec{q})}\right)^2 \mathrm{Im}P(\vec{q},\omega), \tag{2.3.3}$$

where $\Omega$ is the volume of the system $\mathrm{Im}P(\vec{q},\omega)$ is the imaginary part of the electronic polarizability; $\epsilon(\vec{q})$ is the $\omega \to 0$ limit of the dielectric function $\epsilon(\vec{q},\omega)$, which is the relevant limit for this problem.

For the free electron gas, higher order terms $F_l$ with $l > 2$ are known [28] to be quantitatively important, but they do not change the qualitative behaviour. We will ignore such higher order terms here since they can not be analyzed without a microscopic theory for the MFL.

Eq.(2.3.3) provides a suitable starting point for comparing the behaviour of an FL with that of the MFL hypothesis, since the latter is formulated directly in terms of the polarizability $P$,

$$\mathrm{Im}P^{MFL}(q,\omega) = -\frac{2N_F}{\pi}\,\frac{\tanh(\frac{\omega}{2T})}{1+(\frac{\omega}{\omega_c})^2}, \tag{2.3.4}$$

where $N_F$ is the density of states at the Fermi level, and where $\omega_c$ is a frequency cutoff which is of the order of a few hundred meV (a few thousand Kelvin, we take both $\hbar$ and $k_B$ equal to 1). Note that Eq.(2.3.4) is a smooth version of Eq.(2.2.5).

It has already been noted by Zhang *et al.* [2] that there are conceptual difficulties with this ansatz: it does not satisfy the usual $f$-sum rule [7]. This problem can not be remedied by including a $\vec{q}$-dependent function on the right hand side. These conceptual difficulties have, however, also practical implications, in that they make it difficult to estimate the magnitude of the X-ray edge effect or of the temperature dependence of the hopping. As we shall see, when Eq.(2.3.4) is taken at face value, one is lead to expect a measurable temperature dependence of the XPS line shape

(in particular a temperature dependent frequency shift), and a strongly suppressed hopping rate[2].

We compare with the case of an FL. In this case, we have [7]

$$\mathrm{Im}P^{FL}(q,\omega) = -\omega \frac{2\pi^2 N_F^2}{q k_F^2} \Theta(2k_F - q), \tag{2.3.5}$$

so that, upon introducing a cutoff $\omega_c$ which is of the order of $\epsilon_F$

$$F_2^{FL}(t) = -K^{FL} \int_{-\omega_c}^{\omega_c} \frac{d\omega}{\omega} \frac{1 - e^{-i\omega t}}{1 - e^{-\beta\omega}}, \tag{2.3.6}$$

where

$$K^{FL} = \frac{1}{\Omega\pi} \frac{2\pi^3 N_F^2}{k_F^2} \sum_{\vec{q}} \left( \frac{V(\vec{q})}{\epsilon(\vec{q})} \right)^2 \frac{1}{q} \Theta(2k_F - q) = \mathcal{O}(N_F V)^2. \tag{2.3.7}$$

Here $V$ is the order of magnitude of $V(\vec{q})$. For an FL, the dimensionless number $K^{FL}$ is directly related to the phase shift induced by the potential; a typical value for $K^{FL}$ that we get from the experiments is 0.1 [7, 31].

Let us now compare these results to those for an MFL. First of all, experiments show that the high-temperature superconductors show strong Raman scattering. This indicates at least that Eq.(2.3.4) is indeed of the right order of magnitude in the regime relevant for Raman scattering (assuming that the coupling to light is not drastically different from that of other materials). At typical frequency shifts $\omega$ of the order of hundred $cm^{-1}$, i.e. of the order of $10^{-1}E_F$ or less, and wave numbers $q \approx 10^{-3}k_F$, $\mathrm{Im}P^{FL}$ is of order of $N_F$ or somewhat larger. Expression (2.3.4) for $\mathrm{Im}P^{MFL}$ also gives a result of order $N_F$ for $\omega \geq T$.

Using Eq.(2.3.4) in Eq.(2.3.3) gives

$$F_2^{MFL}(t) = -K^{MFL} \int_{-\infty}^{\infty} \frac{d\omega}{\omega^2} \frac{1 - e^{-i\omega t}}{1 - e^{-\beta\omega}} \frac{\tanh(\frac{\beta\omega}{2})}{1 + (\frac{\omega}{\omega_c})^2}, \tag{2.3.8}$$

with

$$K^{MFL} = \frac{2N_F}{\Omega\pi^2} \sum_{\vec{q}} \left( \frac{V(\vec{q})}{\epsilon(\vec{q})} \right)^2 = \mathcal{O}(N_F k_F^3 V^2). \tag{2.3.9}$$

Let us assume that the potential $V$ is of the same order in the high-$T_c$ materials as it is in normal metals. Then we get

$$K^{MFL} = \mathcal{O}\left( \frac{N_F^{MFL} k_F^3}{(N_F^{FL})^2} \right) \cdot K^{FL} = \mathcal{O}(0.1E_F) \approx 1000K. \tag{2.3.10}$$

[2]These difficulties are less manifest when one confines the analysis to properties associated with the self-energy. The strength of the anomalous terms in the self-energy is the product of the prefactor of the polarizability and an interaction parameter. For most properties, only this product is important, but not the separate constants. See e.g. Pelzer, [11] for an estimate of the coupling strength from the conductivity data.

We stress that this is *a very rough* estimate, as it is based on various rather crude approximations; the interaction parameter will depend strongly on the closeness of the relevant site to the copper oxide planes in a high-$T_c$ superconductor, and can be different for the hopping and the X-ray problem. In particular, it seems likely to us that Eq.(2.3.4) overestimates the polarizability considerably for short wave length (note that $\mathrm{Im}P^{FL}$ decreases as $q^{-1}$) ; thus we think our expression overestimates $K^{MFL}$. Moreover, any effects of anisotropy, which should be quite important for the cuprates, have been ignored. We finally note that Zhang *et al.* [2] used $K^{MFL} = 15K$, without justifying this particular choice.

## 2.3.2 Comparison of $F_2^{FL}$ and $F_2^{MFL}$

The integral in Eq.(2.3.6) for an FL with a sharp cutoff can be done [32] but for our numerical studies we choose to evaluate this integral with a smooth cutoff, of the same form as the high frequency cutoff in the MFL expression (2.3.4).

Consider $F_2^{MFL}$, given by Eq.(2.3.8). The contour can be closed in the lower half plane. There are poles in 0, $-i\omega_c$ and $-2n\pi Ti, n = 1, 2, \cdots$. The result for the principal part of the integral is that the real part has a simple expression:

$$\mathrm{Re}F_2^{MFL}(t) = -K^{MFL}\left(\frac{\pi t}{2} - \frac{\pi}{2\omega_c}(1 - e^{-\omega_c t})\right), \tag{2.3.11}$$

and that the imaginary part contains an infinite sum of contributions of the poles $-2n\pi Ti$. These become small for large $n$, and we can sum them numerically until we obtain the desired precision.

The above result agrees with that of Zhang *et al.* for small $t$, but their expression is wrong for large $t$. In particular, their logarithmic correction is absent; this does not affect the results significantly, however. Note that, as Zhang *et al.* already stressed, $\mathrm{Re}F_2^{MFL}$ is temperature independent and decreases linearly for large times. In an FL, on the other hand, $\mathrm{Re}F_2^{FL}$ has strong temperature dependence, and at $T = 0$ it decreases logarithmically for long times; for finite temperature, the asymptotic behaviour is $-K^{FL}\pi tT$.

To obtain the asymptotic behaviour of $\mathrm{Im}F_2^{MFL}$, we note that by using the fact that $(1 - e^{-\beta\omega})^{-1} = 1 - (e^{\beta\omega} - 1)^{-1}$, we can write

$$\mathrm{Im}F_2^{MFL}(t) = -K^{MFL}\int_0^\infty \frac{d\omega}{\omega^2}\,\frac{\sin\omega t\,\tanh(\frac{\beta\omega}{2})}{1 + (\frac{\omega}{\omega_c})^2}. \tag{2.3.12}$$

For $t \to \infty$, $\sin\omega t/\omega \to \pi\delta(\omega)$, and we find

$$\mathrm{Im}F_2^{MFL}(t \to \infty) = -\frac{\pi\beta K^{MFL}}{4}. \tag{2.3.13}$$

By differentiating Eq.(2.3.12) with respect to $t$, we get

$$\left.\frac{d\mathrm{Im}F_2^{MFL}(t)}{dt}\right|_{t=0} = -K^{MFL}\int_0^\infty \frac{d\omega}{\omega}\,\frac{\tanh(\frac{\beta\omega}{2})}{1 + (\frac{\omega}{\omega_c})^2}. \tag{2.3.14}$$

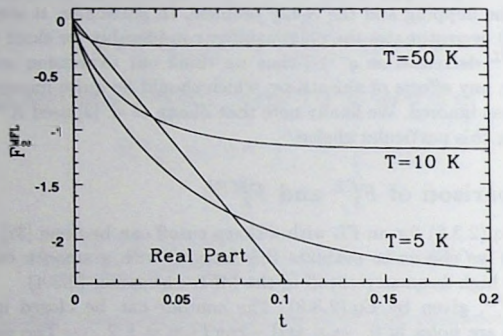



Figure 2-2. Behaviour of $\mathrm{Im}F_2^{MFL}$ and $\mathrm{Re}F_2^{MFL}$ as a function of time, with $K^{MFL} = 15K$ and $\omega_c = 1500K$. The real part of $F_2^{MFL}$ is independent of temperature.

For $\beta\omega_c >> 1$, the integrand behaves as $\omega^{-1}$ over a large part of the $\omega$-interval, and this results in a logarithmic temperature dependence, $\ln \beta\omega_c$. In particular, if we replace the smooth cutoff by a sharp cutoff at $\omega_c$, the integral in Eq.(2.3.14) is of the form well-known in the BCS theory, and we get

$$\left.\frac{d\mathrm{Im}F_2^{MFL}(t)}{dt}\right|_{t=0} = -K^{MFL} \ln 1.13\beta\omega_c. \qquad (2.3.15)$$

With a smooth cutoff, the numerical constant 1.13 is replaced by another constant of order unity.

The behaviour of $\mathrm{Im}F_2^{MFL}$ as a function of time is illustrated in Fig.(2-2) for several temperatures and $K^{MFL} = 15$ Kelvin. We see that the asymptotic value, Eq.(2.3.13) is approached smoothly, the approach being slower at lower temperatures. This is a result of the fact that the $t = \infty$ value diverges as $T^{-1}$, while the initial slope diverges only logarithmically.

### 2.3.3 Consequences for $A_h(\omega)$

Before presenting the results for the MFL case, let us recall the essential features of the spectral density in the case of an ordinary FL, obtained numerically from Eq.(2.3.2) in the approximation $\rho = e^{F_2^{FL}(t)}$. At temperature $T = 0$, the asymptotic $\ln t$ dependence of $\mathrm{Re}F_2^{FL}$ gives rise to the well-known $\omega^{-(1-K^{FL})}$ divergence of the

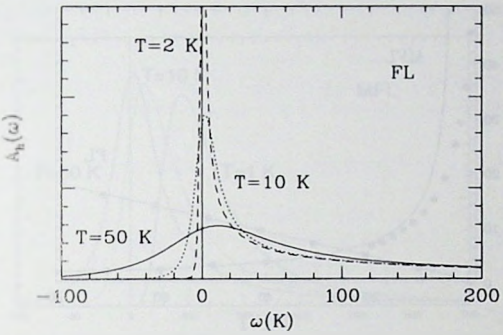



Figure 2-3. Spectral density for the FL case, with $K^{FL} = 0.3$ and $\omega_c = 1500K$. $A_h(\omega)$ has arbitrary units.

spectrum for positive $\omega$, while $A_h(\omega) = 0$ for $\omega < 0$. As illustrated in Fig.(2-3), for finite temperatures the line shape starts to broaden and to become less asymmetric (since $\omega$ is given in units of "temperature", Fig.(2-3) illustrates that the width of the spectrum for $\omega < 0$ is of order $T$). At the same time, there is a small shift in the position of the peak with increasing temperature, and the peak height drops. Since the shift of the peak is small, and since the small $\omega$ behaviour is determined by the asymptotic behaviour of $F_2^{FL}$, we can use the results $\mathrm{Im}F_2^{FL} \approx -K^{FL}\pi/2$ and $\mathrm{Re}F_2^{FL} \approx -K^{FL}\pi tT$ for $t \to \infty$ of the previous section to obtain the following result for $A_h(\omega)$, valid for small $\omega$ and $T$:

$$A_h^{FL}(\omega) \approx \frac{K^{FL}\pi T \cos(K^{FL}\pi/2) + \omega \sin(K^{FL}\pi/2)}{(K^{FL}\pi T)^2 + \omega^2}. \qquad (2.3.16)$$

From this result, the position and width of the peak for small $T$ immediately follow (e.g., for $K^{FL} = 1$ it is easy to see that $\omega_{max} = \pi T$, and that the width goes as $1/T$). In Fig.(2-4), we compare the values of $\omega_{max}$ obtained from the numerical spectra with this approximation for $K^{FL} = 0.3$.

Thus, for an FL, when $T$ decreases, the asymmetric peak becomes narrower, and its position shifts towards zero frequency, while its height increases. The large-$\omega$ behaviour always remains a power law[3]. For a more detailed discussion of the

[3]The starting model is, however, only realistic for small values of $\omega$. See e.g. [7].

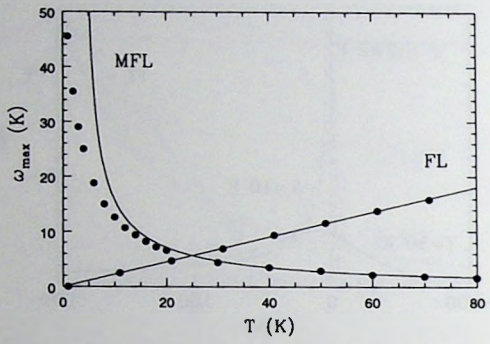



Figure 2-4. Behaviour of the location of the peak as a function of temperature, with $K^{MFL} = 15K$, $K^{FL} = 0.3$ and $\omega_c = 1500K$. The points are from the numerical spectra, the lines are from the approximations discussed in the text.

temperature dependence of the X-ray edge effect in an FL, we refer to Ohtaka and Tanabe [30].

We now turn to the spectral density in the case of an MFL. In Figs (2-5) and (2-6), we plot the spectral density for this case for two values of $K^{MFL}$, $K^{MFL} = 15$ and $K^{MFL} = 400$ (in Kelvin). These figures illustrate that the core hole spectra indeed behave quite differently from that of the FL-case: the broadening of the profile with temperature is quite weak; instead, the most salient feature is the shift of the profile position with temperature. This effect is stronger the larger $K^{MFL}$ is, and is due to the fact that the temperature dependence of $\mathrm{Re}F_2^{MFL}$ and $\mathrm{Im}F_2^{MFL}$ is reversed relative to that of an ordinary FL.

Along similar lines as above, one can derive an approximate expression for the peak region of $A_h(\omega)$ from the $t \to \infty$ asymptotic behaviour of $F_2^{MFL}$ derived in the previous section. The result is

$$A_h^{MFL}(\omega) \approx \frac{K^{MFL}\pi \cos(K^{MFL}\pi/4T) + 2\omega \sin(K^{MFL}\pi/4T)}{(K^{MFL}\pi)^2/4 + \omega^2}. \qquad (2.3.17)$$

This expression is only accurate for small $K^{MFL}$ and large $T$: the first condition is necessary to ensure that $|\rho(t)|$ is not too small in the large time asymptotic regime, and the second condition is necessary to ensure that $\mathrm{Im}F_2^{MFL}$ reaches its asymptotic behaviour relatively quickly, compare Fig.(2-2).

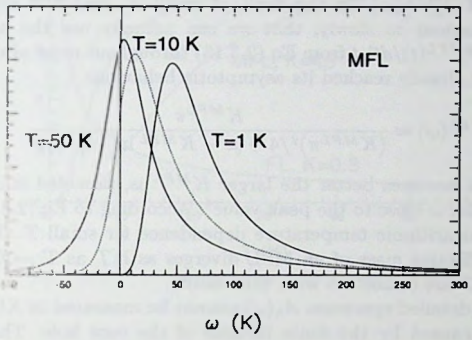



Figure 2-5. Spectral density in arbitrary units for the MFL case, $K^{MFL} = 15K$ and $\omega_c = 1500K$.

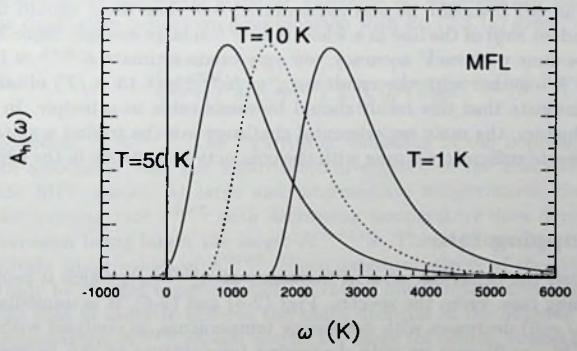



Figure 2-6. Spectral density in arbitrary units for the MFL case, $K^{MFL} = 400K$ and $\omega_c = 1500K$.

In Fig.(2-4) we also show data of $\omega_{max}$ for the MFL spectra obtained numerically, with $K^{MFL} = 15K$. The solid line is the result obtained for $\omega_{max}$ from Eq.(2.3.17); as expected, this approximation is accurate for large $T$, but breaks down at small $T$.

The following approximation helps to understand the physics underlying the above behaviour of $A_h^{MFL}(\omega)$. In the limit of small $T$, $\mathrm{Im}F_2^{MFL}$ crosses over to its asymptotic behaviour so slowly, that we can actually use the approximation $\mathrm{Im}F_2^{MFL}(t) \approx d\mathrm{Im}F_2^{MFL}(t)/dt|_0 t$ from Eq.(2.3.15) throughout most of the the regime where $\mathrm{Re}F_2^{MFL}$ has already reached its asymptotic behaviour.

$$A_h^{MFL}(\omega) \approx \frac{K^{MFL}\pi}{(K^{MFL}\pi)^2/4 + [\omega - K^{MFL}\ln(\frac{1.13\ \omega_c}{T})]^2}. \qquad (2.3.18)$$

This approximation becomes better the larger $K^{MFL}$ is, provided it is still smaller than $\omega_c$, and holds for $\omega$ close to the peak value[4]. According to Eq.(2.3.18), the peak value $\omega_{max}$ has a logarithmic temperature dependence for small $T$. This is due to the fact that the effective mass of an MFL diverges as $\ln T$ as $T \to 0$ (the data in Fig.(2-6) for small $T$ are consistent with this result).

In practice, the detailed spectrum $A_h(\omega)$ cannot be measured in XPS, as a result of the broadening caused by the finite lifetime of the core hole. This broadening can be of the order of an $eV$. As pointed out by Doniach and Sunjic [29], the X-ray edge effect in an FL gives rise to an asymmetric line shape. In principle, the asymmetry is slightly temperature dependent, but often thermal broadening from phonons give a bigger temperature dependence [31]. Our results above show that the line shape in an MFL should be less asymmetric, but that there should be a temperature dependent *shift* of the line as a whole if $K^{MFL}$ is large enough. Since XPS experiments can be done with $meV$ accuracy, our very crude estimate $K^{MFL} \approx 1000$ Kelvin ($\approx 80meV$), together with the result $\omega_{max} \approx K^{MFL}\ln(1.13\ \omega_c/T)$ obtained from Eq.(2.3.18) indicate that this result should be measurable in principle. In the cuprate superconductors, the main experimental challenge may be to find a suitable core level that interacts sufficiently strong with the conduction electrons in the copper oxide planes.

### 2.3.4 The hopping rate

Let us finally discuss the behaviour at zero frequency of $A_h^{MFL}(\omega)$, which is proportional to the hopping rate. From the spectra, Figs (2-5) and (2-6), it is immediately clear that $A_h^{MFL}(\omega = 0)$ decreases with decreasing temperature, in contrast with the FL case where $A_h^{FL}(\omega = 0)$ goes up with decreasing temperature as can be seen in Fig.(2-3). For large temperatures, we can use Eq.(2.3.17) at zero frequency to obtain

$$\nu^{MFL} = 2\Delta^2 \cos[K^{MFL}\pi/4T] \cdot \frac{4}{K^{MFL}\pi}. \qquad (2.3.19)$$

[4]We note that formally the spectrum $A_h^{MFL}(\omega)$ falls off as a Gaussian for frequencies $|\omega - \omega_{max}| >> \omega_c$, due to the fact that $\mathrm{Re}F_2^{MFL}$ goes as $t^2$ for times $t < \omega_c^{-1}$. Of course, this behaviour is just an artifact of the unrealistic behaviour of $\mathrm{Im}P^{MFL}(\omega)$ for $\omega > \omega_c$.

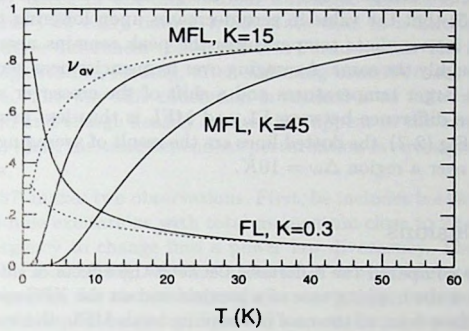



Figure 2-7. Hopping rate as a function of temperature, relative to the high-temperature value.

The analogous expression given by Zhang *et al.*, although not quite right, gives the same qualitative behaviour. Combination with Eq.(2.3.18) gives

$$\frac{\nu^{MFL}(T \to 0)}{\nu^{MFL}(T = \infty)} = \frac{1}{1 + \frac{4}{\pi^2} \ln^2(\frac{1.13\omega_c}{T})}, \quad \text{(low T)} \tag{2.3.20}$$

independent of $K^{MFL}$. The logarithmic vanishing of the hopping rate as $T \to 0$ is again associated with the logarithmic divergence of the effective mass as $T \to \infty$ in the MFL model. At large and intermediate temperatures, the rate of decrease of the hopping rate $\nu^{MFL}$ with decreasing temperature does depend on $K^{MFL}$, the suppression being larger the larger $K^{MFL}$ is. This is illustrated in Fig.(2-7) for two relatively small values of $K^{MFL}$. Since quantum effects of the type considered here can only be studied experimentally at rather low temperatures (less than 50 $K$, say), we see that in order to observe the sharp decrease in the hopping rate as a function of temperature, $K^{MFL}$ should not be too large (of course, it also requires a material that exhibits MFL behaviour down to very low temperatures).

The above analysis corresponds to the ideal case. It is well-known that if there are impurities, there is a distribution of levels, and one has to average $A(\omega)$ over a region around zero frequency to obtain the hopping rate [33] (this may be particularly relevant for the high-$T_c$ materials). As Fig.(2-3) shows, in the FL case, there is a frequency range in the spectrum where the values for a low temperature are higher than at a high temperature, but further away from zero frequency, it is the other

way around. Averaging therefore leads to a flattening of the curve for the hopping rate, especially at low temperatures, where the peaks are narrow. In the MFL case, we saw in Fig.(2-5) that the value in zero decreases upon lowering the temperature, but also that for intermediate temperatures the peak remains nearby, and that its height remains roughly the same. Averaging over frequencies near zero therefore leads to a flattening at larger temperatures and a shift of the crossover region to smaller temperatures. The difference between FL and MFL is thus less pronounced than in the ideal case. In Fig.(2-7), the dotted lines are the result of averaging our numerically obtained spectra over a region $\Delta\omega = 10K$.

### 2.3.5 Conclusions

In this section we compared the difference between the effects of interaction with an FL and an MFL on the hopping rate of a particle and on the XPS spectrum of a deep core hole. As we have seen, in the case of coupling to an MFL, the increasing effective mass of the MFL (or the fact that the quasi-particle spectral weight vanishes as $T \to 0$) causes a decrease of the hopping rate with decreasing temperature, instead of the usual increase. Only when the coupling strength is relatively weak do we expect the characteristic crossover to occur in a temperature regime where, in earlier experiments on normal metals, the quantum effects have turned out to be measurable. In addition, we expect two other effects to adversely affect the experimental testability of the predicted decrease in the hopping rate in high-$T_c$ superconductors. First of all, it will be difficult to disentangle such a decrease from the usual classical activated behaviour ($\propto e^{-E/kT}$) of the diffusion coefficient. Moreover, in high-$T_c$ superconductors with their complicated unit cell, the effect might easily be washed out to a large degree by the hopping of the particle between several nonequivalent sites, at some of which the coupling to the conduction electrons in the $CuO$ planes will be small.

The main feature of the XPS spectrum of a deep core hole interacting with an MFL is found to be a temperature dependent shift of the position of the peak. This effect in principle offers better opportunities for experimental tests. One reason is that, in a photo-emission experiment, the core level of a given atom at a fixed site in the unit cell can be measured. Furthermore, the $meV$ resolution with which the spectra can be measured should allow one to observe this shift, provided $K^{MFL}$ is not much smaller than our rough estimate $K^{MFL} \approx 75meV$ [cf. Eq.(2.3.10)]. We hope our findings will stimulate the experimental search for this effect in photo-emission experiments.

## 2.4 Local disturbances in different anomalous metals

In this section, we consider what happens to the Fermi edge singularity when the conduction electrons properties deviate from the usual Fermi liquid behaviour, in

different ways than the marginal Fermi liquid of the previous section.

For the X-ray effect in one-dimensional Luttinger liquids, different treatments exist in literature. Two early ones [34, 35] state that in a Luttinger liquid, there are only quantitative differences compared to a Fermi liquid. The divergency in the polarization is logarithmic, leading, after summing contributions, to power laws in X-ray line shapes. The way this conclusion is reached is as follows. First, one takes the collective spin and charge density modes that appear in the exact solution of the Luttinger model, then, the time dependent perturbation is envisaged to interact with those excitations.

[36, 37] makes two observations. First, he includes backscattering: the creation of electron-hole excitations with total momentum close to $2k_F$. This causes the logarithmic divergency to change into a power law divergency. The second observation is that this backscattering, which is a matter of density of states of electron-hole excitations, is essential for the orthogonality catastrophe, rather than the (single-particle) non-Fermi liquid behaviour itself. The non-Fermi liquid character of the conduction electrons and the anomalous response to local disturbances (power laws replacing logarithms) must be viewed as having a common origin, rather than one being the cause of the other. This is similar in spirit to our approach of orthogonality in an MFL: the anomalous polarizability is the cause of the vanishing of $Z$ for quasi-particles within the conduction electron system and of the anomalous temperature dependence in X-ray problems.

In contrast to the MFL case, the Luttinger liquid allows for a full treatment of the Kondo effect [38, 39, 40]. The presence of two couplings, for forward scattering and for backward scattering, yields a new structure of flow diagrams, in a poor man's scaling calculation [39]. New features, compared to an FL are that not only antiferromagnetic, but also ferromagnetic Kondo coupling (see Eq.(2.2.8)) flows to the strong-coupling regime, and that the Kondo temperature does not depend logarithmically on the exchange coupling, but as a power law.

Let us make a short digression on models that are outside the scope of this thesis, because the special behaviour is not due to electron-electron interactions, but which do exhibit anomalous features as far as the local properties are concerned. Hopping on quasi-periodic tilings [41], without interactions, yields peculiar single-particle behaviour. Obviously, quasi-particles with quantum number $\vec{k}$ do not exist. Wave functions are neither Bloch states, nor are they exponentially decaying, instead they show algebraic decay. This leads to a drastic change compared to the usual dynamics of excitations over a Fermi surface. This gives rise to a modified Kondo effect, obeying power law behaviour instead of the standard logarithms [41].

Metals with the density of states going to zero with a power law near the Fermi energy give rise to a zero-temperature phase transition at finite coupling constant. This result has been obtained using poor man's scaling [42].

A class of metals with an anomalous Fermi surface is that of the half-metallic magnets. These have a Fermi surface for one spin species only, and a gap for excitations for the other. We did not find any Fermi surface effect [43] that is especially sensitive

to this feature. The X-ray and hopping problems are sensitive to charge excitations, which are, apart from a reduced density of states, not very special in a half-metallic magnet. For the Kondo effect, it is true that only one species of impurity spin can be screened by conduction electrons, but in order to obtain the singular scattering characteristic for the Kondo effect, conduction electrons need to be able to scatter changing their spin, at low energy cost. So no special type of Kondo effect is to be expected in half-metallic magnets.

