Massively parallel exact diagonalization of strongly correlated systems

von

Andreas Dolfen

Diplomarbeit in Physik

vorgelegt der

Fakultät für Mathematik, Informatik und Naturwissenschaften
der Rheinisch-Westfälischen Technischen Hochschule Aachen

im
Oktober 2006

angefertigt am
Institut für Festkörperforschung (IFF)
Forschungszentrum Jülich

communicated by Prof. Dr. Koch

Contents

1 Introduction **1**
- 1.1 The setting . . . 1
- 1.2 Physics of strong correlations . . . 3
- 1.3 Methods . . . 4
- 1.4 Overview . . . 5

2 Hubbard model **7**
- 2.1 The tight-binding approximation (TBA) . . . 7
  - 2.1.1 Theoretical derivation . . . 7
  - 2.1.2 From the TBA to the kinetic energy term of the Hubbard model 9
  - 2.1.3 Features in low dimension . . . 10
    - Density of states . . . 11
    - Two orbitals per cell . . . 12
  - 2.1.4 Many independent particles . . . 15
    - Finite size scaling . . . 16
- 2.2 From the tight-binding approximation to the Hubbard model . . . 18
  - 2.2.1 Important features of the Hubbard model . . . 20
    - Atomic limit with half-filling . . . 21
    - Mott transition . . . 21
    - Distribution of eigenenergies . . . 21
  - 2.2.2 Gutzwiller wave function . . . 24

3 Exact diagonalization **29**
- 3.1 Basis encoding and sparsity . . . 29
- 3.2 The power method . . . 32
- 3.3 The Lanczos method . . . 33
  - 3.3.1 The single-step Lanczos . . . 34
  - 3.3.2 The full Lanczos method . . . 36
    - Obtaining eigenvalues: the first pass . . . 37
    - Computing the ground-state vector . . . 39
  - 3.3.3 Numerical aspects . . . 41
- 3.4 Computing dynamical response functions . . . 42
  - 3.4.1 General introduction . . . 42
  - 3.4.2 Computing the spectral function: the third pass . . . 49

4 Implementation and Checking **55**

4.1 General considerations . . . 55
4.1.1 Example profile of a serial Lanczos run . . . 56
4.1.2 Memory access . . . 56
4.1.3 Matrix-vector multiplication and cache-effects . . . 56
4.1.4 Is a parallelization possible? . . . 58
4.2 Parallelization . . . 59
4.2.1 Shared-memory parallelization with OpenMP . . . 59
4.2.2 Distributed-memory parallelization on BlueGene . . . 61
Direct MPI-2 one-sided communication . . . 62
Vector transposition operation . . . 63
Matrix transpose routines . . . 64
`Alltoall` vs. `Alltoallv` . . . 70
4.3 Checking the results . . . 71
4.3.1 Energy check . . . 71
Band-limit check . . . 71
Lieb-Wu check . . . 71
4.3.2 Green's function checks . . . 72
Sum rules and momenta . . . 72
Band-limit check . . . 74
Atomic-limit check . . . 75

5 Angular-resolved spectral functions and CPT **77**
5.1 Complex boundary conditions . . . 77
5.2 Cluster Perturbation Theory . . . 78
5.2.1 Introduction . . . 78
5.2.2 The CPT method . . . 81
5.2.3 Example calculation for a single site . . . 83
5.2.4 Limits of the CPT . . . 84
5.2.5 Boundary conditions . . . 85
5.2.6 CPT vs. ordinary ED . . . 85
5.2.7 Next-neighbor interaction V . . . 87

6 Metal-insulator transition **89**
6.1 Classical Drude conductivity . . . 89
6.2 Optical conductivity in the Hubbard model . . . 90
6.3 In practice . . . 92
6.4 Mott-band insulator transition . . . 93
6.5 Self-energy . . . 96
6.5.1 Mott and band insulator . . . 97
6.5.2 Metal . . . 98

7 Organics **101**
7.1 Charge-transfer salt TTF-TCNQ . . . 102
7.2 Realistic parameters . . . 103

7.3 TTF-TCNQ in the t-U model 104
7.4 TTF-TCNQ in the t-U-V model 107
7.4.1 Hubbard-Wigner approach 107
7.4.2 Realistic t-U-V model 110
7.4.3 Angular-resolved spectral function with CPT and V 115
7.4.4 TTF and particle-hole symmetry 115
7.5 Accuracy of the results 116

8 Summary **119**

A Speed up and Amdahl's law **121**

B Supercomputers at the Forschungszentrum Jülich **123**
B.1 JUMP 123
B.1.1 Architecture 123
B.2 JUBL 123
B.2.1 Architecture 124
B.3 Run-time restriction 124

C Evaluation of continued fractions **125**

D Particle - hole symmetry **127**

Acknowledgements **129**

Bibliography **131**

1 Introduction

1.1 The setting

At the heart of many modern technologies are the electronic properties of solids. Ubiquitous computing for example strives after integrating small embedded computer systems into everybody's environment. Those devices have to be small and have to work with ever increasing speed. To fulfill these requirements we need to understand the physics of solids, in particular their electronic properties.

Fortunately we do know the Schrödinger equation describing solids in principle exactly. Neglecting relativistic effects it reads

$$i\hbar\frac{\partial}{\partial t}\left|\Psi\right\rangle = H\left|\Psi\right\rangle$$

where

$$H = -\sum_{\alpha=1}^{N_n}\frac{\mathbf{P}_\alpha^2}{2M_\alpha} - \sum_{j=1}^{N_e}\frac{\mathbf{p}_j^2}{2m} - \sum_{j=1}^{N_e}\sum_{\alpha=1}^{N_n}\frac{Z_\alpha e^2}{|\mathbf{r}_j - \mathbf{R}_\alpha|} + \sum_{j<k}^{N_e}\frac{e^2}{|\mathbf{r}_j - \mathbf{r}_k|} + \sum_{\alpha<\beta}^{N_n}\frac{Z_\alpha Z_\beta e^2}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|}$$

and Z_α is the atomic number, M_α the mass, $\mathbf{R}_\alpha$ the position and $\mathbf{P}_\alpha$ the momentum of nucleus α. $\mathbf{p}_j$ and $\mathbf{r}_j$ denote the j^{th} electron's momentum and position and N_e, N_n the number of electrons, nuclei respectively.

If we solve this equation, we will be able to understand current materials and might even design new ones with superior properties. There is, however, a severe problem which makes a brute-force approach infeasible. Let us for example consider an iron atom, neglecting the spin of the electrons. With its $N_e = 26$ electrons the total electronic wave function depends on 26 times 3 coordinates. Choosing a very crude approximation by specifying the wave function on a hyper-cubic grid with 10 points per variable would yield 10^{78} numbers to store and process. Even if we could store one number in a single hydrogen atom, the required memory would weight 10^{51} kg – far more than our home-galaxy, the milky way. Such a memory would be inherently relativistic: to transport a signal from one end of the data storage device to the other would either take ten thousands of years or, if we shrank the device to a manageable size, it would collapse into a black hole. Thus, storing and processing full wave functions is impossible.

Now one might believe that gaining quantitative understanding of solids is hopeless. But do we really need the full wave function? Or are there more efficient ways to determine electronic properties with sufficient accuracy? Thus, a major part of

solid state theory is to answer this question by searching for simulation techniques, which yield reliable results with a minimum of phenomenological input while being applicable to a wide range of problems.

DFT The most successful approach to electronic structure calculations so far is density-functional theory (DFT). A cornerstone of DFT is the Hohenberg-Kohn theorem. Its first statement is that the ground-state energy E of a many-electron system in an external (nuclear) potential is a functional of the electron-density, which can be written as,

$$E\left[\rho, V_{ext}\right] = F\left[\rho\right] + \int d^3r \; V_{ext}(\mathbf{r})\rho(\mathbf{r}) \; ,$$

where $F\left[\rho\right]$ is unknown but universal (in the sense, that it does not depend on V_{ext}, i.e. on the specific system). While the second states that $E\left[\rho, V_{ext}\right]$ is minimized by the ground-state density. This leads to a significant reduction of complexity. We only need to determine the ground state density, which is a function of three coordinates, whereas the wave function depends on $3N_e$ coordinates. In spite of this reduction density-functional theory is in principle exact. In practice, however, we have to rely on approximations since $F\left[\rho\right]$ is unknown. The most common ansatz to find the functional used in "density-functional practice" (DFP) is the Kohn-Sham method. It proceeds by mapping the system of interacting electrons to a system of independent particles T_{ind} in a mean-field Hartree potential plus the exchange-correlation potential under the constraint that the ground-state density of both the interacting and the non-interacting system is the same. The exchange-correlation term describes the difference between the true $F\left[\rho\right]$ and the T_{ind}+ Hartree energy:

$$F\left[\rho\right] = T_{\text{ind}}\left[\rho\right] + \frac{e^2}{2}\int d^3r \; d^3r' \; \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + E_{xc}\left[\rho\right] \; . \tag{1.1}$$

A simple yet very successful approximation to E_{xc} is the local density approximation (LDA). It approximates E_{xc} by

$$E_{xc}\left[\rho\right] = \int d^3r \; \rho(\mathbf{r})\varepsilon_{xc}(\rho(\mathbf{r})) \; ,$$

where $\varepsilon_{xc}(\rho(\mathbf{r}))$ is the exchange and correlation energy per electron of the homogeneous electron gas. Since the Kohn-Sham-Hamiltonian is a full single-particle Hamiltonian with non-vanishing potential, it gives rise to a shell structure, which is e.g. at the heart of the periodic table of elements. The LDA and its generalizations work astonishingly well for materials which have an electronic structure that can be understood by filling single-particle energy levels.

Although there is a wide class of materials which can be sufficiently well described by DFP, it fails to capture the physics of systems with strong correlations[1]. Fortunately there are other means of simplifying the full Schrödinger equation, which might lead to an understanding of these more exotic materials.

[1] This is, however, at least in principle only a matter of finding good methods to construct appropriate functionals.

Model Hamiltonians Instead of investigating the full Hamiltonian, we can replace it with an effective- or model- Hamiltonian, which only contains dominant effects and important single-particle orbitals. In theory one can obtain those model Hamiltonians by integrating out the unwanted part of the Hamiltonian's spectrum, a process called renormalization. In practice, though, this is often difficult to perform for realistic systems and renormalization is done more by intuition.

The simplest model, which describes itinerant electrons and Coulomb repulsion, is the one-band Hubbard model,

$$H = -\sum_{\sigma,ij} t_{ij} c^{\dagger}_{i,\sigma} c_{j,\sigma} + U \sum_{i} n_{i\uparrow} n_{i\downarrow} \ . \tag{1.2}$$

The first term denotes the kinetic energy, where t_{ij} is the Hermitian hopping matrix and the $c_i^{(\dagger)}$ are the (creation)/annihilation operators of Wannier orbitals. It describes the hopping of electrons from the Wannier orbital on site i to the Wannier orbital on site j. In this Hamiltonian only very few orbitals, in many cases only one, per lattice site are considered explicitly. All the others are either neglected or, more accurately, renormalized (downfolded) into the explicit ones. This might be justified, if the resulting band is close to the Fermi level and all other bands are sufficiently far away. This way, we reduce complexity as far as the one-particle properties are concerned. This is quite the opposite compared to Kohn-Sham-DFP. The second term describes the on-site Coulomb potential. Although this is an approximation to the real Coulomb potential the feature that it is a pair interaction is retained which is obviously not the case in any single-particle theories.

1.2 Physics of strong correlations

The Hubbard model is the prototype for studying effects of strong correlations. It describes the interplay between Coulomb and kinetic energy. Evidently, there are two limiting cases:

Band-limit For $U \ll t$ correlations are weak, at least for dimensions greater than one, since the electrons hardly feel each other. They behave as, and at $U = 0$ actually are, independent particles. We get a *band structure* and the system will be metallic (see left part of figure 1.1), unless the band is not completely filled. The kinetic energy term is diagonal in k-space.

Atomic limit If U is large compared to t the Coulomb repulsion dominates. Since it is diagonal in real-space, the movement of a single-electron strongly depends on the positions of the other electrons. This situation cannot be reduced to a single-particle image and those systems are called strongly correlated. For a half-filled chain hopping is forbidden, since it would lead to an energetically expensive double occupation. Thus the system is an insulator, a so-called Mott insulator (cf. figure 1.1).

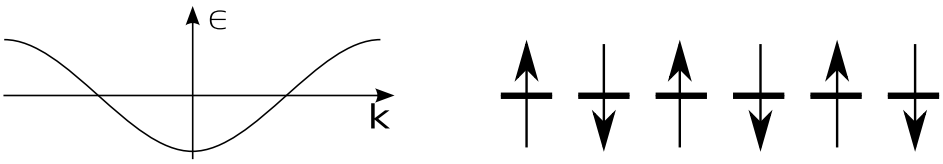


Figure 1.1: Band vs. atomic limit: the left figure shows the band-limit $U/t \ll 1$. The system is a metal, unless the band is completely filled. The right figure shows the atomic limit, i.e. $U/t \gg 1$ for half-filling, thus the systems is an insulator.

Clearly, in-between those limits for a half-filled system there has to be a metal-insulator transition at some value of U/t. At this value the system changes from a metal for small U/t to an insulator for U/t large. In one dimension, however, the situation is special. Here the half-filled system is a Mott insulator for any $U > 0$.

Transition metals oxides Many transition metals have open d- shells. The Hubbard model fits these materials perfectly, since we (a) can restrict our calculations to the d-orbitals, renormalized to take the neglected s- and p- shells in a mean-field way into account. And (b) because the d- shells are rather localized and as a consequence the Coulomb repulsion between the electrons in those shells is quite strong. We should regard it as a pair interaction giving rise to a genuine many-body problem. Examples for those systems are high-T_c cuprates and manganites, showing the effect of colossal magneto resistance (CMR).

Organics There are organic materials, which show effects of strong correlations. Typically there is only a weak overlap between molecular orbitals of different molecules, giving rise to relatively small hopping-matrix elements t, compared to the transition metal oxides. Since the ratio U/t is a measure for the importance of correlations, a relatively small value of t leads to strong correlations as well. Examples for those systems are molecular metals like TTF-TCNQ, showing exotic physics like spin-charge separation. Other organic crystals, the Bechgaard salts like $TMTSF_2PF_6$, are superconductors as are the alkali-doped C_{60} Fullerenes.

1.3 Methods

In this work we will use the Lanczos method to solve extended Hubbard Hamiltonians exactly. Because the size of the Hilbert space grows exponentially with the system size the major challenge is to cope with huge vectors. We therefore work on the very latest supercomputers like the massively parallel Blue Gene system JUBL in Jülich.

The Lanczos algorithm is an iterative method for finding the extremal eigenvalues and eigenvectors of a symmetric or Hermitian matrix, in our case a Hubbard Hamiltonian. It works by building a Krylov subspace, a space spanned by increasing

powers of the matrix H applied to a random state $\mathbf{b}$, i.e. $\{\mathbf{b}, H\mathbf{b}, \ldots, H^m\mathbf{b}\}$, and then diagonalizing the matrix in this subspace. It turns out that relatively small subspaces already yield very good approximations to the lowest eigenpair.

Evidently the matrix-vector multiplication plays a crucial role in this algorithm. Calculating the matrix-vector product is only feasible for large system, if the matrix is sufficiently sparse. Even then it remains the most time-consuming part of this method.Moreover, it is the main problem for implementing an efficient massively parallel Lanczos algorithm, for example for BlueGene.

Having solved the Hamiltonian we have the ground-state wave function. From this we can directly obtain all ground-state expectation values or wave function functionals,

$$\left\langle \hat{O} \right\rangle = \frac{\langle \Psi_0 | O | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle} \, .$$

In addition we can efficiently compute single-particle Green's functions and dynamical response functions. Thus, the Lanczos method is ideal, since we have means to calculate about everything. Everything, that is, in a relatively small system.

Effects of finite system size are for example particularly severe when we are interested in the angular-resolved (k-dependent) spectral functions, for instance to study exotic effects like spin-charge separation. Since our Hubbard system only comprises relatively few sites, typically 10-20, the resolution in k-space is quite low. Introducing complex boundary conditions improves the situation. This way we can achieve an arbitrarily good resolution in k. For each different complex boundary condition, however, we study in principle a different system. This manifests itself for example in shifts in the chemical potential μ. And we thus cannot get reliable results from this method.

A more advanced approach is cluster perturbation theory. To obtain an arbitrary high resolution in k-space, we calculate the one-particle Green's matrix in real-space for a finite cluster. Hereafter we recover the original lattice by considering a superlattice of these clusters, treating the hopping between them perturbatively. This essentially yields the Green's function for the infinite system. The approach is exact in both the strong- ($t/U = 0$) and weak-coupling ($U/t = 0$) limit and it yields quite a good approximation to the spectral functions for each wave vector in the Brillouin zone.

1.4 Overview

In chapter 2 we introduce the Hubbard model and some of its features like the Mott-Hubbard transition or antiferromagnetic correlations.

The following chapter deals with exact diagonalization. We introduce the power method as a simple way of obtaining the ground state of a matrix as well as the more sophisticated Lanczos algorithm. Moreover we show, how to calculate Green's functions using the Lanczos method and in the end describe numerical artifacts, like ghosts.