The last special metal for which we discuss the local properties, as t[illegible] treated in literature, are two dimensional metals exhibiting a (logarithmic) Va[illegible] singularity in the density of states. Gogolin [44, 45] describes how integra[illegible]ge and the usual logarithms in responses to local time dependent potentials a[illegible]ced by a logarithm squared. One result is that the Kondo temperature is en[illegible]

Let us make a final remark on local dynamic disturbances in MF[illegible]ile the absence of a microscopic theory prevents a full treatment of the Kondo [illegible]n such a model, one can discuss the issue on the level of poor man's scaling. D[illegible] special behaviour of the Luttinger liquid, the ferromagnetic coupling being d[illegible]ated by the strong coupling fixed point and the dependences being power law rather than logarithms, also appear in this other case of non-Fermi liquid conduction electrons? The answer is no. As indicated before, the presence of only two Fermi points is the reason for having different flow diagrams, not the non-Fermi liquid behaviour itself. For an MFL in higher dimensions, there is no need to introduce two separate Kondo couplings, one for forward and one for backward scattering. The other important quantity (see [46]), the density of states, does not differ from the normal case to the extent that it yields a completely different flow diagram. Quantitative differences could occur, because the strange temperature dependence in the MFL ansatz, but the smoothness of the density of states as a function of energy guarantees normal behaviour at this level of description.

In chapter 1, we announced that we would treat the question: "If the transport properties show MFL behaviour, instead of FL behaviour, will the response to a dynamic *local* disturbance also be different?". Our conclusions are as follows. The anomalous temperature dependence in the polarizability yields an anomalous temperature dependence in the singular X-ray and hopping effects. While usually, the extent to which a mode is excited depends on temperature, here the mode of Eq.(2.2.5) itself also depends on temperature. Other cases cannot be treated analytically without a microscopic theory for the MFL. The Kondo effect is expected to be only quantitatively different in an MFL, compared to an FL.

As a last remark on local dynamic disturbances and non-Fermi liquid behaviour, we note that in all cases, a common cause leads to singularities in interactions among the conduction electrons themselves (and thus to a breakdown of the usual FL behaviour) and to a change in the response to an external local disturbance. These responses are already singular, in all cases this singularity is enhanced: logarithmic singularities become log squared with a Van Hove singularity in the density of states and power laws in one dimensional Luttinger liquids. Also the suppression near $\omega = 0$

in Figs (2-5) and (2-6) for an MFL is stronger than usual, in an FL the delta function of the free case is replaced by a diverging power law, here it is a temperature dependent decrease for decreasing frequency.

# References


[1] C. M. Varma, P. B. Littlewood, S. Schmitt-Rink, E. Abrahams and A. E. Ruckenstein, Phys. Rev. Lett. **63**, 1996 (1989)

[2] C. Zhang, T. Toyoda and Y. Takahashi, J. Phys.: Condens. Matter **4**, L15 (1992)

[3] The MFL has more recently been considered in the context of the half-filled level, see e.g. B. Marston, Bull. Am. Phys. Soc. **40**, 565 (1995)

[4] van Bemmel and W. van Saarloos, J. Phys.: Condens. Matt. **5**, 1377

[5] See any standard work on Fermi liquid theory, e.g. refs [25, 26, 27] of chapter 1

[6] For a review see e.g. J. Sólyom, Adv. Phys. **28**, 201 (1979)

[7] G. D. Mahan, *Many-Particle Physics*, Plenum, New York (1981)

[8] See e.g. H.J. Schulz, lecture notes Les Houches summer school 1994, cond-mat preprint 9503150

[9] J. Voit, Phys. Rev. B **47**, 6740 (1993)

[10] V. Meden and K. Schönhammer, Phys. Rev. B **46**, 15753 (1992)

[11] F. Pelzer, Phys. Rev. B **44**, 293 (1991)

[12] P. B. Littlewood and C. M. Varma, Phys. Rev. B **45**, 12636 (1992)

[13] Y. Kuroda and C.M. Varma, Phys. Rev. B **42**, 8619 (1990)

[14] E.J. Nicol and J.P. Carbotte, Phys. Rev. B **43**, 1158 (1991)

[15] M.L. Horbach, F.L.J. Vos and W. van Saarloos, Phys. Rev. B **48**, 4061 (1993)

[16] J. Zaanen, M.L. Horbach and W. van Saarloos, unpublished

[17] S. Doniach and E.H. Sondheimer *Green's functions for Solid State Physicists*, Benjamin, Reading (1982)

[18] P. W. Anderson, Phys. Rev. Lett. **18**, 1049 (1967)

[19] J. Kondo, Physica B&C **84B**, 40 (1976); Physica B&C **123B**, 175 (1984); Physica B&C **124B**, 25 (1984)

[20] P.W. Anderson and G. Yuval, Phys. Rev. Lett. **23**, 89 (1969)

[21] K.Yamada and K. Yosida, Progr. Theor. Phys. **59**, 1061 (1978); K.Yamada and K. Yosida, Progr. Theor. Phys. **60**, 353 (1978); K.Yamada and K. Yosida, Progr. Theor. Phys. **62**, 363 (1979)

[22] P. Nozières, Spring college on quantum phases, Trieste (1994)

[23] C. Zhang and Y. Takahashi, Phys. Rev. B **46**, 9247 (1992)

[24] P. B. Littlewood and C. M. Varma, J. Appl. Phys. **69**, 4979 (199

[25] C. Zhang, G. Gumbs and N. Tzoar, Phys. Rev. B **43**, 1463 (199

[26] K. Yamada, A. Sakurai and M. Takeshige, Progr. Theor. Phys. 7 983)

[27] J. Kondo and A. Yoshimori (eds.), *Fermi Surface Effects* Springer (1988)

[28] P. Nozières and C. T. de Dominicis, Phys. Rev. **178**, 1097 (1969)

[29] S. Doniach and M. Šunjić, J. Phys. C **3**, 285 (1970)

[30] K. Ohtaka and Y. Tanabe, Rev. Mod. Phys. **62**, 929 (1990)

[31] P.H. Citrin, G. K. Wertheim and Y. Baer, Phys. Rev. B **16**, 4256 (1977)

[32] C.-O. Almbladh and P. Minnhagen, Phys. Status Solidi B **85**, 135 (1978)

[33] H. Sugimoto, J. Phys. Soc. Japan **55**, 1687 (1986)

[34] T. Ogawa, A. Furusaki and N. Nagaosa , Phys. Rev. Lett. **68**, 3638 (1992)

[35] D. K. Lee and Y. Chen, Phys. Rev. Lett. **69**, 1399 (1992)

[36] A.O. Gogolin, JETP Lett. **57**, 50 (1993)

[37] A.O. Gogolin, Phys. Rev. Lett. **71**, 2995 (1993)

[38] D.-H. Lee and J. Toner, Phys. Rev. Lett. **69**, 3378 (1992)

[39] A. Furusaki and N. Nagaosa, Phys. Rev. Lett. **72**, 892 (1994)

[40] A. Schiller and K. Ingersent, Phys. Rev. B **51**, 4676 (1995)

[41] V.G. Benza and E. Montaldi, J. Phys. A: Math. Gen. **27**, 2299 (1994)

[42] D. Withoff and E. Fradkin, Phys. Rev. Lett. **64**, 1835 (1990)

[43] J. Kondo and A. Yoshimori eds., Fermi surface effects, Springer, Berlin (1988)

[44] A.O. Gogolin, JETP Lett. **57**, 311 (1993)

[45] A.O. Gogolin, Z. Phys. B **92**, 55 (1993)

[46] P.W. Anderson, *Basic notions of condensed matter physics*, Addison-Wesley, Redwood City, (1984)

# 3 Inhomogeneous states

## 3.1 The approximations used in the spin polaron picture

In the previous chapter, we discussed real-space probes, but the issue of interest, the nature of the charge excitations in systems of interacting fermions, was formulated in $\vec{k}$-space. In this chapter, we discuss inhomogeneous states in systems of correlated fermions, in particular spin polarons. Other inhomogeneous states include domain walls and phase separation. The latter issue can be formulated as an instability of the Fermi liquid [1], which is a $\vec{k}$-space notion. Here, by contrast, we shall use a real-space picture to describe the abovementioned structures.

The main purpose of this chapter is to address the question: "If measurements are explained in terms of trapped vacancies in heavy spin polarons, could one perform a direct experimental test to verify this explanation?". Different experimental evidence for their existence will be discussed, our specific proposal aims at testing the metastable character of a state containing polarons. Section 3.2 contains our proposed test[1]. In the present section we give a more elaborate explanation of assumptions which are used in that section. In the last section of this chapter, we shall come back to these assumptions. Relaxing them will require computational tools. That will lead us to the development of the fixed-node quantum Monte Carlo method in later chapters.

In chapter 1, we introduced the quantum antiferromagnet and the issue of doping holes into such a system. The $t$-$J$ Hamiltonian, Eq.(1.3.2), describes this situation. We mentioned that at half-filling in dimension two or bigger, there is Néel order, and that, in the $U = \infty$ Hubbard model (which corresponds to the $t$-$J$ model with $J = 0$), one single hole induces a state that is ferromagnetic throughout the whole system (this result is due to Nagaoka [3]). The discussion in chapter 1 indicates that it is far from trivial to obtain those results, and that their stability against deviations from the ideal limits is often not known. What happens off half-filling at finite $J$ is a complicated problem. In this chapter, we shall only use the fact that opposite tendencies exist, i.e. towards Néel order at half filling, and towards ferromagnetic order upon doping. In a simple theoretical framework, this competition leads to a 'compromise', which is called a spin polaron.

In order to understand the limitations of the spin polaron picture for which we shall propose an experimental test, we need to go into the arguments that lead to the

[1] This section is an adapted version of the paper we published in the Journal of Low Temperature Physics [2].

conclusion that there is a difference in kinetic energy for a single hole, for different spin backgrounds. The approach of Brinkman and Rice [4] (which is an extension of the work by Nagaoka), is to consider the hole dynamics in a *given* classical spin background. One wants to obtain the density of states for a single hole. This can be written in terms of the single particle Green function, which reads

$$G_{ij}^{X}(z) = \sum_{\sigma} \langle X | c_{i\sigma}^{\dagger} \frac{1}{z - \mathcal{H}_{tJ}} c_{j\sigma} | X \rangle, \tag{3.1.1}$$

$X$ denotes a spin configuration and $z$ is the complex frequency. $G_{ij}^{X}$ can be written in terms of the number of paths of a hole that return to the starting point and that restore the spin background. This number is different for different spin backgrounds (e.g. in a ferromagnetic spin background, all paths are background preserving). The conclusion of Brinkman and Rice is that the difference in kinetic energy between different spin backgrounds is a fraction of the hopping amplitude $t$.

This observation is the basis for the arguments about the existence of spin polarons. While Nagaoka already proved that for $J = 0$, the ground state is ferromagnetic and the argument of Brinkman and Rice indicates that the energy difference with an antiferromagnetic or paramagnetic state is large, the situation is more complicated for finite $J$. The background itself tends to order antiferromagnetically. The compromise is to make a ferromagnetic region around the hole, in which the hole is localized. Further away, the spins order antiferromagnetically.

The existence of the spin polaron in the Hubbard model for $U >> t$ was suggested in the late seventies. While the original approaches have been reconsidered and refined in the last few years as a result of the increased interest of the behaviour of a hole in an antiferromagnetic background (see section 3.3), for our purposes it suffices to follow the line of argument of the original approaches. Montambaux *et al.* [5] have listed the assumptions that are used to develop these, in addition to the assumptions used by Brinkman and Rice. It aims at solid $^3He$, in which the motion of a vacancy is modeled by the motion of a hole in a Hubbard or $t$-$J$ model. The assumptions are

1. The vacancies are independent.
2. A vacancy can be viewed as a particle in a continuum.
3. The ferromagnetic volume is assumed to be a sphere (or a circle in 2d).
4. The vacancy is completely localized within this volume.
5. The spin polarization is uniform and takes its maximum value within the sphere, it is zero outside of it.
6. The magnetic coupling between the magnetic lattice and the spins at the surface of the sphere is neglected.
7. Lattice distortion around the vacancy is neglected.

With these assumptions, the problem is reduced to solving the Schrödinger equation for a particle in a deep well, and balancing the kinetic energy, which is lower for larger polarons, against the energy it costs to break down the spin background, which is higher for larger polarons. The possible experimental ramifications we discuss in the next section are based on this picture. In section 3.3, we shall come back to both the assumptions of Brinkman and Rice and those listed above.

## 3.2 Spin polarons in solid $^3He$: Suggestions for further experiments

In this section, a proposal is made for a simple way to test experimentally whether anomalies in the low temperature thermodynamic properties of solid $^3He$ are due to *metastable trapped* vacancies in spin polarons. The idea is to perform susceptibility measurements before and after applying a strong magnetic field for some time, with the aid of a SQUID or NMR. If this *magnetic annealing* reduces the susceptibility, it provides evidence for the existence of metastable trapped vacancies in the original sample. With this procedure it may also be possible to clarify the issue of the existence of a 'vacancy solid'. Finally, the effects of the tendency of vacancies to cluster, eventually leading to phase separation, are discussed.

### 3.2.1 Motivation for choosing $^3He$ to study spin polarons

In this section we make a simple but, to our knowledge, new proposal for a way to experimentally investigate the existence of ferromagnetic spin polarons in solid $^3He$. Arguments in favour of the formation of these ferromagnetically ordered regions around vacancies were already put forward almost 30 years ago by Andreev [6] and Héritier and Lederer [7] on the basis of the description of the motion of a vacancy in solid $^3He$ in terms of a Hubbard model (in this case, the usual hopping parameter $t$ is the hopping amplitude of vacancies, and $U$ is of the order of the energy of a vacancy-interstitial pair). The basic idea underlying these arguments is the result of Brinkman and Rice [4] that the density of states for the vacancy in a given spin background is narrower in the paramagnetic and antiferromagnetic phases than in the ferromagnetic phase. As a result, in the limit $t/U \ll 1$ in which the effective spin interaction $J = 2t^2/U$ is much smaller than $t$, a vacancy can lower its kinetic energy by surrounding itself with a ferromagnetically ordered region — a spin polaron. Although a model with higher order interactions is needed to describe the spin ordering of solid $^3He$ in the milliKelvin range [8], this basic feature of polaron formation for $t/J \gg 1$ remains in all earlier approaches. A particularly thorough discussion of the theoretical work on spin polarons in $^3He$ as well as of the experimental evidence for polarons up to 1982 was given by Montambaux *et al.* [5]. Their main conclusion was that some anomalies in the static susceptibility and in the specific heat between a few mK and $40 \sim 50$ mK could well be consistent with the existence of polarons, but that the

number of vacancies needed to explain the data they consider (which were taken away from the melting line), was larger than one would expect. Possibly, they suggested, these vacancies are trapped in the $^3He$ crystals; some experiments that they discussed indeed suggest that the number of vacancies decreases upon (thermal) annealing.

To our knowledge, a systematic experimental investigation [9] of spin polarons has not been undertaken, even though the polaron concept has occasionally been invoked to explain certain data (as will be discussed later, it is believed that the susceptibility data by Kirk *et al.* [10] show evidence of a partial melting of the crystal [11, 12], not of the existence of spin polarons, as was suggested earlier [13]). From a different perspective, Lengua and Goodkind [14] have recently also stressed the need to understand the behaviour of vacancies in solid $He$ at higher temperatures.

From the theoretical side, there is renewed interest in the old subject of magnetic textures around holes in the Hubbard model, in the context of high $T_c$ materials. These textures include polarons, domain walls [15] and vortices [16]. The question of phase separation of holes, which had already been introduced in the discussion of polarons in $He$ [17], was recently discussed for such models by Emery *et al.* [18]. There are several reasons why it is useful to study these textures in $^3He$. First, the long range Coulomb interaction that could cause the actual behaviour in high $T_c$ materials to be different from that of the Hubbard model (e.g. it could prevent phase separation of holes), does not exist between vacancies in $^3He$. Also, the ratio $U/t$ as estimated for $^3He$ [5] is a factor five larger than for the high $T_c$ materials [19] or any other material for which the Hubbard model yields an appropriate description of the electronic degrees of freedom. In fact the extent to which the one band Hubbard model incorporates the essential features of the high $T_c$ materials is still subject of debate [19]. Therefore, studying solid $^3He$ as an example of a correlated Fermi system is very attractive as it appears to provide the *cleanest experimental realization* of a Hubbard type model in the extreme large $U$ limit. This is the parameter region in which polarons could exist, while for smaller $U/t$, domain walls are energetically more favourable. One should keep in mind however one important difference between most theoretical analyses of the Hubbard model and related models and of the studies of polarons in $^3He$: the former are mostly focussed on the $T = 0$ ground state properties in two dimensions, while the latter have the most relevance in the paramagnetic phase at temperatures $T$ above the Néel ordering temperature $T_N$. In this case the formation of a ferromagnetic polaron results from the balance of the reduced kinetic energy of the vacancy and the decrease in spin entropy associated with the ferromagnetic spin ordering, while at $T = 0$ the kinetic energy of the vacancy is balanced with the increase in exchange energy resulting from the breaking of the antiferromagnetic bonds.

In the Brinkman-Rice [4] approach on which the polaron theory [6, 7, 5] is based, quantum fluctuations in the antiferromagnetic order are ignored, and as a result, the vacancy motion in an antiferromagnetic background is found to be diffusive. Recent numerical and theoretical results [20] on the $t$-$J$ model in with $t/J = \mathcal{O}(1)$ show that the quantum fluctuations make the motion coherent, as a quasiparticle band of

width of order $J$ is found. If these findings extrapolate to smaller values of $J$ then they do not necessarily, however, invalidate the earlier arguments for the existence of polarons for very small $J/t$: in this limit the decrease in kinetic energy of the hole in a region of ferromagnetic order ($\sim t$) always exceeds the kinetic energy of the coherent quasi-particle motion in the antiferromagnetic or paramagnetic background ($\sim J$). Moreover, there have been arguments [20] that the bandwidth for the hole motion is is exponentially small for $t/J \gg 1$, which is the limit relevant for $^3He$.

Our main purpose in this section is to advocate a simple way to test experimentally whether *metastable* trapped vacancies are present in a sample. The point is that polarons around vacancies have been invoked to explain anomalies in low-temperature thermodynamic properties sufficiently away from the melting line, and that more vacancies are needed than one would expect in equilibrium [21]. In this picture, the reason for the excess of vacancies is that polarons are effectively very heavy, and are trapped in the crystal when the temperature is reduced [5]. This scenario can be tested in the following way. We propose to perform systematic measurements in the temperature and molar volume range of the susceptibility [22, 23] and specific heat [24, 25] measurements the behaviour of which was interpreted in terms of polarons by Montambaux *et al.* [5]. If these measurements show behaviour consistent with the earlier experiments, and hence with the existence of metastable trapped vacancies, a strong magnetic field (of several Tesla) should be applied for some time (we shall estimate how long later). Since their mobility is a strongly increasing function of the polarization (as the vacancies can move coherently when the polarization approaches unity) this allows the density of vacancies to approach the equilibrium density at the temperature at which the measurement is done. Then, after having switched off the magnetic field, the measurement should be repeated. If the anomalies are then reduced, one can conclude that their occurrence in the first measurement was indeed due to polarons around metastable trapped vacancies. This procedure comes down to *magnetic annealing*. Furthermore, if polarons are found to be stable rather than metastable according to these measurements, we estimate that they can be oriented in fields of the order of 0.5 Tesla (depending, of course, on the temperature). This translates into a nonlinear susceptibility that should start to decrease in fields higher than these values, an effect that should be easily measurable.

Following the suggestion of Castaing and Nozières [26], the melting of highly polarized solid $^3He$ has been used as a way to obtain strongly polarized liquid $^3He$. In a number of experiments of this type, a relatively sharp break in the polarization dependence of the melting pressure has been observed [27]. Again, it has been speculated that this behaviour might be associated with the existence of polarons in the solid phase [5, 28]. After having discussed in subsection 3.2.2 in which experiments in zero field the existence of spin polarons could be visible and how the results would change after magnetic annealing if there was an excess of vacancies before, in subsection 3.2.3 we shall briefly comment on the particular suggestion of the existence of a 'magnetic vacancy solid' of Bouchaud and Lhuillier [28]. It seems appropriate to test whether the effects in the melting behaviour for which this new thermody-

namic phase is invoked, could be due to (trapped) vacancies. Therefore, resolving the question of the existence of polarons in low temperature $^3He$ crystals would at least settle the issue of the appropriateness of using polarons as one of the ingredients for an explanation of the change of the melting behaviour that is observed.

The predictions for the effects of a diminishment of the number of vacancies are rather clear-cut. In subsection 3.2.4, we analyze the behaviour that could arise because holes in the Hubbard model have the tendency to cluster. Clustering could be especially important if there was no excess of vacancies to begin with, or if the applied field in magnetic annealing is too weak or is applied for too short a time for the vacancies to leave the crystal. These effects are more subtle, and possibly difficult to test directly in experiments. Yet it is important to discuss them because we need to know how they affect the interpretation of measurements before and after magnetic annealing with respect to the number of vacancies. Moreover, it would be of fundamental importance if one would be able to demonstrate experimentally that clustering exists.

## 3.2.2 Spin polarons and tests for their existence

### Spin polarons

The properties of spin polarons in $^3He$ are discussed extensively in the paper by Montambaux *et al.* [5]. A recent review can be found in the book by Dobbs [29]. The main idea is that a vacancy has most freedom to move, and thus the lowest kinetic energy, in a ferromagnetic region [3, 4]. The background is paramagnetic in the case of $^3He$ above $T_N$ (or, in the analogous case of the Hubbard model at $T = 0$, antiferromagnetically ordered). This leads to a competition, and in the simple continuum picture of a particle in a deep well with sharp walls of section 3.1, one has to minimize an expression for the free energy of the form [6, 7, 5]

$$F = \frac{a}{R^2} + bR^d, \tag{3.2.1}$$

where the first term arises because the vacancy is localized in a ferromagnetic region of size $R$, and the second term accounts for the fact that the spin background is broken down over a $d-$dimensional volume. Eq.(3.2.1) applies at $T = 0$ with the free energy replaced by the energy if one has holes in a $t$-$J$ or large $U$ Hubbard model; then $a \sim t$ and $b \sim J$. For the case of $^3He$ above $T_N$, the competition is between ferromagnetic ordering of the spins around vacancies and the paramagnetic background. Then the main contribution to the second term in Eq.(3.2.1) comes from the entropy difference, while the first term still represents the vacancy kinetic energy of order $t$, the vacancy hopping amplitude. From Eq.(3.2.1) one finds $R_{min} \sim \left(\frac{a}{b}\right)^{\frac{1}{d+2}}$; for the $t$-$J$ or large-$U$ Hubbard model with $J = 2t^2/U$, this gives $R_{min} \sim \left(\frac{t}{J}\right)^{\frac{1}{d+2}}$, which illustrates that the continuum picture becomes self-consistent for large $t/J$. The estimate of Montambaux *et al.* [5] for $^3He$, after putting in the geometrical factors

and accounting for several corrections to the simple picture implied by Eq.(3.2.1), is that such a polaron contains 30-50 spins. Much work has been done on the Hubbard model and $^3He$ since; the feature that is of importance here, the preference of a hole to have a ferromagnetic surrounding, remains in all these more sophisticated approaches, although the exponent with which the radius scales as a function of $J$ remains subject of debate [20]. It is important to emphasize that the size of a polaron (if it exists) is not necessarily the same in all experiments, since we believe $t$ to be strongly dependent on the molar volume [30]. Moreover, the estimate $t \approx 50$ mK that Montambaux *et al.* used [5] appears somewhat low; we believe $t$ is likely to be substantially larger in most experiments [31], and this would tend to give larger polaron sizes.

### Thermodynamic properties

The fact that polarons act as clusters with giant moments has consequences for thermodynamic properties. For the susceptibility one finds [5]

$$\chi = \frac{\mu^2}{k_B(T+\Theta)}(1-Nx) + x\frac{N^2\mu^2}{3k_BT}, \tag{3.2.2}$$

where $x$ is the vacancy concentration and $N$ is the number of spins per polaron (it is assumed that there is one vacancy per polaron; we shall come back to this point in subsection 3.2.4). In addition, there is a loss of entropy due to the alignment of the nuclear spins. The expression given in [5] for the case of three dimensions is

$$\Delta S = -xk_B \ln 2 \left(\frac{\pi t}{k_BT_N + 2k_BT\ln 2}\right)^{3/5}, \tag{3.2.3}$$

In this expression the term between brackets is the number of spins per polaron $N$ (in terms of the general formula this term is of the form $\left(\frac{a}{b}\right)^{\frac{d}{d+2}}$). These relations have been used by Montambaux *et al.* [5] to explain the enhanced static susceptibility between 1 and about 5 mK, found by Bernier and Delrieu [22] and by Prewitt and Goodkind [23] in a SQUID measurement, and the anomalies in the entropy measurements of Halperin [24] and in the specific heat data of Dundon and Goodkind [25], both below 20 mK.

We note that NMR measurements by Kirk *et al.* [10] show an anomaly opposite to those of Bernier and Delrieu [22] and of Prewitt and Goodkind [23], i.e. a *decrease* in the susceptibility upon lowering the temperature. Kumar and Sullivan [13] have argued that the particular form of NMR used in those experiments could well probe only the first contribution to the susceptibility in Eq.(3.2.2), as the polaron contribution from the second term could be shifted out of the NMR frequency window due to local field effects. If this assumption is made, one can fit the data of Kirk *et al.* in terms of a polaron model. However, Bernier and Suaudeau [11] have put forward as a more likely explanation of Kirk's data a partial melting of the solid, as

the experiments were done close to the melting line, and since the susceptibility of the liquid is much smaller than that of the solid. Later measurements of Bernier *et al.* [12] strongly support this suggestion. We will therefore concentrate our discussion only on those measurements that show an *increase* of the susceptibility (or specific heat), an effect that can not be attributed to a partial melting of the sample.

We finally note that for polarons consisting of $N$ spins, the magnetization will saturate at fields large enough that $N\mu H/k_BT \gtrsim 4$. For $N$ of the order of 30 and $T$ of the order of 5 $mK$, the crossover occurs for fields of the order of 0.3 Tesla. If polarons are larger, the resulting decrease of the susceptibility will of course occur at even smaller field values. Hence, if one still observes an enhancement of the susceptibility after the magnetic annealing discussed below, one can test whether this is due to the existence of stable polarons by measuring whether the nonlinear susceptibility indeed decreases for fields larger than about 0.5 Tesla.

### Magnetic annealing and the detection of a diminishment of the number of vacancies

Although Montambaux *et al.* [5] may have overestimated the number of vacancies in the experiments the data of which they analyzed [21], the vacancy concentration they needed to account for the anomalies is consistent for the various measurements, but higher than would be expected for thermally activated vacancies at the temperatures ($T \lesssim 40$ mK) at which the experiments are performed. (Except very close to the melting line, vacancy activation energies are of the order of 1 K, so that at low temperatures, one expects only a very low fraction of thermally activated vacancies.) The reason why there could be many vacancies in a metastable state is the following. Once the magnetic polarons form at a certain temperature and pressure, they are trapped in the lattice: because many spins are involved, the polarons are effectively very heavy [20]. Iordanskiĭ [32] has discussed the possibility of polaron motion by the propagation of spin waves through the ferromagnetic region. This gives a larger mobility than if only motion by flipping spins at the boundary is considered, but it is much lower than the mobility of a bare vacancy. As already noted by Montambaux *et al.* [5], a magneto-elastic coupling to the lattice would make the polarons even heavier [33], and possibly make them self-trapped.

This seems a plausible argument for the explanation of the anomalies mentioned, and it is consistent with a number of observations, discussed by Montambaux [5], that their occurrence depends on the way the sample is prepared. We propose to test this scenario directly as follows. If, in a certain sample, a susceptibility increase (or change in the specific heat) is found, a large magnetic field should be applied for some time, while the temperature is kept fixed. If the sample becomes almost completely polarized, the reason why the excess of vacancies was trapped is taken away, because it is a *magnetic* trap. After sufficient time to reach the equilibrium number of vacancies, the field should be switched off. Then the original measurement should be repeated after the sample is depolarized again. If the previously measured effect has diminished, this would indicate that the proposed explanation is correct.

If one does not find a reduction, the assumption that *metastable* trapped vacancies are the cause of the effects would be proven incorrect. As discussed in the previous subsection, in that case one should investigate the nonlinear susceptibility.