In chapter 4 we discuss our efficient implementation of the Lanczos method. Examining the functions, that consume the bulk of execution time, shows that they can be parallelized. We first describe our shared-memory implementation based on OpenMP. The main part of the chapter gives implementation details of our massively parallel code for BlueGene like supercomputers. Finally, we describe how we tested our implementations.

In chapter 5 we discuss techniques for overcoming the limitations posed by the finite size of our clusters. We introduce how to use complex boundary conditions to get a higher resolution in k-space and study a second far superior method – called cluster perturbation theory.

In the final chapters we apply our methods to physical problems. In chapter 6 we use a characterization technique developed by Kohn, which enables us to distinguish a metal from an insulator from the ground state alone, to look at band and Mott insulator transitions.

And in the final chapter we investigate the properties of the one-dimensional organic conductor TTF-TCNQ. Using a simple Hubbard model description we find Luttinger liquid signatures in the angular-resolved spectral function computed with cluster perturbation theory. With a more realistic extended Hubbard model we can resolve a long standing problem in the understanding of the experimental width of the photoemission spectra.

2 Hubbard model

The Hubbard model, introduced by John Hubbard [1],[2],[3], Martin C. Gutzwiller [4] and Junjiro Kanamori [5], is the simplest many-body model, which cannot be reduced to an effective single-particle system. Although it is a very simple model it has rich physics and is used to study phenomena of correlated electrons like high-T_c superconductivity, Mott transitions in transition metal oxides, or one dimensional organic conductors. Unfortunately the exact solution is not known aside from the ground state in the one dimensional model [6], hence we depend on approximations and/or computer simulations.

The kinetic energy part of the Hubbard model is derived from the tight-binding approximation (see e.g. chapter 11 of [7]) of solid state theory. It is an independent-particle theory and thus contains no many-body features. We will study this approximation in the first part of the chapter.

The second part will take the Coulomb electron-electron repulsion into account, leading to the complete Hubbard model. Whereas other methods like mean-field theory replace this interaction with an average one, retaining an effective single-particle picture, the Hubbard model incorporates the Coulomb potential as a pair interaction leading to a genuine many-particle problem.

2.1 The tight-binding approximation (TBA)

The tight-binding approximation (TBA) is used to describe tightly bound electrons in solids. Tightly bound means that the electrons are rather localized at the nuclei, i.e. the electron density is concentrated around the nuclei. There is only a small overlap of the electron's wave functions of neighboring atoms. We can consider the atoms almost as isolated and hence the splitting of atomic energy levels is relatively weak, leading to narrow bands. This situation can be found in transition metals' d- and f- shells to a good approximation.

2.1.1 Theoretical derivation

The picture of almost isolated atoms suggests an ansatz of atomic orbitals, which are exponentially localized. Assuming that $\phi_\nu(\mathbf{r} - \mathbf{R}_i) = \phi_{i\nu}$ is the ν^{th} atomic orbital centered around the nucleus at $\mathbf{R}_i$. In order to satisfy the Bloch condition we

construct the wave function as [1]

$$\phi_{\mathbf{k}\mu}(\mathbf{r}) = \sum_i e^{i\mathbf{k}\cdot\mathbf{R}_i}\phi_\mu(\mathbf{r}-\mathbf{R}_i) \ . \tag{2.1}$$

The one-body Hamiltonian of the system looks like

$$h^0 = -\frac{\hbar^2}{2m}\nabla^2 + V^{ion}(x) = -\frac{\hbar^2}{2m}\nabla^2 + \sum_i v_i(x) \ , \tag{2.2}$$

where $v_i(x)$ is the i^{th} nuclear potential centered at $\mathbf{R}_i$. Let us take a look at its representation in atomic orbitals. The diagonal elements $\langle\phi_{i\nu}|h^0|\phi_{i\mu}\rangle$ look like,

$$\left\langle \phi_{i\nu}\middle| -\frac{\hbar^2}{2m}\nabla^2 + v_i(x) + \sum_{j\neq i} v_j(x)\middle|\phi_{i\mu}\right\rangle = \varepsilon_\nu\delta_{\mu\nu} + \left\langle \phi_{i\nu}\middle|\sum_{j\neq i} v_j(x)\middle|\phi_{i\mu}\right\rangle \ ,$$

where the first term of the right hand side yields the main contribution. The off-diagonal elements are given by,

$$\left\langle\phi_{i\nu}\middle|h^0\middle|\phi_{j\mu}\right\rangle = \varepsilon_\nu\left\langle\phi_{i\nu}|\phi_{j\mu}\right\rangle + \left\langle \phi_{i\nu}\middle|\sum_{l\neq i} v_i(x)\middle|\phi_{j\mu}\right\rangle = \varepsilon_\nu\left\langle\phi_{i\nu}|\phi_{j\mu}\right\rangle - t_{ij\nu\mu} \ ,$$

where $\langle\phi_{i\nu}|\phi_{j\mu}\rangle \neq 0$, since the atomic orbitals are non-orthogonal. These terms are, however, usually very small and are therefore often neglected. Except for nearest-neighbors, we expect also $t_{ij\mu\nu}$ to be very small. As a further simplification we assume, that the bandwidth is small compared to the difference of atomic energy levels. Then, the hopping between bands can be neglected, i.e. $t_{ij\mu\nu} = t_{ij\mu}\delta_{\mu\nu}$.

Let us calculate the energy expectation value of the Bloch function, equation (2.1),

$$\epsilon_{\mathbf{k}\nu} = \frac{\sum_{i,j} e^{i\mathbf{k}\cdot(\mathbf{R_i}-\mathbf{R_j})}\left\langle\phi_{i\nu}|h^0|\phi_{j\nu}\right\rangle}{\left\langle\phi_{\mathbf{k}\nu}|\phi_{\mathbf{k}\nu}\right\rangle} \ ,$$

yielding

$$\epsilon_{\mathbf{k}\nu} = \varepsilon_\nu - \sum_{ij} t_{ij\nu}e^{i\mathbf{k}\cdot(\mathbf{R_i}-\mathbf{R_j})} \ ,$$

where the overlap of the atomic orbitals at different sites in the denominator are neglected, and thus the Bloch waves are considered orthogonal. For only nearest and second nearest neighbor hopping the resulting energy is given by

$$\epsilon_{\mathbf{k}\nu} = \varepsilon_\nu - t_\nu\sum_{\text{n.n.}} e^{i\mathbf{k}\mathbf{R}_i} - t'_\nu\sum_{2^{nd}\text{ n.n.}} e^{i\mathbf{k}\mathbf{R}_j} \ , \tag{2.3}$$

[1] This is the simplest ansatz possible. If the band width is wider than the energetic difference between the atomic levels, we have to make a more complex ansatz of linear combinations of atomic orbitals (LCAO).

where $t_\nu = t_{01\nu} = t_{10\nu}$ and $t'_\nu = t_{02\nu} = t_{20\nu}$. $\mathbf{R}_i$ and $\mathbf{R}_j$ denote nearest, second nearest neighbor vectors respectively, as seen from a representative lattice site $\mathbf{R}_0$.

The ansatz (2.1), the simplest possible, is not used in practice. First, the atomic orbitals are not orthogonal with respect to the orbitals of other atoms and second the ansatz of a simple atomic orbital is usually too oversimplified. One would rather use linear combinations of atomic orbitals (LCAO) as ansatz for $\phi_{i\nu}$ retaining, however, the problems of non-orthogonality. For application with the Hubbard model an "inverse" approach is used.

2.1.2 From the TBA to the kinetic energy term of the Hubbard model

Assuming we have the solution of some Hamiltonian H^0 in terms of band structure energies $\epsilon_{\mathbf{k}\nu}$ and corresponding Bloch wave functions $\phi_{\mathbf{k}\nu}$. Fourier transforming the Bloch waves leads to Wannier states

$$\varphi_{i\mu}(\mathbf{x}) = \frac{1}{\sqrt{L}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{x}} \phi_{\mathbf{k}\mu}(\mathbf{x}) \ . \tag{2.4}$$

The on-site energies and hopping-matrix elements can be deduced from the $\epsilon_{\mathbf{k}\nu}$ via Fourier transformation, as well. We get

$$\begin{aligned} -t_{ij\nu\mu} &= \left\langle \varphi_{i\nu} \middle| h^0 \middle| \varphi_{j\mu} \right\rangle = \frac{1}{L} \sum_{\mathbf{k},\mathbf{k}'} e^{-i\mathbf{k}'\cdot\mathbf{r}_j} e^{i\mathbf{k}\cdot\mathbf{r}_i} \left\langle \phi_{\mathbf{k}\nu} \middle| h^0 \middle| \phi_{\mathbf{k}'\mu} \right\rangle \\ &= \frac{1}{L} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot(\mathbf{r}_j - \mathbf{r}_i)} \epsilon_{\nu\mathbf{k}} \delta_{\mu\nu} \ , \end{aligned} \tag{2.5}$$

and for the on-site energies,

$$\varepsilon_{i\nu} = t_{ii\nu} \ . \tag{2.6}$$

In second quantized form the Hamiltonian reads,

$$H^0 = -\sum_{\sigma,i\neq j,\nu} t_{ij\nu} c^\dagger_{i\nu\sigma} c_{j\nu\sigma} + \sum_{\sigma,i,\nu} \varepsilon_{i\nu} c^\dagger_{i\nu\sigma} c_{i\nu\sigma} \ , \tag{2.7}$$

where $c^{(\dagger)}_{i\nu\sigma}$ creates/annihilates an electron in the ν^{th} Wannier state at lattice site $\mathbf{R}_i$. This representation has the advantage of being independent from the number of particles.

Like the Bloch waves the Wannier states are orthonormal, since the unitary Fourier transform conserves the scalar product. But unlike Bloch waves the Wannier functions are usually localized like atomic orbitals. Localization means that for a Wannier state centered at $\mathbf{R_i}$, the density decays exponentially for $|\mathbf{r} - \mathbf{R}_i| \to \infty$.

General results concerning the localization of Wannier functions are difficult to obtain [8]. Actually this problem has been called "one of the few basic questions of the quantum theory of periodic solids in the one-electron approximation which is not

completely solved" [9]. What has been shown so far is the existence of exponentially localized Wannier functions for isolated, simple bands in any dimension [9], as well as for complex bands in perturbation theory [10] and in the tight-binding limit [10].

Though Wannier functions do not have a direct physical interpretation they are very useful especially if they are exponentially localized. There exist different recipes for constructing "maximally localized" Wannier functions. Marzari and Vanderbilt for instance construct them by exploiting the arbitrariness of the phase. Before Fourier transforming the Bloch waves $\phi_{\mathbf{k}\mu}$ are multiplied with a k-dependent phase factor $e^{i\Phi_{\mathbf{k}}}$ in such a way, that after the Fourier transform the spatial width of the Wannier functions is minimized [11].

Whenever phenomena occur in solids in which electrons or their effects are spatially localized Wannier functions are the representation of choice. This is the case in the Hubbard model. We will see that the Hubbard interaction term is local. Thus a description in terms of Wannier functions is preferable.

2.1.3 Features in low dimension

In this part we will look at the features of the TBA in one and two dimensions. Starting point is the band structure equation (2.3). We consider a chain of single Wannier orbitals per site in one dimension with next-neighbor hopping t and second next-neighbor hopping t'. Setting the on-site energies to zero (2.3) yields

$$\epsilon_k^{1d} = -2t\cos(ka) - 2t'\cos(2ka)\ , \tag{2.8}$$

where a denotes the distance in coordinate space between neighbor sites and k is restricted to the first Brillouin zone, i.e. $k \in (-\pi/a, \pi/a]$. For different values of t'/t this band structure is plotted in figure 2.1.

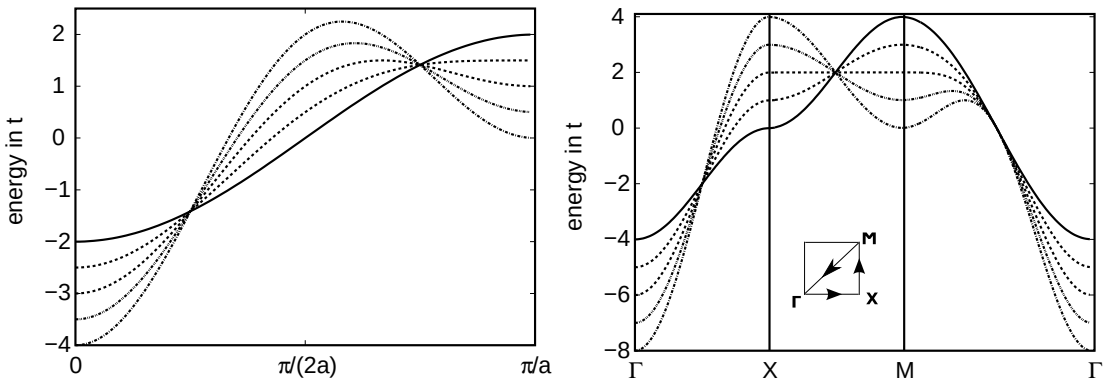


Figure 2.1: Band structure of a linear chain (left) and a square lattice (right) in TBA for different values of t'/t (from upper to lower at $k = 0$, Γ respectively $t'/t = \{0.00,\ 0.25,\ 0.50,\ 0.75,\ 1.00\}$).

Similarly we obtain for a square lattice in two dimensions, with nearest-neighbors at $\mathbf{R}_{1,2} = \pm a\,(1,0)^T$, $\mathbf{R}_{3,4} = \pm a\,(0,1)^T$ and second nearest neighbors at $\mathbf{R'}_{1,2} = \pm a/\sqrt{2}\,(1,1)^T$

and $\mathbf{R}'_{3,4} = \pm a/\sqrt{2}\,(1,-1)^T$,

$$\epsilon_{\mathbf{k}}^{2d} = -2t\left(\cos(k_x a) + \cos(k_y a)\right) - 2t'\left(\cos\left(\frac{a(k_x+k_y)}{\sqrt{2}}\right) + \cos\left(\frac{a(k_x-k_y)}{\sqrt{2}}\right)\right), \tag{2.9}$$

where $\mathbf{k}$ is restricted to the first Brillouin zone, i.e. $k_i \in (-\frac{\pi}{a}, \frac{\pi}{a}]\ \ \forall i$. Examples for several values of t'/t are shown in figure 2.1. Moreover figure 2.2 shows the energy surface and its iso-energy contours for $t' = 0$ and $t'/t = -0.25$.

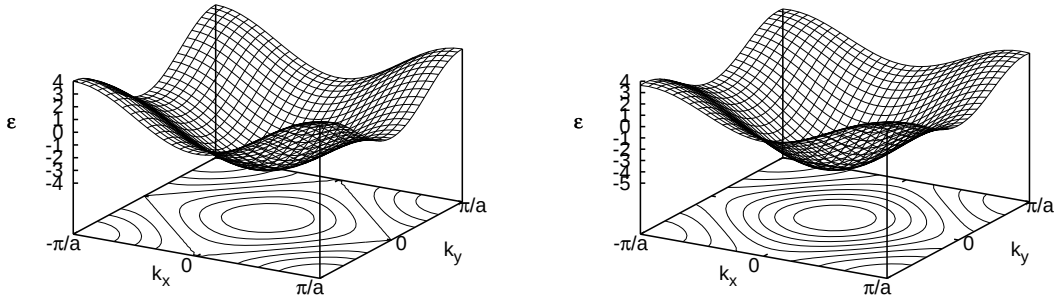


Figure 2.2: Energy surface and iso-energy lines for the two dimensional system with $t' = 0$ (left) and $t'/t = -0.25$ (right).

Density of states

The density of states in k-space for an arbitrary dimension n is constant. Each k-state consistent with the boundary conditions occupies a volume of $\Delta k = (2\pi/L)^n$ in k-space and the density is therefore given by its inverse. This can be derived by considering a particle in a n dimensional hypercube with side length L. More interesting is the density of states as function of the energy. If L is sufficiently large the sum over k-states can be regarded as an integral, i.e.

$$\lim_{L\to\infty} \frac{1}{L^n} \sum_{\mathbf{k}} \to \int \frac{d^n\mathbf{k}}{(2\pi)^n}$$

Thus, the density of states is given by

$$D(E) = \int_{1BZ} \frac{d^n k}{(2\pi)^n}\, \delta(E - \epsilon_{\mathbf{k}})\ . \tag{2.10}$$

This integral often has to be evaluated numerically. In the case of a one-dimensional monotonous band structure, we can however obtain it with a quite simple method.

With the density in k-space

$$D^{1d}(k)\, dk = \frac{1}{2\pi}\, dk$$

we obtain using our tight-binding dispersion relation (2.8) for nearest-neighbor hopping the density of states as

$$D^{1d}(E)\,dE = 2\frac{1}{2\pi}\left|\frac{dk}{dE}\right|dE = \frac{1}{2\pi a}\frac{dE}{|t\sin(ka)|} = \frac{1}{2\pi a}\frac{dE}{|t\sin(\arccos(E/2t))|}\,. \quad (2.11)$$

This density of states is plotted in the left plot of figure 2.3. At the Brillouin zone boundaries, where the slope of the dispersion relation is vanishing, i.e. $dE/dk \to 0$, the density diverges.

In two dimensions the density of states is given by

$$D^{2d}(E) = \frac{K\left(1-(\frac{E}{4t})^2\right)}{2\pi^2 a^2 t}\,, \quad (2.12)$$

where K is the complete elliptic integral of the first kind. This function is plotted in the left plot of figure 2.3. Again we observe a singularity in the density of states. These singularities are named after van-Hove and arise whenever $|\nabla_{\mathbf{k}} E| \to 0$.