How strong does the field have to be for such a magnetic annealing, and how long does it need to be applied? It is difficult to give a precise answer to this question, but the following estimates indicate that annealing in a field of several Tesla at a few mK should suffice:

1. In the Brinkman-Rice picture, the effective hopping amplitude $t_{eff}$ of a polaron is equal to the bare vacancy hopping amplitude $t$ times the overlap of the two spin states which differ by the shift of the polaron by one lattice unit. If the background is paramagnetic, the overlap of these two states is of order $\left(\frac{1}{2}\right)^{N_s}$, where $N_s$ is the number of spins on the surface of the polaron. Hence the effective hopping amplitude $t_{eff} \approx \left(\frac{1}{2}\right)^{N_s} t$ is indeed extremely small in the absence of a field, if the polaron is not too small ($N_s \gtrsim 20$).

   If the background has a polarization $\Delta$ in the presence of a field, the effective hopping becomes $t_{eff} \approx \left(\frac{1+\Delta}{2}\right)^{N_s} t$. This shows that if the polarization becomes appreciable (60 % or more), the hopping amplitude and hence the mobility of the polarons becomes large enough to leave the crystal. Note that a polarization of about 60 % can be obtained in a field of about 7 Tesla at 6 mK [34]. With fields of the order of 10 Tesla, one can get a $\Delta$ of order 90 % at about 4 mK.

2. In addition to the above polarization dependence of the effective single polaron hopping amplitude, the polaron mobility will also be increased by a field due to two collective effects. First of all, the increase of polarization will tend to increase the attractive force between polarons [5]; this effect, which Montambaux *et al.* estimate will become significant at fields of a few Tesla in the mK range, will enhance clustering of polarons. Secondly, in fields of a few Tesla, the size of a polaron will increase, since the entropy difference between the ferromagnetic spin state in the polaron and the polarized state outside it, decreases. For example, in fields of the order of 7 Tesla at 4 mK, the entropy difference per spin is reduced by a factor of two [34], and this increases the number of spins in the polaron by some 50 %. Depending on the original concentration of vacancies and the field, such an increase may lead to the breakdown of the independent polaron picture. For example, if the vacancy concentration is of order $10^{-3}$, and if each vacancy gives rise to a zero field polaron of some 50 spins, the total fraction of spins in polarons in the above field increases to 6 %; at even somewhat higher values, such an estimate would lead to a percolating polaron network. Clearly, in this regime, the independent polaron picture is not justified.

3. It is useful to consider the other limit also. In a very strong field, where the polarization is almost 100 %, the concept of a spin polaron has no meaning.

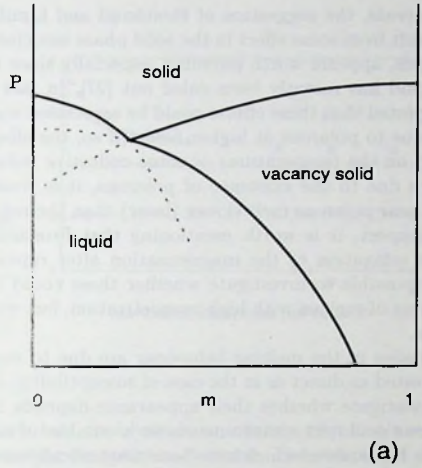
At low temperatures, vacancies then move coherently with a mean free path $l$ which will be much larger than the lattice spacing $a$: $\frac{l}{a} \gg 1$. For their diffusion coefficient $D_\nu$ we then have $D_\nu = \left(\frac{l}{\hbar} \cdot a\right) l = \frac{l}{a}\left(\frac{l}{\hbar} \cdot a^2\right) \approx 4 \cdot 10^{-7} \cdot \frac{l}{a} \frac{cm^2}{s}$. As a result, for a sample size of the order of centimeters, a reasonable fraction of the vacancies can diffuse to the walls in an hour if $\frac{l}{a} > 10$.

Based on these observations, we believe that annealing in fields of several Tesla for a good fraction of an hour should be sufficient to test the idea of *magnetic annealing*.

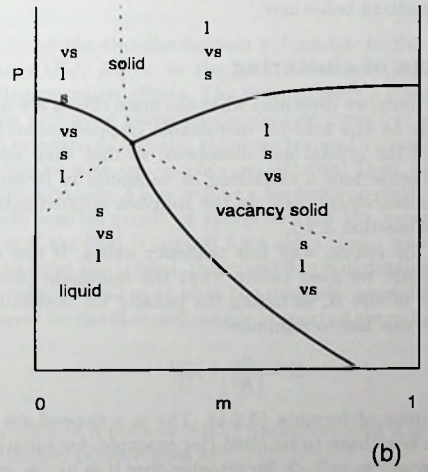
Unfortunately, other sensitive experimental probes of the formation of polarons besides measurements of the specific heat and susceptibility (either with NMR or with a SQUID) do not seem to exist. Spin polarons in magnetic semiconductors [35] can be probed with magneto-electric effects, but we lack the coupling to charge in $^3He$. Likewise, magneto-optical effects will be much weaker in $^3He$ than in materials like $EuTe$ [36, 37] because we deal with nuclear spins instead of electron spins. A different possibility to test for the existence of spin polarons in $^3He$ are the resonances predicted by Meĭerovich [38], which are due to the curvature of the wave function of the vacancy at the origin, when it is localized in a polaron. Such experiments appear to be difficult, however, and, to our knowledge, they have not been performed.

### 3.2.3 Melting behaviour, spin polarons and the vacancy solid

As an explanation of rather sharp breaks in the dependence of the melting pressure on the polarization during rapid melting of $^3He$ [27], Bouchaud and Lhuillier [28] assumed the existence of a new phase: the magnetic vacancy solid. The possibility of having zero point vacancies in $He$ crystals near the minimum of the melting curve where the formation energy of vacancies appears to go to zero [39] has often been speculated about [40], especially for $^4He$ [41]. Bouchaud and Lhuillier build on this and the idea of spin polarons by speculating that for nonzero magnetization $m$ of the solid there might be a stable vacancy solid $^3He$ phase. The phase diagram they propose is sketched in Fig.(3-1 a). There are two reasons why we consider this proposal not tenable. First of all, the phase diagram is thermodynamically inconsistent. As pointed out to us by Nozières [42], if one tries to keep track of the relative stability of the three phases involved in Fig.(3-1), one finds out that it is not possible to get the order correct everywhere. The lines cross in an inconsistent way: between two lines that separate two phases, there should be the extrapolation of a line that separates a different set of phases, so that upon circling the triple point, one crosses a dashed line each time after a full line is crossed. A thermodynamically consistent way to draw the phase diagram is sketched in Fig.(3-1 b), but this phase diagram would not explain the breaks in the data for the melting pressure as a function of polarization. Our second reason to question the proposed explanation of these data is that the measurements are performed relatively far away from the minimum of the melting curve, where the formation energy of vacancies tends to zero, and so where zero point vacancies are most likely to be found, if they do exist.




(a)




(b)

Figure 3-1. a) The phase diagram as proposed by Bouchaud and Lhuillier. b) Thermodynamically consistent phase diagram; in each part of the diagram, the relative order of the energies of the phases is indicated, the lowest position corresponding to the lowest energy. In a), a consistent labeling is impossible.

In spite of these caveats, the suggestion of Bouchaud and Lhuillier [28] that the observed anomalies result from some effect in the solid phase associated with polarons near (trapped) vacancies, appears worth pursuing, especially since the possibility of an anomaly in the liquid has recently been ruled out [27]. In fact Montambaux *et al.* [5] had already suggested that these effects could be associated with the saturation of the magnetization due to polarons at higher fields (if so, the effect would have to be strongly dependent on the temperature) or some collective polaron effect. If the solid is inhomogeneous due to the existence of polarons, it is conceivable that the ferromagnetic regions near polarons melt slower (later) than the regions with a lower polarization. In this respect, it is worth mentioning that Frossati *et al.* [34] have observed kinks in the relaxation of the magnetization after rapid melting of $^3He$ crystals. It should be possible to investigate whether these could be interpreted in terms of the slow melting of regions with high magnetization, but we shall not pursue that investigation here.

Whether the anomalies in the melting behaviour are due to *metastable* trapped vacancies can not be tested as direct as in the case of susceptibility. It would be interesting, however, to investigate whether their appearance depends on sample preparation. For example, one could do measurements on a number of samples which are grown in the same way, but some of which have been magnetically annealed. In particular, if phase separation of vacancies occurs (as has been predicted theoretically [18]), this might affect the melting behaviour.

### 3.2.4 The effects of clustering

In the previous subsections, we described what the main effects are when the mobility is increased sufficiently by the field for metastable trapped vacancies to be able to move to the surface of the crystal and disappear, so that their concentration goes down. We will now discuss how a clustering of vacancies in larger polarons would affect the proposed measurements [43], in the simplest approximation based on the picture discussed in subsection 3.2.2.

The first issue is the reason why this tendency exists. If one calculates in the simple continuum picture we used before what the minimum (free) energy is of $i$ vacancies in a polaron of size $R$, including the penalty for change in spin energy or entropy in this region, one has to minimize

$$E_i = \frac{a_i}{R^2} + bR^d, \tag{3.2.4}$$

which is a generalization of formula (3.2.1). The $a_i$'s depend on which levels (of spinless fermions in a box) have to be filled (for example, for square polarons in two dimensions, we obtain $a_2 : a_1 = 7 : 2$, for circular ones it is $a_2 : a_1 = (2.41^2 + 3.83^2) : (2.41^2)$). One finds that the total energy of more vacancies in a polaron is always lower than the total energy of separate polarons with fewer vacancies each. The fact that the background has to be broken down only once, is more important than the need to fill higher levels. (As an illustration we show in Fig.(3-2) how the energy *per*

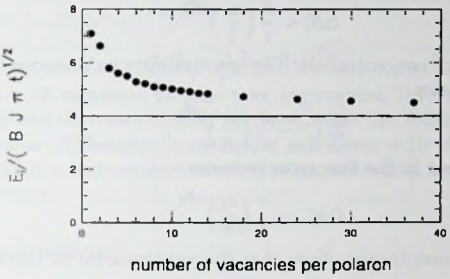



Figure 3.2. Energy per hole, for $i$ holes in an infinitely deep square well in two dimensions, in the $t$-$J$ model, illustrating the tendency to cluster. The prefactor of $\sqrt{BJ\pi t}$ tends to four for large $i$, which is the value in the case of phase separation (see Emery *et al.* [18]).

*hole* depends on $i$, for the two-dimensional $t$-$J$ model. In this case the hole energy is proportional to $\sqrt{tBJ}$, $BJ$ is, in the notation of ref. [18], the magnetic energy per bond, including quantum effects. The limit of large $i$ gives the same value as is obtained using the picture of phase separation (see Emery *et al.* [18]).

Suppose in an experiment initially, there are trapped vacancies, mainly in separate ferromagnetic polarons, with one vacancy per polaron. Then, by an applied magnetic field, their mobility increases. There will be a tendency to cluster, because that lowers the energy. If there was no excess of vacancies (but the vacancies were trapped in their positions), or if the field is applied for a short time , which gives the vacancies insufficient time to leave the crystal, this clustering will dominate. If this is the case, fewer but bigger polarons will exist after the magnetic annealing.

The consequences for the thermodynamic properties are as follows. From Eq.(3.2.4) we find

$$R_i \sim \left(\frac{a_i}{b}\right)^{\frac{1}{d+2}}, \tag{3.2.5}$$

where $R_i$ is the radius of a polaron with $i$ vacancies, and $d$ is the dimension. Then the number of spins per polaron goes as

$$N_i \sim \left(\frac{a_i}{b}\right)^{\frac{d}{d+2}}, \tag{3.2.6}$$

and thus we can see what happens to the thermodynamic properties. For the entropy

loss due to polarons, we find

$$\Delta S_i \sim \frac{x}{i}\left(\frac{a_i}{b}\right)^{\frac{d}{d+2}}, \tag{3.2.7}$$

where $x$ is the vacancy concentration. The susceptibility enhancement in the second term of Eq.(3.2.2) is now

$$\Delta\chi_i^+ \sim \frac{x}{i}\left(\frac{a_i}{b}\right)^{\frac{2d}{d+2}}, \tag{3.2.8}$$

while the diminishment in the first term becomes

$$\Delta\chi_i^- \sim \frac{x}{i}\left(\frac{a_i}{b}\right)^{\frac{d}{d+2}}, \tag{3.2.9}$$

just as the entropy contribution. Note that the enhancement of the first term $\Delta\chi^+$ always dominates.

First we consider what happens if the polarons remain small, that is, if we go from single polarons to bipolarons. In this case, it is important that the second level in the 'well' is much higher than the first, and that makes the polarons so much bigger that the effect in $\Delta\chi^+$, which is proportional to $N^2$, increases, despite the diminishment of the *number* of polarons. The effects proportional to $N$ become a little bit smaller. For square polarons in two dimensions we obtain $\Delta\chi_2^+ = \frac{1}{2}\cdot\frac{7}{2}\Delta\chi_1^+ = 1.75\cdot\Delta\chi_1^+$, and $\Delta S_2 = \frac{1}{2}\cdot\left(\frac{7}{2}\right)^{\frac{1}{2}}\Delta S_1 = 0.94\cdot\Delta S_1$, the latter factor also applies to $\Delta\chi^-$. The numbers for circular polarons are approximately the same. In three dimensions we obtain $\Delta\chi_2^+ = \frac{1}{2}\cdot\left(\frac{9}{3}\right)^{\frac{6}{5}}\Delta\chi_1^+ = 1.87\cdot\Delta\chi_1^+$, and $\Delta S_2 = \frac{1}{2}\cdot\left(\frac{9}{3}\right)^{\frac{3}{5}}\Delta S_1 = 0.97\cdot\Delta S_1$. The precise numbers depend on the shape of the polarons, but the trend is that going from single to bipolarons gives an enhancement of $\Delta\chi^+$, while $\Delta\chi^-$ and $\Delta S$ do not change much.

The effect on the resonances in high frequency NMR measurements, predicted by Meĭerovich [38] is qualitatively as follows. The resonances arise because a single vacancy in a polaron is situated primarily at the centre, which causes the effective exchange interaction to be inhomogeneous. The frequencies depend on how the wave function is curved at the origin. If there are more vacancies in a polaron, the first effect is that the polaron becomes bigger, which makes the curvature smaller and the resonance frequencies lower. Furthermore, the second vacancy occupies a state with a wave function that is not centered at the origin. The exchange will be much more homogeneous over the polaron, which makes the frequencies lower. It is likely that, because polarons of different shapes and with different amounts of vacancies per polaron form, no clear resonances will be observed at all, if clustering occurs.

Finally, we discuss what happens if very many vacancies go into big ferromagnetic regions. The filling of higher levels has a different effect than in the case of bipolarons, because the number of levels with a certain energy proliferates for higher energies. We can use the picture of a Fermi pressure for $a_i$: $a_i \sim \epsilon_F \sim (i)^{\frac{2}{d}}$. So we obtain for $i$ large

$$N_i \sim (i)^{\frac{2}{d+2}}. \tag{3.2.10}$$

Thus

$$\Delta\chi_i^+ \sim (i)^{\frac{d-1}{d+2}-1}, \tag{3.2.11}$$

and

$$\Delta\chi_i^- \sim \Delta S_i \sim (i)^{\frac{2}{d+2}-1}. \tag{3.2.12}$$

So one sees that, if vacancies go into very big regions, $\Delta\chi^+$ remains the same in two dimensions and decreases slightly in three, while the other two effects decrease in both dimensions. Although the particular exponents with which these quantities scale may change in a more sophisticated approximation [20], we expect the tendency to cluster to remain.

### 3.2.5 Conclusions

In this section, we hope to have convinced the reader that it should be possible to convince or disprove the existence of spin polarons in $^3He$ crystals with existing experimental techniques: systematic experiments, in particular SQUID or NMR susceptibility measurements (or measurements of the specific heat), together with the proposed magnetic annealing, should be able to show whether polarons exist, and if so, whether they are associated with metastable trapped vacancies. If the magnetic annealing does not give rise to a change in behaviour, one should study the nonlinear susceptibility. Unravelling this issue may also help understand the origin of some of the anomalies seen upon melting of polarized $^3He$.

## 3.3 Theoretical improvements

Clearly, the theoretical picture used above has many limitations. Before embarking on more fundamental issues, let us make one remark within the context of the polaron picture. One of the assumptions listed in section 3.1 was that the shape of a polaron is spherical. We discussed the issue of clustering using square (cubic in 3d) polarons. The geometrical aspect is a matter of prefactors, not of different scaling laws. To obtain a low energy for small numbers of holes in a polaron there is a new possibility: elongated polarons. The argument is as follows.

The first step in calculating the size and energy of a polaron is to solve the Schrödinger equation in an infinitely deep potential well of the shape one wishes to consider. E.g., this involves sines and cosines for rectangular polarons and Bessel functions for spherical polarons. After that, one fills the one particle levels up to the number of holes per polaron one wants to consider. This gives an energy that depends on the size of the polaron. Then, one calculates the amount of magnetic energy lost, which is proportional to the volume of the polaron. The sum of the two energies is minimized.

For two holes per polaron, the reason why a rectangular polaron yields a lower energy than a square one, is that the second single particle level needs to have a node in the wave function in one direction only. The kinetic energy of the two holes strongly

depends on the size of the polaron in that direction, but the size in the direction in which the single particle wave functions do not have nodes except at the boundary has less influence. There is less need to destroy the spin background to gain delocalization energy in that direction. Optimizing the two lengths of a two dimensional rectangular polaron yields an energy for two holes per polaron of $6.30\sqrt{BJ\pi t}$. Like in Fig.(3-2), we compare prefactors of $\sqrt{BJ\pi t}$. The aspect ratio for which this minimum is reached, is 1.6.

A further observation is that one easily sees that a circular polaron is more favourable than a square one. Relatively little background is broken down for a certain gain in delocalization energy. Together, this suggests that an elliptic polaron would be very favourable for two holes per polaron. Solving the Schrödinger equation with as boundary condition that the wave function should go to zero on an ellipse, is done best in elliptic coordinates. For one coordinate, the solution involves Mathieu functions. Using that information and balancing against the loss of magnetic energy over an ellipse, we found numerically that the minimum energy is $6.2\sqrt{\pi t}$ for an aspect ratio of 1.3. This is indeed lower than the value for two holes on a circular polaron, which is $6.39\sqrt{BJ\pi t}$. To summarize, the prefactors for two holes per polaron are 6.63 for a square, 6.39 for a circle, 6.30 for a rectangle and 6.20 for an ellipse.

We now turn to structures other than spin polarons. A concept related to the spin polaron is the spin bag [45]. In a spin bag, a hole is localized in a region with reduced antiferromagnetic order, rather than with ferromagnetic order. For intermediate coupling, there is more numerical evidence for the existence of spin bags than for spin polarons [45].

We now come back to the assumptions in polaron theory mentioned in section 3.1, other than than assumptions about geometry. The remainder of this section summarizes results from the literature that shed light on the issue of the possible existence of spin polarons. We shall not go into couplings to the lattice, we remain in the realm of correlated fermions. All other assumptions, the independent hole approximation, the continuum picture and the assumption of full magnetization would be relaxed if one could treat the full quantum many-body problem on a lattice. Here, we indicate the trends that arise when the assumptions are relaxed. To begin with, most of the further discussion will be about the lattice Hamiltonians, not the continuum approximation. No a priori assumptions about the magnetization are made.

The first complication in a more thorough treatment is quantum mechanics. The Brinkman-Rice argument presented in section 3.1 is in terms of classical spin backgrounds. The fact that hole motion destroys spin backgrounds other than a ferromagnetic one is less conclusive with quantum fluctuations added. These can also restore the spin background. Hole motion is inhibited to a lesser extent in a quantum antiferromagnet than in a classical one. For the polaron picture, moreover, it is of importance that the Néel background that is broken down, is not fully magnetized to begin with, due to quantum fluctuations. All this indicates that the polaron is less stable if quantum fluctuations are added.

As far as clustering is concerned, $J$ was assumed to be small. The tendency to

phase separate also exists for large $J$. There is no polaron picture in that limit, but the presence of holes simply breaks antiferromagnetic bonds. More holes next to each other break less bonds than separate holes. While in the two limits phase separation indeed occurs, the situation is less clear for intermediate $J$. A further complication is that a long-range Coulomb interaction frustrates phase separation. Tendencies towards charge density waves as well as towards superconductivity exist (for a discussion of these effects, see e.g. ref [44]). Taking a finite $J$ implies that notions like Nagaoka's theorem, or the result of Brinkman and Rice, cannot be used, and that the picture of polarons with sharp walls is not justified. In other words, a finite $J$ and quantum fluctuations nessecitate the treatment of the full problem.

A fruitful approach to the problem is to do Hartree-Fock calculations. This is on a lattice and does involve quantum mechanics, but not all fluctuations are included. Decoupling in real-space and performing self-consistent calculations yields various inhomogeneous structures, other than polarons, depending on the coupling strength and on the doping. A *domain wall* is a structure in which charge is accumulated in one-dimensional structures, which separate domains without charge and antiferromagnetic magnetization with opposite orientation on both sides.

Domain walls were found as mean-field solutions in the three-band Hubbard model by Zaanen *et al.* [46]. Further work includes continuum approximations by Schulz [47]. Inui and Littlewood [48] showed that in the two-dimensional one-band Hubbard model, one obtains vertical domain walls for small $U/t$. For $U/t > 3.6$, one obtains diagonal walls as structures with lowest energy, polarons are found for $U/t > 8$.

The structures we discuss, spin polarons as well as domain walls, show similarities to the Su Schrieffer Heeger model [49], in which doped holes distort the lattice background. In chapter 5, we shall encounter another example of a competition which leads lo a localized breakdown of a background, namely the spin soliton in the Kondo lattice model. A generic feature is that the undoped state has a gap for single-particle excitations, while the doped system exhibits midgap states (see chapter 5).

Hartree-Fock tends to ascribe too much stability to inhomogeneous structures. One reason is that it overestimates the Néel order in the antiferromagnetic domains. The fluctuations left out in a Hartree-Fock calculation would reduce this magnetization. While no strict assumptions about the degree of magnetization are made, as was done in the description of section 3.2, the difference in kinetic energy in domains and in non-antiferromagnetic regions is still overestimated.

A different approach in which quantum effects are not fully treated, is to consider the $t - J_z$ model. The spins are Ising-like. Without further mean-field assumptions, quantum fluctuations are reduced. This model is more tractable numerically than the full problem. Spin polarons are found [50], as well as domain walls [51]. Again, one does not know whether these exist in the full quantum problem.

One way to improve on this situation is to perform Monte Carlo calculations, which in principle yield results that are exact up to stochastic accuracy. Such a calculation would treat the full many-body problem and thus none of the assumptions

used in the polaron picture would be maintained as a requirement of a possible solution.

In the following chapter, we shall discuss that the sign problem hampers the possibilities. The so-called fixed-node approximation, however, does not suffer from the sign problem. More fluctuations than in Hartree-Fock are included, but not all possible ones. Fixed-node Monte Carlo calculations would be a test for the stability against quantum fluctuations of inhomogeneous Hartree-Fock solutions.

As said before, to study polarons and domain walls, one certainly would like to start from a lattice description which is certainly much better than the simple continuum picture used before for polarons. This gives additional complications in fixed-node Monte Carlo. Our efforts in the next chapters are aimed at applying the fixed-node idea to lattices. We shall present results for an inhomogeneous solution in the Kondo Lattice model. As far as spin polarons and other inhomogeneous structures in the Hubbard model are concerned, we shall not present fixed-node calculations here, we just mention that these would be an interesting follow-up of Hartree-Fock calculations.

The main conclusion of this chapter is that a direct experimental test for the presence of spin polarons is possible. A magnetic annealing procedure would clarify the tenability of explanations of other thermodynamic measurements in terms of trapped vacancies in heavy spin polarons. An important step towards lifting the limitations would be achieved if full quantum mechanical calculations on lattice many-fermion problems would be possible.

## References

[1] A.H. Castro Neto, cond-mat preprint 9410005

[2] H.J.M. van Bemmel and W. van Saarloos, J. Low Temp. Phys. **94** , 551 (1994)

[3] Y. Nagaoka, Solid State Commun. **3**, 409 (1965); Phys. Rev. **147**, 392 (1966)

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

[5] G. Montambaux, M. Héritier, and P. Lederer, J. Low Temp. Phys. **47**, 39 (1982)

[6] A.F. Andreev, JETP Lett. **24**, 564 (1976)

[7] M. Héritier and P. Lederer, J. Phys. (Paris) Lett. **38**, L-209 (1977)

[8] See e.g. M. Roger, J.H. Hetherington and J.M. Delrieu, Rev. Mod. Phys. **55**, 1 (1983) ; and M.C. Cross and D.S. Fisher, Rev. Mod. Phys. **57**, 881 (1985)

[9] The need for a systematic study is particularly important in view of the fact that it is difficult to compare different experiments directly , as the vacancy activation energy, the hopping amplitude $t$ and effective spin interaction all depend very sensitively on the molar volume.

[10] W.P. Kirk, Z. Olejniczak, P.S. Kobelia, and A.A.V. Gibson, in Proceedings of the Seventeenth International Conference on Low Temperature Physics, edited by U. Eckern *et al.* (North-Holland Amsterdam, 1984), Vol. BE1, p.273.

[11] M.E.R. Bernier and E. Suaudeau, Phys. Rev. B **38**, 784 (1988)

[12] M.E.R. Bernier, M. Bassou, M. Chapellier and M. Rotter, Phys. Rev. B **40**, 8700 (1989)

[13] P. Kumar and N.S. Sullivan, Phys. Rev. Lett. **55**, 963 (1985); see also C. Jedrzejek and W.P. Kirk, Phys. Rev. Lett. **57**, 2599 (1986), and P. Kumar and N.S. Sullivan, Phys. Rev. Lett. **57**, 2600 (1986)

[14] G.A. [illegible] and J.M. Goodkind, J. Low Temp. Phys. **79**, 251 (1990)

[15] J. [illegible] and O. Gunnarsson, Phys. Rev. B **40**, 7391 (1989)

[16] See e.g. A.R. Bishop, F. Guinea, P.S. Lomdahl, E. Louis and J.A. Vergés, Europhys. Lett. **14**, 157 (1991), and references therein.

[17] M.V. Feigel'man, JETP Lett. **27**, 462 (1978)

[18] V.J. Emery, S.A. Kivelson, and H.Q. Lin, Phys. Rev. Lett. **64**, 475 (1990); see also various articles in E. Sigmund and K.A. Müller, eds, *Phase Separation in Cuprate Superconductors*, Springer-Verlag, Berlin (1994)

[19] Typical values are $U/t \simeq 8-10$ for the high $T_c$ materials (see e.g. P. Fulde and P. Horsch, Europhysics News **24**, 73 (1993) for an introductory discussion), which is a factor of 5 smaller than one estimates for $^3He$ [5].

[20] For a review, see Yu Lu, Su Zhao-Bu and Li Yan-Bin, Chin. J. Phys. **31**, 579 (1993)

[21] As discussed later in this chapter, Montambaux *et al.* [5] may have *overestimated* the number of vacancies needed to explain anomalies in the susceptibility and specific heat measurements, since it is likely that they have *underestimated* the vacancy hopping amplitude somewhat.

[22] M. Bernier and J.M. Delrieu, Phys. Lett. **60A**, 156 (1976)

[23] T.C. Prewitt and J.M. Goodkind, Phys. Rev. Lett. **39**, 1283 (1977)

[24] W. P. Halperin, Ph-D thesis, Cornell University (1975)

[25] J. M. Dundon and J. M. Goodkind, Phys. Rev. Lett. **32**, 1343 (1974)

[26] B. Castaing and P. Nozières, J. Phys. (Paris) **40**, 257 (1979)

[27] For a review, see G. Bonfait, L. Puech and A. Schuhl, in *Helium Three*, Eds. W. P. Halperin and L. P. Pitaevskii (North Holland, Amsterdam, 1990), p. 881. Note that the recent experimental results by S. A. J. Wiegers, P. E. Wolf and L. Puech, Phys. Rev. Lett. **66**, 2895 (1991) and **68**, 3817 (1992), virtually rule out the possibility that these observed anomalies result from a change in behavior in the equilibrium properties of the liquid phase.

[28] J.P. Bouchaud and C. Lhuillier, Europhys. Lett. **3**, 481 (1987)

[29] E.R. Dobbs, *Solid Helium Three*, Clarendon Press, Oxford (1994)

[30] Montambaux *et al.* [5] assume the hopping amplitude $t$ to depend very weakly on the molar volume. From the molar volume dependence of $T_N$ a strong dependence $J \sim V^{20}$ is well documented. We see no reason not to expect $t$ to have a similarly strong molar volume dependence.