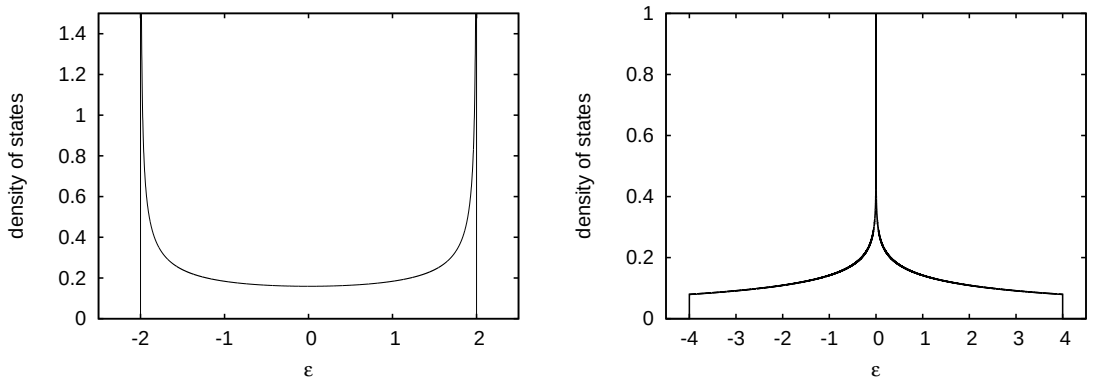


Figure 2.3: Density of states (DOS) in one (left) and two (right) dimension(s).

Two orbitals per cell

Consider a one-dimensional system with two Wannier orbitals per unit cell, denoted by + and − with an energy difference of Δ. For symmetry reasons we define the on-site energies as $\varepsilon_{\mp} = \mp\Delta/2$. Hopping occurs only between nearest neighbors. This situation is sketched in figure 2.4. Hence, the Hamiltonian describing the system

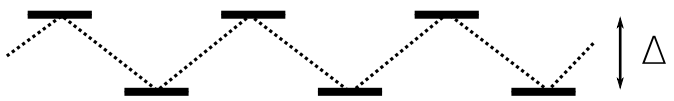


Figure 2.4: Two orbitals per cell, separated by Δ in energy.

looks like

$$H = -t \sum_i \left(c^\dagger_{i,+} c_{i,-} + h.c. \right) - t \sum_{\langle ij \rangle} \left(c^\dagger_{j,-} c_{i,+} + h.c. \right) + \sum_i \frac{\Delta}{2} \left(n_{i,+} - n_{i,-} \right) \ , \quad (2.13)$$

where i denotes the unit cell and $n_{i,\nu} = c^\dagger_{i,\nu} c_{i,\nu}$ the occupation of its orbital $\nu = \{+,-\}$.

To solve this Hamiltonian we again make an ansatz of Bloch waves,

$$\phi_{\mathbf{k}} = \sum_j e^{ikR_j} u_j \ , \quad (2.14)$$

where $R_j = 2aj$ is the coordinate vector of the i^{th} unit cell and $u_i = \alpha\varphi_{i,-} + \beta\varphi_{i,+}$ is a (unit cell) periodic function. We have to solve $\langle \phi_k | H | \phi_k \rangle$, where k is restricted to $k \in (-\frac{\pi}{2a}, \frac{\pi}{2a}]$. With the unitary transformation $\hat{U}^\dagger$, which contains the phase factors of the Bloch wave,

$$\phi_{\mathbf{k}} = \hat{U}^\dagger \mathbf{u} \ , \quad (2.15)$$

we shift the k-dependence into the Hamiltonian, yielding

$$\langle \phi_k | H | \phi_k \rangle = \left\langle \mathbf{u} \middle| \hat{U} H \hat{U}^\dagger \middle| \mathbf{u} \right\rangle = \langle \mathbf{u} | H_k | \mathbf{u} \rangle \ .$$

$\hat{U} H \hat{U}^\dagger$ and $\mathbf{u}$ look like

$$\begin{pmatrix} \ddots & \ddots & & & & \\ \ddots & -\Delta/2 & -te^{ika} & 0 & 0 & \\ & -te^{-ika} & \Delta/2 & -te^{ika} & 0 & \\ & 0 & -te^{-ika} & -\Delta/2 & -te^{ika} & \\ & 0 & 0 & -te^{-ika} & \Delta/2 & \ddots \\ & & & & \ddots & \ddots \end{pmatrix}, \begin{pmatrix} \vdots \\ \alpha \\ \beta \\ \alpha \\ \beta \\ \vdots \end{pmatrix} .$$

$\mathbf{u}$ contains $N \to \infty$ 2-vectors, where N is the number of unit cells in the chain. The 2-vectors actually are the $u_i = \alpha\varphi_{i,-} + \beta\varphi_{i,+}$ and thus are for all i the same due to periodicity.

We can therefore construct a replacement Hamiltonian,

$$H_k = -t \left(c^\dagger_+ c_- + h.c. \right) - t \left(e^{-ik2a} c^\dagger_- c_+ + h.c. \right) + \frac{\Delta}{2} \left(n_+ - n_- \right) \ , \quad (2.16)$$

which in matrix elements of the Wannier basis looks like

$$H_k = \begin{pmatrix} -\Delta/2 & -t\left(1 + e^{-ik2a}\right) \\ -t\left(1 + e^{ik2a}\right) & \Delta/2 \end{pmatrix} . \quad (2.17)$$

The replacement Hamiltonian actually yields exactly the same physics. It is k-dependent but operates only on a two dimensional space.

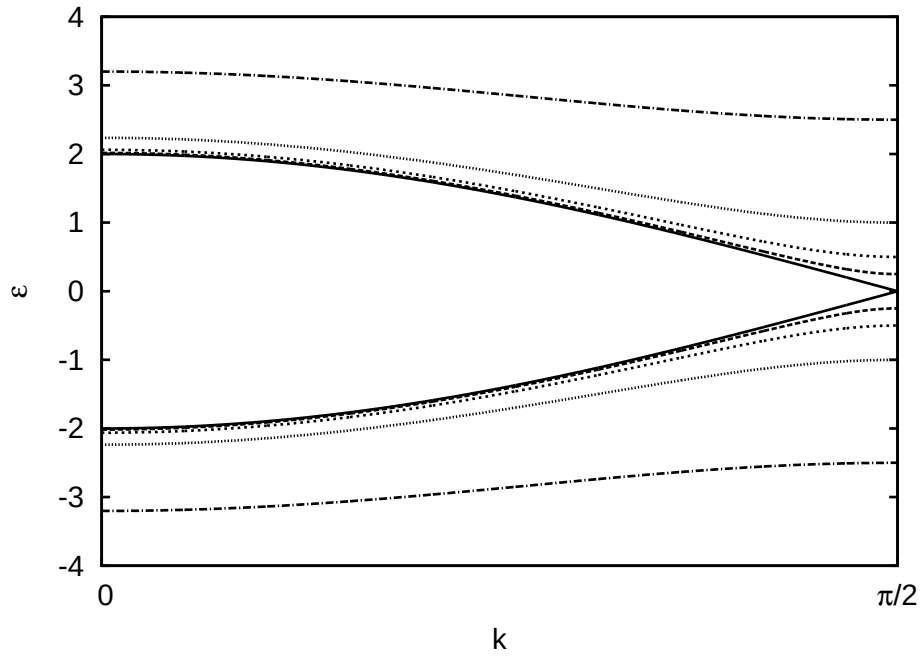


Figure 2.5: Band structure for the two band model in one dimension for various values of $\Delta/t = \{5, 2, 1, 0.5, 0\}$ (upper to lower in upper band).

By diagonalization we obtain two bands with energy dispersion,

$$\epsilon_{\nu k} = \mp\sqrt{4t^2 \cos^2 ka + \left(\frac{\Delta}{2}\right)^2} \,, \tag{2.18}$$

for $k \in (-\frac{\pi}{2a}, \frac{\pi}{2a}]$. The resulting band structure is plotted in figure 2.5 for various values of Δ/t.

For $\Delta = 0$, the equation (2.18) coincides with equation (2.8) for $t' = 0$. Aside from the fact that we have different unit cells and therefore different Brillouin zones. In the case of two orbitals per cell, we have a halved Brillouin zone but a second band. It is the second half of the original band folded into the reduced Brillouin zone. We retain our original band for one site per unit cell by folding it out in the unreduced zone. Mathematically this means for an arbitrary number L of equivalent sites in the unit cell,

$$\epsilon_k^{\text{full}} = -2t \cos(ka) \,, \tag{2.19}$$

with

$$k = \frac{\pi}{La}\nu + \tilde{k} \,,$$

where $\tilde{k}$ is restricted to $\tilde{k} \in (-\frac{\pi}{La}, \frac{\pi}{La}]$.

2.1.4 Many independent particles

Up to now we treated a single electron only. If the many-body Hamiltonian is a separable sum of single-particle ones, we can regard the system as a Fermi gas and obtain the solution of the full problem from the solution of the single-particle. For the TBA Hamiltonian (2.7) this is for instance the case. Let $\{|\phi_i\rangle\}_i$ denote the orthonormal single-particle eigenbasis with eigenenergies ϵ_i.

Assume we have N_e electrons which occupy the spin orbitals of the set $\{|\phi_i\rangle\,\chi_\sigma\}_i$. For Fermions the total wave function has to be anti-symmetrized. This is ensured by the construction with Slater determinants, which are in the real-space basis given by

$$\langle \mathbf{x}, \sigma|\psi\rangle = \frac{1}{\sqrt{N_e!}} \begin{vmatrix} \langle x_1,\sigma_1|\phi_{i_1,\uparrow}\rangle & \langle x_2,\sigma_2|\phi_{i_1,\uparrow}\rangle & \dots & \langle x_n,\sigma_n|\phi_{i_1,\uparrow}\rangle \\ \langle x_1,\sigma_1|\phi_{i_2,\uparrow}\rangle & \langle x_2,\sigma_2|\phi_{i_2,\uparrow}\rangle & \dots & \langle x_n,\sigma_n|\phi_{i_2,\uparrow}\rangle \\ \vdots & \vdots & & \vdots \\ \langle x_1,\sigma_1|\phi_{i_n,\uparrow}\rangle & \langle x_2,\sigma_2|\phi_{i_n,\uparrow}\rangle & \dots & \langle x_n,\sigma_n|\phi_{i_n,\uparrow}\rangle \\ \langle x_1,\sigma_1|\phi_{i_1,\downarrow}\rangle & \langle x_2,\sigma_2|\phi_{i_1,\downarrow}\rangle & \dots & \langle x_n,\sigma_n|\phi_{i_1,\downarrow}\rangle \\ \langle x_1,\sigma_1|\phi_{i_2,\downarrow}\rangle & \langle x_2,\sigma_2|\phi_{i_2,\downarrow}\rangle & \dots & \langle x_n,\sigma_n|\phi_{i_2,\downarrow}\rangle \\ \vdots & \vdots & & \vdots \\ \langle x_1,\sigma_1|\phi_{i_n,\downarrow}\rangle & \langle x_2,\sigma_2|\phi_{i_n,\downarrow}\rangle & \dots & \langle x_n,\sigma_n|\phi_{i_n,\downarrow}\rangle \end{vmatrix} , \tag{2.20}$$

where $\left|\phi_{i_j\sigma}\right\rangle = \left|\phi_{i_j}\chi_\sigma\right\rangle$ are the occupied spin orbitals indexed by $j \in [1,\dots,N_e]$.

Equivalently this state can be represented in second quantization by

$$|\psi\rangle = |n_{1\uparrow}n_{1\downarrow}n_{2\uparrow}n_{2\downarrow}\dots\rangle = \prod_{\sigma,i}\left(c^\dagger_{i\sigma}\right)^{n_{i\sigma}}|0\rangle \ , \tag{2.21}$$

where $c^\dagger_{i\sigma}$ creates an electron in the spin orbital $\phi_{i,\sigma}$ and $|0\rangle$ denotes the vacuum state. Anti-symmetry is ensured by the commutation rules for Fermions, i.e.

$$\left\{c_{i\sigma}, c^\dagger_{j\sigma'}\right\} = \delta_{ij}\delta_{\sigma\sigma'} \tag{2.22}$$

$$\left\{c_{i\sigma}, c_{j\sigma}\right\} = 0\,. \tag{2.23}$$

$n_{i\sigma}$ represents the occupation of the spin orbitals and is therefore restricted to 0 or 1.

The eigenenergies of the full many-body Hamiltonian are given by summing over the eigenenergies ϵ_i of the occupied spin orbitals. In the ground-state only the spin orbitals of lowest energy are filled.

Spin-assigned Slater determinants Often we calculate expectation values of observables which have no explicit spin dependence. For these observables it can be shown that using spin-assigned wave functions (see for example chapter A11 6-3.2 of [12]) yields the same results. They are however calculated more easily, since the Slater determinants become smaller. Spin-assigning means that we give the first $N_\uparrow$

particles the value $\sigma =\uparrow$ and the next $N_\downarrow$ the value $\sigma =\downarrow$. This leads to a block-diagonal structure of the determinant (2.20) and thus to a product of determinants, one for the up and one for the down electrons:

$$\langle \mathbf{x}|\psi\rangle = \left\langle \mathbf{x}, \{\underbrace{\uparrow,\ldots,\uparrow}_{N_\uparrow},\underbrace{\downarrow,\ldots,\downarrow}_{N_\downarrow}\}\middle|\psi\right\rangle = \frac{1}{\sqrt{N_e!}}$$
$$\begin{vmatrix} \langle x_1|\phi_{i_1,\uparrow}\rangle & \ldots & \langle x_{N_\uparrow}|\phi_{i_1,\uparrow}\rangle \\ \langle x_1|\phi_{i_2,\uparrow}\rangle & \ldots & \langle x_{N_\uparrow}|\phi_{i_2,\uparrow}\rangle \\ \vdots & \vdots & \vdots \\ \langle x_1|\phi_{i_n,\uparrow}\rangle & \ldots & \langle x_{N_\uparrow}|\phi_{i_n,\uparrow}\rangle \end{vmatrix} \times \begin{vmatrix} \langle x_{N_\uparrow+1}|\phi_{i_1,\downarrow}\rangle & \ldots & \langle x_N|\phi_{i_1,\downarrow}\rangle \\ \langle x_{N_\uparrow+1}|\phi_{i_2,\downarrow}\rangle & \ldots & \langle x_N|\phi_{i_2,\downarrow}\rangle \\ \vdots & \vdots & \vdots \\ \langle x_{N_\uparrow+1}|\phi_{i_n,\downarrow}\rangle & \ldots & \langle x_N|\phi_{i_n,\downarrow}\rangle \\ \cdot \end{vmatrix} \tag{2.24}$$

Finite size scaling

For single particle theory, the k-space representation is optimal. The Hamiltonian is diagonal and we can restrict our calculations to a single unit cell. Consider two electrons in a unit cell consisting of a single orbital. The direct Coulomb repulsion between them can be taken into account accurately. They however see mirror images of themselves on the periodically extended system, giving rise to non-physical interactions.



Figure 2.6: (a) open, (b) periodic boundary conditions in one dimension.

When the Coulomb interaction becomes important and should be considered as a pair interaction, a real-space representation in terms of Wannier functions is preferable. Since we cannot handle infinitely large matrices, we approximate the infinite physical system with a finite cluster. This of course gives rise to finite size effect, which need to be studied systematically.

Let us consider a d dimensional bulk system. Its extensive quantities scale proportional to L^d where L denotes a characteristic length of the system. In the bulk the division of an extensive quantity by L^d yields a constant. Deviation from this constant behavior is due to the system's finite size. The surface for instance scales with L^{d-1} and thus the leading order of the finite size correction is in general proportional to $1/L$.

For the bulk, all physical properties must be invariant under translations. We therefore introduce periodic boundary conditions (also called Born-von Karman-

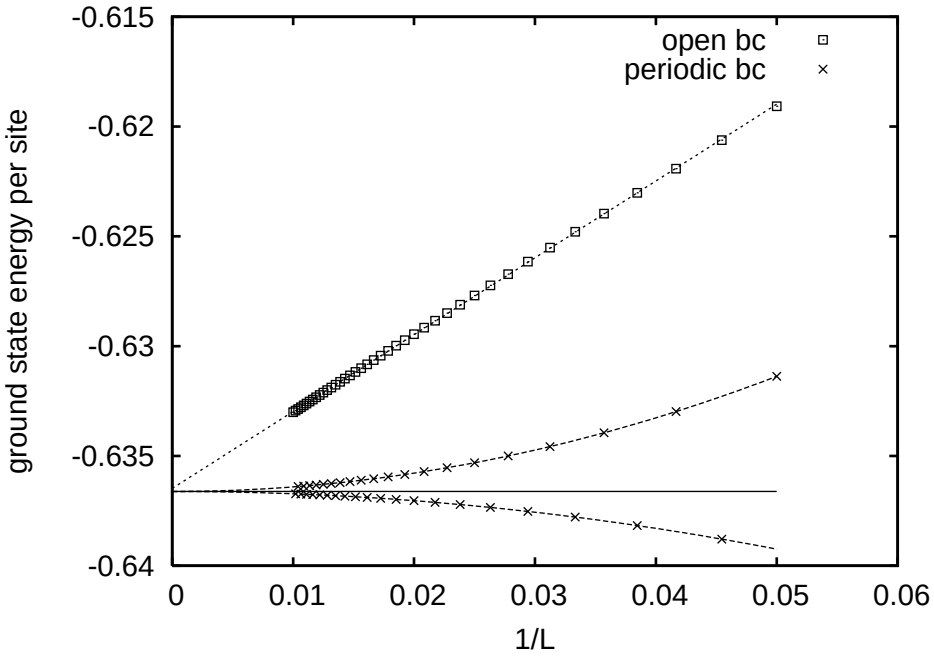


Figure 2.7: Finite size scaling: ground-state energy of one dimensional half-filled systems in TBA with varying (even) length and open (upper graph)/periodic (lower graphs) boundary conditions. The OBC systems show a linear scaling due to surface effects. With PBC theses are gone and the next order correction is dominant. Extrapolated to infinite system yields the same value for both system types as denoted by the horizontal line, $E_{\infty,0} = -2a/\pi$.

boundary condition) and expect, that they reduce finite size effects. Periodic boundary conditions mean, that the single particle wave functions satisfies $\phi(\mathbf{r}) = \phi(\mathbf{r}+\mathbf{L}_j)$, where $\mathbf{L}_j$ is the length of the system in dimension j. In one dimension for example it can be regarded as closing a chain to a ring (cf. figure 2.6), or in two as closing a plane to a torus.

Figure 2.7 shows the ground-state energy of a one-dimensional half-filled system in TBA against $1/L$. For open boundary conditions (OBC), i.e. treating the cluster as a finite system, we obtain a linear correction in $1/L$ (surface correction). PBC calculations indeed show better results. The surface effects are gone and the next higher effects proportional to $1/L^2$ dominate. Moreover the absolute deviation from the infinite system's ground state energy (solid line) is considerably smaller. In the limit of $1/L \to 0$ boundary conditions become irrelevant and all extrapolations from finite systems meet at the exact energy for the infinite system $E_{\infty,0} = -2a/\pi$.

There are actually two parabolas for the PBC systems. The upper one represents open-shell, whereas the lower one closed-shell systems. In closed shell systems each occupied single-particle orbital in the ground state is doubly occupied. Thus the

resulting ground-state wave function is unambiguously determined. In open-shell systems on the other hand an energy shell is not completely filled and thus there are several linear combinations for the ground state. Hence, the ground state is degenerate.

2.2 From the tight-binding approximation to the Hubbard model

In the single-particle Hamiltonian h^0 we have studied the electron-electron repulsion was neglected. In the full Hamiltonian for a single particle, the electron-electron interaction spoils the separability. The Hamiltonian reads,

$$H = \sum_{i}^{N_e} h^0 \left(\nabla_i, \mathbf{x}_i\right) + \frac{1}{2} \sum_{i \neq j}^{N_e} v^{\text{el}-\text{el}}(r_i, r_j) \ , \tag{2.25}$$

where r_i is the position of the i^{th} electron. Much of the electron-electron interaction can be incorporated into the single-particle part of the Hamiltonian, leading to an effective (nuclear) potential.

This effective Hamiltonian looks like [13],

$$\tilde{H}^0 = \sum_{i=1}^{N_e} \left(h^0(\nabla_i, \mathbf{x}_i) + v^{\text{eff}}(\mathbf{x}_i, \rho) \right) \ , \tag{2.26}$$

where v^{eff} is a functional of the ground-state density ρ_0. In order to calculate ρ_0 we need the solution of equation (2.26), which in turn needs ρ_0. Hence this (Kohn-Sham) system needs to be solved self-consistently.

As result we obtain the band energies $\epsilon_{\mathbf{k}\nu}$ and the corresponding Bloch waves, which via Fourier transformation yield the on-site energies ε_i, the hopping matrix elements $t_{ij\nu}$ and the Wannier states. With the localized Wannier states we treat the residual interaction

$$\tilde{v}_{ij} = v^{\text{el}-\text{el}}(r_i, r_j) - \frac{v^{\text{eff}}(\mathbf{x}_i, \rho) + v^{\text{eff}}(\mathbf{x}_j, \rho)}{N_e} \ , \tag{2.27}$$

which was neglected in equation (2.26). It regards the Coulomb interaction as, what it really is in reality, a pair interaction. This might lead to an understanding of far-reaching ordering phenomena like antiferromagnetism, superconductivity or Mott insulation. The full electron Hamiltonian thus reads,

$$H = \tilde{H}^0 + \tilde{V}^{\text{el}-\text{el}} \ , \tag{2.28}$$

where $\tilde{H}^0$ is given by equation (2.26) and $\tilde{V}^{\text{el}-\text{el}}$ by

$$\tilde{V}^{\text{el}-\text{el}} = \sum W_{ijkl}^{\alpha\beta\gamma\delta} c_{\alpha i\sigma}^{\dagger} c_{\beta j\sigma'}^{\dagger} c_{\gamma k\sigma'} c_{\delta l\sigma} \ , \tag{2.29}$$

where the matrix $W_{ijkl}^{\alpha\beta\gamma\delta}$ is,

$$W_{ijkl}^{\alpha\beta\gamma\delta} = \int d^3x \int d^3x' \; \tilde{v}(x,x')\phi_{\alpha i}^*(x)\phi_{\beta j}^*(x')\phi_{\gamma k}(x')\phi_{\delta l}(x) \; . \tag{2.30}$$

If we have well localized Wannier functions, we expect many terms of equation (2.29) to be very small. The standard approximation is to neglect all contributions of (2.29) except the local one, i.e.

$$\sum_{iss'} W_{iiii} c_{i\sigma}^\dagger c_{i\sigma'}^\dagger c_{i\sigma'} c_{i\sigma} = U \sum_i n_{i\uparrow} n_{i\downarrow} \; , \tag{2.31}$$

where $U = 2W_{iiii}$. The normal-ordered operators in (2.29) ensure that the unphysical self-interaction term ($\sigma = \sigma'$) drops out. The local interaction parameter U is called Hubbard-U. This harsh approximation was justified by Hubbard for the $3d$-transition metals [1]. The order of magnitude of the Hubbard-U is about $U \approx 20$ eV. The largest of the neglected terms, belonging to the nearest neighbor interaction V is about $V \approx 6$ eV, without taking screening into account. He estimated that screening would further reduce V to $V \approx 2 - 3$ eV. It thus is an order of magnitude smaller than U and can be neglected.

The interactions neglected in (2.31) can be classified into two types: direct terms and exchange terms. The direct terms are given by interaction matrix elements $V_{ij} = W_{ijji}$ and can be included in the standard Hubbard model by terms like

$$V = \sum_{i \neq j} V_{ij} n_i n_j \; , \tag{2.32}$$

where n_i is given by $n_i = n_{i\uparrow} + n_{i\downarrow}$. We will treat those terms later, when we look into organic conductors. The exchange terms are obtained by W_{ijij} interaction matrix elements and give rise to inter-site magnetic couplings (cf. [14]).

To conclude we give some equivalent representations of the Hubbard model. In real space it is given by

$$H = -\sum_{\sigma, i \neq j, \nu} t_{ij,\nu} c_{i\nu\sigma}^\dagger c_{j\nu\sigma} + \sum_{\sigma,i,\nu} \varepsilon_i c_{i\nu\sigma}^\dagger c_{i\nu\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} \; , \tag{2.33}$$

or rather after Fourier transforming with $c_{\mathbf{k}\nu\sigma} = \dfrac{1}{\sqrt{L}} \sum_i c_{\mathbf{R}_i\nu\sigma} e^{i\mathbf{k}\cdot\mathbf{R}_i}$ in k-space

$$H = \sum_{\mathbf{k}\sigma\nu} \epsilon_{\mathbf{k}\nu} c_{\mathbf{k}\nu\sigma}^\dagger c_{\mathbf{k}\nu\sigma} + \frac{U}{L} \sum_{\mathbf{k},\mathbf{k}',\mathbf{q},\nu} c_{\mathbf{k}\nu\uparrow}^\dagger c_{\mathbf{k}-\mathbf{q},\nu\uparrow} c_{\mathbf{k}'\nu\downarrow}^\dagger c_{\mathbf{k}'+\mathbf{q},\nu\downarrow} \; . \tag{2.34}$$

Sometimes it is also given in mixed Bloch and Wannier representation

$$H = \sum_{\mathbf{k}\sigma\nu} \epsilon_{\mathbf{k}\nu} c_{\mathbf{k}\nu\sigma}^\dagger c_{\mathbf{k}\nu\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} \; , \tag{2.35}$$

which stresses that the kinetic energy part is diagonal in k-space, whereas the interaction part is diagonal in real space.

2.2.1 Important features of the Hubbard model

The Hubbard model describes the interplay of potential and kinetic energy. Depending on the ratio of U/t either the Coulomb interaction or the kinetic energy dominates.

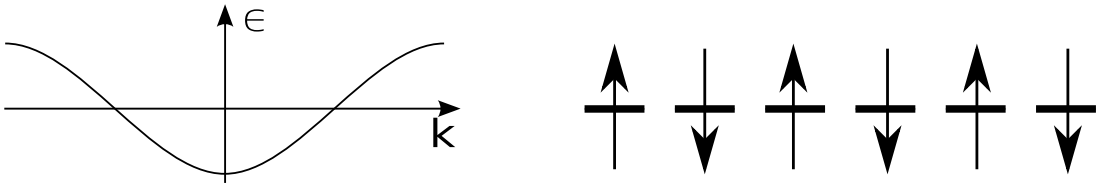


Figure 2.8: Band vs. atomic limit: The left side shows the cosine dispersion relation of the band-limit. If the band is not completely filled the system is metallic. The right side shows a ground state of the atomic limit for a half-filled system. The system is an antiferromagnetic Mott-insulator.

In case of $U/t \ll 1$ and systems of dimension[2] $n > 1$ we retain the results of the TBA with small corrections and thus the well-known cosine band (cf. figure 2.8). The system therefore is metallic unless the band is completely filled. Electrons are delocalized all over the system and their wave functions can to good approximation be written as Slater determinants.

If $U/t \gtrapprox 1$ the movement of one electron is influenced by the locations of all other electrons in the system since double occupation of a site is energetically expensive and thus they try to avoid each other. Hence we cannot describe this system by an independent-electron approximation. Wave functions for those systems are difficult to find and generally have to be computed numerically. This regime is said to be correlated.

If the ratio $U/t \gg 1$ it is even called strongly correlated. Electrons try to distribute themselves as uniformly as possible on the sites in order to minimize the Coulomb energy. In the case of a non-degenerate half-filled system this would lead to exactly a single electron per site. Hopping hardly takes place and can be treated as a perturbation. So we decompose equation (2.33) as

$$H = H_0 + H_p$$

where

$$H_0 = U \sum_i n_{i\uparrow} n_{i\downarrow}$$

denotes the (unperturbed) Coulomb potential and

$$H_p = -\sum_{\langle ij \rangle, \sigma} t_{ij} c^\dagger_{i,\sigma} c_{j,\sigma}$$

[2]Lieb and Wu showed [6], that in a one dimensional half-filled Hubbard chain for $U > 0$ the systems become Mott insulators.

the hopping (perturbation). H_0 is diagonal in real-space. This suggests a local picture. All sites are (almost) independent of each other. Hence this limit is called atomic limit. The energy only depends on the number of doubly occupied sites. Let $|D, j\rangle$ denote a basis vector of a configuration with D doubly occupied orbitals. j is a number which indexes different configurations with energy UD, thus:

$$H_0 \, |D, j\rangle = UD \, |D, j\rangle \ .$$

Atomic limit with half-filling

How is the degenerate ground state of a half-filled system in the atomic limit affected if we turn on hopping? In first order,

$$E^{(1)}_{0,k}\delta_{lk} = \langle 0l|H_p|0k\rangle \ ,$$

there is no correction to the energy, since the hopping creates one doubly occupied and one empty site and thus connects independent subspaces with $D = 0$ and $D = 1$. The second order correction leads to the effective Heisenberg Hamiltonian in the subspace with $d = 0$:

$$H = \sum_{\langle ij \rangle} \frac{4t_{ij}^2}{U} \left(\mathbf{S_j S_i} - \frac{1}{4} \right) \tag{2.36}$$

Since the prefactor $J = 4t_{ij}^2/U$ is positive this system is antiferromagnetic. An illustrative explanation for this behavior is the following: If the system was in a ferromagnetic ground state, increasing t/U would not lead to hopping, since it would be forbidden by the Pauli principle. In an antiferromagnetic ground state virtual hopping occurs and gives rise to a decrease in energy. This situation is depicted in the right sketch of figure 2.8.

Mott transition

Such a system is not only antiferromagnetic but also a Mott insulator. This means that for small U/t it is a metal. Between both limits there certainly must be, at some critical U_c/t, a transition between the metallic and insulating phase. This transition is called Mott transition. Lieb and Wu [6] showed analytically, that in one dimension a half-filled cluster always is an insulator for $U > 0$.

Figure 2.9 shows the spectral function for a one-dimensional chain with 12 sites and half-filling for different values of U. For $U = 0$ there is spectral weight at the Fermi level and therefore the system is a metal. For all other values of U a gap opens at the Fermi level, giving rise to two bands called Hubbard bands. Since there is no spectral weight at the Fermi level the system is an insulator.

Distribution of eigenenergies

Not only the ground state is of interest but also the distribution of the excitation energies. We restrict our considerations here to a one dimensional half-filled Hubbard

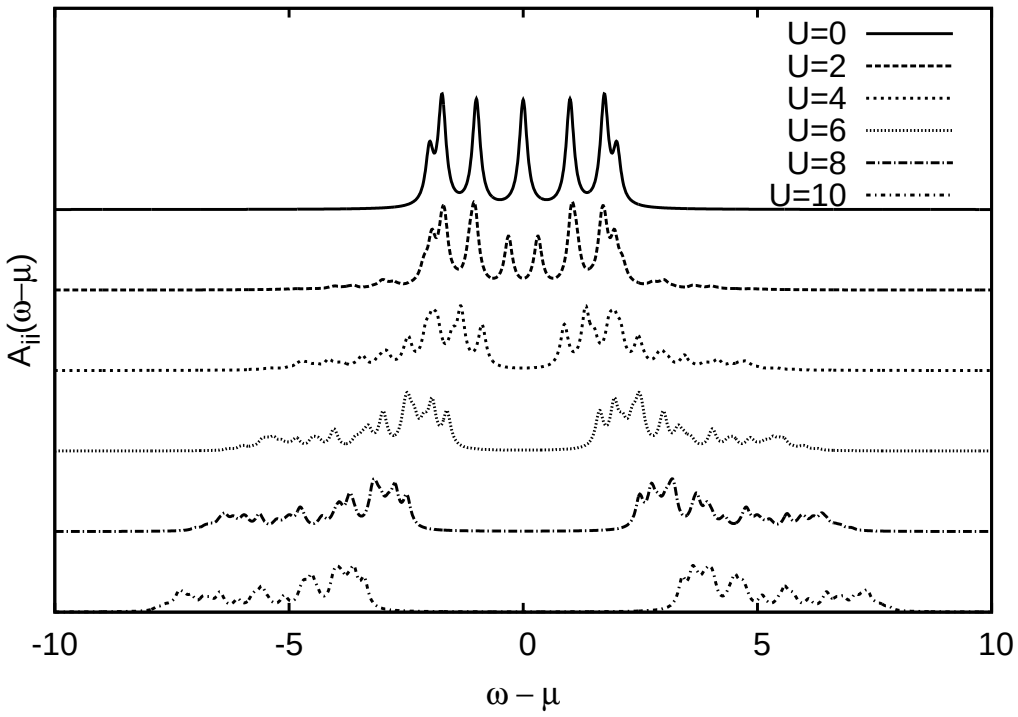


Figure 2.9: Mott insulator in one dimension for a 12 sites half-filled Hubbard chain (U as in image) with $t = 1$.

chain, for $U \gg t$. To study the ground state and the lowest-lying excitation energies we resort to the antiferromagnetic Heisenberg model, which is equivalent to the Hubbard model in this situation (cf. chapter 2.2.1).

In the ground state each site is singly occupied and neighboring spins have opposite spin directions due to J being positive. This configuration is twofold degenerate since flipping all spins yields the same energy. Flipping a single spin increases the energy by $2J$ since the flipped spin has two parallel neighboring spins. Flipping more spins increases the number of Js to be paid. The width of the lowest lying energy band is thus proportional to J and therefore to U^{-1}. The next band has one double occupancy leading to an energy offset of U. In general the D^{th} band has the energy offset $D \cdot U$. Moreover there are D vacant sites in the D^{th} band and thus hopping becomes possible. This leads to a bandwidth proportional to $t \gg J$ (cf. chapter 2.1.3), which is actually wider than the one of the first band. If the bandwidth becomes greater than U, bands may overlap.