[31] For the experiments in K.O. Keshishev, Sov. Phys. JETP **45**, 273 (1977) and in K.O. Keshishev and A.É. Meĭerovich, Sov. Phys. JETP **45**, 1027 (1977), on $^4He$, one estimates a total bandwidth of order 4 K and hence a hopping amplitude $t$ of a few hundred mK. The value of $t$ for $^3He$ would typically be expected to be larger than for $^4He$.

[32] S.V. Iordanskiĭ, JETP Lett. **26**, 171 (1978)

[33] Note also that heavy objects are more susceptible to localization by disorder. See N.F. Mott, *Metal-Insulator Transitions* Taylor & Francis, London (1974)

[34] See e.g. G. Frossati, Japanese J. Appl. Phys. **26**, supplement 1833 (1987)

[35] For a review, see e.g. E.L. Nagaev, J. Magn. Magn. Mat. **110**, 39 (1992)

[36] É.L. Nagaev and É.B. Sokolova, Sov. Phys. Solid State **21**, 767 (1979)

[37] J. Vitins and P. Wachter, Solid State Commun. **13**, 1273 (1973)

[38] A.É. Meĭerovich, JETP Lett. **25**, 455 (1977)

[39] See e.g. M. Chapellier, M. Bassou, J.M. Delrieu, M. Devoret and M.S. Sullivan, J. Low Temp. Phys. **59**, 45 (1985)

[40] A.F. Andreev, in Progress in Low Temperature Physics **8**, ed. D.F. Brewer (North Holland, Amsterdam, 1982). The possible consequences of zero point vacancies in $^4He$ crystals were recently also discussed by A.F. Andreev, in Lectures at the University of Leiden 1992 (unpublished)

[41] For a review of the theory and of recent experimental investigations of zero point vacancies in $^4He$, see M.W. Meisel, Physica B **178**, 121 (1992). To our knowledge, all experiments on supersolid $He$ have been confined to $^4He$, the most accurate

being those by P.G. van de Haar, C.M.C.M. van Woerkens, M.W. Meisel, and G. Frossati, J. Low Temp. Phys. **86**, 349 (1992) and E.D. Adams, M.W. Meisel, E. Suaudeau and J.S. Xia, Bull. Am. Phys. Soc. **35**, 1080 (1990)

[42] P. Nozières, private communication. M.E. Fisher has pointed out to us that these constraints on the angles at a triple point are an example of Schreinemaker's rules. See J.C. Wheeler, J. Chem. Phys. **61**, 4474 (1974) for details.

[43] Although it has been mentioned before that vacancies would tend to cluster into larger polarons —see [5] [17] [18] and also the paper by M.C. Cross and D.S. Fisher in [8]— the consequences of this for the susceptibility do not appear to have been worked out before.

[44] V.J. [illegible] and S.A. Kivelson, Physica C **209**, 597 (1993)

[45] J.R. [illegible] X.G. Wen, and S.C. Zhang, Phys. Rev. Lett. **60**, 944 (1988)

[46] J. Zaanen and O. Gunnarsson, Phys. Rev. B **40**, 7391 (1989)

[47] H.J. Schulz, J. Phys. France **50**, 2833 (1989)

[48] M. Inui and P.B. Littlewood, Phys. Rev. B **44**, 4415 (1991)

[49] W.P. Su, J.R. Schrieffer, and A.J. Heeger, Phys. Rev. Lett. **42**, 1698 (1979)

[50] P. Prelovšek and I. Sega, Phys. Rev. B **49**, 14241 (1994)

[51] T. Barnes, E. Dagotto, A. Moreo, and E.S. Swanson, Phys. Rev. B **40**, 10977 (1989)

# 4 Fixed-node Monte Carlo on lattices

The remainder of this thesis is devoted to Monte Carlo methods. The properties of models of correlated fermions are difficult to extract from the Hamiltonians, despite their simple appearance. In the previous chapter, we resorted to approximations to discuss the tendency to form inhomogeneous structures. It is of primary importance to be able to test the effect of the approximations made, and to see which features remain once the restrictions are lifted. In principle, Monte Carlo methods give exact results, with only statistical uncertainties.

However, for fermions the sign problem interferes. Both in finite temperature and ground state algorithms, positive and negative contributions to each measured quantity appear, which tend to cancel each other out [1]. This yields large uncertainties. In this chapter we shall supply an answer to the question we posed in chapter 1: "Which approximations should be made in order to avoid the sign problem in stochastic simulations aiming at the ground state properties of lattice many-fermion systems?".

For continuum models, an existing diffusion Monte Carlo method makes an approximation that circumvents the difficulty of cancellations due to signs, in the context of ground state algorithms [2]. This so-called fixed-node Monte Carlo method for continuum problems will be discussed in subsection 4.2.1. The main purpose of this chapter is to generalize this fixed-node idea to lattice problems. We designed a new prescription, which could be proven to retain the variational nature of the continuum method. This will be presented in subsection 4.2.2.

While this chapter discusses the principle of the method, the next chapter contains an application, and is therefore a test of the performance of the Monte Carlo programs based on the ideas presented here. In the last chapter, we consider an ingredient of lattice fixed-node calculations, the sign structure of wave functions. In this last chapter, we shall take the same approach as in developing the Monte Carlo method, and seek to find lattice equivalences of results known for the continuum case.

## 4.1 A perspective on Monte Carlo methods

While our own contributions are in the field of diffusion Monte Carlo, to be discussed in the next section, this section discusses the sign problem in different Monte Carlo methods.

Monte Carlo methods have a stochastic character. The reason why this is useful for computing integrals is the following. Other methods use a grid of equally spaced points. For a given number of points, the spacing goes as $N^{-1/d}$, and the corresponding inaccuracy as $N^{-2/d}$, where $d$ is the dimension of the space over which one integrates. If the points are chosen stochastically, the central limit theorem tells that the distribution of the sum will be a Gaussian with width proportional to $N^{-1/2}$,

irrespective of $d$. For large $d$, this scaling of the error with $N$ is more favourable in the stochastic case than in the regular grid case.

In the field of correlated fermions, performing integrals over high dimensional spaces is often needed. Calculating the properties of a proposed variational wave function requires summing over configuration space. In $d$ space dimensions and with $N$ particles, the relevant dimension is $dN$. Often, a particular function with one or more variational parameters is chosen in order to approximate the ground state. Minimizing the variational energy as a function of the parameters requires many high-dimensional integrals, one for every set of parameters. If a Monte Carlo method is used for the integrals, the procedure is called variational Monte Carlo [3].

Although our main interest will be in ground state properties, we will also discuss briefly how the sign problem arises in finite temperature Monte Carlo approaches to correlated fermions. Finite temperature properties [3] are derived from the partition sum

$$Z = \mathrm{Tr} e^{-\beta \mathcal{H}}. \tag{4.1.1}$$

The fact that in non-trivial problems the terms in $\mathcal{H}$ do not commute, makes it difficult to compute the exponent. Breaking up $\beta$ in small slices, $\beta = L\Delta\tau$, enables one to approximate the partition sum as the trace of a product, in which the kinetic part and the potential part appear in separate exponents. In the notation of ref. [4] this becomes

$$Z = \mathrm{Tr} e^{-\beta \mathcal{H}} \approx \mathrm{Tr} \prod_{l=1}^{L} e^{-\Delta\tau \mathcal{H}_0} \exp\left[-\Delta\tau \left[U \sum n_{i\uparrow} n_{i\downarrow} - \mu \sum (n_{i\uparrow} + n_{i\downarrow})\right]\right]. \tag{4.1.2}$$

The error which arises from the fact that the kinetic part and the potential part do not commute, is proportional to $\Delta\tau^2 tU$. This decomposition is named after Trotter and Suzuki.

Depending on which complete set of states is used to evaluate the trace, the method is called *Path integral Quantum Monte Carlo* (position basis, the position of each numbered particle is specified on every time slice) or *Checker board Quantum Monte Carlo*, sometimes also called *World-line Quantum Monte Carlo* (occupation number basis) [3].

In a third method, the problematic term with four fermion operators, which appears in an exponent, is eliminated by using a discrete version of the Hubbard-Stratonovich transformation :

$$e^{-\Delta\tau U n_{i\uparrow} n_{i\downarrow}} = \frac{1}{2} \mathrm{Tr}_\sigma exp\left[\lambda\sigma(n_{i\uparrow} - n_{i\downarrow}) - \frac{\Delta\tau U}{2}(n_{i\uparrow} + n_{i\downarrow})\right] \tag{4.1.3}$$

with $\lambda = \mathrm{arccosh}(\exp(\Delta\tau U/2))$. The auxiliary field is a spin, $\sigma = \pm 1$. The trace over the fermion operators can be taken explicitly, because only bilinear forms in those operators appear in the right hand side of Eq.(4.1.3).

The remaining problem consists of performing a trace over the new spin variables. It is of the form

$$Z = \mathrm{Tr}_\sigma \det O_\uparrow \det O_\downarrow, \tag{4.1.4}$$

the determinants are the result of integrating out the fermions.

In the sum over Ising spins, the product over determinants serves as Boltzmann weight. Only at half-filling, the product can be proven to be positive, otherwise the absolute value is taken as weight and one needs to keep track of the average sign. The observation in early work, ref. [4], that the average sign does not go to zero rapidly in their calculation, has borne out not to hold for the temperatures and lattice sizes one needs to consider for the high-$T_c$ problem [5]. The fact that in Eq.(4.1.4) contributions of different sign appear, leads to cancellations. The large statistical uncertainties this gives are the essence of the *sign problem*.

A separate method involving a Hubbard-Stratonovich transformation exists for measuring ground state properties [6]. We shall first concentrate on the *Green Function Monte Carlo*, or *Diffusion Monte Carlo* method for the ground state. This is the method in which nodes of wave functions, in a real-space (or rather configuration space) domain, are relevant. This is the method in which the fixed-node approximation was put forward to avoid the sign problem, a notion we shall extend to lattice models. Later, in section 4.3, we shall compare our diffusion Monte Carlo method with work in which the sign problem is avoided in the Hubbard-Stratonovich ground state algorithm (also called Projector Monte Carlo).

## 4.2 Diffusion Monte Carlo

In diffusion Monte Carlo, the task is to solve a partial differential equation in a high-dimensional space, rather than to perform an integral as in the abovementioned methods. In order to compute ground state properties, one does not take the limit $\beta \rightarrow \infty$ in a calculation involving the partition sum. Rather, one tries to find the lowest-energy solution of the Schrödinger equation. The formal equivalence between the Schrödinger equation and the diffusion equation in imaginary time is taken advantage of.

The idea is to start from a 'trial state', and to apply a projection operator to it, which filters out the ground state. In the continuum case, one often uses an exponential operator, while for lattice problems it is convenient to use an operator that is linear in $\mathcal{H}$, because its matrix elements are easy to calculate. In the absence of an exponential, the machinery of the Trotter-Suzuki formalism of Eq.(4.1.2) is not needed.

The projection operator $\mathcal{F} = 1 - \tau(\mathcal{H} - w)$ is iterated. After $n$ iterations the trial state has evolved to

$$|\psi^n\rangle = \mathcal{F}^n|\psi_T\rangle = [1 - \tau\,(\mathcal{H} - w)]^n\,|\psi_T\rangle. \tag{4.2.1}$$

Choosing $w$ close to the ground state energy, and taking $\tau$ small enough, ensures that eventually, the ground state will dominate. The implementation of the projection operator has a stochastic character.

This projection is implemented in configuration space $\{R\}$, which should not be confused with the original real-space: a point in $\{R\}$ specifies the positions of all

the labeled fermions in real-space. The properties of the system are sampled by an ensemble of random walkers, which initially represents the trial state, and evolves towards the true ground state according to Eq.(4.2.1).

We give a more detailed description of the evolution of the ensemble, and of the way expectation values of operators are measured. For the wave function at point $R$ we write $\psi^n(R) = \langle R|\psi^n\rangle$, and for the Hamiltonian operator $H(R,R') = \langle R|\mathcal{H}|R'\rangle$. To obtain good statistics, *importance sampling* is used. The Green function

$$G(R,R') = \psi_T(R)F(R,R')\psi_T^{-1}(R') \tag{4.2.2}$$

is introduced. The mixed estimate for the expectation value of an operator $\mathcal{O}$ then becomes [7]

$$\langle\psi^n|\mathcal{O}|\psi_T\rangle\langle\psi^n|\psi_T\rangle^{-1}, \tag{4.2.3}$$

where

$$\langle\psi^n|\mathcal{O}|\psi_T\rangle = \sum_{\mathcal{R}} O(R_n)\prod_{i=1}^{n} G(R_i,R_{i-1})\psi_T^2(R_0). \tag{4.2.4}$$

Here $O(R) \equiv \langle R|\mathcal{O}|\psi_T\rangle\psi_T^{-1}(R)$ is the local value of $\mathcal{O}$, and $\mathcal{R} = \{R_0, R_1, \ldots, R_n\}$ is a path in configuration space. The denominator in Eq.(4.2.3) is equal to Eq.(4.2.4) with $\mathcal{O}$ equal to the identity operator. In this way, one samples more in regions with large contributions. The product in Eq.(4.2.4) must be interpreted as a product along the path of a walker, the sum as a sum over different walkers.

The stochastic interpretation of the expression (4.2.4) for expectation values, is as follows. If we split the Green function in Eq.(4.2.2) in a part $P$ which satisfies $\sum_R P(R,R') = 1$ and a remaining part $m$, which we can choose to depend only on $R'$,

$$G(R,R') = P(R,R')m(R'), \tag{4.2.5}$$

we have a random walk interpretation. The stochastic matrix $P$ defines the transition probabilities, and $m$ is the 'multiplicity' of the walkers. The path of a walker is selected stochastically, with the transition probabilities equal to $P(R,R')$, and the product of all $m$'s along the path serving as weight in the measurements.

As discussed by Trivedi and Ceperley [8] and Hetherington [9], in order to obtain a proper statistics, the ensemble of walkers is updated regularly by letting walkers die or multiply according to their accumulated multiplicity. This is a separate stochastic process. So, the process involves steps of random walkers in configuration space, corresponding to the kinetic terms in the Hamiltonian, and branching, corresponding to the local energy of the position the walker is in, which depends on potential terms.

We illustrate this with a concrete example. $H(R,R')$ (and hence $P$) connects points in configuration space. For example, for the Hubbard model with nearest neighbour hopping amplitude $t$, $H(R,R') = -t$ if $R'$ and $R$ differ by moving over one labeled electron by one lattice unit in real-space, and $H(R,R') = UN_D$ if $R' = R$, with $N_D$ the number of doubly-occupied sites. In the continuum case, the kinetic term is not a hopping integral, but involves derivatives with respect to the coordinates of the particles.

### 4.2.1 The fixed-node approximation in the continuum case

Both in the lattice and in the continuum case, the above interpretation works well in practice, without further assumptions, if all Green functions (Eq.(4.2.2)) are positive. If this is the case, all transition probabilities and multiplicity factors in Eq.(4.2.5) are positive. This is not the case in all systems, however. For fermions, many-particle wave functions are required to be antisymmetric, and points in configuration space in which the wave function has either sign necessarily exist. In Eq.(4.2.2), one sees that this leads to the existence of negative Green functions. In chapter 6 we shall discuss that systems without fermions can also show this problem, in the presence of frustration.

If, being in configuration $R'$, configurations $R$ for which $G(R,R')$ is negative exist, it is not possible to find a proper diffusion Monte Carlo process. To keep a random walk interpretation, one could ascribe transition probabilities to all steps which are proportional to $|G(R,R')|$. But then it follows from the fact that one is making a stochastic implementation of the product in Eq.(4.2.4), and from the splitting of the Green function in a transition probability and a multiplicity factor $m$ in Eq.(4.2.5), that a negative $m$ is introduced along the path of a walker, if a step to an $R$ for which $G(R,R')$ is negative is chosen. This is the origin of the sign problem in diffusion Monte Carlo. Both in the numerator and in the denominator of Eq.(4.2.3), contributions of both signs appear.

The basic idea of fixed-node Monte Carlo is to eliminate the possibility to perform steps which have $G(R,R') < 0$. In the continuum case, in the absence of nonlocal interactions, kinetic terms always have the same sign. Moves of a random walker correspond to off-diagonal terms in the projection operator, which we take, for the sake of the argument, to be the same as in the lattice case: $\mathcal{F} = 1-\tau(\mathcal{H}-w)$, although we really consider the continuum case. Since these off-diagonal terms have a prefactor $-\tau$, and since the kinetic terms in $\mathcal{H}$ always have a negative sign, the criterion not to perform a step, which is $G(R,R') < 0$, always corresponds to the simple criterion that $\psi_T(R)$ and $\psi_T(R')$ are of unequal sign. So, in the continuum case, whether steps are allowed or not is determined by the signs of the *trial* wave function. Connected regions in configuration space of one specific sign will be called 'nodal regions'. On a lattice, hopping terms of either sign can occur (see chapter 6), and the criterion $G(R,R') < 0$ for not performing steps is more general than forbidding the trial wave function to change sign in a step. The latter criterion is only equivalent if all off-diagonal terms in $\mathcal{H}$ are negative.

In practice, the calculation proceeds as follows. As formulated in ref. [10], the Schrödinger equation is solved in the separate nodal regions of the trial wave function. While treating a random walker in a specific nodal region, the solution is assumed to be zero in neighbouring regions. Such an assumption is needed, because importance sampling introduces a drift in the continuum diffusion Monte Carlo process. For a finite time step $\tau$, moves of a random walker to a different region occur. If this happens, the random walker is deleted. This implements the assumption that the solution must be zero outside the nodal region in which the random walker moves. This proce-

dure ensures that no random walkers exist that carry a minus sign in the product of Eq.(4.2.4). It should be noted that importance sampling keeps random walkers away from regions in which the wave function is small. In a continuous configuration space, this is always the case close to nodes of the wave function. Especially for small time steps, moves over nodes, which lead to the disappearance of the walker, do not often happen.

In the limit of small time steps, i.e. $\tau \to 0$, steps in a continuous configuration space also become infinitesimally small. In this limit, the above assumption that the wave function is zero outside the nodal region one is working in, effectively means that the wave function should vanish *on the boundary*. In this limit the assumption becomes consistent: the only place the solution should go to zero is the nodal boundary, i.e. the location of the nodes of the trial wave function. This is the criterion while working on both sides of the boundary: solving in one region, considering the other to be the neighbour, and vice versa. By contrast, for a *finite* step, an assumption is made about the solution in a certain nodal region while working in another region, which is not true while working in that region itself. If a certain boundary point $R$ is considered as point *just over* the boundary, $\psi_T(R)$ is assumed to be zero. If one solves in the nodal region of $R$ itself, and $R$ is viewed as a point *just before* the boundary, $\psi_T(R)$ is *not* assumed to be zero. This inconsistency is *fundamental* in the lattice case.

Let us discuss the consequences of leaving steps out of the Monte Carlo process for the solution one finds. In the continuum case, one effectively solves for the ground state of the Hamiltonian in a configuration space with perfectly absorbing walls between nodal regions of the trial wave function. The solution must have the same nodes as the trial wave function. From this formulation, it immediately follows that the method obeys a variational principle. The true Hamiltonian is used, and its lowest energy state on a restricted Hilbert space is calculated.

The main points in the above are that in the continuum case, the fact that nodes are fixed at specified positions, corresponds to having random walks in a configuration space with *absorbing* boundaries, and that the assumptions made are consistent in the limit of small time steps, only because this implies that steps in a continuous configuration space are also small.

### 4.2.2 The fixed-node approximation in the lattice case

In this subsection, we present our proposed method in the lattice case, and its interpretation. First, we illustrate why simple alternatives are not reliable.

In a lattice problem, not only real-space but also configuration space consists of discrete points. The 'node' of a wave function can be in between two points of configuration space, i.e. the wave function can be positive on one point and negative on a neighbouring point. Taking smaller time steps does not lead to smaller steps in configuration space. In many cases, the only possible moves in configuration space correspond to letting one numbered particle hop to a neighbouring site in real-space,

neighbouring meaning that the sites are connected by a hopping term in the Hamiltonian.

While it is clear which steps should not be performed in the lattice diffusion Monte Carlo process in order to avoid the sign problem, these have $G(R, R') < 0$, it is not immediately clear how the fixed-node idea should be implemented. The simplest implementation is the direct equivalence of the continuum procedure: eliminating these steps, without compensating factors [11]. This procedure is not variational. We illustrate this by trying two different interpretations.

In the importance sampling procedure, having zero probability to jump to a configuration of different sign could be accomplished by assuming the trial wave function to be *zero* outside the nodal region in which one is working. However, since one assumes, in this way, $\psi_T(R)$ to have different values depending on the nodal region in which one is working, this interpretation boils down to taking a trial wave function which is not single-valued, and thus inconsistent. In the continuum case, the same problem occurs, but, as we have discussed, in that case the assumption becomes consistent in the limit of small time steps, unlike the lattice case.

The second way to interpret the avoidance of steps outside the nodal region, is to work with a *truncated Hamiltonian*. This is the interpretation we shall adopt from now onward. Steps that lead to the sign problem, i.e. steps with $G(R, R') < 0$, are left out, without further changes in the Hamiltonian [11]. The difficulty with this interpretation is, that one samples a different Hamiltonian than the original one. Moreover, it is clear that positive contributions to the energy are left out, because they are, in the importance sampling process, of the form $-t\ \psi_T(R)/\psi_T(R')$, with $R$ and $R'$ in different nodal regions.

The truncated Hamiltonian has an interpretation in terms of boundary conditions. It is the direct equivalence of the continuum interpretation. The following simple example illustrates what will happen if the truncated Hamiltonian is used. Consider a Hamiltonian with only nearest neighbour hopping in one dimension. Assume that the trial wave function has a node in between sites 0 and 1, and consider the Schrödinger equation that determines the value of the wave function at site 1, using the *true* Hamiltonian:

$$-t(\psi_0 + \psi_2) = E\psi_1. \tag{4.2.6}$$

In order to have a diffusion interpretation, we write

$$-t(\psi_0 + \psi_2 - 2\psi_1) = (E + 2t)\psi_1. \tag{4.2.7}$$

Truncating the Hamiltonian corresponds to assuming $\psi_0$ to be zero when solving for $\psi_1$. The absorbing wall in configuration space is at the point just outside the nodal region of $\psi_1$. When solving for $\psi_0$, on the other hand, the absorbing wall is at site 1. Fig.(4-1) illustrates the consequences of solving the effective Hamiltonian. The dotted lines indicate the trend to go to zero on sites just outside the region.

Taking into account that the bond between sites 0 and 1 has been cut in the truncated Hamiltonian, one sees that for the bonds that are left, the wave function that is obtained seems smoother than the true wave function, and thus the contribution

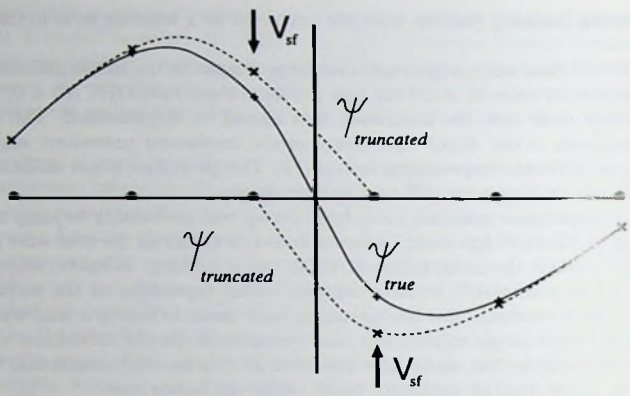



Figure 4-1. Illustration of the effect of truncating the Hamiltonian and of the working of the sign flip potential. If the drawn line goes trough the values of the ground state of the true Hamiltonian, the dotted line gives the trend for the ground state of the truncated Hamiltonian. The arrows indicate the effect of the sign flip potential.

to the energy of the truncated Hamiltonian is lower than in the true wave function. Moreover, as discussed before, positive contributions corresponding to the removed steps are omitted in the importance sampling process. So, the energy obtained by calculating $\langle\psi_{\text{trunc}}|\mathcal{H}_{\text{trunc}}|\psi_{\text{trunc}}\rangle$, where $\psi_{\text{trunc}}$ is the ground state wave function of the truncated Hamiltonian, as is effectively done, is certainly lower than the true ground state energy of the full Hamiltonian.

Fig.(4-1) illustrates a few more issues. Firstly, the wave function $\psi_{\text{trunc}}$ tends to be larger at the boundary than it should be, but the effect is not as strong as when reflecting boundary conditions are used, which would lead to wave functions with zero *slope* on both sides. Secondly, in the limit of small lattice spacing, the wave function $\psi_{\text{trunc}}$ belonging to the truncated Hamiltonian will coincide with that belonging to the true Hamiltonian, $\psi_0$. Incorrect boundary conditions only affect the continuum problem in a range in configuration space that goes to zero if the time step $\tau$ goes to zero, while the range is unaltered by changing $\tau$ in the lattice problem, because steps in configuration space are always of the same size.

An and van Leeuwen [11] not only used the truncated Hamiltonian, they also introduced a correction term. In the importance sampling process, a walker is in a certain configuration $R'$, and moves to $R$ are possible in the true Hamiltonian. Some do not lead to sign flips. These are possible moves and give a contribution to the 'local energy' in the importance sampling process, equal to $-t\psi_T(R)/\psi_T(R')$. Other terms *do* lead to sign flips. The idea of the correction term is not to perform these

steps, but to add $-t\psi_T(R)/\psi_T(R')$ to the energy, as a correction. This correction term clearly is positive.

Simply adding these contributions to the energy, as a correction term [11], is a good thing to do intuitively, but still corresponds to *sampling* the wave function belonging to the truncated Hamiltonian, not the true Hamiltonian. A proof that this procedure gives an upper bound for the ground state energy of the true Hamiltonian had not been established, at the time it was used in ref. [11].

We already discussed that the smoothness of the wave function $\psi_{\text{trunc}}$ *away* from the nodes makes the energy $E_{\text{trunc}} = \langle\psi_{\text{trunc}}|\mathcal{H}_{\text{trunc}}|\psi_{\text{trunc}}\rangle$ lower than the energy for which one tries to obtain an upper bound, which is $E = \langle\psi_0|\mathcal{H}|\psi_0\rangle$. With the correction term added, the other reason why $E_{\text{trunc}}$ yields too low an energy, is no longer valid. Terms that had been omitted *are* taken into account. Moreover, we also discussed that the wave function as obtained by sampling $\mathcal{H}_{\text{trunc}}$ is larger than the true wave function near the nodes. So , by using $\mathcal{H}_{\text{trunc}}$ to determine the transition probabilities, one includes more positive corrections than one should. So it is not immediately clear whether the energy one obtains as $E_{\text{trunc}} + \textit{corrections}$ is higher or lower than the true ground state energy. One contribution (of the truncated Hamiltonian) is too low, the other (of the correction term) is too high. Later, after having presented a proof for an upper bound for a better procedure than the above, we shall see that the procedure involving correction terms *is variational indeed*.

We now present our proposal for a variational way to perform fixed-node quantum Monte Carlo calculations for lattice problems. Steps are left out, just as in the proposal to use a truncated Hamiltonian, but they are replaced by potential terms, in such a way that the method becomes variational. The potential terms have the same value as the abovementioned correction term, but the essential difference will be that they will be treated as a potential in the Monte Carlo diffusion process, not just as a correction to the energy.

Steps that are left out satisfy

$$\langle R|H|R'\rangle\psi_T(R)\psi_T(R') > 0. \tag{4.2.8}$$

An effective Hamiltonian is constructed: the off-diagonal terms are

$$\begin{aligned}\langle R|H_{\text{eff}}|R'\rangle &= \langle R|H|R'\rangle \quad (\text{if } \langle R|H|R'\rangle\psi_T(R)\psi_T(R') < 0)\\ &= 0 \qquad\qquad (\text{otherwise}),\end{aligned} \tag{4.2.9}$$

the diagonal terms are

$$\langle R|H_{\text{eff}}|R\rangle = \langle R|H|R\rangle + \langle R|V_{\text{sf}}|R\rangle, \tag{4.2.10}$$

where the 'sign flip' potential that replaces the hops that satisfy Eq.(4.2.8) is given by

$$\langle R|V_{\text{sf}}|R\rangle = \sum_{R'}^{\text{sf}}\langle R|H|R'\rangle\frac{\psi_T(R')}{\psi_T(R)}. \tag{4.2.11}$$

Note that the entire prescription can be applied for continuum problems with non-local interactions as well. These non-local terms arise naturally if one demands that the conduction electrons have wave functions which are orthogonal to the core electron wave functions. Non-local pseudo potentials ensure this orthogonality.