By the use of combinatorics it is possible to calculate the number of states in each band. Let us consider a half-filled L site system with even value of L and $N_\uparrow = N_\downarrow = L/2$. In the lowest band each site is occupied. There are $\binom{L}{N_\uparrow} = \binom{L}{L/2}$ ways to distribute the identical up-spin particles. The full configuration is unambiguously determined by this distribution, since all other sites are occupied by the identical

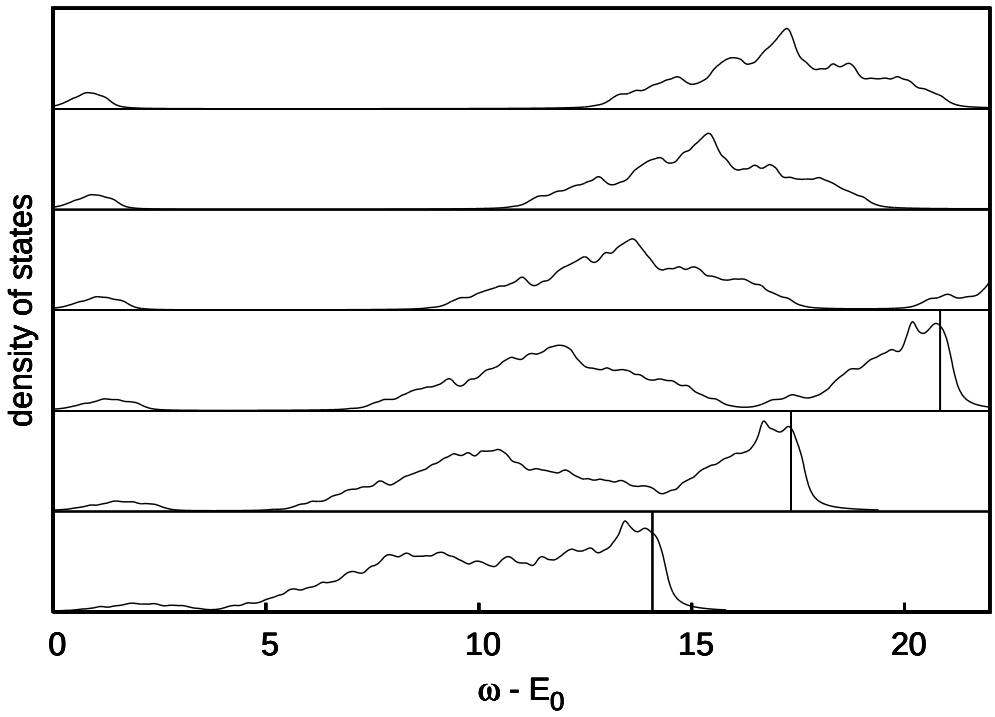


Figure 2.10: Density of many-particle configuration states of an 8 site system with half-filling for different $U = \{6, 8, 10, 12, 14, 16\}$. The perpendicular lines in the lowest three plots show the cut-off of the calculation. Contributions above are unphysical and only stem from Lorentzian broadening.

$\downarrow$-spin particles, or more formally $\binom{L-N_\uparrow}{N_\downarrow} = \binom{N_\downarrow}{N_\downarrow} = 1$. Thus the number of states in the lowest band is

$$\binom{L}{L/2} .$$

The second band again yields $\binom{L}{N_\uparrow} = \binom{L}{L/2}$ ways to distribute the up-electrons. One site is doubly occupied, leading to another factor $N_\uparrow$. So there are $L - N_\uparrow$ sites left to distribute the rest, i.e. $N_\downarrow - 1$, of the down electrons or a single hole. Thus leading to the number of states in the second band

$$\binom{L}{N_\uparrow} N_\uparrow \binom{L - N_\uparrow}{N_\downarrow - 1} = \binom{L}{N_\uparrow} N_\uparrow N_\downarrow = \binom{L}{L/2} L^2/4 .$$

For illustration purposes let us consider an 8 site system. The lowest lying band has 70 and the next one 1120 states. The density of states is plotted in figure 2.10 for $U/t = \{6, 8, 10, 12, 14, 16\}$. The bandwidth of the first band becomes narrower with increasing U due to $J = 4t^2/U$. Figure 2.11 shows the expected proportionality.

The second lowest band is broader since it is proportional to t and hardly changes with varying U. One also observes the linear energy offset in U, which, shifts the center of weight of the D^{th} band by $D \cdot U$.

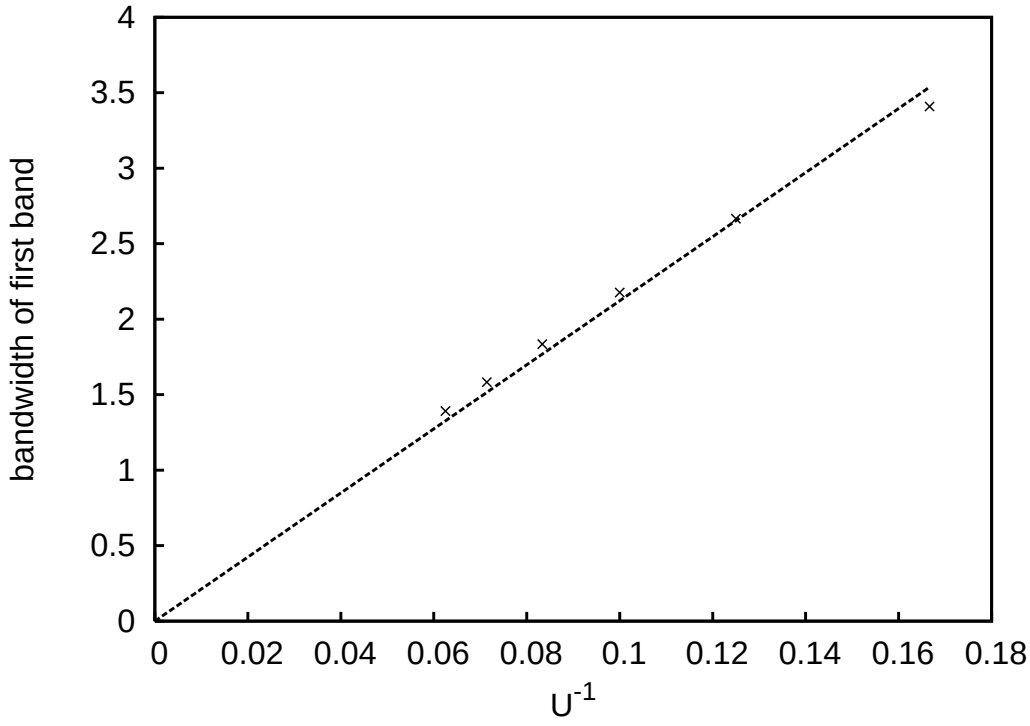


Figure 2.11: Bandwidth of lowest band in a half-filled system. It has a $1/U$ dependence.

The lowest three plots in figure 2.10 also show, that the third band superimposes with the second one. This is a result of the bandwidth of those two bands being larger than the shift of their center of gravity.

These calculations were carried out using ARPACK[3]. The first 2000 eigenvalues are calculated and in order to plot their peaks, they are broadened by a Lorentzian function. The energy of the 2000^{th} eigenvalue is denoted by a perpendicular line in the lower three plots. The weight above this line is a consequence of the finite broadening and yields no real states.

2.2.2 Gutzwiller wave function

Because of the Pauli principle there are four different states a lattice site can have. The site is either empty $|\cdot\rangle$, occupied by a single electron with either up $|\uparrow\rangle$ or down $|\downarrow\rangle$ spin, or by two electrons with opposite spins $|\uparrow,\downarrow\rangle$. Thus the Fock space of a system with L lattice sites has the dimension 4^L. Even for few sites, this Hilbert space becomes "galactic" (cf. table 3.1). Since the Hamiltonian conserves the number of particles and their spin, we can restrict our calculations to a Hilbert space with a

[3]A FORTRAN linear algebra software packages, which contains subroutines for diagonalizing large matrices. See http://www.caam.rice.edu/software/ARPACK/

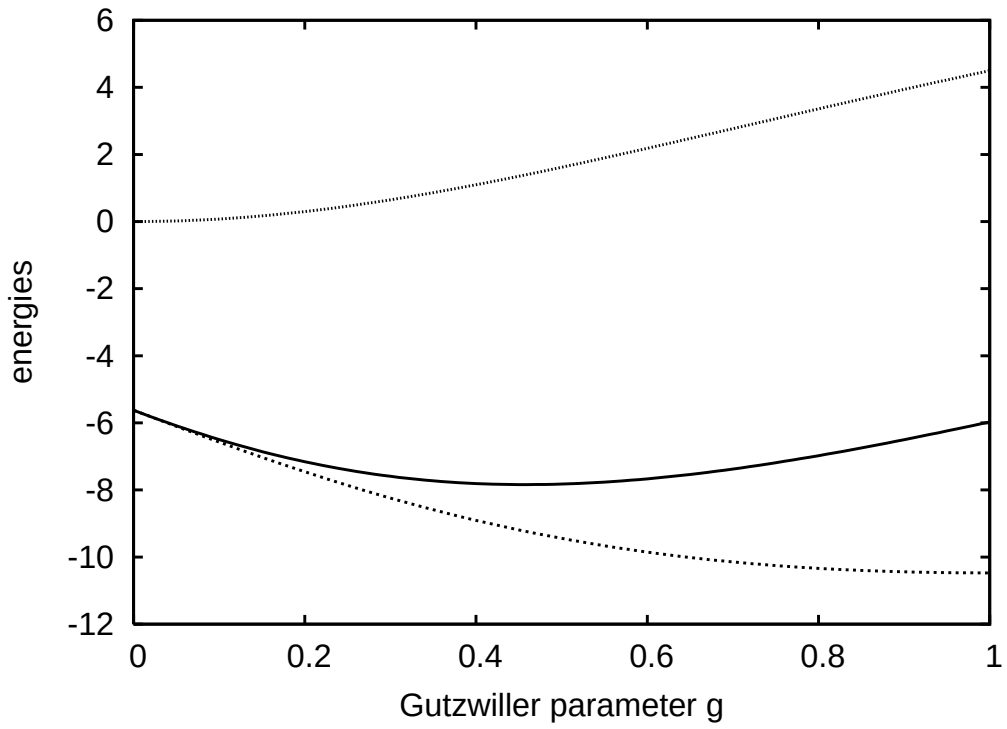


Figure 2.12: Variation of Gutzwiller wave function for various values of Gutzwiller parameter g in a system comprising 10 sites and 3 electrons of either spin for $t = 1$ and $U/t = 5$. The graphs show the g-dependent expectation values of the Coulomb potential $\langle V \rangle_g$ (dotted line), the total energy $\langle H \rangle_g$ (solid line) and the kinetic energy $\langle T \rangle_g$ (dashed line). For $g = 0$ there are no double occupancies and thus $\langle V \rangle_0 = 0$. For $g = 1$ $\langle T \rangle_1$ is minimized. In between is the optimal value of $g = g_{opt}$ which minimizes the total energy $\langle E \rangle_{g_{opt}}$.

fixed number of particles of either spin-type $N_e = N_\uparrow + N_\downarrow$.

In the absence of electron-electron interaction the ground-state wave function is a Slater determinant, hence there are no correlations. One might argue, however, that even the Pauli principle is a correlation effect: for spin-like particles it indeed impedes double occupancies and thus particles somehow feel each other. It is astonishing that Slater determinants efficiently remedy this problem. Hilbert space is a huge place and Slater determinants only comprise $N_e^2 = N_e \cdot \#O$ degrees of freedom, where $\#O = N_e$ is the number of occupied spin orbitals. Moreover for spin-like particles this effect also solves the local Hubbard interaction term! But as soon as a second spin type or next-neighbor repulsion enter the scene, no such easy solutions are known and we have to take more degrees of freedom into account.

A practical way to obtain an approximation to the ground-state wave function is

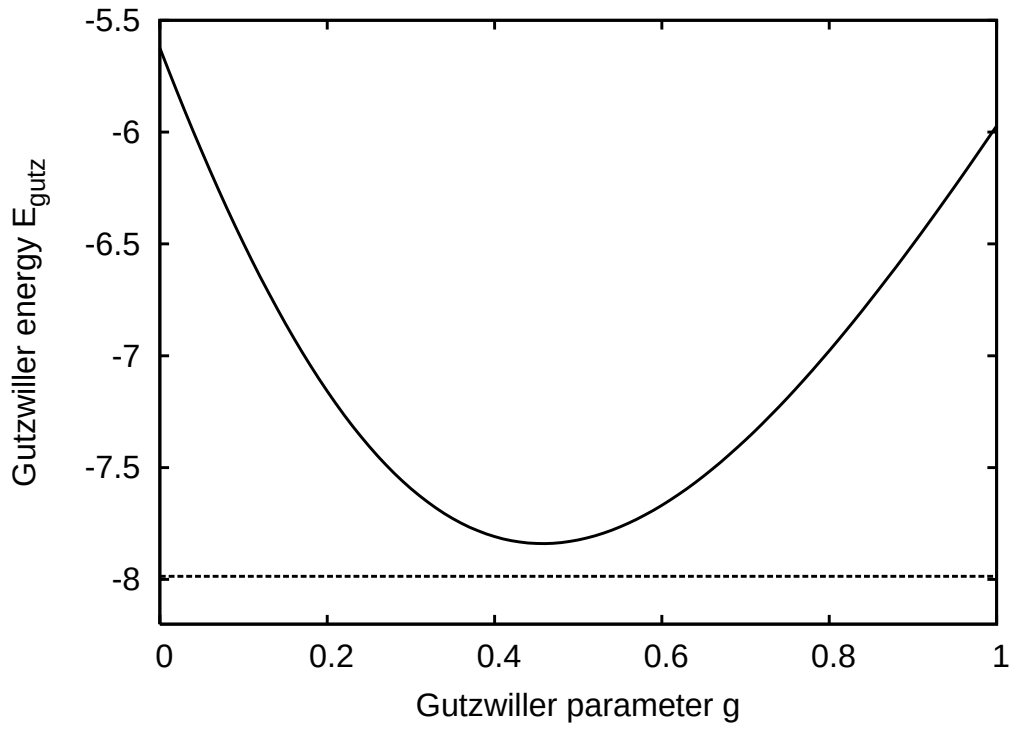


Figure 2.13: Energy variation of Gutzwiller wave function for various values of parameter g in a system comprising 10 sites and 3 electrons of either spin for $t = 1$ and $U/t = 5$. The Gutzwiller wave function approximates the actual ground state quite well (dashed line is actual ground-state energy).

the use of the variational principle. It states that

$$\frac{\langle \Psi_g | H | \Psi_g \rangle}{\langle \Psi_g | \Psi_g \rangle} = E_g \geq E_0 \ , \tag{2.37}$$

where E_0 is the ground-state energy of H and g a set of variational parameters for the trial wave functions $|\Psi_g\rangle$. By minimizing this expression and possibly enlarging the space spanned by $|\Psi_g\rangle$ we can systematically approach the ground state $\min_g(E_g) \approx E_0$. A good ansatz for the $|\Psi_g\rangle$ requires physical intuition. As an example we will briefly discuss the Gutzwiller wave function [4] for the Hubbard model (2.33).

In case of $U = 0$, i.e. no correlations, the probability of finding a σ electron at a site i is $n_\sigma = N_\sigma/L$. Hence the probability for this site to be empty or doubly occupied is $(1 - n_\uparrow)(1 - n_\downarrow)$ or $n_\uparrow n_\downarrow$ respectively. For a half-filled system, all four configurations have the same weight. When we slowly increase the value of U double occupancies become increasingly expensive and hence their weight is suppressed. For $N_e \leq L$ in the limit of $U \to \infty$ the weight of doubly occupied sites goes to zero.

Gutzwiller therefore proposed the following ansatz,

$$|\Psi_\eta\rangle = \frac{1}{\Gamma} \prod_i^L (1 - \eta n_{i\uparrow} n_{i\downarrow}) |\Phi_0\rangle \ , \tag{2.38}$$

where $|\Phi_0\rangle$ is the Slater determinant of the uncorrelated electrons, Γ a normalization factor and $\eta \in [0:1]$ the variational parameter.

As one can directly verify equation (2.38) can equivalently be written as

$$|\Psi_g\rangle = \frac{1}{\Gamma} g^{\sum_i n_{i\uparrow} n_{i\downarrow}} |\Phi_0\rangle \ , \tag{2.39}$$

where $g = 1 - \eta$.

For $g = 1$ (or equivalently $\eta = 0$) we obviously retain the ground state of the system for $U = 0$, whereas for $g = 0$ ($\eta = 1$) all double occupancies are suppressed. Thus g regulates the interplay between kinetic and Coulomb energy. This can be seen from figure 2.12. For $g = 1$ $\langle T \rangle_1$ is minimized. $\langle V \rangle_1$, however, is quite large. For $g = 0$ the roles are exchanged. In-between there must be a minimum of the total energy at a g_{opt}.

In figure 2.13 the expectation value $\langle H \rangle_g$ is plotted in comparison to the actual ground-state energy. We see, that Gutzwiller's wave function is quite a good approximation as far as the ground-state energy is concerned. Studies by Millis and Coppersmith [15] point out, that Gutzwiller states are always metallic. However, we know from Lieb and Wu [6] that in one dimension for $U > 0$ the system always is an insulator.

Gutzwiller also introduced an approximation to calculate $\langle \Psi_g | H | \Psi_g \rangle = E_g$, the Gutzwiller approximation. In this approximation a half-filled system can go over to an insulating state at a critical value of U_c. It is the famous Brinkman-Rice transition [16], which is, however, an artifact of the Gutzwiller approximation. The approximation to the variational wave functions describes the physics correctly. An overview of the Gutzwiller approximation can be found in [17].