The justification of the above prescription is at three different levels. Only the last argument that we shall give is conclusive. The second is wrong, we discuss it, because it seemed to be an appealing interpretation. We point out the flaw in this argument about 'lever rule boundary conditions', to avoid confusion.

### An intuitive argument

Again, we consider the case of a negative hopping term as an illustration. A forbidden step corresponds to a change of the sign of the trial wave function. This means, in Fig.(4-1), that the solution is assumed to be negative on the right hand side. The hopping term connecting the two sides in the true Hamiltonian will make the wave function smooth. However, the link is cut in the truncated Hamiltonian. The corresponding wave function $\psi_{\text{trunc}}$, is larger in absolute value on both sides of the node. One sees, however, that the bond that is cut, is replaced by a positive contribution to the potential on both sides. This also tends to suppress the wave function close to the node. So, it leads to the same tendency as the corresponding term in the original Hamiltonian.

This effect on the wave function is the fundamental difference with the implementation using correction terms [11], not acting as potentials. In that prescription, the *truncated* Hamiltonian fully determines the sampling, and thus the resulting wave function. Since the correction terms had the same value as the potential terms, the second procedure in An and Van Leeuwen [11] corresponds to sampling $\langle \psi_{\text{trunc}} | \mathcal{H}_{\text{eff}} | \psi_{\text{trunc}} \rangle$, while the new procedure samples $\langle \psi_{\text{eff}} | \mathcal{H}_{\text{eff}} | \psi_{\text{eff}} \rangle$. Up to this point in the discussion, no bounds have been proven to be obeyed by either of these values.

### Solving with 'lever rule' boundary conditions?

The second argument in favour of the prescription we gave for the way to deal with steps that would cross nodes if they would not be left out of the Hamiltonian, is in terms of boundary conditions. The way of reasoning is appealing, and was our own motivation to introduce the effective Hamiltonian[1]. There is a flaw in the argument, however, which necessitates the rigorous proof that the lattice fixed-node method obeys a variational principle, which will be presented later.

Our original argument was as follows. The truncated Hamiltonian corresponds to having absorbing walls *at the points just outside the nodal region*. This was illustrated in Fig.(4-1). These are not the correct places in which the wave function should go to zero.

Our original idea was to define the location of the node in the 'linear interpolation' sense, i.e. if two neighbouring points in configuration space have opposite signs in

[1]This is the reasoning in the letter in which the method was put forward [12].

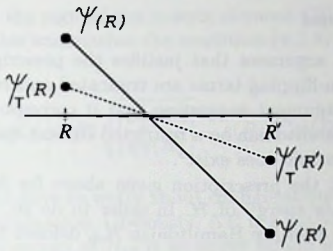



Figure 4-2. Lever rule.

a wave function would identify the location of the node in the way depicted in Fig.(4-2). The idea was to solve the Schrödinger equation *under the restriction* that the location of the node should remain the same, in this 'lever rule' sense. Solving for the ground state of the true Hamiltonian on a restricted Hilbert space would automatically lead to a variational principle: the true ground state energy on the full Hilbert space is lower than or equal to the lowest energy on a restricted space.

The 'lever rule' requirement would be

$$\frac{\psi(R')}{\psi(R)} = \frac{\psi_T(R')}{\psi_T(R)} \ , \tag{4.2.12}$$

only at the boundary. Using this, one would, for a forbidden step from $R'$ to $R$, replace the term $-t\psi(R)$ in the Hamiltonian by $\left[-t \cdot \frac{\psi_T(R)}{\psi_T(R')}\right] \psi(R')$. So the sign flip potential in $\mathcal{H}_{\text{eff}}$ above, would *follow* from the boundary condition of Eq.(4.2.12). Not all pairs of points across a boundary could be given a prescribed ratio, for more interconnected points, some ratios depend on others. The freedom of the solution to differ from the trial solution would eventually depend on the number of 'free points', i.e. points that are away from boundaries in every direction in the high dimensional configuration space. These would embody the freedom to find a solution that is lower in energy than the trial solution.

The flaw in this argument is the following. Using the effective Hamiltonian is not the same as imposing the lever-rule boundary condition. The ground state of $\mathcal{H}_{\text{eff}}$ does not necessarily obey the relation Eq.(4.2.12) we wanted to impose: working out simple examples on small systems for which the ground state of the effective Hamiltonian, obtained from a specific trial wave function, can be calculated exactly, one finds[2]

$$\frac{\psi_{\text{eff}}(R')}{\psi_{\text{eff}}(R)} \neq \frac{\psi_T(R')}{\psi_T(R)}. \tag{4.2.13}$$

Therefore, there is no basis for the argument "the method is variational because it corresponds to solving under a restriction".

[2]This has been presented in ref. [13] and also in the thesis by D.F.B. ten Haaf.

### Proof for an upper bound

Here we present the final argument that justifies the prescription for the effective Hamiltonian, in which sign-flipping terms are truncated and replaced by Hamiltonians. While the previous argument suggesting that it corresponds to solving for the ground state of the true Hamiltonian on a restricted Hilbert space is wrong, a general proof for a variational principle does exist[3].

We want to show that the prescription given above for $\mathcal{H}_{\text{eff}}$ leads to an upper bound for the ground state energy of $\mathcal{H}$. In order to do so, we use the truncated Hamiltonian $\mathcal{H}_{\text{trunc}}$, and a *sign-flip* Hamiltonian $\mathcal{H}_{\text{sf}}$, defined by

$$\mathcal{H} = \mathcal{H}_{\text{trunc}} + \mathcal{H}_{\text{sf}}, \tag{4.2.14}$$

$$\mathcal{H}_{\text{eff}} = \mathcal{H}_{\text{trunc}} + \mathcal{V}_{\text{sf}}. \tag{4.2.15}$$

$\mathcal{V}_{\text{sf}}$ is the sign-flip potential, for which the matrix elements are given by (4.2.11), and $\mathcal{H}_{\text{sf}}$ contains only the off-diagonal elements of $\mathcal{H}$ which are put to zero in the effective Hamiltonian. We now take *any* state

$$|\psi\rangle = \sum_R |R\rangle\psi(R), \tag{4.2.16}$$

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

$$\begin{aligned}\Delta E &= \langle\psi|(\mathcal{H}_{\text{eff}} - \mathcal{H})|\psi\rangle \\ &= \langle\psi|(\mathcal{V}_{\text{sf}} - \mathcal{H}_{\text{sf}})|\psi\rangle.\end{aligned} \tag{4.2.17}$$

$\Delta E$ can be written explicitly in terms of the matrix elements of $V_{\text{sf}}$ and $H_{\text{sf}}$:

$$\Delta E = \sum_R \psi(R)^* \left[\langle R|V_{\text{sf}}|R\rangle\psi(R) - \sum_{R'}\langle R|H_{\text{sf}}|R'\rangle\psi(R')\right]. \tag{4.2.18}$$

We rewrite this expression in terms of the matrix elements of $H$:

$$\Delta E = \sum_R \psi(R)^* \left[\sum_{R'}^{\text{sf}}\langle R|H|R'\rangle\frac{\psi_{\text{T}}(R')}{\psi_{\text{T}}(R)}\psi(R) - \sum_{R'}^{\text{sf}}\langle R|H|R'\rangle\psi(R')\right]. \tag{4.2.19}$$

In this double summation each pair of configurations $R$ and $R'$ occurs twice. We combine these terms and rewrite (4.2.19) as a summation over pairs:

$$\Delta E = \sum_{(R,R')}^{\text{sf}} \langle R|H|R'\rangle \left[|\psi(R)|^2\frac{\psi_{\text{T}}(R')}{\psi_{\text{T}}(R)} + |\psi(R')|^2\frac{\psi_{\text{T}}(R)}{\psi_{\text{T}}(R')} - \psi(R)^*\psi(R') - \psi(R')^*\psi(R)\right]. \tag{4.2.20}$$

[3]This proof has been published by Ten Haaf *et al.* [13].

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 condition (4.2.8) is satisfied, we can finally write $\Delta E$ as

$$\Delta E = \sum_{(R,R')}^{\text{nf}} |\langle R|H|R'\rangle| \left| \psi(R) \sqrt{\left|\frac{\psi_T(R')}{\psi_T(R)}\right|} - s(R,R')\psi(R') \sqrt{\left|\frac{\psi_T(R)}{\psi_T(R')}\right|} \right|^2 . \quad (4.2.21)$$

Note that we do not have to worry about configurations $R$ where $\psi_T(R) = 0$: they do not occur in this summation. Obviously, $\Delta E$ is positive for any wave function $\psi$. Thus the ground state energy of $\mathcal{H}_{\text{eff}}$ is an upper bound for the ground state energy of the original Hamiltonian $\mathcal{H}$.

Now the GFMC method can calculate the exact ground state energy $E_{\text{eff}}$ and wave function of $\mathcal{H}_{\text{eff}}$, without any sign problem. Assuming the trial function $\psi_T$ has the correct symmetry (for example, for fermions one takes a wave function that is antisymmetric), then $\psi_{\text{eff}}$ will carry the same symmetry and hence:

$$E_{\text{eff}} \geq \langle \psi_{\text{eff}}|H|\psi_{\text{eff}}\rangle \geq E_0, \quad (4.2.22)$$

where the second inequality follows from the usual variational principle. Hence the fixed-node energy is an upper bound to the true ground state energy.

One can easily verify that $\mathcal{H}|\psi_T\rangle = \mathcal{H}_{\text{eff}}|\psi_T\rangle$, and thus one can be sure that the GFMC procedure improves on the energy of the trial wave function:

$$E_{\text{eff}} \leq \langle \psi_T|H_{\text{eff}}|\psi_T\rangle = \langle \psi_T|H|\psi_T\rangle. \quad (4.2.23)$$

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 this case, the effective Hamiltonian has the same ground state energy $E_0$, and the same ground state $|\psi_0\rangle$, as this would make it possible to find the true ground state by varying the trial wave function in some way. In Eq.(4.2.21) we substitute $\psi_0$ for $\psi_T$. In order to have $\Delta E$ equal to zero, each individual term in the summation (4.2.21) has to vanish, thus leading to

$$\psi(R) \sqrt{\left|\frac{\psi_0(R')}{\psi_0(R)}\right|} - s(R,R')\psi(R') \sqrt{\left|\frac{\psi_0(R)}{\psi_0(R')}\right|} = 0, \quad (4.2.24)$$

or,

$$\frac{\psi(R)}{\psi(R')} = s(R,R') \left|\frac{\psi_0(R)}{\psi_0(R')}\right| = \frac{\psi_0(R)}{\psi_0(R')} \quad (4.2.25)$$

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 the coefficient positive.

In summary, we introduced an effective Hamiltonian, consisting of a truncated Hamiltonian and a sign flip potential. A proof was established for the relation

$$\langle\psi|H_{\text{eff}}|\psi\rangle \geq \langle\psi|H|\psi\rangle, \tag{4.2.26}$$

for *any* $\psi$.

The best prescription for a Monte Carlo process is to sample $\mathcal{H}_{\text{eff}}$ and to measure its energy. This gives $\langle\psi_{\text{eff}}|H_{\text{eff}}|\psi_{\text{eff}}\rangle$. The proof implies that the old procedure of An and Van Leeuwen [11] involving correction terms, is also variational: it corresponds to measuring $\langle\psi_{\text{trunc}}|H_{\text{eff}}|\psi_{\text{trunc}}\rangle$. Since the proof is valid for any $\psi$, this is higher than the energy of $\mathcal{H}$ in the same wave function. Since $\psi_{\text{eff}}$ is, by definition, the ground state of $\mathcal{H}_{\text{eff}}$, the new bound is always better than the old prescription:

$$\langle\psi_{\text{eff}}|H_{\text{eff}}|\psi_{\text{eff}}\rangle \leq \langle\psi_{\text{trunc}}|H_{\text{eff}}|\psi_{\text{trunc}}\rangle. \tag{4.2.27}$$

## 4.3 A related method

In a recent letter, Zhang *et al.* [14] compare their so-called 'constrained path' method, which we shall briefly describe, with our lattice fixed-node method. Here, we wish to consider the similarities and differences of the two methods.

The method of Zhang *et al.* [14] uses a Hubbard-Stratonovich transformation. As elucidated by Fahy and Hamann [15], the Hubbard-Stratonovich auxiliary-field approach to the projection of the ground state of an interacting fermion system is equivalent to solving a differential equation with diffusion, drift and branching terms on a manifold of Slater determinants. The main difference with diffusion Monte Carlo is that the latter is formulated on configuration space.

Zhang *et al.* [14] impose a fixed-node-like constraint on the walks in the space consisting of Slater determinants of single-particle levels. Moves corresponding to a change of sign in the trial wave function are forbidden. No compensating terms are introduced in the Hamiltonian.

In ref. [14], a number of advantages of 'constrained path' algorithm quantum Monte Carlo are listed. However, two merits of that method that are mentioned, namely built-in antisymmetry and the overcompleteness of the set of basis states, are not unique. Our trial states are antisymmetric as well, and, as Fahy and Hamann remark, *both* methods work in an overcomplete basis: all of configuration space rather than $1/N!$ of it in the Green function approach, and all Slater determinants rather than a finite set of them in the auxiliary-field approach. The remark that off-diagonal expectation values are easier to calculate in the constrained path algorithm than in the Green function approach may be true. We do not have experience in applying the former method. We do know, however, that it is not impossible to calculate off-diagonal correlations in GFMC: we shall present results for such correlation functions in the Kondo lattice model, calculated with our method, in the next chapter.

The main advantage of the constrained path algorithm is that, in contrast with Green function Monte Carlo in configuration space, the random walk in the Slater

determinant space becomes *continuous* as the step in imaginary time goes to zero, $\Delta\tau \to 0$, regardless of the discrete nature of the original system. Like in the case of Green function Monte Carlo in a continuous space, one can impose the constraint of eliminating moves *without compensating potentials*. Despite the notion of discreteness caused from the finiteness of the number of single-particle levels, the fact that while the Hubbard-Stratonovich fields are sampled stochastically, the changes in the single particle levels are evaluated *exactly*, gives the method its continuous character indeed.

In both formalisms, an approximation that avoids the sign problem is implemented. Both are variational. Depending on the problem one wishes to study, one should choose one of the two. While the constrained-path formalism is likely to be reliable if a description in terms of single-particle levels is appropriate, lattice fixed-node Monte Carlo appears to be the best candidate if the problem is to study correlations defined in real space [16]. An example is the effect of a Gutzwiller factor (to be discussed in chapter 5) in a trial wave function. Such wave functions are easily evaluated in real space. All tests we shall present in chapter 5 are for the lattice fixed-node Monte Carlo method we developed in section 4.2.

## Appendix: Complex phases

In this appendix, we consider the question what can be done along the lines of lattice FNMC, if the problem is *complex*, in an essential way. Cases for which this applies include the dispersion of excitations, the criteria for a model showing metallic or superconducting behaviour [17], since these are in terms of twisted boundary conditions, and cases in which a magnetic field is present.

A first note is that while in the real case, nodes are lines in configuration space, in the complex case, the nodes of the real part of a function and those of the imaginary part are both lines, but the true nodes of the function are the intersection of those, and thus they consist of a set of points in configuration space. For measured quantities, the real part of a function is relevant.

Writing $\psi(R)$ as $|\psi(R)| \exp[i\phi(R)]$ and demanding that $\phi(R)$ should be the same as $\phi_T(R)$ would be the lattice equivalent of the work of Ortiz *et al.* [18]. Their approach is called the fixed-phase method. The restriction leads to extra terms in the equation for $|\psi(R)|$.

It is more in the spirit of imposing not more restrictions than needed for avoiding cancellations, to allow contributions to measured quantities to come from *sectors* of the complex plane. If only $\psi_T$ can be complex, and $\mathcal{H}$ is real, and if one is sampling a real quantity, the criterion is that walkers should remain in the right or in the left half plane. Since in the product $\prod_{R_i} G(R_i, G(R_{i-1}))$, with $G(R_i, G(R_{i-1}) = \psi_T(R_i)F(R_i, G(R_{i-1})\psi_T^{-1}(R_{i-1})$, every $\psi_T(R_i)$ is counteracted by the $\psi_T^{-1}(R_{i-1})$ of the next step, the phase of the walker is the phase of $\psi_T(R)$, if the walker is in $R$, at that moment. So the criterion to leave steps out is that the real part changes sign, or, in other words, that the phase of the trial wave function enters the other half plane.

It is less clear what the *potential* that replaces these steps should be. Let us follow

the proof that the prescription of Eq.(4.2.11), for the real case, yields an upper bound, in order to be able to discuss what the problems with the generalization are.

If we would take as replacing potential in the complex case, for a forbidden step from $R$ to $R'$, $V_{RR'} = \psi_T(R)/\psi_T(R')$, we would retain Eq.(4.2.19), but it would be a *complex* number. Neither the number nor its real part obeys bounds. For the real part only, the step from Eq.(4.2.19) to Eq.(4.2.21) is not possible because $\mathrm{Re}\,(\psi_T(R)/\psi_T(R')) \neq 1/\mathrm{Re}\,(\psi_T(R')/\psi_T(R))$.

If we take either $V_{RR'} = |\psi_T(R)/\psi_T(R')|$, or $V_{RR'} = \mathrm{Re}\psi_T(R)/\mathrm{Re}\psi_T(R')$, we *would* have $V_{RR'} = 1/V_{R'R}$, and therefore the step to Eq.(4.2.21) is possible: one can write the energy difference as a complete square.

So, like in the real case, any potential that satisfies $V_{RR'} = 1/V_{R'R}$ gives an upper bound. In the real case, however, we have more. If the ratio $\psi(R)/\psi(R')$ is the same as the ratio $\psi_T(R)/\psi_T(R')$, the energy difference is *zero*. For the complex case, it is not possible to reach this situation: in the equivalent of Eq.(4.2.[illegible]) both the real and the imaginary part of $\psi$ would have to satisfy a condition. And there is not enough freedom in the choice of $V_{RR'}$ to fulfill those. So, one can construct effective Hamiltonians that obey bounds, but one does not have zero difference for any wave function.

## References


[1] For a review and recent references, see e.g. W. von der Linden, Phys. Rep. **220**, 53 (1992); or H. de Raedt and W. von der Linden, in *The Monte Carlo Method In Condensed Matter Physics*, edited by K. Binder, Springer, Berlin (1992)

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

[3] M. Suzuki, editor, *Quantum Monte Carlo Methods in Condensed Matter Physics*, World Scientific, Singapore (1993)

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

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

[6] S. Sorella, E. Tosatti, S. Baroni, R. Car, and M Parrinello, Int. J. Mod. Phys. B **1**, 993 (1988)

[7] While mixed estimators of this form arise most naturally in GFMC calculations, they are only exact if $\mathcal{O}$ commutes with $\mathcal{H}$, see e.g. ref. [10].

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

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

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

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

[12] H.J.M. van Bemmel, D.F.B. ten Haaf, W. van Saarloos, J.M.J. van Leeuwen, and G. An, Phys. Rev. Lett. **72**, 2442 (1994)

[13] D.F.B. ten Haaf, H.J.M. van Bemmel, J.M.J. van Leeuwen, and W. van Saarloos, Phys. Rev. B **51**, 13039 (1995)

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

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

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

[17] D.J. Scalapino, S.R. White, and S. Zhang, Phys. Rev. B **47**, 7995 (1993)

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

# 5 FNMC for the 1D Kondo Lattice Model

The main goal of this chapter is to address the question posed in chapter 1: "How does the fixed-node quantum Monte Carlo method for lattice fermions perform in practice?". The effectiveness of the fixed-node quantum Monte Carlo method for lattice fermions is tested by performing calculations on small 1D Kondo lattices. The implementation of this method for the Kondo Lattice model is discussed in detail. We compare the fixed-node upper bound for the ground state energy with exact-diagonalization results from the literature. How different trial states lead to different results for correlation functions, is illustrated on spin-spin correlations in small 1D lattices for which exact results are known. The stochastic projection on the ground state leads to 'best estimates' for the fixed-node values for the correlation functions that do not depend sensitively on the input trial wave function. The practical complications that arise when one has a many-Slater-determinant trial state are discussed, in the context of obtaining the spin gap of the 1D Kondo Lattice model. The lowest-energy spin excitation is confirmed to be a running spin soliton with wave vector $\pi$, as was concluded earlier from Gutzwiller-projected mean-field calculations.

## 5.1 Introduction

In this chapter, we report on tests of the lattice fixed-node Monte Carlo (FNMC) procedure for fermions [1] on small 1D Kondo Lattice models (KLM), for which results from other methods are available for comparison. As discussed in the previous chapter, FNMC involves an approximation that avoids the sign problem in the context of Green Function Monte Carlo (GFMC). Different Monte Carlo techniques have been applied to the 1D KLM. Finite temperature properties were obtained using the world-line algorithm [2] and using the grand-canonical method involving a Hubbard-Stratonovich transformation [3]. The ground state method that uses Hubbard-Stratonovich fields, developed by Sorella *et al.* [4], has been applied to the 1D KLM as well [5]. All three Monte Carlo methods suffer from the sign problem, even in 1D. This observation makes the 1D KLM a suitable testing ground for our lattice FNMC method.

The KLM is one of the most widely studied models for correlated fermions. As discussed in chapter 1, it can be obtained as the strong-coupling limit of the Periodic Anderson model, which aims at capturing some of the essential physics of heavy-fermion materials. In the limit of strong on-site repulsion among the $f$-electrons, a picture emerges of localized $f$-electrons interacting with a conduction band (in chapter 1, we introduced the Periodic Anderson model and the KLM in more detail).

Apart from the Monte Carlo calculations mentioned earlier, exact-diagonalization and variational results are available. We shall give these results alongside with the results of our FNMC calculations.

The lattice version of FNMC gives, like the continuum version, upper bounds for the ground state energy [6]. The proof has been given in chapter 4. It improves upon a trial wave function by allowing fluctuations consistent with the given signs, but it is not exact because fluctuations across nodes of the trial wave function are not allowed, in order to avoid the sign problem. One question of interest is how close to the exact ground state energy the FNMC energy comes. Furthermore, we compare the FNMC results for some correlation functions (which do not obey bounds), with exact results. The comparisons are for six sites, with coupling constant $J$ equal to 0.2 and 1.0, for which exact-diagonalizations have been performed by Otsuka and Ueda [7] and by Otsuka [5], respectively. The sizes for which exact diagonalization results exist are small, because the size of the Fock space is $8^N$, with $N$ the number of sites. A further comparison is with the work of Wang *et al.* [8], who performed Gutzwiller-projected mean-field calculations for the spin gap in the 1D KLM. We shall discuss the mean-field calculations and the Gutzwiller-projection in section 5.3.

This chapter is organized as follows. Firstly, we briefly discuss the 1D KLM and the reason why sampling it with unrestricted random walks leads to the sign problem. Mean-field results, which will be used as input for the FNMC calculations later, are presented in section 5.3. In section 5.4, details of the implementation for the KLM are given. Section 5.5 gives the comparison with exact results for small lattices. In section 5.6, the running spin excitation is discussed.

## 5.2 The Kondo Lattice Model

The Kondo Lattice Hamiltonian is given by

$$\mathcal{H} = -t \sum_{\langle ij \rangle} \left( c^\dagger_{i\sigma} c_{j\sigma} + h.c. \right) - \mu \sum_i n_{ic} + J \sum_i \vec{S}_{if} \cdot \vec{S}_{ic}. \tag{5.2.1}$$

The two kinds of electrons, denoted by $c$ for the conduction band and $f$ for localized levels, have a spin-spin interaction. For the $f$s, the constraint is that there has to be precisely one $f$-electron at every site.

We seek to write the Hamiltonian in a form that is convenient for GFMC calculations. If one would treat the $f$s as spins, which are not antisymmetrized but form a dynamical background for the conduction electrons, the total number of up-spin (and of down-spin) *fermions* would not be conserved by the Hamiltonian. It is not convenient to use this representation in a GFMC calculation. We use the constraint of one $f$-electron per site, the identity

$$\vec{S}_{if} = \frac{1}{2} f^\dagger_{i\sigma} \vec{\tau}_{\sigma\sigma'} f_{i\sigma'}, \tag{5.2.2}$$

and the equivalent identity for the $c$s. The components of $\vec{\tau}_{\sigma\sigma'}$ denote the three Pauli matrices. We obtain

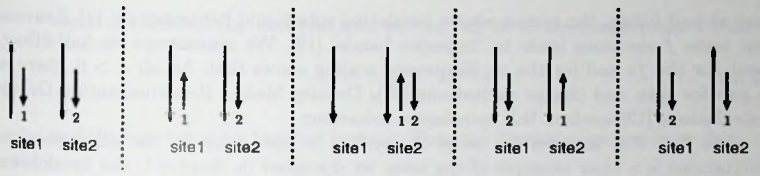



Figure 5-1. A sequence of processes that lets two conduction electrons of equal spin pass each other in 1D. Large arrows denote $f$-electrons, small ones denote $c$-electrons.

$$\mathcal{H} = -t \sum_{\langle ij \rangle \sigma} \left( c^{\dagger}_{i\sigma} c_{j\sigma} + h.c. \right) - \frac{J}{2} \sum_{i} \left( c^{\dagger}_{i\sigma} f_{i\sigma} \right) \left( f^{\dagger}_{i\sigma'} c_{i\sigma'} \right) - \mu' \sum_{i} n_{ic}, \qquad (5.2.3)$$

where the new chemical potential $\mu'$ equals $\mu - J/4$. This is now fully in fermionic language. Therefore, Slater-determinants of fixed dimension can be used.

In different fermionic models, e.g. the Hubbard model, statistics is not really important in one dimension, and Monte Carlo calculations can be performed without encountering the sign problem. The underlying reason is that, in 1D and with hard-core conditions (for fermions, hard-core conditions are implied by antisymmetry), the particles that have to be symmetrized or antisymmetrized cannot pass each other. Therefore, the ordering is fixed. This ordering is not possible in higher dimensions.

For the 1D KLM, there is no ordering of conduction electrons of equal spin. The presence of the term $\sim J$ in $\mathcal{H}$ makes the following series of processes possible, which is depicted in Fig.(5-1). An up $c$-electron can disappear at one site, due to a simultaneous flip of a $c$-$f$ pair (from $c$-up, $f$-down to $c$-down, $f$-up). Then, the $c$-electron has spin down, and nothing prevents it from occupying the same position as an up $c$-electron, after hopping ($\sim t$). It can pass such an electron and, on a different site, flip back to *up*. So, eventually, compared to the original situation, two up conduction electrons have been interchanged. This can also happen in a GFMC simulation, and one needs to take into account that two configurations that differ by the interchange of two numbered electrons must have opposite weights in the wave function. This is the reason why even the 1D KLM exhibits a sign problem, as noted empirically in ref. [5], for example. In the case of 6 sites, $J = 1.0$, which is studied in ref. [5], it appears that the sign problem is not very severe, and that ground state properties can be obtained.

The 1D KLM has been studied in different regimes. If the number of $f$-electrons is equal to the number of sites, and the carrier concentration is low, there is a ferromagnetic state [9]. The reason for this behaviour is the same in the KLM and in the Hubbard (and $t$-$J$) model that we discussed mainly in chapter 3: conduction electrons have the most freedom to move on a ferromagnetic spin background. In weak coupling, at larger density but below half-filling, one obtains a paramagnetic state [10],

and at half-filling, the system shows insulating spin-liquid behaviour [3, 11]. Leaving out some $f$-electrons leads to 'impurity bands' [12]. We concentrate on half-filling, both for the $f$s and for the $c$s. Finite-size scaling shows that, for all $J > 0$, there is a gap for spin and charge excitations [11]. Density Matrix Renormalization Group calculations [13] confirm this spin-liquid behaviour.

The fact that interactions cause the system to have gaps for the single particle excitations is a clear example of the issue we discussed in chapter 1: the breakdown of Fermi liquid theory in 1D. If there would be Fermi liquid quasi-particles, there would be no gap for excitations. The presence of gaps also indicates that low-energy collective excitations are absent. If there are no Goldstone modes, this implies that there is no magnetic order of any kind, no true long-range order and no algebraic long-range order. So, the study of the energy spectrum of the 1D K.L.M. can be viewed as an example of the general questions we discussed in the introductory chapter: does a model show ordering in $\vec{k}$-space, and does it order in real-space?

## 5.3 Mean-field calculations

In this section, we discuss aspects of mean-field analyses of the K.L.M. Our main interest is in the numerical self-consistent real-space calculation using the so-called Kondo decoupling, which will be used later as input for the Monte Carlo calculations. In one of the papers [8] with which we shall compare our FNMC results, this is considered as the sole possibility. Different possible decouplings as well as different calculation methods, which we considered in order to check whether this is the correct way to proceed, are discussed as well.