3 Exact diagonalization

Eigenvalue problems for huge matrices arise in many fields of modern natural and engineering sciences. The problems are as diverse as for instance structural analysis in civil engineering or stability analysis of electronic networks. IBM's HITS ("hypertext induced topic selection") [18] or Google's famous PageRank algorithm [19], which are used to rate search results, work by diagonalizing huge matrices. PageRank seems to be the largest application of the Lanczos method.

Huge matrices are only manageable if they are sparse, i.e. have few nonzero elements. A commonly used definition for sparsity, attributed to Wilkinson, is that a matrix is sparse "whenever it is possible to take advantage of the number and location of its nonzero entries". The real-space Hubbard Hamiltonian is sparse, as we will see in the first part of this chapter. Thus only the nonzero elements and their positions need to be stored, leading to a huge reduction in storage requirements. Moreover the matrix-vector multiplication is an $\mathcal{O}(n)$ operation instead of an $\mathcal{O}(n^2)$.

For sparse matrices iterative algorithms, which heavily rely on matrix-vector products, are the methods of choice. In this chapter we will discuss the power and the more advanced Lanczos method to diagonalize such matrices. It turns out, that the Lanczos method is not only capable of giving us the ground-state eigenpair of a Hamiltonian, but also provides an efficient way of computing dynamical correlation functions. We will deal with this at the end of the chapter.

3.1 Basis encoding and sparsity

Choice of basis set To diagonalize the Hamiltonian, we need a basis set. Both, the Hubbard Hamiltonian in its real-space and k-space representation,

$$H = -\sum_{\sigma,i\neq j,\nu} t_{ij,\nu} c^\dagger_{i\nu\sigma} c_{j\nu\sigma} + \sum_{\sigma,i,\nu} \varepsilon_i c^\dagger_{i\nu\sigma} c_{i\nu\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} \tag{3.1}$$

$$H = \sum_{\mathbf{k}\sigma\nu} \epsilon_{\mathbf{k}\nu} c^\dagger_{\mathbf{k}\nu\sigma} c_{\mathbf{k}\nu\sigma} + \frac{U}{L} \sum_{\mathbf{k},\mathbf{k}',\mathbf{q},\nu} c^\dagger_{\mathbf{k}\nu\uparrow} c_{\mathbf{k}-\mathbf{q},\nu\uparrow} c^\dagger_{\mathbf{k}'\nu\downarrow} c_{\mathbf{k}'+\mathbf{q},\nu\downarrow} \,, \tag{3.2}$$

are natural candidates. In k-space the kinetic energy is diagonal, as is the Coulomb energy in real-space. Taking a closer look at the non-diagonal parts, clearly suggests the use of a real-space configuration basis. In the tight-binding picture we have hopping to only near, possibly even only to nearest neighbors. Thus there are only very few nonzero elements, i.e. the Hamiltonian is sparse. Storage requirements and computing time only scale with $\mathcal{O}(n)$, where $n = \dim(\mathcal{H})$. In k-space representation,

the Coulomb interaction term contains a sum over three indices, all ranging over number of sites. This term is clearly non-sparse.

Hilbert space dimension The Hubbard Hamiltonian in its second quantized form is independent of the actual number of particles in the system. It can be diagonalized in Fock space. For L sites the Fock space's dimension is 4^L, since each site can either host a single electron, two electrons or one of either spin. This Fock space, however, grows exponentially and even for small number of sites a solution is difficult to obtain. A single many-body wave function of a 18 sites system requires 512 GB of data to store it in double precision. Even though the Hubbard Hamiltonian is independent of the number of particles, it conserves charge and spin. There are no terms which can flip a spin or effectively create/annihilate an electron. Thus, in a sub-Fock space with fixed spin up and down electron numbers, $N_\uparrow$ and $N_\downarrow$, the Hamiltonian is block-diagonal and we can restrict our calculations to these subspaces, the Hilbert spaces $\mathcal{H}$. Its dimension is given by,

$$\dim(\mathcal{H}) = \binom{L}{N_\uparrow} \cdot \binom{L}{N_\downarrow} . \tag{3.3}$$

This mitigates the problem somewhat (cf. table 3.1). A wave function's data storage requirements for 18 sites now depends on the number of particles in the system. At half-filling the Hilbert space is largest, and the wave function needs 17 GB. At third-filling only 2 GB are needed.

Basis setup Suppose the cluster in arbitrary dimensions consists of L sites and hosts $N_e = N_\uparrow + N_\downarrow$ electrons. The basis vectors in real-space-occupation-number representation are given by,

$$|n_{1\uparrow} \dots n_{L\uparrow}\rangle \otimes |n_{1\downarrow} \dots n_{L\downarrow}\rangle = |n_{1\uparrow} \dots n_{L\uparrow}, n_{1\downarrow} \dots n_{L\downarrow}\rangle \ ,$$

where $n_{i\sigma}$ is the occupation of state i with a spin-σ electron, thus either 0 or 1. As an example consider a four sites system with two up and a single down electron. A state of this system is, e.g.

$$|0101, 0100\rangle \ ,$$

which in second quantization reads,

$$c^\dagger_{2,\uparrow} c^\dagger_{4,\uparrow} c^\dagger_{2,\downarrow} |0\rangle \ .$$

To represent such a state on a computer we need to store two bits per site, namely their spin up/down occupation. Hence we can efficiently represent the complete state with two binary arrays or equivalently two integers. The first one holds the up electron, the second one the down electron configuration. Therefore the bit length of the integers must be equal or greater than the number of sites. The restriction of the calculation on Hilbert spaces with fixed $N_\uparrow$ and $N_\downarrow$ translates into integers whose

Table 3.1: Dimension of Hilbert space $\dim(\mathcal{H})$ and corresponding computer memory requirements for a single many-body wave function of Hubbard models with L sites and $N_\uparrow + N_\downarrow$ electrons. The first group of numbers gives the dimensions for half-filling, where the Hilbert space is largest. The second group shows how the dimension grows with the filling.

L	$N_\uparrow$	$N_\downarrow$	dim of Hilbert space	memory
2	1	1	4	
4	2	2	36	
6	3	3	400	
8	4	4	4 900	
10	5	5	63 504	
12	6	6	853 776	6 MB
14	7	7	11 778 624	89 MB
16	8	8	165 636 900	1263 MB
18	9	9	2 363 904 400	17 GB
20	10	10	34 134 779 536	254 GB
20	1	1	400	
20	2	2	36 100	
20	3	3	1 299 600	9 MB
20	4	4	23 474 025	179 MB
20	5	5	240 374 016	1833 MB
20	6	6	1 502 337 600	11 GB
20	7	7	6 009 350 400	44 GB
20	8	8	15 868 440 900	118 GB
20	9	9	28 210 561 600	210 GB
20	10	10	34 134 779 536	254 GB

binary representations have exactly N_σ bits set. It is convenient to introduce integer labels for those configuration integers. With the label of the up $i_\uparrow$ and the label of the down $i_\downarrow$ electron configuration given, the full state is unambiguously determined. We can denote this state i with a tuple $i \equiv (i_\uparrow, i_\downarrow)$.

Lin developed an efficient and convenient way of establishing the one-to-one correspondence between the single-spin configurations and the set of label integers. His two-table method is described in [20]. Our own implementation uses a slightly different approach. For both spin types an integer array is set up, which holds all integers, whose binary representation has exactly N_σ bits set. These arrays are set up by looping over all integers in $\left[0, 2^L - 1\right]$. If an integer meets the criterium it is appended to the array. The index can now be regarded as label for the state.

Such a table for the example above of a 4 sites chain with $N_\uparrow = 2$ and $N_\downarrow$ may look like:

index	tuple	up	dn	index	tuple	up	dn
0	(0,0)	0011	0001	12	(0,2)	0011	0100
1	(1,0)	0101	0001	13	(1,2)	0101	0100
2	(2,0)	0110	0001	14	(2,2)	0110	0100
3	(3,0)	1001	0001	15	(3,2)	1001	0100
4	(4,0)	1010	0001	16	(4,2)	1010	0100
5	(5,0)	1100	0001	17	(5,2)	1100	0100
6	(0,1)	0011	0010	18	(0,3)	0011	1000
7	(1,1)	0101	0010	19	(1,3)	0101	1000
8	(2,1)	0110	0010	20	(2,3)	0110	1000
9	(3,1)	1001	0010	21	(3,3)	1001	1000
10	(4,1)	1010	0010	22	(4,3)	1010	1000
11	(5,1)	1100	0010	23	(5,3)	1100	1000

Note that the basis is set up in such a way, that for a fixed $i_\downarrow$, all up electron configurations follow. Moreover, we know from the Hamiltonian that each hopping term only affects either up or down electrons. Thus it only changes one label in the tuple for a given state. These two observations will become important in the implementation, see chapter 4.

Reading an element in an array at a given position i is an $\mathcal{O}(1)$ operation. This important lookup (label $i \equiv (i_\uparrow, i_\downarrow) \rightarrow$ configuration) is therefore very efficient. Moreover we only need $\sum_\sigma \binom{L}{N_\sigma}$ integer elements.

The reverse lookup is not as important. That is why we choose to use binary searching on the configuration arrays to save memory. It would need 2^L elements. Binary searching is possible since the elements are ordered by construction. Hence, the reverse lookup is a $\mathcal{O}(n \log n)$ operation.

Our basis handling data structures thus need the two arrays, containing the spin configurations, and amount to a memory consumption of $\sum_\sigma \binom{L}{N_\sigma} \times$ sizeof(int).

3.2 The power method

The matrix, we want to diagonalize, is sparse and therefore iterative approaches are the methods of choice. A very basic iterative method is the so-called power method. Consider an arbitrary state $|\psi\rangle \in \mathcal{H}$, which must not have vanishing overlap with the ground-state vector. Expanding $|\psi\rangle$ in terms of the eigenbasis $|i\rangle$ of H yields

$$|\psi\rangle = \sum_i C_i |i\rangle \ .$$

Let E_i be the eigenvalue of $|i\rangle$. Applying a power m of the matrix H to $|\psi\rangle$ yields

$$H^m |\psi\rangle = \sum_i C_i E_i^m |i\rangle \ . \tag{3.4}$$

For m sufficiently large the state $|i\rangle$ with the largest absolute eigenvalue $|E_i|$ and $C_i \neq 0$ dominates the sum and hence the corresponding eigenvector $|i\rangle$ is projected out. In practice the calculated vector is normalized after each matrix-vector multiplication, i.e.

$$|\psi_n\rangle = H^n |\psi\rangle / \|H^n |\psi\rangle\| .$$

The eigenvalue with the largest modulus may not be the ground state, we are interested in. If $|E_{max}| > |E_{min}|$ we can shift the spectrum by replacing H with $\hat{H} = H - \rho\mathbb{1}$. If $\rho > 0$ is sufficiently large, $|E_{min} - \rho| > |E_{min} - \rho|$ holds and the method converges to the ground-state vector of H.

This method for obtaining the ground-state eigenvalue and its vector is in principle exact, simple and easily implemented. However convergence is rather slow, see for example [21] or look at the comparison in figure 3.1. Moreover, if the ground state is not the eigenvector, whose eigenvalue has the largest modulus, we must already have information about the spectrum of the matrix in order to find an appropriate ρ. If ρ is chosen too large the convergence speed suffers considerably.

A more advanced way of calculating the ground-state eigenpair is the Lanczos method. Still both methods are intimately connected to each other. The n^{th} state

$$|\psi_n\rangle = H^n |\psi\rangle / \|H^n |\psi\rangle\|$$

is a linear combination in the space spanned by

$$\{|\psi\rangle, H|\psi\rangle, H^2|\psi\rangle, \ldots, H^{m-1}|\psi\rangle\} .$$

Diagonalizing the matrix in this subspace typically yields a better linear combination, and thus a better approximation to the ground state.

3.3 The Lanczos method

At the heart of the Lanczos method is the concept of invariant subspaces [21]. The central idea is, that we diagonalize the matrix H in a small subspace of the Hilbert space. If this subspace is an invariant subspace, we will see that the eigenpairs of the smaller Hamiltonian yield a subset of the eigenpairs of H. We choose a Krylov subspace, which ensures convergence to the extremal eigenpairs. Thus, to obtain the ground-state eigenpair, we diagonalize the Hamiltonian in this subspace.

We start with an introduction to invariant subspaces. Let $\mathcal{H} = \mathbb{C}^n$ be the Hilbert space of our system and $H \in \mathbb{C}^{n\times n}$ its Hamiltonian. $\mathcal{K}_m$ denotes an m-dimensional subspace of $\mathcal{H}$. Then $\mathcal{K}_m$ is called an invariant subspace of $\mathcal{H}$ if

$$H\mathcal{K}_m = \{H\mathbf{k} \mid \mathbf{k} \in \mathcal{K}_m\} \subseteq \mathcal{K}_m .$$

m eigenvectors of H, for instance, span an m dimensional invariant subspace.

Let the columns of $Q = (q_1, \ldots, q_k)$ be a basis of a k-dimensional invariant subspace $\mathcal{Q}_k$ of H. For $Hq_i \in \mathcal{Q}_k$, there must be a unique vector t_i such that

$$Hq_i = Qt_i, \quad i \in [1, k] .$$

If we define T to be $T = (t_1, \ldots, t_k)$, then

$$HQ = QT$$

and therefore the matrix T is the representation of H on $\mathcal{Q}_k$ with respect to basis Q.

If $\mathbf{x} \in \mathcal{Q}_k$ it can be expanded in basis Q as $\mathbf{x} = Q\mathbf{z}$. If, moreover, $H\mathbf{x} = \lambda\mathbf{x}$, then

$$\begin{array}{rcl} H\mathbf{x} & = & \lambda\mathbf{x} \\ = \; HQ\mathbf{z} & = & \lambda Q\mathbf{z} \\ = \; Q\,T\mathbf{z} & = & Q\,\lambda\mathbf{z}\,. \end{array} \tag{3.5}$$

This list of equations can be read in both directions. Hence we see, why invariant subspaces are important for the Lanczos method. They establish a one-to-one mapping between the eigenpairs of T and a subset of the eigenpairs of the matrix H.

Of course in practice the Lanczos method does not generate an exact invariant subspace, instead we will have an approximate one which means:

$$HQ - QT = G\,.$$

Or equivalently $(H+E)Q = QT$ with $E = -GQ^\dagger$. If $\|E\| = \|G\|$ is sufficiently small and $H + E$ is well conditioned, then an eigenpair of T is a good approximation to an eigenpair of H.

The Lanczos method generates a Krylov subspace. It is defined as the space spanned by

$$\mathcal{K}_{m,\mathbf{b}} = \{\mathbf{b}, H\mathbf{b}, H^2\mathbf{b}, \ldots, H^{m-1}\mathbf{b}\}\,,$$

where $\mathbf{b}$ denotes an arbitrary non-zero normalized starting vector[1]. This vector must not be orthogonal to the ground-state vector. The sequence $\mathcal{K}_m$ is called a Krylov sequence, where we dropped the explicit dependence on the start vector to simplify the notation.

Orthogonalization on $\mathcal{K}_m$ yields an orthogonal basis in the Krylov subspace. The matrix H in this basis turns out to be tridiagonal and can be diagonalized by standard means. The resulting eigenvalues of the tridiagonal matrix yield a very good approximation to a fraction of the eigenvalues of the original matrix H, if the Krylov space is large enough. It shows that even very small Krylov spaces yield very good results. Convergence is very fast. In order to get an understanding why this is the case, we will first discuss the so called single-step Lanczos technique.

3.3.1 The single-step Lanczos

A general way to obtain the ground state is the variational method. The basic idea is to minimize the wave-function functional,

$$E\,[\psi] = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}\,. \tag{3.6}$$

[1] Often $\mathbf{b}$ is chosen randomly, especially if no a priori information about the ground state is available. For Hubbard models, one may start from a Gutzwiller wave vector (cf. 2.2.2).

The vector $|\psi_0\rangle$, which minimizes (3.6), is the ground state of H. The Lanczos method exploits this idea which can be seen when considering a slightly modified Lanczos method [22], the single-step Lanczos technique. It works as follows:

We choose a random vector $|\phi_0\rangle$, leading to a first energy value $E_0 = E[\phi_0]$ with (3.6). In analogy to the gradient descent method we look for the direction of steepest descent in the energy functional, given by the functional derivative

$$\left.\frac{\delta E[\psi]}{\delta \psi}\right|_{\phi_0} = \frac{2}{\langle\phi_0|\phi_0\rangle}\left(H - E[\phi_0]\right)|\phi_0\rangle \ . \tag{3.7}$$

To include this direction we thus consider the additional vector $\left|\tilde{\phi_0}\right\rangle = H|\phi_0\rangle$. This way the Krylov subspace is built.

In contrast to the power method or ordinary gradient descent methods we include the new dimension in our variational basis and thus have two degrees of variational freedom. To simplify the calculation we orthonormalize the new vector with respect to the starting vector, yielding

$$\langle\phi_1|H|\phi_0\rangle\,|\phi_1\rangle = H|\phi_0\rangle - \langle\phi_0|H|\phi_0\rangle\,|\phi_0\rangle \ .$$

If $|\phi_1\rangle$ turns out to be zero we started from an exact eigenstate of H which already is an invariant subspace. A normalized linear combination in this two dimensional subspace is in general given by,

$$|\psi_\theta\rangle = \cos(\theta)\,|\phi_0\rangle + \sin(\theta)\,|\phi_1\rangle \ .$$

Minimizing the expectation value (3.6) of this linear combination with respect to the free parameter θ always yields a better approximation to the ground-state eigenvalue than both E_0 or $E_1 = \langle\phi_1|H|\phi_1\rangle$, unless $|\phi_0\rangle$ was already an eigenvector. This is a direct consequence of the variational principle. Alternatively, we can diagonalize the matrix

$$T = \begin{pmatrix} \alpha_0 & \beta_1 \\ \beta_1 & \alpha_1 \end{pmatrix} ,$$

where

$$\alpha_i = \langle\phi_i|H|\phi_i\rangle \qquad\qquad \beta_1 = \langle\phi_1|H|\phi_0\rangle \ .$$

The lower eigenstate of this matrix is equal to $|\phi_{\theta^*}\rangle$, with θ^* chosen in such a way that (3.6) is minimized.