### 5.3.1 Three approaches

- We first adopt a $\vec{k}$-space approach. In the Kondo decoupled Hamiltonian, the field $b_i \equiv \langle f^\dagger_{i\sigma'} c_{i\sigma'} \rangle$ obtains an expectation value. Using this decoupling, the Hamiltonian reads

$$\begin{aligned} \mathcal{H} &= -t \sum_{\langle ij \rangle} \left( c^\dagger_{i\sigma} c_{j\sigma} + h.c. \right) - \frac{J}{2} \sum_i b_i f^\dagger_{i\sigma} c_{i\sigma} - \frac{J}{2} \sum_i b^*_i c^\dagger_{i\sigma} f_{i\sigma} \\ &\quad + \frac{J}{2} \sum_i |b_i|^2 - \mu' \sum_i n_{ic}. \end{aligned} \tag{5.3.1}$$

  A more sophisticated way of obtaining this effective Hamiltonian, is to write down the partition function, introduce Hubbard-Stratonovich fields to decouple the interaction term, and to make a saddle point approximation for the resulting path integral [14, 8]. The Hubbard-Stratonovich fields correspond to the $b_i$s, and the saddle point approximation to letting those fields be constant in time.

  Having obtained Eq.(5.3.1), we first consider a BCS-like approach for homogeneous states, in which two transformations are applied that diagonalize the

effective Hamiltonian. The two spin sectors can be treated separately. Define

$$c_k = \frac{1}{\sqrt{N}} \sum_j e^{-ikj} c_j \tag{5.3.2}$$

and the equivalent for the $f$-operators. Then the Hamiltonian becomes ($b_i = b$)

$$\mathcal{H} = \sum_k \epsilon_k c^\dagger_{k\sigma} c_{k\sigma} - \frac{J}{2} b \sum_k f^\dagger_{k\sigma} c_{k\sigma} - \frac{J}{2} b \sum_k c^\dagger_{k\sigma} f_{k\sigma} + \frac{J}{2} N |b|^2. \tag{5.3.3}$$

where

$$\epsilon_k = -2t \cos k - \mu'. \tag{5.3.4}$$

Now define, just as in BCS,

$$\begin{pmatrix} c_k \\ f_k \end{pmatrix} = \begin{pmatrix} u_k & -v_k^* \\ v_k & u_k^* \end{pmatrix} \begin{pmatrix} \gamma_{1k} \\ \gamma_{2k} \end{pmatrix} \tag{5.3.5}$$

with $|u_k|^2 + |v_k|^2 = 1$.

One can write the Hamiltonian in terms of the $\gamma$-operators. The requirement that the non-diagonal terms vanish yields an equation for $v_k$ and $u_k$. The diagonal terms eventually become

$$\sum_{k\sigma} \frac{1}{2} \left( \epsilon_k + \sqrt{\epsilon_k^2 + b^2 J^2} \right) \gamma^\dagger_{1k} \gamma_{1k} - \sum_{k\sigma} \frac{1}{2} \left( \sqrt{\epsilon_k^2 + b^2 J^2} - \epsilon_k \right) \gamma^\dagger_{2k} \gamma_{2k}. \tag{5.3.6}$$

In the ground state, all $\gamma_{2k}$s are occupied, the $\gamma_{1k}$s are unoccupied. Because $b$ is defined as $\langle c^\dagger_{i\sigma} f_{i\sigma} \rangle$, there is a self-consistency equation for $b$. The charge gap is the minimum excitation energy. This is equal to the $\gamma_1$-energy at $k = 0$, minus the $\gamma_2$-energy at $k = \pi$, this difference equals $\sqrt{4 + |b|^2 J^2} - 2$. Solving the gap equation numerically, we get values for $b$, and values for the charge gap as a function of $J$ which are indistinguishable from those in Fig.(5-2), which are obtained by a real-space approach. It should be noted, however, that the true values for the charge gap, obtained by performing Gutzwiller-projected mean-field [8], Density Matrix Renormalization group [13] and exact-diagonalization [11] calculations, are substantially larger than the mean-field values. It is clearly necessary to go beyond mean-field.

- We wish to know whether different decouplings are relevant. Lacroix and Cyrot [14] studied this issue in mean-field approximation, using Green functions. Remaining in the real-space description of the decoupled Hamiltonian, Eq.(5.3.1), one can easily verify by writing down the equations of motion, that the Green functions are

$$G^\sigma_f(\omega) = \sum_k \left( \omega - E_0 - \frac{1}{4} Jn - \frac{J^2 b^2}{4(\omega - \epsilon_k - \frac{1}{4} J)} \right)^{-1} \tag{5.3.7}$$

for the $f$s, and

$$G_c^\sigma(\omega) = \sum_k \left( \omega - \epsilon_k - \frac{1}{4}Jn - \frac{J^2 b^2}{4(\omega - E_0 - \frac{1}{4}Jn)} \right)^{-1} \tag{5.3.8}$$

for the cs; $n$ is the number of conduction electrons per atom and $E_0$ corresponds to an additional term in the Hamiltonian, $E_0 f^\dagger f$, which means that the $f$-level is at a different energy than the middle of the conduction band.

In order to calculate the spectrum, Lacroix and Cyrot assume a specific form for the density of states of the conduction band, in the free problem. The true density of states is known, of course, but for convenience they take

$$\rho_0(\epsilon) = \begin{cases} \frac{1}{2D} & \text{when } -D < \epsilon < D \\ 0 & \text{otherwise} \end{cases} \tag{5.3.9}$$

Using

$$\rho(\omega) = \frac{1}{\pi} \mathrm{Im} G(\omega). \tag{5.3.10}$$

the density of states as a function of $\omega$, corresponding to the decoupled Hamiltonian, is calculated. The spectrum appears to have a gap.

The conclusion in ref. [14] that, at half-filling, in mean-field approximation, the system is in a Kondo state for large $J$ and in an antiferromagnetic state for small $J$ is also obtained using Eq.(5.3.9) for the density of states. Next we verify that this is not an artifact of the assumed density of states for the free conduction band.

- We performed self-consistent mean-field calculations for different types of decoupling, keeping the hopping term instead of assuming a certain density of states for the conduction band. Using a decoupled Hamiltonian, one needs to diagonalize a $2N$ by $2N$ matrix, where $N$ is the number of sites. For every site, one has a $c$ and an $f$-operator.

  In the matrix that needs to be diagonalized in the case of Kondo decoupling , the $c$-$c$ block has $\mu'$ (see Eq.(5.2.3)) on the diagonal, and $-ts$ if the two corresponding sites are connected. The $f$-$f$ block has only zeroes, the $c$-$f$ and $f$-$c$ blocks have $-Jb_i/2$ on the diagonals. One starts with arbitrary $b_i$s, diagonalizes the matrix, fills the number of levels that correspond to the situation one considers (for half-filling, $N$ levels are filled). Then, one calculates the $b_i$s in a state with these levels filled. This procedure is iterated, inserting the new $b_i$s and diagonalizing until self-consistency is reached.

  In the mean-field solution, the constraint $n_{if} = 1$ is not satisfied. In the paper by Wang *et al.* [8] we shall compare our FNMC results with, a *Gutzwiller-projection* is performed, which means that configurations that do not fulfill the constraint, are projected out of the mean-field result. Note that the Gutzwiller-projection is different than in the Hubbard model: in the KLM one wishes

to *completely* eliminate configurations that do not obey the constraint, in the Hubbard model one suppresses the amount of doubly occupied sites (in the $t$-$J$ model, there is a constraint again). In this section, we restrict ourselves to mean-field calculations, but we shall come back to the issue of Gutzwiller-projection, in the context of FNMC calculations.

In order to find possible magnetic states, different decouplings are made. Wang *et al.* [8] only use Kondo decoupling, but Lacroix and Cyrot [14] discussed that magnetic order is also relevant, for small $J$. We performed self-consistent calculations in which $\langle f_{i\uparrow}^{\dagger} f_{i\uparrow} \rangle$ obtains an expectation value. Both ferromagnetic and antiferromagnetic ordering were considered. We find that for small $J$, an antiferromagnetic state is favourable, while Kondo ordering has the lowest energy for large $J$. This conclusion is in agreement with Lacroix and Cyrot [14]. Taking the hopping term as it is, without making assumptions for the free electron density of states, the same conclusion is reached as in their calculation.

We conclude that the different mean-field approaches are consistent. It follows that for large $J$, Kondo decoupling is the most relevant one. More sophisticated calculations showed [8, 13, 11] that the antiferromagnetic state for small $J$ is absent in the true solution. Long-range antiferromagnetic order would lead to Goldstone modes. One finds gaps for every finite $J$, however. Antiferromagnetic correlations exist, and they have a large correlation length, but no long-range order. This observation constitutes a further justification to use only Kondo decoupled trial states in the FNMC calculations.

### 5.3.2 The gaps

In the literature, various elementary excitations are considered. These are classified according to symmetry. At half-filling, the KLM is particle-hole symmetric. Not only the total spin is conserved by the Hamiltonian, also a pseudo-spin operator $I$ is a conserved quantity [13]. It is defined by $I^{+} = \sum_i (-1)^i (c_{i\uparrow}^{\dagger} c_{i\downarrow}^{\dagger} - f_{i\uparrow}^{\dagger} f_{i\downarrow}^{\dagger})$, $I^{-} = (I^{+})^{\dagger}$, and $I^{z} = \sum_{i\sigma} (c_{i\sigma}^{\dagger} c_{i\sigma} + f_{i\sigma}^{\dagger} f_{i\sigma} - 1)/2$. The $z$-component of this operator is the total charge.

The lowest excited state appears to be a spin triplet [13, 8], the excitation energy is

$$\Delta_S = E(S = 1, I = 0) - E_0(S = 0, I = 0), \tag{5.3.11}$$

where $E_0$ is the energy of the ground state. $E(S = 1, I = 0)$ can be viewed as the lowest energy *in that sector*. The charge gap is larger, it is given by

$$\Delta_C = E(S = 0, I = 1) - E_0(S = 0, I = 0). \tag{5.3.12}$$

It appears [13] that the first excited state with the same symmetry as the ground state, i.e. with $S = 0$ and $I = 0$, is lower in energy than the charge excitation.

Since the lowest-energy excitation is the spin triplet, we shall concentrate on calculating $\Delta_S$ and the corresponding wave function, in most of the discussion.

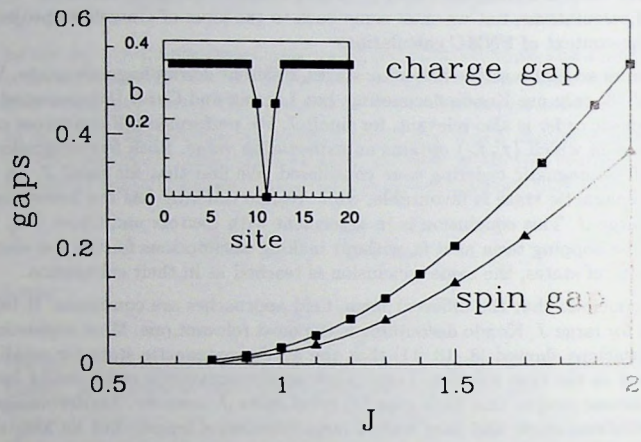



Figure 5-2. Mean-field spin and charge gap as a function of $J$. The inset shows the field $b_i$ as a function of position, in a self-consistent mean-field solution with a spin excitation at a fixed position, for $J = 1.0$. The total size of the 1D lattice is 40 sites.

The mean-field charge gap is obtained as the difference between the highest filled level and the lowest empty level, in the self-consistent solution described in the previous subsection. It is plotted in Fig.(5-2). The values agree with the mean-field results of Wang *et al.* [8] and with those obtained using the first approach of the previous subsection.

The mean-field *spin gap* is calculated as follows. An inhomogeneous self-consistent mean-field solution is obtained, numerically: if one of the $b_i$s is set to zero to begin with, it remains zero in the iteration process. In the inset of Fig.(5-2), the resulting $b_i$s are plotted as a function of position. The corresponding mean-field energy is calculated. One needs to avoid double counting: the decoupled Hamiltonian Eq.(5.2.3) includes a term $+\frac{J}{2}\sum_i |b_i|^2$ that needs to be taken into account.

The spectrum of single-particle energies is interesting. While at half-filling, the interaction causes the conduction band to split into two bands, separated by a large gap, the localized structure appears to have two 'midgap states', exactly in between the two bands, at zero energy. This is a common feature of localized structures: it also appears for spin polarons and domain walls in the Hubbard model (see chapter 3)

and if the electrons couple to the lattice, as in the Su Schrieffer Heeger model.

The difference in energy between the uniform and the inhomogeneous solution is the spin gap, it is plotted as a function of $J$ in Fig.(5-2). The values agree with the mean-field results of Wang *et al.* [8]. Their values for the gaps after Gutzwiller-projection are substantially larger. These, moreover, agree with exact-diagonalization results. In the next section, we shall investigate what happens to mean-field results after applying the fixed-node Monte Carlo method.

## 5.4 The FNMC method for the KLM

### 5.4.1 Principles

The general principles of FNMC for lattice fermions [1, 6] have been laid out in chapter 4; here the KLM serves as a specific example with which the method is illustrated and tested explicitly.

We use the Kondo decoupling scheme, presented in the previous section, in two different ways. In the first, we iterate in the mean-field program, until self-consistency is reached. The second way to treat the decoupling is used if we wish to compare different trial states. Different values for $\langle f^{\dagger}_{i\sigma'} c_{i\sigma'} \rangle$ are inserted and, after diagonalizing the decoupled Hamiltonian only once, different wave functions, which have different new values for $\langle f^{\dagger}_{i\sigma'} c_{i\sigma'} \rangle$, are obtained. For the FNMC calculations, the different trial states provide different sign structures, other correlations possibly change while filtering out the best ground state consistent with the sign structure. Taking different trial states is an example of 'de-optimizing' the trial wave function, a common procedure for checking whether the results are relatively independent of the starting point (see e.g. ref. [5]).

The GFMC method proceeds as follows. The mean-field result is the starting point. One chooses an ensemble of random walkers in configuration space that represents the trial wave function. A single random walker consists of a set of up and down electrons, numbered from 1 to $N_{\mathrm{up}}$ and from 1 to $N_{\mathrm{down}}$ respectively, with a specification of the position of each electron, and of the state it is in: $c$ or $f$.

The ensemble is chosen by generating configurations at random, and then comparing the weight squared with a random number, in order to decide whether that configuration should be accepted as a member of the ensemble or not. Since all walkers are made to satisfy the constraint of one $f$-electron per site, the ensemble is automatically Gutzwiller-projected.

The weight of a certain configuration in the trial wave function can be calculated since the mean-field single-particle levels carry the information: the number $\phi(j, i, +/-)$ denotes how much $c/f$ (last variable: + is $c$, − is $f$) there is at site $i$, in level $j$. The trial wave function consists of a Slater-matrix for the up electrons, and a Slater-matrix for the down electrons. If a configuration has its up electron number Inumuc(i) at position $i$, and in band $c$ (there is a separate array Inumuf (and Inumdc, Inumdf for the downs)), then the element of the Slater-matrix SMup($j, Inumuc(i)$)

equals $\phi(j, i, +1)$. The $+1$ denotes the $c$ character, $j$, which labels the levels, runs from 1 to $N_{\text{up}}$, the number of *filled* levels. Thus, the complete matrices are filled, both for the up and for the down sector; the product of the determinants gives the weight of the configuration.

This way of representing the trial wave function is suitable for the order in which the operators in the interaction term appear in Eq.(5.2.3), and for the decoupling we have chosen in the mean-field calculation. The operators between brackets in Eq.(5.2.3), represent intermediate steps, in a Monte Carlo diffusion process. These correspond to changes within one spin sector. It is, therefore, natural to have numbered electrons of a certain spin and to allow changes from $c$ to $f$ and vice versa, rather than to have numbered electrons with the $c/f$ character fixed and letting the spin change. Both representations are equivalent, but our choice allows working with Slater-determinants of fixed size.

Let us be more specific about what Eq.(4.2.1) and Eq.(4.2.2), with $\mathcal{F}$ given by Eq.(5.2.3), mean for the KLM. Since $\mathcal{F}$ contains $\mathcal{H}$, the following things can happen, in one time step. $R'$ can go to a configuration with one $c$-electron hopped to a neighbouring position (the term $\sim t$ in Eq.(5.2.3)). The second possibility is a simultaneous spin flip: the term $\sim J$ allows a configuration which has, at a certain site, a pair $c$-up, $f$-down, to change into $c$-down, $f$-up (or vice versa). The third possibility is that nothing happens in a time step: $R'$ remains $R'$. The relative probabilities are given by Eq.(4.2.2). The fact that Eq.(4.2.2) can become negative interferes, but otherwise what one should do is make a list of all possible steps ($\sim t$, $\sim J$, staying) and calculate the ratios $\psi_T(R)/\psi_T(R')$ of the trial wave function for every configuration $R$ that can be reached from $R'$ in one step. Since for every $R$, one electron at most changes its state (site or $c/f$) per spin sector, this ratio can be calculated in a number of operations that are linear in the size of the system, if one already has the transposed inverse of the Slater-matrices available.

### 5.4.2 Implementation

The structure of the program that implements the lattice fixed-node method in the case of the KLM is as follows. Input to the program are the single-particle wave functions that are the result of a mean-field calculation, and an ensemble of random walkers. The FNMC program has two tasks: first to let the ensemble evolve according to Eq.(4.2.1), in such a way that it will eventually be a sample of the ground state. The second task is to *measure* the properties of the ground state.

The evolution of the ensemble (first approaching the ground state, then accumulating statistics) proceeds as follows.

The first stochastic process is choosing moves of a random walker. One treats one walker at the time. The two Slater-matrices are constructed, as well as their transposed inverses. Then, all possible moves are considered. With the help of the transposed inverse matrices, the ratios $\psi_T(R)/\psi_T(R')$ are calculated. This operation corresponds to a dot-product, the time needed to compute the dot-product is linear

in the system size.

For a hop, only one ratio needs to be calculated, for a flip, there is a change both in the up and in the down determinant (as discussed before, we choose a representation in which a numbered electron of a certain spin keeps its number, in a flip, and remains of the same spin; it is the label $c/f$ that is interchanged). If $\psi_T(R)/\psi_T(R')$ is negative, the corresponding contribution to the sign-flip potential is

$$|\psi_T(R)/\psi_T(R')|, \tag{5.4.1}$$

for a hop ($t =$ 1), and

$$J|\psi_T(R)/\psi_T(R')|, \tag{5.4.2}$$

for a flip.

If $\psi_T(R)/\psi$ [illegible] tive, the step to $R$ is allowed. To give it the appropriate probability, we [illegible] remental array, in which the extra part corresponding to this step is eq[illegible] $R)/\psi_T(R')|$ or $J\tau|\psi_T(R)/\psi_T(R')|$. The part that corresponds to stay[illegible] ame position in configuration space has length

$$1 - \tau(U_{\text{pot}} + V_{\text{sf}} - w), \tag{5.4.3}$$

$U_{\text{pot}} + V_{\text{sf}}$ is the *potential* energy of the *effective* Hamiltonian. The sum of Eqs (5.4.1), (5.4.2), and (5.4.3) gives $m(R')$ (see Eq.(4.2.5), the multiplicity factor arising in this time step.

This number $m(R')$ serves as scale: a random number between 0 and 1 is picked and multiplied with $m(R')$. The resulting number gives a position in the incremental array, the corresponding move will be performed. In this way, all possible moves (hopping, flipping, and staying in the same configuration) have a probability to occur as prescribed by the lattice FNMC idea.

It is efficient to continue with the same walker for a number of time steps: after a step, *updating* the transposed inverses of the Slater-matrices costs a computation time linear in the size of the system (again because only one electron changes its position or $c/f$ per spin sector), while the time for constructing new ones is cubic in the size of the system. Only after, say, $N_{\text{time}}$ steps, one takes the next walker.

After having treated all walkers, the total multiplicities, or 'weights', accumulated up to that time, determine the number of copies to be made of each walker. This is the second stochastic process, it is called 'branching'. The product of multiplicity factors along the path gives the weight of a walker. The number of copies to be made of a certain walker, is the integer part of its weight, plus a random number between 0 and 1. In most cases, we tune $w$, $\tau$ and $N_{\text{time}}$ such that all weights are between 0 and 2. The value of $w$ gives the average weight; it should be close to the measured energy, because otherwise the total number of walkers will decrease (if $w$ is chosen too low), or increase (if $w$ is chosen too high). $N_{\text{time}}$ and $\tau$ determine the width of the distribution of weights. Some walkers will disappear, some will be doubled, others stay the same. In the next time steps, all walkers will have weight 1 again, and they will move independently.

In this way, in a region of low 'local energy', more walkers will appear (this corresponds to an increase of the amplitude of the wave function). This process of *diffusion with branching* will eventually yield an ensemble that represents the ground state of the *effective* Hamiltonian.

The first stage of the diffusion process is a thermalization. During the corresponding evolution of the ensemble, the parameter $w$ in $\mathcal{F} = 1 - \tau(\mathcal{H} - w)$, which should be close to the ground state energy, is adjusted self-consistently, i.e. it follows the *measured* energy.

If the energy does not drop anymore, one measures the properties of the ground state of the effective Hamiltonian. The *local value* of an operator $\mathcal{O}$ is obtained as a *mixed estimator* $\langle\psi_T|\mathcal{O}|\psi_0\rangle$, the usual quantity being measured in GFMC. It is given by

$$O_{\text{local}}(R) = \langle R|\mathcal{O}|\psi_T\rangle/\psi_T(R) = \left[\sum_{R'}\langle R|\mathcal{O}|R'\rangle\langle R'|\psi_T\rangle\right]/\psi_T(R) \tag{5.4.4}$$

Some correlation functions are diagonal: if one is in $R$, the operator $\mathcal{O}$ only connects to $R$. These are easy to compute: no ratios of trial wave functions need to be calculated. If there are off-diagonal terms, one does need to calculate the corresponding $\psi_T(R')/\psi_T(R)$s, this costs extra computer time. Among the correlation functions we shall calculate are off-diagonal spin-spin correlations, i.e. correlation functions involving $S^+$ and $S^-$.

Apart from the system parameters and the parameters in the projection operator, $\tau$ and $w_{\text{start}}$, the following parameters need to be specified when the program is executed: The starting number of walkers is $N_{\text{base}}$. In one *interval*, all walkers are propagated during $N_{\text{time}}$ time steps, before branching. The first $N_{\text{therm}}$ such *intervals* constitute the *thermalization process*. After this, statistics is accumulated in $N_{\text{block}}$ blocks of $N_{\text{intv}}$ intervals each. In principle, the blocks are treated as independent measurements, but it should be checked whether these are sufficiently independent indeed. A larger $N_{\text{intv}}$ makes them more independent. When calculating the error-bar, possible correlations among the blocks should be taken into account. We shall see this in practical examples.

## 5.5 Results for $J = 0.2$ and $J = 1.0$

The first test of the method on the KLM is a comparison with exact diagonalization results by Yamamoto and Ueda [7]. The coupling constant is $J = 0.2$, the system size is six sites with periodic boundary conditions. The trial wave functions we use are obtained by inserting a value for $\langle f^\dagger_{i\sigma'}c_{i\sigma'}\rangle$ in the mean-field program, and iterating only once. This means that we diagonalize the bilinear Hamiltonian with the given number for $\langle f^\dagger_{i\sigma'}c_{i\sigma'}\rangle$, and that we take the lowest-energy single particle levels for filling the Slater-matrices.

In Fig.(5-3), we plot the fixed-node energy against the value of $\langle f^\dagger_{i\sigma'}c_{i\sigma'}\rangle$ in the trial wave function (so this is not the *inserted* value of $\langle f^\dagger_{i\sigma'}c_{i\sigma'}\rangle$ but the value in the

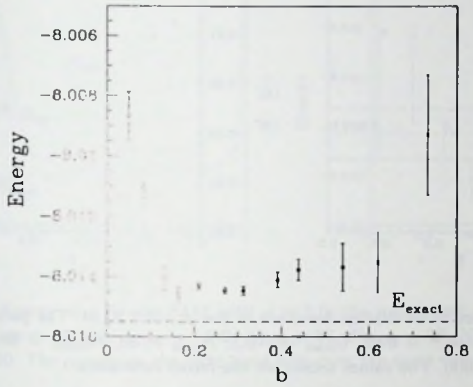



Figure 5-3. Energy of a six-site KLM with $J = 0.2$; the parameters used in the Monte Carlo calculation are $\tau = 0.01$, $w_{\text{start}} = -8.3$, $N_{\text{time}} = 30$, $N_{\text{therm}} = 20$, $N_{\text{intv}} = 5$, $N_{\text{block}} = 400$, $N_{\text{base}} = 1000$.

wave function obtained after diagonalizing with that value).

The fixed-node energies are above the exact ground state energy, as they should be, according to the proof that they provide an upper bound for that exact energy. The second observation is that the minimum is very flat. The best estimate for the ground state energy that can be obtained, apparently does not depend much on the way the value $\langle f^{\dagger}_{i\sigma'} c_{i\sigma'} \rangle$ affects the sign structure of the wave function, which is the only property that is not free to change in the Monte Carlo diffusion process, and therefore establishes the only obstacle for projecting on the true ground state. Thirdly, the statistical fluctuations are smaller close to the optimal value of $\langle f^{\dagger}_{i\sigma'} c_{i\sigma'} \rangle$ than for $\langle f^{\dagger}_{i\sigma'} c_{i\sigma'} \rangle$s for which the fixed-node energy is higher. Note that, in general, the presence of a gap in the excitation spectrum is helpful in achieving statistical accuracy [15]. The 1D KLM does have a gap.

Next, we turn to correlation functions. All correlation functions are of the type $\vec{S}^{i}_{\text{band1}} \cdot \vec{S}^{j}_{\text{band2}}$, where 'band' is $c$ or $f$ and $i$ and $j$ are either the same site, or nearest neighbours. Note that the dot-product not only includes the $S^z S^z$ term, which is diagonal and easy to calculate, but also the $x$ and $y$ components (or $S^+, S^-$). We shall treat those correlation functions more extensively for $J = 1.0$. For $J = 0.2$, we observe that Fig.(5-4) indicates that the mixed-estimators we obtain are consistent

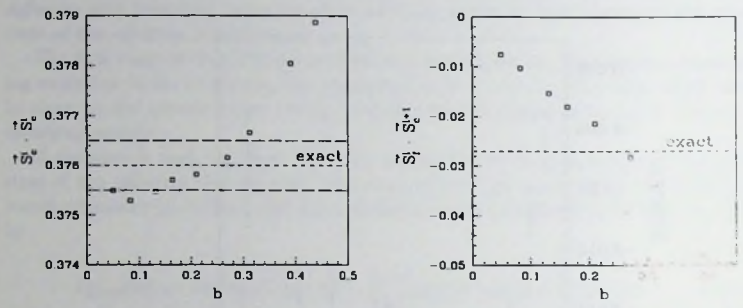



Figure 5-4. Two examples of correlation functions in the $J = 0.2$ KLM. The parameters used in the FNMC program are $\tau = 0.03$, $w_{\text{start}} = -8.3$, $N_{\text{time}} = 30$, $N_{\text{therm}} = 30$, $N_{\text{av}} = 5$, $N_{\text{block}} = 200$, $N_{\text{base}} = 1000$. The values shown are the mixed estimators.

with the exact results of Yamamoto and Ueda: around the value of $\langle f^{\dagger}_{i\sigma'} c_{i\sigma'} \rangle$ where the lowest energies are obtained, the curves for the correlation functions cross the exact values.

For $J = 1.0$ we follow the same procedure as for $J = 0.2$ and compare with exact results obtained by Otsuka [5]. The upper line in Fig.(5-5) gives the energy measured in the *starting* ensemble, the lower line is the fixed-node value, i.e. after projection. Like in the case of $J = 0.2$, the latter curve is very flat, while the starting values depend strongly on the input wave function. The picture on the right in Fig.(5-5) shows the same data on a different scale, on which one can see that the flat part of the picture on the left really has a minimum. The exact diagonalization result $E = -8.561616$ is also indicated in the picture. The main conclusion is that the FNMC method is able to come close to the exact energy despite a rather high starting value, but never reaches the exact energy. Again, the statistical fluctuations are smallest close to the minimum in fixed-node energy.

In Figs (5-6) and (5-7), we present the results for on-site correlation functions and correlation functions involving different sites, respectively. The mixed-estimator is directly measured in the FNMC program. The 'best estimate' is obtained by assuming that the trial state is close to the ground state and neglecting quadratic terms in the difference [16]:

$$\langle \psi_{\text{eff}} | \mathcal{O} | \psi_{\text{eff}} \rangle \approx 2 \cdot \langle \psi_{\text{eff}} | \mathcal{O} | \psi_{\text{T}} \rangle - \langle \psi_{\text{T}} | \mathcal{O} | \psi_{\text{T}} \rangle, \tag{5.5.1}$$

In Figs (5-6) and (5-7), three values are plotted: the variational value, the mixed estimator and the best estimate of Eq.(5.5.1). For all correlation functions considered, the latter curve is the most flat one, near the value of $\langle f^{\dagger}_{i\sigma'} c_{i\sigma'} \rangle$ for which the minimum

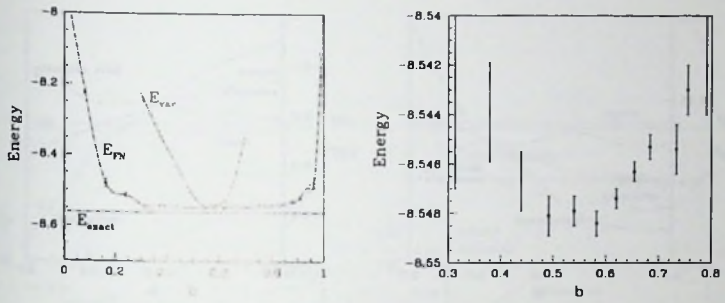



Figure 5-5. Energy of a six-site KLM with $J = 1.0$; the parameters used in the Monte Carlo calculation are $\tau = 0.003$, $w_{start} = -9.9$, $N_{time} = 20$, $N_{therm} = 20$, $N_{intv} = 1$, $N_{block} = 10000$, $N_{base} = 1000$. The picture on the right hand side gives the same values on a different scale.

in energy is obtained. Thus, a reliable value for the correlation function in fixed-node approximation is obtained. This value is not the same as the exact-diagonalization result. There is also no bound to be obeyed. Closeness to the exact value and flatness with respect to the parameters of the input wave function are the main merits of these results.

## 5.6 FNMC calculation for the spin soliton

The lowest-energy excitation above the $S = 0$ ground state of the half-filled KLM has total spin $S = 1$ [8]. In a mean-field calculation, one is able to obtain a self-consistent solution with the extra spin concentrated on a few sites [8]. We reproduced Wang's calculation, it was presented in the previous section. Wang *et al.* [8] proceed by performing a Gutzwiller-projected mean-field calculation and by writing down a wave function that implies a picture of a tight-binding model for the motion of the entire structure:

$$|\psi_q\rangle = \sum_{x_c} \exp(iqx_c)|\psi_{x_c}\rangle, \tag{5.6.1}$$

with $|\psi_{x_c}\rangle = \mathcal{P}_G|\psi^{mf}_{x_c}\rangle$ the Gutzwiller-projected local triplet state with the centre of the soliton located at $x_c$. As we shall confirm below, the minimum of the energy-dispersion is at wave number $q = \pi$.

We follow the general strategy of investigating the robustness of mean-field results by using the mean-field wave function as trial wave function in a FNMC calculation. To obtain the spin-gap in fixed-node approximation, we perform calculations both in

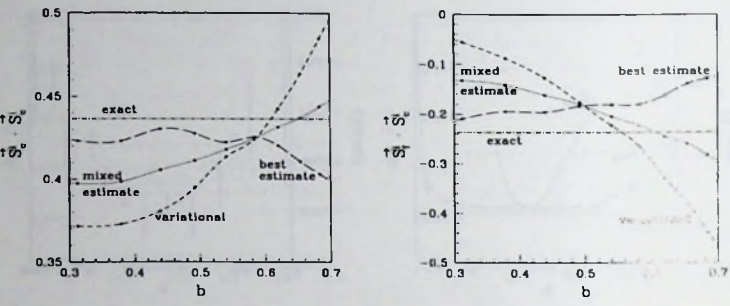



Figure 5-6. On-site correlations in the $J = 1.0$ KLM. The parameters used in the Monte Carlo calculation are $\tau = 0.003$, $w_{\text{start}} = -9.9$, $N_{\text{time}} = 20$, $N_{\text{therm}} = 20$, $N_{\text{intv}} = 1$, $N_{\text{block}} = 200$, $N_{\text{base}} = 1000$.

the $S = 0$ and in the $S = 1$ sector. GFMC does not always project on the ground state, only on the lowest state that has a component along the trial-state. Here, we use this to our advantage: the total spin is conserved by the Hamiltonian, and, therefore, if one starts in the $S = 1$ sector, one remains in the $S = 1$ sector. Comparing the lowest energies in both sectors gives the gap. Note that the charge gap could also be calculated using this method, because $I$ is a conserved quantity. The neutral singlet excitation [13], which has $I = S = 0$, would be difficult to find, because it is in the same sector as the ground state.

A special feature of GFMC is that random walkers can be made to satisfy the constraint $n_{if} = 1$, which is not satisfied in the mean-field result. A separate explicit Gutzwiller-projection is not needed.

Formula (5.6.1) as it stands, seems to indicate that not only different *signs* occur, but also different complex *phases*. Because of inversion symmetry, however, one can combine $q$ and $-q$ and write

$$|\psi_q\rangle = \sum_{x_c} \cos(qx_c)|\psi_{x_c}\rangle, \tag{5.6.2}$$

which is a real problem again. (Ideas concerning complex lattice problems and possible generalizations of the fixed-node method, have been discussed in the Appendix of chapter 4).

We perform FNMC calculations for a system of 20 sites with periodic boundary conditions. The first practical problem is that this takes computer time: for each

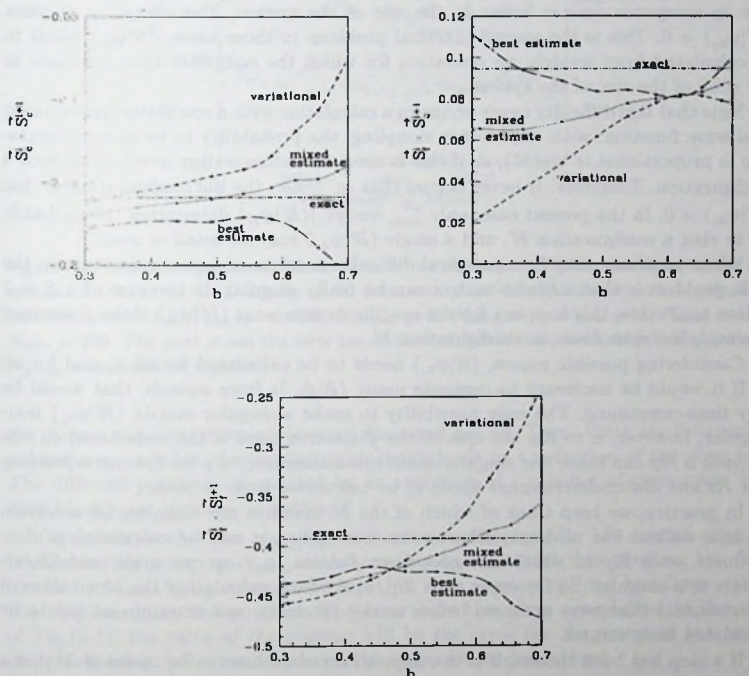



Figure 5-7. Correlations on different sites in the $J = 1.0$ KLM. The parameters used in the Monte Carlo calculation are $\tau = 0.003$, $w_{\text{start}} = -9.9$, $N_{\text{time}} = 20$, $N_{\text{therm}} = 20$, $N_{\text{intv}} = 1$, $N_{\text{block}} = 200$, $N_{\text{base}} = 1000$.

possible step, 20 ratios of determinants need to be calculated, not just one, since

$$\frac{\psi_T(R)}{\psi_T(R')} = \frac{\sum_{x_c} \cos(qx_c)\langle R|\psi_{x_c}\rangle}{\sum_{x_c} \cos(qx_c)\langle R'|\psi_{x_c}\rangle} = \frac{\sum_{x_c} \cos(qx_c)\langle R'|\psi_{x_c}\rangle \frac{\langle R|\psi_{x_c}\rangle}{\langle R'|\psi_{x_c}\rangle}}{\sum_{x_c} \cos(qx_c)\langle R'|\psi_{x_c}\rangle}. \quad (5.6.3)$$

The factors $\frac{\langle R|\psi_{x_c}\rangle}{\langle R'|\psi_{x_c}\rangle}$ can be obtained as dot-products again, in most cases. The cost in computer time is linear in the size of the system. The exception is when $\langle R'|\psi_{x_c}\rangle = 0$. This is the second practical problem: in those cases, $\langle R|\psi_{x_c}\rangle$ needs to be calculated from scratch, an operation for which the computer time increases as the cube of the size of the system.

Note that this difficulty never occurs in a calculation with a one Slater-determinant trial wave function: with importance sampling, the probability to be in a configuration is proportional to $\psi_T(R')$, so if this is zero, a random walker never visits such a configuration. Therefore, it never occurs that in $\frac{\psi_T(R)}{\psi_T(R')}$, the *old* configuration $R'$ has $\langle R'|\psi_{x_c}\rangle = 0$. In the present case, only $\sum_{x_c} \cos(qx_c)\langle R'|\psi_{x_c}\rangle$ determines the probability to visit a configuration $R'$, and a single $\langle R'|\psi_{x_c}\rangle$ may be small or zero.

While smallness may be a practical difficulty in terms of numerical accuracy, the main problem is that a Slater-matrix can be really singular. In the case of a $S = 1$ soliton trial state, this happens for the specific determinant $\langle R'|\psi_{x_c}\rangle$ if the $f$-electron at site $x_c$ has spin *down*, in configuration $R'$.

Considering possible moves, $\langle R|\psi_{x_c}\rangle$ needs to be calculated for all $x_c$ and for all $R$. If it would be necessary to compute many $\langle R|\psi_{x_c}\rangle$s from scratch, that would be very time-consuming. The only possibility to make a singular matrix $\langle R'|\psi_{x_c}\rangle$ non-singular, however, is to flip the spin of the $f$-electron (and of the $c$-electron) on site $x_c$. Such a flip can make one singular matrix non-singular, *only* for the corresponding new $R$s *one* Slater-determinant needs to be calculated from scratch.

In practice, we keep track of which of the 20 matrices are singular, for a certain random walker. For all hops, all non-zero new $\langle R|\psi_{x_c}\rangle$s can be calculated as dot-products, each flip in which one goes from $f$-down to $f$-up can make *one* Slater-matrix non-singular. So for every such flip, apart from calculating the new values of determinants that were non-zero before as dot-products, *one* determinant has to be calculated from scratch.

If a step has been chosen, if it is a hop, all matrices have to be updated. If it is a flip going from $f$-up to $f$-down, one determinant becomes singular. If it is a flip with $f$-down to $f$-up, one matrix becomes non-singular, and its transposed inverse needs to be calculated, for facilitating the calculation of transition probabilities for *further* steps. Except for this more complicated calculation of $\frac{\psi_T(R)}{\psi_T(R')}$, the FNMC program is the same as described before.

Fig.(5-8) shows how a run proceeds. First, one observes projection on the ground state: the measured energy drops. Note that the starting value does not need to be equal to the mean-field energy: the starting ensemble does not fully represent the mean-field wave function, because only sites containing one $f$-electron are present in the ensemble, while the mean-field wave function violates this constraint. This is

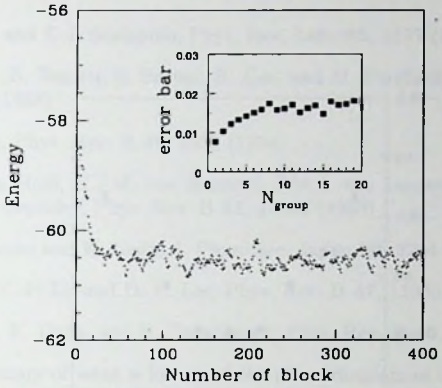



Figure 5-8. How a run proceeds: The energies of the blocks in a FNMC calculation for a $k = \pi$ soliton in a KLM on a 1D lattice of 20 sites with $J = 4.0$, calculated with Monte Carlo parameters $\tau = 0.01$, $w_{start} = -90$, $N_{therm} = 3$, $N_{time} = 1$, $N_{intv} = 3$, $N_{block} = 400$, $N_{base} = 200$. The inset shows the error bar of the last 300 blocks, and calculated by grouping $N_{group}$ measurements together.

the 'automatic Gutzwiller-projection' discussed before. Then, there are fluctuations around a mean value. Accumulating statistics leads to a reduction of the error bars. The different points are separated by an evolution of a number of time steps, after each block, new statistics is gathered.

One observes that the measurements are not independent. The correct error bar is obtained by grouping $N_{\text{group}}$ measurements together and thus dividing the $N_{\text{total}}$ blocks in $N_{\text{total}}/N_{\text{group}}$ measurements. After this regrouping, one has fewer values, but they are more independent. The error bars this gives are plotted in the inset of Fig.(5-8): the value of the plateau will be the error bar we report in the energy dispersion curve (all error bars in energies are obtained this way).

The FNMC dispersion of the spin soliton, for $J = 4.0$, on 20 sites, is presented in Fig.(5-9). Since the $S = 0$ value is $E_{S=0}^{FNMC} = -63.423(5)$, and the minimum of the dispersion is at $E_{S=1}^{FNMC} = -60.47(3)$, the gap we obtain in FNMC approximation is $\Delta_S^{FNMC} = 2.95(3)$, which is the same as Wang *et al.* obtain in Gutzwiller-projected mean-field approximation (and therefore also the same as the exact result, since these values coincide [8]). So, while the energies in both the $S = 0$ and the $S = 1$ sector drop, the *difference* between the $S = 0$ and the $S = 1$ ground state energies is the same in FNMC and in Gutzwiller-projected mean-field approximation. (The gaps in mean-field approximation without Gutzwiller-projection, i.e. in our trial wave functions, are substantially smaller, see Wang *et al.*.)

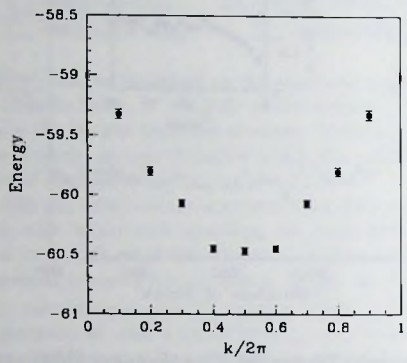



Figure 5-9. Dispersion of a running soliton in a $J = 4.0$ KLM of 20 sites. The parameters used in the Monte Carlo calculations are $\tau = 0.005$, $w_{\text{start}} = -90$, $N_{\text{time}} = 1$, $N_{\text{therm}} = 3$, $N_{\text{intv}} = 3$, $N_{\text{block}} = 200$, $N_{\text{base}} = 200$.

## 5.7 Conclusions

We applied the lattice FNMC method to the 1D KLM. We observe that, for small lattices, energies are obtained which imply a large improvement over the starting values in the case of de-optimized trial wave functions. All energies are above the true ground state energy, as they should be. Consistent values for correlation functions in FNMC approximation can be obtained, as we illustrated by the flatness of the best-estimate curves when the trial-state is varied.

To be able to calculate a dispersion of an excitation, a conserved quantity is needed. If this is present, the lowest energies in different sectors can be compared. We calculated the energy of an $S = 1$ excitation. If inversion symmetry is present, considering *signs* suffices, complex phases are not needed to calculate the dispersion. The FNMC spin gap agrees with Gutzwiller-projected mean-field results

## References

[1] H.J.M. van Bemmel, D.F.B. ten Haaf, W. van Saarloos, J.M.J. van Leeuwen, and G. An, Phys. Rev. Lett. **72**, 2442 (1994)

[2] M. Troyer and D. Würtz, Phys. Rev. B **47**, 2886 (1993)

[3] R.M. Fye and D.J. Scalapino, Phys. Rev. Lett. **65**, 3177 (1990)

[4] S. Sorella, E. Tosatti, S. Baroni, R. Car, and M. Parrinello, Int. J. Mod. Phys. B **1**, 993 (1988)

[5] H. Otsuka, Phys. Rev. B **49**, 1557 (1994)

[6] D.F.B. ten Haaf, H.J.M. van Bemmel, J.M.J. van Leeuwen, W. van Saarloos, and D.M. Ceperley, Phys. Rev. B **51**, 13039 (1995)

[7] K. Yamamoto and K. Ueda, J. Phys. Soc. Japan **59**, 3284 (1990)

[8] [illegible], X.-P. Li, and D.-H. Lee, Phys. Rev. B **47**, 11935 (1993)

[9] [illegible], K. Ueda, and H. Tsunetsugu, Phys. Rev. B **46**, 175 (1992)

[10] For a summary of what is known of the phase diagram of the 1D KLM off half-filling, see e.g. K. Ueda, T. Nishino, and H. Tsunetsugu, Phys. Rev. B **50**, 612 (1994)

[11] H. Tsunetsugu, Y. Hatsugai, K. Ueda and M. Sigrist, Phys. Rev. B. **46**, 3175 (1992)

[12] P. Schlottmann, Phys. Rev. B. **46**, 998 (1992)

[13] C.C. Yu and S.R. White, Phys. Rev. Lett. **71**, 3866 (1993)

[14] C. Lacroix and M. Cyrot, Phys. Rev. B. **20**, 1969 (1979)

[15] See e.g. S. Goedecker and K. Maschke, Phys. Rev. B. **44**, 10365 (1991)

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

# 6 Outlook for the lattice FNMC method

In this concluding chapter, we consider the possibilities opened by the lattice FNMC method, and the questions that remain.

In the application of the lattice FNMC method to the 1D KLM, we saw that correlations are built up, in the projection, even if the trial wave function itself does not have the correct correlations. We also found that a running spin excitation has considerably less energy than a fixed one. In combination with the absence of a sign problem in FNMC, these observations could be of interest for the study of spin polaron separation and domain walls in the two-dimensional Hubbard model, the topic of chapter 3. For the study of those structures, one needs to be able to treat large systems. The absence of the sign problem makes that possible. If one decides to spend computer time on handling a many-Slaterdeterminant trial wave function for a running structure in the Hubbard model, the presence of a sign problem would cause the calculation to be impossible. Studying the energy-dispersion of a polaron-band, or of a domain-wall band, would be an interesting future application of the lattice FNMC method.

Possible applications are not restricted to lattice fermions, the type of system the method was designed for. The FNMC-method involving sign flip potentials also has the potential to become a useful tool for continuum problems with non-local potentials [1], and for frustrated lattice boson and spin models. For frustrated spin models, in the context of finite temperature Monte Carlo, it has been observed that the severeness of the sign problem depends on the basis in which the problem is formulated [2]. In such a case, it would be interesting to study how this works out in diffusion Monte Carlo, using lattice FNMC.

An extension that has been studied before for continuum models, is the 'release node' idea [3]. If one starts with the fixed-node approach, one could continue with *unrestricted* random walks, after projection on the lowest-energy state of the effective Hamiltonian has been completed. If one is able to come close enough to the true ground state using FNMC, the sign problem that is introduced in the 'release-node' process is, depending on the problem, possibly benign enough to obtain true ground state properties. Testing this idea for lattices should be tried.

Problems involving complex phases were already mentioned in chapter 4. For these problems in lattice systems, no practical calculations along FNMC-lines have been performed yet. The same is true for the extension of the lattice fixed-node idea to finite temperatures. A continuum method for finite temperatures, which involves fixing the nodes of the *density matrix*, has been developed, and applied to $^3He$, by Ceperley [4].

Apart from these possible new applications, gaining insight in the physical meaning of the signs of the wave function would be desirable, since it is the most impor-

tant ingredient of the FNMC method. The specific effect of the sign structure can be studied by FNMC, because this method allows other correlations to change in the projection process, while the signs are fixed.

Especially for fermions, little is known about the signs of the wave function [5]. It is clear that for a metallic phase, changes of sign occur on a length scale set by $k_F$ [5]. It is also known that the half-filled Hubbard model will, for large $U$, behave as a Heisenberg model. We shall see that for this case, the signs are also known. It would be interesting to know the evolution of the sign structure as a function of $U$. This is left for future research, here we concentrate on what is known about signs of wave functions of lattice models.

The few known general theorems for continuum systems need to be reconsidered for models defined on lattices. Boson and spin systems with competing interactions do not always have a ground state wave function of one sign only. In this chapter, we use this to illustrate that the ground state wave function of lattice fermions with competing hopping terms will often not obey the *tiling property* [9] which we shall explain in the next section.

In recent years, it has become clear that frustrated quantum spin systems can, depending on the interactions, have a sign problem. This is illustrated e.g. by the square lattice Heisenberg model with both antiferromagnetic nearest-neighbour interaction $J_1$ and next nearest neighbour interaction $J_2$. As discussed in chapter 1, for $J_2 = 0$, the ground state of the nearest neighbour Heisenberg model is known to have antiferromagnetic order [6]; for $J_1 = 0$ this model decomposes into two sublattices which each are ordered antiferromagnetically. When $J_1$ and $J_2$ are of the same order of magnitude, these two types of order strongly compete; it appears from a large number of recent studies [7] that in between the large $J_1$ and large $J_2$ ordered phases this frustration gives rise to a disordered state over some range of interactions $J_2/J_1$. The ground state wave function in this regime almost certainly has nontrivial signs as a result of a violation of the so-called Marshall-Peierls sign rule [8].

As we will discuss later, the Heisenberg model with nearest neighbour and next nearest neighbour interactions can be mapped onto a model of hard-core bosons with nearest neighbour and next nearest neighbour hopping terms. The recent results quoted above for the Heisenberg model therefore imply that this frustrated boson lattice model has a ground state wave function with nontrivial signs, which underscores the possible difference between lattice bosons and bosons in continuum space, whose ground state wave function is always nodeless [9].

In section 6.1, we use the above-mentioned results for the frustrated Heisenberg model or the equivalent boson model to shed light on the difference in sign structure of lattice models and continuum models. In particular, both the positiveness of the ground state wave function of bosons and the corresponding tiling property of fermions, which means that there are no more nodes than required by fermion antisymmetry, only hold on a lattice if the off-diagonal terms are of one sign. We also discuss how doping affects the usual Marshall-Peierls sign rule for a spin system.

## 6.1 The consequences of frustration for the sign structure

In continuum problems, the kinetic term in the Hamiltonian is quasi-local, i.e., it only depends on spatial derivatives, not on differences between values of the wave function in points finite distances apart. Moreover, the sign of this kinetic term is fixed. A lattice model, either for bosons or for fermions, has a discrete configuration space and the simplest kinetic term is

$$\mathcal{H} = -t \sum_{i\delta} \left( c^\dagger_{i+\delta} c_i + h.c. \right), \tag{6.1.1}$$

where $\delta$ runs over all nearest neighbours of $i$ and where we take $t > 0$. We focus on the off-diagonal terms, because the (potential) terms which are diagonal in this representation do not have an effect on the signs of the wave function. The hopping term can be rewritten:

$$\begin{aligned} -t \sum_{i\delta} \left( c^\dagger_{i+\delta} c_i + h.c. \right) &= t \sum_{i\delta} \left( c^\dagger_{i+\delta} - c^\dagger_i \right) \cdot \left( c_{i+\delta} - c_i \right) - t \sum_{i\delta} \left( c^\dagger_{i+\delta} c_{i+\delta} + c^\dagger_i c_i \right) \\ &= t \sum_{i\delta} \left( c^\dagger_{i+\delta} - c^\dagger_i \right) \cdot \left( c_{i+\delta} - c_i \right) - 2tz \sum_i c^\dagger_i c_i, \end{aligned} \tag{6.1.2}$$

where $z$ is the number of neighbours. The operator $t \sum_{i\delta} \left( c^\dagger_{i+\delta} - c^\dagger_i \right) \cdot \left( c_{i+\delta} - c_i \right)$ is positive definite, and can be thought of as the discrete version of the kinetic energy operator in a continuum. With the kinetic part written in this form, it is clear that the ground state wave function will be as smooth as possible, as in the continuum case. For bosons, this means that the ground state wave function will be of one sign: for given absolute values of the wave function, the kinetic term is minimized if we take the wave function of which all signs are equal [9]. In the appendix, we follow a somewhat different reasoning to obtain the proof of a more general theorem, which contains this statement about the bosonic ground state. For fermions, the wave function must be antisymmetric. That the ground state wave function must be as smooth as possible now implies that the *tiling property* [5] holds: in the presence of a single hopping term like Eq.(6.1.1), there are *no more* nodes than required by fermion antisymmetry.

The above-mentioned properties of the ground state wave function of lattice boson and fermion systems, which hold if there is one hopping term only, are the same as in the continuum case. On a lattice, there can be competing off-diagonal terms. Let us first discuss the consequences for quantum spins. Consider the antiferromagnetic $J_1 - J_2$ spin-1/2 Heisenberg model on a bipartite lattice:

$$\mathcal{H} = J_1 \sum_{\langle ij \rangle} \vec{S}_i \cdot \vec{S}_j + J_2 \sum_{[ij]} \vec{S}_i \cdot \vec{S}_j, \tag{6.1.3}$$

where $\langle ij \rangle$ denotes nearest neighbour pairs and $[ij]$ denotes next-nearest neighbour pairs. If $J_2 = 0$, we have the Heisenberg model, for which the so-called Marshall-Peierls sign rule holds [10, 11] . It is derived as follows. Define an Ising state $|\mu\rangle$ as a

configuration of spins. If one writes the ground state as

$$|\psi_0\rangle = \sum_{\mu} C_\mu |\mu\rangle, \tag{6.1.4}$$

then the variational energy becomes (up to a constant term)

$$E_{var} = \frac{1}{2} J_1 \sum_{\mu} C_\mu \sum_{\mu'}{}' (C_{\mu'} - C_\mu). \tag{6.1.5}$$

The second summation is over all configurations that can be reached from $|\mu\rangle$ with one simultaneous flip on neighbouring sites. Clearly, for given $|C_\mu|$ and $|C_{\mu'}|$, and $J_1 > 0$, the lowest value is obtained if $C_\mu$ and $C_{\mu'}$ have opposite [illegible], there is a consistent sign rule for the ground state: if one assigns a fact[illegible] one for each down spin on one sublattice, one has all $C_\mu$ and $C_{\mu'}$ of states co[illegible] through a single hop of opposite sign. The proof breaks down if the lattice [illegible]partite, or if there is a competing, further-range interaction, e.g. a next-n[illegible]ghbour interaction with $J_2 > 0$ in Eq.(6.1.3).

The spin-1/2 model (6.1.3) can be mapped on a lattice boson m[illegible] by taking $S_i^+ = b_i^\dagger, S_i^- = b_i$ and $S_i^z = n_i - \frac{1}{2}$ on one sublattice and $S_i^+ = -b_i^\dagger, S_i^- = -b_i$, $S_i^z = n_i - \frac{1}{2}$ on the other. Here $b_i^\dagger$ and $b_i$ are the creation and annihilation operators of hard-core bosons, i.e. in addition to the usual boson commutation relations they satisfy $\left(b_i^\dagger\right)^2 = 0$, $(b_i)^2 = 0$. With this transformation the Heisenberg model becomes

$$\begin{aligned} \mathcal{H} &= -\frac{J_1}{2} \sum_{\langle ij \rangle} \left( b_i^\dagger b_j + b_i b_j^\dagger \right) + \frac{J_2}{2} \sum_{[ij]} \left( b_i^\dagger b_j + b_i b_j^\dagger \right) + \\ &+ J_1 \sum_{\langle ij \rangle} \left( n_i - \frac{1}{2} \right) \left( n_j - \frac{1}{2} \right) + J_2 \sum_{[ij]} \left( n_i - \frac{1}{2} \right) \left( n_j - \frac{1}{2} \right). \end{aligned} \tag{6.1.6}$$

This simple transformation immediately illustrates the following points:

1. For $J_2 = 0$, the Marshall-Peierls sign rule is equivalent to the earlier observation that for a lattice boson Hamiltonian with a simple hopping term of type (6.1.1), the ground state wave function is of one sign only.

2. If there is a $J_2 > 0$, for which in the spin formulation the Marshall-Peierls sign rule can not be proven to hold, the corresponding boson model has competing hopping terms, such that the ground state wave function can not be proven to be positive. One term favours smoothness of the wave function, the other term favours large variations. Even stronger, the results of Retzlaff *et al.* [8] demonstrate that in a range $J_2/J_1$ near 1, the boson ground state wave function *does* have nontrivial signs, and that Monte Carlo simulations of (6.1.3) and (6.1.6) in this range will suffer from a sign problem.