This new and improved ground-state approximation (Ritz pair) can be used as starting vector for the next single-step Lanczos pass yielding again a refined ground-state approximation. The whole procedure can be repeated until the ground-state vector or its energy is sufficiently well converged.

Figure 3.1 shows the convergence to the ground-state energy of this single-step Lanczos method (dashed line) in comparison to the full Lanczos method and the power method (dotted line). The single-step Lanczos method clearly is superior to the power method. Moreover it shares the advantage of needing only two large vectors and of having the latest ground-state approximation directly at hand.

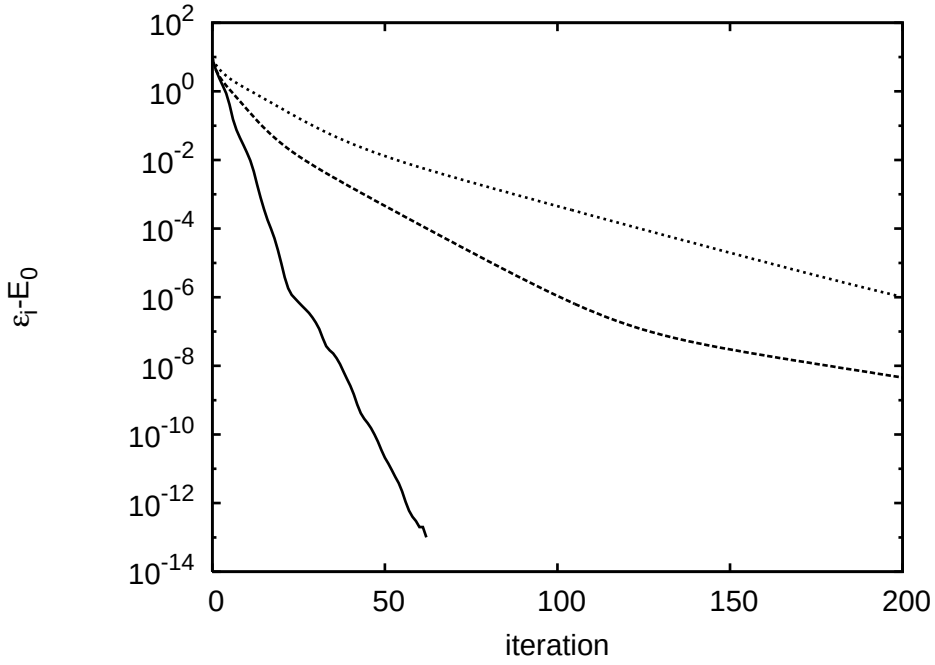


Figure 3.1: Comparison of power method (dotted line), single-step (dashed line) and full Lanczos method (solid line): convergence to the ground-state energy for a small half-filled Hubbard system (10 sites; $\dim(\mathcal{H}) = 63504$) with $U = 5t$. The full Lanczos method is by far the fastest method to converge.

3.3.2 The full Lanczos method

But we need not to restrict the Lanczos iterations to a single one. It can be generalized to larger subspaces, increasing the variational degrees of freedom. Hence, we expect the minimization to be more efficient.

The first step is exact as in the algorithm described above: we again start from a random vector $\mathbf{b} = |\phi_0\rangle$, perform the matrix vector multiplication and normalize the result with respect to the starting vector, i.e.

$$\langle\phi_1|H|\phi_0\rangle\,|\phi_1\rangle = H\,|\phi_0\rangle - \langle\phi_0|H|\phi_0\rangle\,|\phi_0\rangle \ . \tag{3.8}$$

But instead of stopping the iteration and diagonalizing H in this two-dimensional subspace we start a new iteration yielding

$$\langle\phi_2|H|\phi_1\rangle\,|\phi_2\rangle = H\,|\phi_1\rangle - \langle\phi_1|H|\phi_1\rangle\,|\phi_1\rangle - \langle\phi_1|H|\phi_0\rangle\,|\phi_0\rangle \ . \tag{3.9}$$

In general the $n+1$ Krylov vector can be written recursively as:

$$\beta_{n+1}\,|\phi_{n+1}\rangle = H\,|\phi_n\rangle - \alpha_n\,|\phi_n\rangle - \beta_n\,|\phi_{n-1}\rangle \ , \tag{3.10}$$

where $n \in 2, \ldots, m$ and

$$\alpha_n = \langle\phi_n|H|\phi_n\rangle\,, \quad \beta_{n+1} = \langle\phi_{n+1}|H|\phi_n\rangle\,.$$

An equivalent expression for β_{n+1} is $\beta_{n+1} = \|H|\phi_n\rangle - \alpha_n|\phi_n\rangle - \beta_n|\phi_{n-1}\rangle\|$. It is often used in practice since it is numerically superior (refer to [23] and [24]).

From equation (3.10) we see that

$$\langle\phi_m|H|\phi_n\rangle = 0 \quad \text{for } |n-m| > 1\,, \tag{3.11}$$

and hence the matrix H in this subspace with the basis $\{|\phi_n\rangle\}_n$ is a tridiagonal with matrix elements β_i and α_i,

$$T = \begin{pmatrix} \alpha_0 & \beta_1 & 0 & 0 & 0 & \dots \\ \beta_1 & \alpha_1 & \beta_2 & 0 & 0 & \dots \\ 0 & \beta_2 & \alpha_2 & \beta_3 & 0 & \dots \\ 0 & 0 & \beta_3 & \alpha_3 & \beta_4 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{pmatrix}\,. \tag{3.12}$$

This matrix T can be easily diagonalized with standard means, like for example library routines found in LAPACK. Let (λ_i, z_i) be an eigenpair of T. Because of equations (3.5) the so-called Ritz pair (λ_i, Qz_i) is an approximation to the corresponding eigenpair of H.

The reason why this method is so powerful is due to the fact that relatively few Lanczos iterations m are needed to converge reasonably well towards the ground-state energy. Kaniel and Paige showed that this is the case under quite general conditions. Typically m is of the order of a few dozens. Therefore only a few matrix-vector multiplications are needed.

Obtaining eigenvalues: the first pass

If we are solely interested in the ground-state eigenvalue we need only two Lanczos vectors, as in the single-step Lanczos method above. The first pseudo code, listing 3.1, assumes a method "mult" of the matrix data type which calculates the matrix-vector products and returns the result. The size of the arrays α and β is m and $m-1$ respectively, where m denotes the maximum number of Lanczos iterations. **b** denotes the starting vector. **b** as well as **q** are of the dimension of the Hilbert space n. Since n increases rapidly with the system size the storage requirements for these vectors determine how large systems we can treat.

```
1   q[0:n-1]= 0;
    β0 = j = 1;
3   j = 0;
    while (βj ≠ 0):
5       if j ≠ 0:
            for i in (0:n−1):
7               t = bi; bi = qi/βj; qi = −βj t;
            end;
9       end;
```

$q = q + H.\text{mult}(b)$;
11 $j = j + 1; \alpha_j = \langle b|q\rangle\,; q = q - a_j b; \beta_j = \|q\|_2;$
end;

Listing 3.1: Pseudo code of the Lanczos algorithm for the first pass to obtain the ground-state energy. It is based on the Lanczos algorithm from ([25] see algorithm 9.2.1) using a modified Gram-Schmidt method for superior numerical stability. b contains the normalized starting vector.

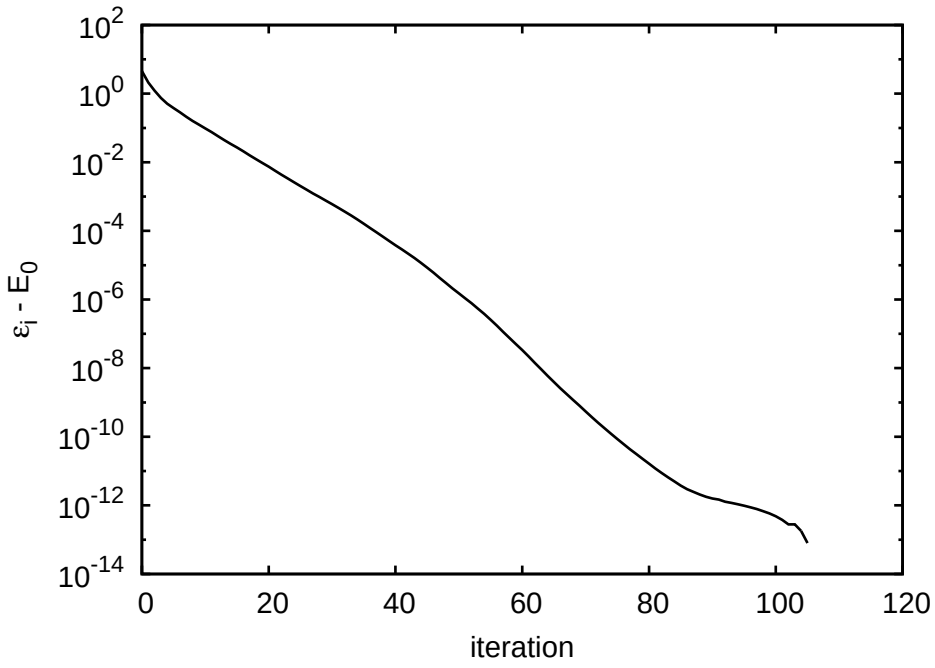


Figure 3.2: Convergence of a large Lanczos run. The plot shows the exponential-like convergence to the ground-state energy for a 20 sites Hubbard system with 6 electrons of either spin ($U = 1.96$, $t = 0.18$). Despite the size of the Hilbert space ($\dim(\mathcal{H}) = 1\,502\,337\,600$) only about 100 iterations are needed for convergence.

In theory, i.e. with exact arithmetics, the algorithm terminates after $i \leq \dim(\mathcal{H})$ because of line 4. This means that an invariant subspace has been reached. In practice this condition will never be true due to numerical inaccuracies, hinting at the break down of orthogonality.

The loss of orthogonality which goes hand in hand with the convergence to an eigenvalue leads to duplicate as well as spurious eigenvalues (ghosts). These effects will be dealt with in chapter 3.3.3. Waiting until an invariant subspace is reached is infeasible anyway, because it would most probably become too huge. We can, however, stop the iteration when the ground-state energy is sufficiently well converged.

Therefore we check in each iteration if

$$\|\lambda_0(m) - \lambda_0(m-1)\|/\|\lambda_0(m)\| < \varepsilon \ ,$$

where ε is an arbitrarily small value.

Figure 3.1 shows the exponential convergence of the Lanczos algorithm for the ground-state energy of a small Hubbard chain (8 sites with half-filling, i.e. $\dim(\mathcal{H}) = 63504$) in comparison to the other discussed methods. Only about $m \approx 60$ iterations are needed, to converge to the ground-state energy with numerical accuracy. Even for considerably larger system m remains of the same order. An example for such a large system is shown in figure 3.2. It is a 20 sites Hubbard chain with $U = 1.96$, $t = 0.18$ and 6 electrons of either spin. The resulting Hilbert space has a dimension of $\dim(\mathcal{H}) = 1\,502\,337\,600$ and still only about 100 iterations are needed to obtain the ground-state energy to numerical accuracy.

Computing the ground-state vector

In the so-called second pass the Lanczos algorithm builds an eigenvector, typically the ground-state vector. If we stored all orthogonal Krylov vectors in the first pass we could simply calculate the Ritz vector, i.e. a good approximation to the ground-state vector, $\mathbf{r}$ by

$$\mathbf{r} = \sum_{j=0}^{m} y[j]\mathbf{K}_j \ , \tag{3.13}$$

where m is the number of iterations, $\mathbf{y}$ the ground-state vector of T and the $\mathbf{K}_j$ the m orthonormal Krylov vectors. This is, put differently, the basis transformation of the ground-state vector from the subspace to the full Hilbert space.

To avoid having to store all Krylov vectors in the first pass we restart the iterations from the same start vector $\mathbf{b}$. In each iteration we obtain the Krylov vector we already had in the first first pass and can thus successively calculate the ground-state vector by equation (3.13). Therefore a third Lanczos vector is needed. The corresponding code is shown in listing 3.2.

```
1   q[0:n-1]= 0;
    r[0:n-1]= 0;
3   for (j = 0, j < m, j++ ):
        if j ≠ 0:
5           for i in (0:n−1):
                t = b_i; b_i = q_i/β_j; q_i = −β_j t;
7           end;
        end;
9       r = r + y[j] b;
        q = q + H.mult(b);
11      q = q − a_j b;
```

end;

Listing 3.2: Pseudo code of second pass of Lanczos algorithm. b contains the start vector of the first pass, y the eigenvector corresponding to the lowest eigenvalue of the tridiagonal matrix of the first pass. After termination, r contains the ground-state Ritz vector, i.e. an approximation to the ground-state vector of H.

In order to test whether the ground-state vector is accurate enough we calculate the residual norm given by

$$r = \|E_0\mathbf{r} - H\mathbf{r}\| \,, \tag{3.14}$$

where E_0 denotes the ground-state energy. In practice residual norms r of the magnitude of $10^{-5} - 10^{-7}$ can be achieved even for huge systems like 20 sites with 7 electrons of either spin, resulting in $\dim(\mathcal{H}) = 6\,009\,350\,400$. If the accuracy is not high enough it is possible to refine the vector by using the single-step Lanczos method discussed in chapter 3.3.1. Figure 3.3 shows the convergence of the residual norm if a matrix is diagonalized by the single-step technique alone.

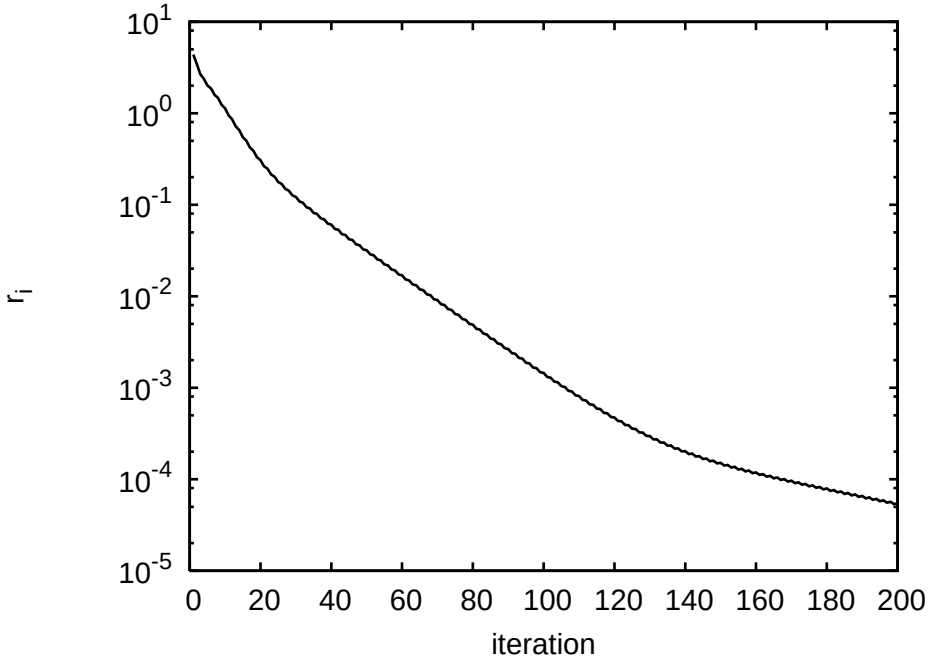


Figure 3.3: Convergence of residual norm, r_i with the single-step Lanczos method for a half-filled Hubbard model (10 sites; $\dim(\mathcal{H}) = 63\,504$) with $U = 5t$.

Note that when performing the Lanczos method for systems having a degenerate ground state one vector from this eigenspace will be obtained. The actual vector depends on the starting vector.

3.3.3 Numerical aspects

With exact arithmetics all Lanczos vectors are mutually orthonormal, i.e.

$$\langle \phi_m | \phi_n \rangle = \delta_{mn} \ .$$

But in real calculations the mutual orthogonality is lost due to numerical inaccuracies. Only local orthogonality is retained, i.e. only Lanczos vectors i, j with $0 < i, j \leq m$ and $|i - j|$ small are mutually orthonormal. For larger values of $|i - j|$ this does not hold; possibly not even linear independence.