3. The existence of a lattice boson ground state wave function with nontrivial signs shows that there is every reason to believe that fermion lattice models with competing off-diagonal hopping terms generally will have more domains of different sign of the wave function in configuration space, than dictated by fermion antisymmetry. In other words, we generally expect the tiling property *not* to hold in models with competing hopping terms, such as e.g. the $t - t' - J$ model [12].

Note that additional terms in a quantum spin model can lead to exactly solvable models if they are precisely tuned. We mentioned this in chapter 1. For these models, which have valence bond states as exact ground state, one knows all signs. Since the separate terms are rather independent, there is some freedom to choose the signs in different terms. However, the generic effect of adding terms is introducing frustration, leading to a spin-liquid state of unknown sign structure.

In chapter 1, we discussed that both frustration and doping affect the magnetic order. In addition to the treatment of the consequences of frustration for the sign structure, we consider the consequences for this property of the other deviation from the simplest Heisenberg model.

Doping one hole into a $t$-$J_z$ model poses no problem. The sign of the Ising state is inconsequential for the magnetic energy. For the hopping term, choosing all prefactors positive yields the lowest energy. As discussed in chapter 3, the $t$-$J_z$ model is a simple model in which inhomogeneous solutions seem to occur.

Doping holes into a *quantum* spin model is more complicated, as far as the signs are concerned. In one dimension, a consistent sign rule can be devised [13]. In addition to factors $i$ attributed to spins that deviate from the spin direction they should have on their sublattice, it involves factors $i$ for empty sites. In higher dimensions, it is not possible to find a consistent sign rule [14].

In the $t - t' - J$ model, Gooding *et al.* [12] argue that hole doping and electron doping correspond to different sign relations among the interaction parameters. For the *single-particle* dispersion, they find that the minimum is at a different $\vec{k}$-vector for electron doping than for hole doping. Since the wave-vector determines the signs, this is a clear example of the fact that the signs in the Hamiltonian determine the signs in the wave function.

Apart from the abovementioned simple cases, little is known about the sign structure of doped magnets.

## 6.2 Conclusions

We have studied the consequences of the fundamental non-locality of the off-diagonal terms in lattice problems on the sign structure of many-particle wave functions. We have demonstrated that for hoppings of different sign, which can be present on a lattice, the simplest sign rules for bosons (the ground state wave function is positive), spins (Marshall-Peierls sign rule) and fermions (tiling property) cannot be proven to

hold. The antiferromagnetic Heisenberg model with next nearest neighbour interactions provides an example where these sign rules indeed do *not* hold. We indicated how lattice FNMC calculations can be performed, which avoid the sign problem. This method leads to upper bounds for the energy, for each sign structure that is put in. So, this method not only avoids cancellations, and therefore may be a good method to perform Monte Carlo simulations in highly frustrated spin problems, it is also able to separate the problem of finding the correct sign structure from finding other correlations. The extent to which this approach is practical will have to be tested on concrete examples.

## Appendix: A lemma on the signs on a lattice

In this appendix, we illustrate the complications due to the discreteness of configuration space for lattice problems on a lemma for the continuum case due to Lieb and Mattis [10]:

If there are two functions $\psi_1$ and $\psi_2$ such that the nodal surface of $\psi_1$ encloses a region $R_1$ which neither contains, nor is intersected by any nodal surface of $\psi_2$, then $E_1 > E_2$.

In the continuum case, it is immediately clear what 'nodal surface' means: the surface of co-dimension one on which the wave function vanishes. The proof of the lemma [10] uses the form of the kinetic energy operator in the continuum case and Green's theorem. For the lattice case, we need to phrase the issue in a different language.

In general, there are points with opposite signs in the wave function which are connected by hops in the Hamiltonian. The 'nodal surface' is only defined as 'in between those points', there is no location of the node on the lattice of configuration points itself (except if the wave function vanishes on a point $R$). Because of the discreteness, the problem is non-local.

Let us assume that the Hamiltonian only contains nearest-neighbour hopping of one sign $(-t)$ and find the precise equivalent of the lemma for the lattice case.

Consider the Schrödinger equation: the eigenfunction $\psi_1$ satisfies

$$(\mathcal{H}\psi_1)(R) = E_1\psi_1(R), \tag{6.2.1}$$

and similarly for $\psi_2$. $(\mathcal{H}\psi_1)(R)$ satisfies

$$(\mathcal{H}\psi_1)(R) = -t \cdot \sum_{R'}{}' \psi_1(R'), \tag{6.2.2}$$

where $\sum'_{R'}$ means summing over all $R'$ that can reach $R$ in one hop.

Now multiply Eq.(6.2.1) by $\psi_2(R)$, and the corresponding equation for $\psi_2(R)$ by $\psi_1(R)$. The resulting equations are subtracted. This yields

$$-t\psi_2(R)\cdot\sum_{R'}{}'\psi_1(R')+t\psi_1(R)\cdot\sum_{R'}{}'\psi_2(R')=(E_1-E_2)\cdot\psi_1(R)\cdot\psi_2(R). \qquad (6.2.3)$$

Now sum both sides over $R\in R_1$ and then split the sum over $R'$ on the left hand side in a sum over $R'\in R_1$ and a sum over $R'\notin R_1$, these are the points that can be reached from $R_1$ by one hop over the boundary:

$$\begin{aligned} & t\sum_{R\in R_1}\left[\sum_{R'\in R_1}{}'\psi_1(R)\cdot\psi_2(R')-\psi_2(R)\cdot\psi_1(R')\right]+ \\ & +t\sum_{R\in R_1}\left[\sum_{R'\notin R_1}{}'\psi_1(R)\cdot\psi_2(R')-\psi_2(R)\cdot\psi_1(R')\right] \\ & =0+t\sum_{R\in R_1}\left[\sum_{R'\notin R_1}{}'\psi_1(R)\cdot\psi_2(R')-\psi_2(R)\cdot\psi_1(R')\right] \\ & =(E_1-E_2)\cdot\sum_{R\in R_1}\psi_1(R)\cdot\psi_2(R). \end{aligned} \qquad (6.2.4)$$

Since by assumption $\psi_1(R)\cdot\psi_2(R)>0$, it is clear that $(E_1-E_2)>0$, if for all $(R,R')$ at the boundary

$$\psi_1(R)\cdot\psi_2(R')>\psi_2(R)\cdot\psi_1(R'). \qquad (6.2.5)$$

So the equivalent of the abovementioned lemma for the lattice is true if there is only nearest neighbour hopping of negative sign, and if the nodes of $\psi_2$ are outside those of $\psi_1$. This means that either $\psi_2(R')>0$, or the node of $\psi_2$ is in between the same pair of points as the node of $\psi_1$, but further outside the region $R_1$ in the linear interpolation sense:

$$\frac{\psi_1(R)}{\psi_1(R')}>\frac{\psi_2(R)}{\psi_2(R')}. \qquad (6.2.6)$$

# References


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

[2] T. Munehisa and Y. Munehisa, Phys. Rev. B **49**, 3347 (1994)

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

[4] D.M. Ceperley, Phys. Rev. Lett. **69**, 331 (1992)

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

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

[7] See e.g. A.V. Dotsenko and O.P. Sushkov, Phys. Rev B **50**, 13821 (1994). This paper contains an extensive list of references to the earlier literature.

[8] K. Retzlaff, J. Richter and N.B. Ivanov, Z. Phys. **B93**, 21 (1993); N.B. Ivanov and J. Richter, J. Phys. Condens. Matter **6**, 3785 (1994). See also J. Richter, N.B. Ivanov and K. Retzlaff , cond-mat preprint 9407041.

[9] R.P. Feynman, *Statistical Mechanics*, Benjamin (1972)

[10] D.C. Mattis, *The Theory of Magnetism I*, Springer, Berlin (1981), section 4.11.

[11] W.J. Caspers, *Spin Systems*, World Scientific, Singapore (1989)

[12] See e.g. R.J. Gooding, K.J.E. Vos, and P.W. Leung, Phys. Rev. B **50**, 12866 (1994) and references therein.

[13] Z.Y. Weng, Phys. Rev. B **50**, 13837 (1994)

[14] Z.Y. Weng, Y.C. Chen, and D.N. Sheng, cond-mat preprint 9505124 (1995)

# Samenvatting

In deze samenvatting wordt zonder formules uiteengezet waar dit proefschrift over gaat. Eerst bespreken we de termen in de titel, daarna gaan we in op de ideeën achter de verschillende hoofdstukken.

De vertaling van de titel van dit proefschrift is 'Reële-ruimte-aspecten van gecorreleerde fermionen'. Fermionen zijn deeltjes die voldoen aan het Pauli-uitsluitingsprincipe. Dit houdt in dat twee fermionen nooit tegelijkertijd al hun eigenschappen gemeen kunnen hebben. Ze kunnen dus niet op hetzelfde moment dezelfde spin hebben (de 'spin' geeft aan of de energie hoger of lager wordt als een magneetveld wordt aangelegd, de spin van een fermion kan 'op' of 'neer' zijn) én op dezelfde positie zitten.

In een metaal zitten heel veel elektronen. Dat zijn fermionen. De experimenteel gevonden eigenschappen van de meeste metalen kun je goed beschrijven als je ervan uitgaat dat er fermionen zijn, dat die beïnvloed worden door het rooster van ionen en dat ze aan bovengenoemd uitsluitingsprincipe voldoen, maar dat die fermionen verder in essentie geen krachten op elkaar uitoefenen. Dit wordt het 'Fermi-vloeistofbeeld' genoemd. Voor sommige nieuwe, vreemde metalen gaat dat niet meer op. Bij temperaturen die zo hoog zijn dat supergeleiding niet meer optreedt, gedragen de materialen die als 'hoge-temperatuur-supergeleiders' worden aangeduid zich anders dan gewone metalen. Dit gedrag heet 'Niet-Fermi-vloeistof-gedrag'. Het abnormale gedrag wordt veroorzaakt door interacties tussen de fermionen. De naam 'gecorreleerde fermionen' wordt gebruikt omdat de fermionen zich, door de aanwezigheid van interacties, als collectief, ofwel 'gecorreleerd', gedragen. In een berekening aan een model voor gecorreleerde fermionen kun je de fermionen niet één voor één behandelen, maar moet je de eigenschappen van alle fermionen tegelijk bepalen. Dat is de grote moeilijkheid.

In hoofdstuk 1 wordt uiteengezet wat bekend is over twee extreme gevallen. Het verschil tussen die twee geeft ook aan wat wordt bedoeld met 'reële-ruimte-aspecten'. Voor beide gevallen ga je uit van een rooster van posities (ionen) waarop de fermionen kunnen zitten. Volgens het Pauli-uitsluitingsprincipe kunnen per positie vier situaties bestaan. De positie is leeg, er is een fermion met spin op, er is een fermion met spin neer, of er zijn twee fermionen, één met spin op en één met spin neer.

Het ene extreme geval is dat de laatste mogelijkheid vervalt omdat er een heel grote afstotende interactie is tussen fermionen op één en dezelfde positie. Als er nu *gemiddeld* één fermion per positie is, zal er in deze limiet van grote afstoting ook echt op elke positie precies één fermion zitten. Alleen de spinrichting kan variëren. Zulke 'quantum-spinsystemen' zijn het gemakkelijkst te beschrijven in termen van de toestand per positie (spin op of spin neer) in de reële ruimte. Een belangrijke vraag is nu of deze spins een gemiddelde magnetisatie zullen hebben.

In het andere extreme geval zijn interacties onbelangrijk, en telt alleen dat fermionen niet vastzitten op hun positie, maar naar naburige posities kunnen springen. Dit

probleem is het best te beschrijven door een Fourier-transformatie uit te voeren, dat wil zeggen een beschrijving te nemen in termen van *golven* die op het rooster passen. De term 'reële ruimte' vormt dus een tegenstelling met 'Fourier- of golvenruimte'.

Je kunt een systeem van gecorreleerde fermionen krijgen als je uitgaat van het laatste geval van onafhankelijke fermionen, door de interacties te laten toenemen. Je krijgt ook een systeem van gecorreleerde fermionen als je uitgaat van quantum-spins en dan een kleiner totaal aantal fermionen dan één per positie neemt óf de afstotende interactie laat afnemen. In beide gevallen kunnen de fermionen bewegen. Er zijn twee vragen. De eerste is: "Als het quantum-spinsysteem magnetische ordening vertoonde, blijft dat dan zo als er minder fermionen zijn dan posities?". De tweede vraag is: "Gedragen de fermionen zich nog op soortgelijke wijze als onafhankelijke fermionen, als je interacties meeneemt in je beschrijving?". Deze kwesties vormen de achtergrond voor ons eigen onderzoek.

In hoofdstuk 2 gaan we dieper in op het Fermi-vloeistofbeeld. We bouwen een soort experiment waarin op één bepaalde plek plotseling een verstoring wordt aangebracht. Het bijzondere is dat *alle* fermionen dat merken. Dat komt omdat een *lokale* verstoring invloed heeft op alle Fourier-modes, ofwel op alle golven, en dat is precies de beschrijving van onafhankelijke fermionen: elk heeft zijn eigen Fourier-mode. Dat dit leidt tot singulier gedrag, zichtbaar in het experiment niet de lokale verstoring, is een oud resultaat in de Fermi-vloeistoftheorie.

Nu beschouwen we ook *marginale Fermi-vloeistoffen*. Deze term duidt een fenomenologische beschrijving van geleidingselektronen aan, waarin de toestanden met een bepaalde energie wel van dezelfde soort zijn als in een Fermi-vloeistof, maar waarin deze toestanden niet *lang* bestaan: ze hebben een eindige levensduur. Deze aanname wordt niet onderbouwd door een model met een specifieke interactie tussen de elektronen. De aanname geeft een goede beschrijving van onder andere de geleiding van de metallische fase, dat is de fase boven de supergeleidende overgangstemperatuur, van hoge-temperatuur-supergeleiders. De vraag is nu of een marginale Fermi-vloeistof zich wat betreft de reactie op een lokale verstoring anders gedraagt dan een gewone Fermi-vloeistof. We verduidelijken een door anderen gemaakte observatie, namelijk dat in het marginale geval de snelheid waarmee een zwaar, geladen deeltje door het materiaal loopt naar nul gaat voor lage temperaturen. We doen dat door dit specifieke experiment te beschouwen als speciaal geval van een grotere klasse van experimenten. Het blijkt dat het spectrum van een met Röntgenstraling uit het materiaal geschoten elektron *verschuift* als de temperatuur daalt, als de geleidingselektronen een marginale Fermi-vloeistof vormen. In tegenstelling tot het gewone Fermi-vloeistof geval verandert de *vorm* van het spectrum niet erg als de temperatuur verandert.

Hoofdstuk 3 gaat over spinpolaronen. Als er niet één fermion per positie in het rooster is, kunnen de lege posities zich verplaatsen en zo de magnetische ordening van de omgeving beïnvloeden. Een oude, niet erg goed onderbouwde verklaring van enkele metingen in vast helium ($^3He$) is in termen van ferromagnetische ('alle spins op') gebiedjes rond een lege positie, die in een antiferromagnetische ('de spins om en om op en neer') achtergrond zitten. Een lege positie met zijn ferromagnetische gebiedje heet

een spinpolaron. Een essentieel ingrediënt in de verklaring van onder andere susceptibiliteitsmetingen is dat er meer lege posities zijn dan men in een evenwichtstoestand zou verwachten. Dat zou komen doordat er bij hoge temperaturen veel lege posities zijn, die, als je gaat afkoelen, elk hun eigen ferromagnetisch gebiedje vormen, dat ze vanaf dat moment moeten meeslepen. De lege posities worden zwaar en langzaam, en kunnen het materiaal niet verlaten. Onze bijdrage aan de discussie of dit het juiste scenario is, is een idee voor een experimentele test. Stel dat een stukje materiaal de bijzondere eigenschappen vertoont die worden toegeschreven aan een overmaat aan spinpolaronen. Ons idee is het dan gedurende enige tijd bloot te stellen aan een sterk magneetveld. De lege plekken kunnen dan bewegen en het materiaal verlaten. Dan zou het veld weer uitgeschakeld moeten worden. Daarna zou het materiaal zich weer normaal gedragen, als spinpolaronen inderdaad een belangrijke rol speelden bij het oorspronkelijke abnormale gedrag van het betreffende stukje materiaal.

In hoofdstuk 4 beschrijven we de theoretische fundamenten van het belangrijkste onderwerp van dit proefschrift: de vaste-nulpuntenbenadering in de quantummechanische Monte Carlo-methode voor het berekenen van grondtoestandseigenschappen van modellen van gecorreleerde fermionen. De grondtoestand is de toestand met de laagste energie. Het is moeilijk deze te vinden omdat de Schrödinger-vergelijking voor alle deeltjes tegelijk moet worden opgelost, en er bovendien voor fermionen een beperkende voorwaarde is, namelijk dat de oplossing antisymmetrisch moet zijn. Dat laatste betekent dat de oplossing dezelfde absolute waarde en tegengesteld teken heeft als twee deeltjes worden verwisseld. Met Monte Carlo-methoden bereken je de eigenschappen voor een zo groot mogelijk aantal deeltjes. Eigenlijk moet je de toestand berekenen in een hoog-dimensionale ruimte, waarin de posities van *alle* deeltjes worden gespecificeerd. Je kunt numeriek maar een beperkt aantal punten behandelen. Het blijkt efficiënt te zijn deze punten met behulp van random getallen te kiezen, de naam Monte Carlo is gekozen vanwege de analogie met het casino.

We concentreren ons nu op het vaste-nulpuntenprincipe. Een Monte Carlo-berekening begint met een benaderde oplossing. Daarna wordt in principe de echte grondtoestand bepaald. Voor fermionen is er een probleem. Alle getallen die je wilt weten krijg je als het verschil tussen twee heel grote getallen. De antwoorden zelf zijn klein. Deze gang van zaken leidt tot grote onnauwkeurigheden in de antwoorden. Omdat het gaat om het van elkaar aftrekken van twee getallen heet dit het 'tekenprobleem'.

Met het vaste-nulpuntenprincipe maak je een benadering, zó dat het tekenprobleem wordt omzeild. Je zegt dat de oplossing die je zult vinden positief is als de benaderde oplossing positief is, en negatief als de benaderde oplossing negatief is. De absolute waarde mag veranderen, en kan dus verbeteren, dat wil zeggen, vanaf het startpunt dichter bij de echte grondtoestand komen. Je antwoord is niet langer exact als je het vaste-nulpuntenprincipe gebruikt, want je hebt de klasse van mogelijke oplossingen beperkt, maar het antwoord is wel beter dan het startpunt. Als je het zo doet, blijkt het tekenprobleem niet op te treden.

Voor continue systemen bestond er een implementatie van het vaste-nulpuntenprincipe, die aan een belangrijke eis voldoet. Die eis is dat de energie die gevonden

wordt met de gemaakte benadering *boven* de werkelijke grondtoestandsenergie ligt. Je wilt dat omdat je zoekt naar de toestand met de laagste energie, en je geen lage waarde voor de energie wilt vinden *alleen door de benadering die je maakt*. Als je een variationeel principe hebt, dat wil zeggen als de gevonden waarde zeker boven de werkelijke grondtoestandsenergie ligt, weet je dat als je een lage energie vindt, de bijbehorende toestand ook werkelijk de grondtoestand goed benadert.

Voor roosterproblemen bestond geen Monte Carlo-methode die aan de ene kant door het maken van een benadering het tekenprobleem omzeilt en die aan de andere kant variationeel is. De moeilijkheid schuilt in het feit dat de waarde van een functie op een rooster van teken kan wisselen zonder *op een roosterpunt* de waarde nul aan te nemen. Vanwege het discrete karakter gaat de waarde soms direct van positief naar negatief, op twee naburige punten. In hoofdstuk 4 bespreken we hoe deze punten moeten worden behandeld. Door sprongen van fermionen te vervangen door *potentialen*, dat wil zeggen extra bij de energie van een configuratie van fermionen op te tellen getallen, simuleren we een *effectief* model, waarvan kan worden aangetoond dat de energie ervan in *elke* toestand hoger is dan de energie van het oorspronkelijke model. Onze implementatie is dus variationeel.

Hoofdstuk 5 bevat een test van deze methode. We passen onze nieuwe 'quantummechanische Monte Carlo-methode met vaste-nulpuntenprincipe voor roosterfermionsystemen' toe op het ééndimensionale Kondo-roostermodel. Dit is een model voor twee soorten met elkaar interagerende elektronen, die we $c$ en $f$ noemen. Naast sprongen van de $c$-elektronen omvat dit model ook gelijktijdige veranderingen van de spin van twee elektronen, dat wil zeggen dat een $c$-elektron met spin op en een $f$-elektron met spin neer die samen op een roosterpositie zitten, kunnen overgaan naar $c$-neer, $f$-op. Dit model wordt gebruikt voor het beschrijven van een klasse van materialen die als *heavy fermions* wordt aangeduid. Voor kleine roostertjes zijn exacte antwoorden bekend, waar we de resultaten van onze methode mee vergelijken.

In hoofdstuk 6, tenslotte, bespreken we wat bekend is over de tekens van toestanden en speciaal wat er bijzonder is aan roosters. Omdat deze eigenschap van toestanden een essentieel ingrediënt is van de vaste-nulpuntenmethode, is kennis hierover van belang voor de vraag in hoeverre de methode gebruikt kan worden. Eén van de conclusies is dat de roosterversie van de vaste-nulpuntenmethode, die ontwikkeld is met roosterfermionen in het achterhoofd, ook van nut kan zijn voor het bestuderen van andere systemen waarin het tekenprobleem optreedt, met name voor modellen van gefrustreerde quantum-spins, waarbij 'gefrustreerd' wil zeggen dat verschillende termen in het model een verschillende magnetische ordening bevorderen.

Laten we tenslotte terugkeren naar de titel. We hebben de afzonderlijke termen behandeld, nu moet het verband duidelijk zijn. De verzamelnaam 'Reële-ruimte-eigenschappen van gecorreleerde fermionen' is gekozen omdat we in hoofdstuk 2 *lokale* verstoringen behandelen, in hoofdstuk 3 gelocaliseerde structuren zoals spinpolaronen, en in de resterende hoofdstukken een vorm van Monte Carlo die zich afspeelt in de configuratieruimte, waarin van alle deeltjes de positie in de reële ruimte is gespecificeerd.

# Curriculum Vitae

Ik ben op 9 juli 1968 geboren te Nijmegen. In 1986 heb ik het VWO-diploma behaald aan het 'Nederrijn College' te Arnhem. Van 1986 tot 1991 heb ik theoretische natuurkunde gestudeerd aan de Katholieke Universiteit Nijmegen. Als student-assistent heb ik werkcolleges verzorgd in de natuurkunde, voornamelijk mechanica, voor eerstejaars studenten scheikunde. In augustus 1991 ben ik afgestudeerd bij de afdeling 'Theoretische Vaste-stoffysica', op een scriptie over de optica van incommensurabele kristallen. Het werk was deels theoretisch, deels experimenteel.

In september 1991 begon ik als assistent in opleiding aan het Instituut-Lorentz voor theoretische natuurkunde. Mijn bijdrage aan het onderwijs bestond uit het verzorgen van opgaven en uitwerkingen bij het vak 'Theorie van de gecondenseerde materie'. De scholen en conferenties die ik heb bijgewoond zijn: Winterschool Supergeleiding, Limeland 1991, Enrico Fermi zomerschool, Varenna (Italië) 1992, Landau zomerschool, Chernogolovka (Rusland) 1993, met voordracht 'Spin polarons in solid $^3He$', ICTP Spring College on Quantum Phases, Triëst (Italië) 1994, met voordracht 'Fixed-node Quantum Monte Carlo Method for Lattice Fermions', en tenslotte de APS March Meeting, San Jose (USA) 1995, met voordracht 'Fixed-node Quantum Monte Carlo Method for Lattice Fermions: application to the 1D KLM'.

Het in de AIO-periode gedane onderzoek is beschreven in dit proefschrift. Verder heb ik in Leiden een discussiegroep voor studenten en promovendi in de theoretische natuurkunde opgericht en gedurende anderhalf jaar gecoördineerd. Als voorbereiding op het leraarschap heb ik de cursus 'Leren en Communiceren' gevolgd.

# List of publications

1. H.J.M. van Bemmel and W. van Saarloos, *Temperature dependence of the x-ray photoemission line shape and of the hopping rate in a marginal Fermi liquid*, J. Phys.: Condens. Matt. **5**, 1377 (1993)

2. H.J.M. van Bemmel and W. van Saarloos, *Spin polarons in solid $^3He$: Suggestions for further experiments*, J. Low Temp. Phys. **94** , 551 (1994)

3. H.J.M. van Bemmel, D.F.B. ten Haaf, W. van Saarloos, J.M.J. van Leeuwen, and G. An, *Fixed-node quantum Monte Carlo method for lattice fermions*, Phys. Rev. Lett. **72**, 2442 (1994)

4. 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)

5. 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.

# STELLINGEN

behorende bij het proefschrift

*Real-space aspects of correlated fermions*

van H.J.M. van Bemmel

1. De in [1] gebruikte implementatie van het vaste-nulpuntenprincipe in Monte Carlo-simulaties van systemen van fermionen op een rooster geeft een bovengrens voor de grondtoestandsenergie.

   [1] *G. An en J.M.J. van Leeuwen, Phys. Rev. B* **44**, *9410 (1991)*

2. Bij het bestuderen van de mogelijkheid dat niet-halfgevulde systemen van gecorreleerde fermionen *domeinwanden* vormen is het noodzakelijk de met stochastische technieken uit een homogene gemiddeld-veldoplossing geprojecteerde toestand te onderzoeken op de aanwezigheid van domeinwandcorrelaties.

3. Ook als de symmetrie van een gemoduleerd kristal voorschrijft dat de tensor $\gamma_{ijk}$, die conventionele optische activiteit beschrijft, overal nul is, kan in principe optische draaiing optreden. Deze is het gevolg van de plaatsafhankelijkheid van de diëlektrische tensor en zal een dispersie vertonen die evenredig is met de dispersie van de dubbelbreking in het kwadraat vermenigvuldigd met $(\lambda_{\text{modulatie}}/\lambda_{\text{licht}})^2$.

4. Het is mogelijk analytische berekeningen van de subtiele eigenschappen van golffuncties van niet-interagerende, springende elektronen op een quasi-periodiek ééndimensionaal rooster of op een fractaal rooster met niet-gehele fractale dimensie aan te vullen met numerieke renormalisatiegroepberekeningen in de reële ruimte. De dichtheidsmatrix-formulering met overlappende blokken in [2] is het meest geschikt voor de berekening van de bandenstructuur van fermionen, ook op anomale roosters.

   [2] *S.R. White en R.M. Noack, Phys. Rev. Lett.* **68**, *3487 (1992)*

5. De in de reële ruimte geformuleerde argumenten voor de afwezigheid van magnetische ordening zoals beschreven in [3] zijn, in tegenstelling tot het oorspronkelijke argument van Mermin en Wagner [4], ook toepasbaar op spinsystemen gedefinieerd op fractale roosters.

   [3] *G. Parisi,* Statistical field theory, *Addison-Wesley, Redwood City, (1988)*

   [4] *N.D. Mermin en H. Wagner, Phys. Rev. Lett.* **17**, *1133 (1966)*

6. Een antiferromagnetisch Heisenberg-model van $S = 1$ quantum-spins, gedefinieerd op een bipartiete Koch-curve met spectrale dimensie tussen één en twee, vertoont geen magnetische ordening in de grondtoestand, in tegenstelling tot het tweedimensionale geval. De vraag of quantumfluctuaties leiden tot de aanwezigheid van een Haldane-gap in het spectrum, zoals het geval is in één dimensie, kan in principe worden beantwoord met behulp van numerieke dichtheidsmatrix-renormalisatiegroepberekeningen.

7. De traditionele vorm van doceerles en hoorcollege maakt de student tot een passieve consument.

8. In een lijst met aankondigingen van oraties zou ofwel in alle gevallen voor de academische titels 'De Heer' of 'Mevrouw' moeten staan, ofwel elke aanduiding van het geslacht van de hoogleraar achterwege moeten blijven.

9. Spreekwoorden zijn ongenuanceerd.

10. De betekenis van 'verdubbeling' van woorden is verschillend voor zelfstandig en bijvoeglijk naamwoorden. 'Oud-oud' is extreem oud, maar als het antwoord op de vraag 'Welke kaas?' luidt 'kaas-kaas', dan wordt de meest gangbare vorm van kaas bedoeld, namelijk jong belegen.

Hans van Bemmel
27 september 1995

RIJKSUNIVERSITEIT TE LEIDEN
BIBLIOTHEEK INSTITUUT-LORENTZ
Postbus 9506 - 2300 RA Leiden
Nederland