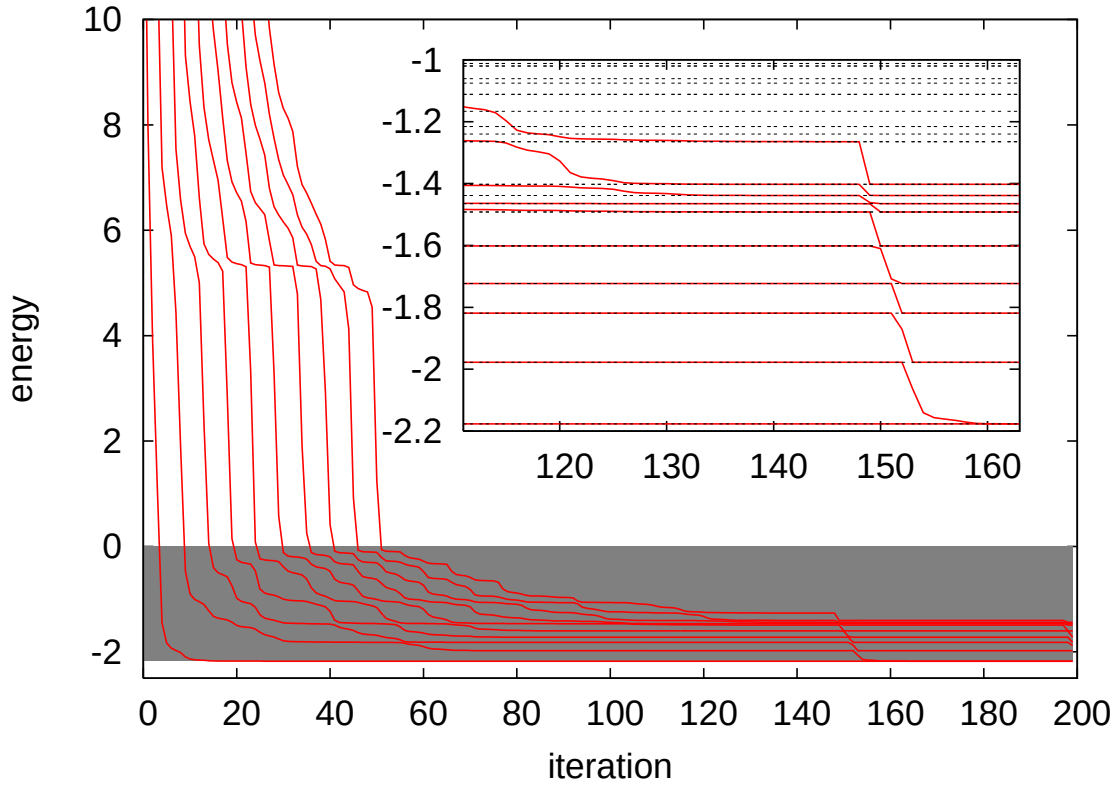


Figure 3.4: 200 Lanczos iterations for a small half-filled Hubbard system with 8 sites ($U = 10t$). The light gray region in the plot shows the first band (exactly calculated with ARPACK), consisting of 70 energy levels. The red lines show the first ten Lanczos eigenvalues against the number of iterations. The inset shows a magnified region of the original plot. Here the dashed black lines denote the actual energy levels. We observe for example that the ground state is duplicated ($\approx$ 155 iteration). Note, however, that convergence is already reached after 20 iterations.

This loss of mutual orthogonality can give rise to duplicate eigenvalues, see figure 3.4, and spurious eigenvalues. According to Paige [23], this happens in finite precision

arithmetics as soon as an eigenvalue is (almost) converged.

An intuitive explanation is again given by the analogy between the method of gradient descent and the single-step Lanczos method. If an eigenpair is almost converged, the application of H yields almost the same vector. This translates into the image of gradient descent to a saddle point where the norm of the gradient becomes very small. In the Lanczos picture we orthogonalize two almost identical vectors, an operation which is ill-conditioned. After normalization this new vector has the same weight as all the others but it points in a more-or-less random direction which in turn breaks orthogonality as well as probably linear independence. As a result, we obtain more and more duplicate eigenvalues of T which correspond to a single Ritz value of H. This applies to the eigenvectors as well. Several eigenvectors of T yield the same Ritz vector. However, for the convergence to the ground state this does not matter since due to the variational principle we can, once the ground state is reached, neither go lower nor higher.

Lanczos already was aware of the problem of the loss of mutual orthogonality. His simple suggestion was to orthogonalize with all preceding Lanczos vectors at each step (full orthogonalization). This however would spoil the appealing features of this algorithm, namely that we only have to store three Lanczos vectors. For large vectors the method would often have to be restarted due to storage constraints. Restarting however leads to slow convergence as observed with the single-step Lanczos technique.

Other approaches are the semi-orthogonalization methods. The main idea is to reorthogonalize only when a loss in orthogonality is detected. The various methods differ in the choice of the vectors, which are used in the reorthogonalization step. Selective reorthogonalization for example performs the reorthogonalization with respect to all nearly converged Ritz vectors [26].

3.4 Computing dynamical response functions

The Lanczos method can also be used as an efficient tool for computing dynamical response functions. We start with a short introduction to the Green's function formalism and then discuss how to actually extract them and other response functions from the Lanczos method.

3.4.1 General introduction

Green's function in time domain The causal single-particle Green's function $G_{i\sigma j\sigma'}(t,t')$ in a many-particle system governed by a Hamiltonian H is defined in such a way that we can interpret $G_{i\sigma j\sigma'}(t,t')$ as the probability amplitude of finding an electron ($t > t'$) or a hole ($t' > t$) of spin σ in state i at time t, when at time t' an electron/hole of spin σ' has been added to the ground state of H in state j. Thus,

$$G_{i\sigma j\sigma'}(t,t') = -i \left\langle \psi_0 \middle| \mathcal{T}(c_{i\sigma}(t) c^\dagger_{j\sigma'}(t')) \middle| \psi_0 \right\rangle \ , \tag{3.15}$$

where $|\psi_0\rangle$ is the ground state, σ the spin index and $\mathcal{T}$ denotes the time ordering operator, that arranges the operators according to time with the operator, whose value of t is the largest, at the left. Each permutation yields a minus sign. The $c_{i\sigma}^{(\dagger)}(t)$ are the creation/annihilation operators in the Heisenberg picture, i.e.

$$c_{i\sigma}^{\dagger}(t) = e^{iHt} c_{i\sigma}^{\dagger} e^{-iHt} \ , \tag{3.16}$$

where $c_{i\sigma}^{\dagger}$ is the ordinary (Schrödinger picture) creation operator. If the Hamiltonian H is not time dependent the Green's function only depends on the difference $t - t'$. Thus we can w.l.o.g. set $t' = 0$, only retaining t as argument. In this case the energy of the system is conserved. To simplify the notation we suppress the spin index from here on.

We now consider a finite system with N electrons. Inserting the eigenstates of the system with a missing or additional electron $\left|\psi_n^{N\pm1}\right\rangle$, we obtain the spectral representation from (3.15) with (3.16),

$$G_{ij}(t) = \begin{cases} -i \sum_n \left\langle \psi_0 \middle| c_i \middle| \psi_n^{N+1} \right\rangle \left\langle \psi_n^{N+1} \middle| c_j^{\dagger} \middle| \psi_0 \right\rangle e^{i(E_0 - E_n^{N+1})t} & : \quad t > 0 \\ i \sum_n \left\langle \psi_0 \middle| c_j^{\dagger} \middle| \psi_n^{N-1} \right\rangle \left\langle \psi_n^{N-1} \middle| c_i \middle| \psi_0 \right\rangle e^{i(E_n^{N-1} - E_0)t} & : \quad t < 0 \end{cases} \ , \tag{3.17}$$

where the $E_n^{N\pm1}$ are the energy eigenvalues of H to the eigenvectors $\left|\psi_n^{N\pm1}\right\rangle$. This is the spectral- or Lehmann representation.

Density matrix From G_{ij} we obtain the two particle density matrix of the system by,

$$\rho(i,j) = \lim_{t \to 0^-} \Im \, G_{ij}(t) = \left\langle \psi_0 \middle| c_j^{\dagger} c_i \middle| \psi_0 \right\rangle \ , \tag{3.18}$$

where 0^- is an infinitesimal small negative value to ensure proper ordering of the operators. The density matrix can be employed to compute arbitrary one-particle observables, by evaluating the trace of the product $\mathcal{O}\rho$. Thus,

$$\langle \mathcal{O} \rangle = \mathrm{Tr} \left[\rho \mathcal{O} \right] \ . \tag{3.19}$$

Typical density matrices for one-dimensional half-filled Hubbard chains are shown in figure 3.5. In a translationally invariant system the density matrix only depends on the modulus of the difference $|i - j|$ just like the Green's function. According to [8] for a one-dimensional half-filled system ρ_{0j} can be to a good approximation written as

$$\rho_{0i} = \frac{\sin\left(\pi i/2\right)}{\pi i} e^{-\gamma |i|} \ , \tag{3.20}$$

where γ denotes a decay constant, which is zero for $U = 0$. We indeed observe this behavior in figure 3.5. For values of $U > 0$, the density matrix is damped exponentially, leading to local physics.

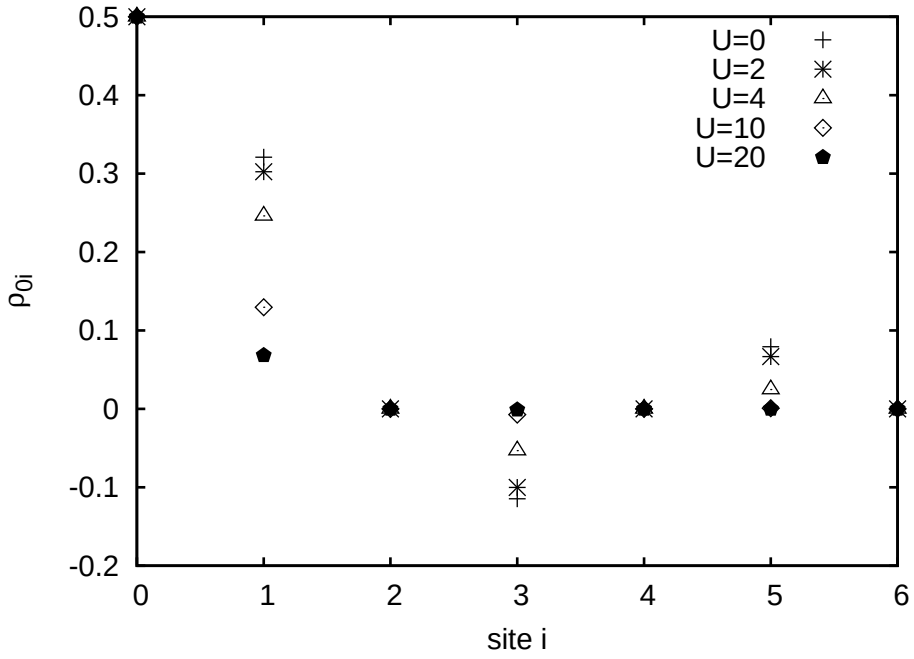


Figure 3.5: Density matrix of the σ electrons in a 14 sites half-filled Hubbard chain with $t = 1$ and different values of U. $\rho_{00} = 0.5$ yields the particle density.

Green's function in energy domain The Fourier transform from time to frequency of (3.17) , i.e.

$$G_{ij}(\omega) = \int_{-\infty}^{\infty} dt\; e^{i\omega t} G_{ij}(t) \;,$$

leads to the spectral representation

$$G_{ij}(\omega) = \sum_n \frac{\left\langle \psi_0 \middle| c_i^\dagger \middle| \psi_n^{N-1} \right\rangle \left\langle \psi_n^{N-1} \middle| c_j \middle| \psi_0 \right\rangle}{\omega + (E_n^{N-1} - E_0) - i\eta} + \sum_n \frac{\left\langle \psi_0 \middle| c_i \middle| \psi_n^{N+1} \right\rangle \left\langle \psi_n^{N+1} \middle| c_j^\dagger \middle| \psi_0 \right\rangle}{\omega - (E_n^{N+1} - E_0) + i\eta} \;. \tag{3.21}$$

The first term of this equation describes the extraction of an electron, a process similar to photoemission. In a real experiment a photon impinges upon the target material and ejects an electron, whose kinetic energy is measured. The second term describes the injection of an electron into the system, similar to inverse photoemission where a photon is emitted. G_{ij}, however, neglects the electron-photon interaction.

More technically, for each eigenstate $\left|\psi_n^{N\pm1}\right\rangle$ in the $N \pm 1$ particle system there obviously is a pole in the lower/upper complex half-plane at $\mp(E_n^{N\pm1} - E_0)$ with spectral weight $\left\langle \psi_0 \middle| c_i \middle| \psi_n^{N+1} \right\rangle \left\langle \psi_n^{N+1} \middle| c_j^\dagger \middle| \psi_0 \right\rangle$, $\left\langle \psi_0 \middle| c_i^\dagger \middle| \psi_n^{N-1} \right\rangle \left\langle \psi_n^{N-1} \middle| c_j \middle| \psi_0 \right\rangle$ respectively.

The poles in the lower half-plane, i.e. belonging to the inverse photoemission, start at $\mu^{N+1} = E_0^{N+1} - E_0$ and increase in energy whereas the poles in the upper half-plane start at $\mu^N = E_0 - E_0^{N-1}$ and decrease in energy. In the limit of $N \to \infty$, $\mu^N = \mu^{N+1} = \mu$ and μ is called chemical potential.

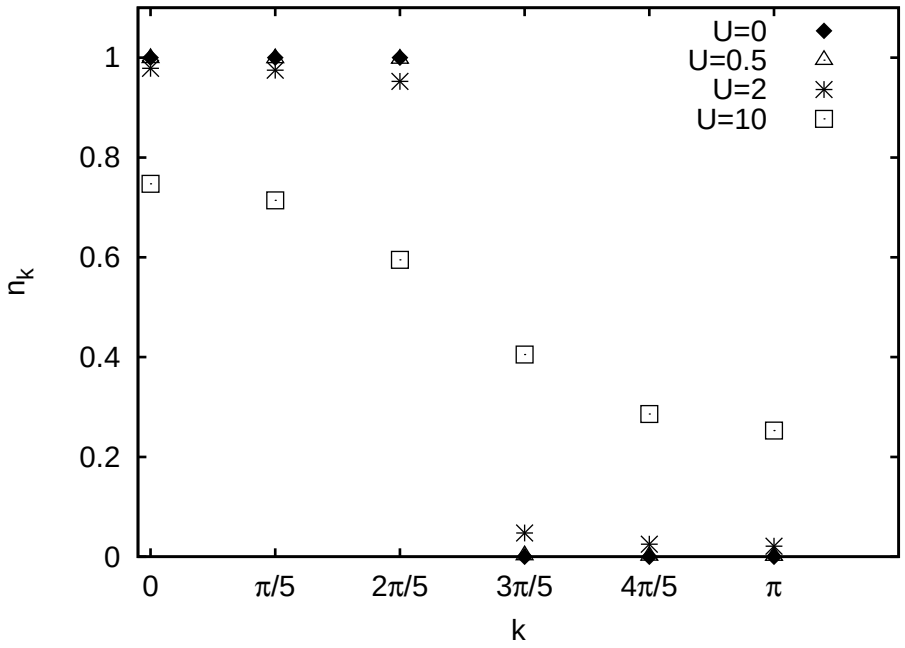


Figure 3.6: Momentum distribution of a 10 sites one-dimensional half-filled Hubbard chain for $t = 0$ and various values of U. For $U = 0$ we retain the Fermi distribution for $T = 0$.

Green's function in k-space To obtain the Green's function in k-space we need to Fourier transform the particle operators. Writing the momentum operators in their coordinate space representation yields,

$$c_{\mathbf{k}} = \sum_i \frac{1}{\sqrt{V}} e^{-i\mathbf{k}\cdot\mathbf{R}} c_i \ . \tag{3.22}$$

In a translationally invariant system the Green's function is diagonal in k-space. In this case it is given by,

$$G_{\mathbf{k}}(\omega) = \sum_n \frac{\left|\left\langle \psi_0 \middle| c_{\mathbf{k}}^\dagger \middle| \psi_n^{N+1} \right\rangle\right|^2}{\omega - (E_n^{N+1} - E_0) + i\eta} + \sum_n \frac{\left|\left\langle \psi_0 \middle| c_{\mathbf{k}}^\dagger \middle| \psi_n^{N-1} \right\rangle\right|^2}{\omega + (E_n^{N-1} - E_0) - i\eta} \ . \tag{3.23}$$

The momentum distribution is given – in analogy to the density matrix – by the Fourier transformation of (3.23) into the time domain and evaluating the expression at $t = 0^-$. This yields,

$$\langle n_{\mathbf{k}} \rangle = \Im\, G_{\mathbf{k}}(0^-) = \left\langle \psi_0 \middle| c_{\mathbf{k}}^\dagger c_{\mathbf{k}} \middle| \psi_0 \right\rangle \ . \tag{3.24}$$

In the case of non-interacting particles, all single-particle states are filled up to the Fermi wave vector k_F. Thus the momentum distribution is $\left\langle \psi_0 \middle| c_{\mathbf{k}}^\dagger c_{\mathbf{k}} \middle| \psi_0 \right\rangle =$

$\Theta(|\mathbf{k}-\mathbf{k}_F|)$, where Θ is Heaviside's step function. A typical momentum distribution in a one-dimensional half-filled Hubbard chain for different values of U is shown in figure 3.6. In [8] Koch and Goedecker showed, that the momentum distribution in this system can be written as

$$\langle n_{\mathbf{k}} \rangle = \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{\cos(ka)}{\sinh(\gamma)}\right) \tag{3.25}$$

to a good approximation, where γ is the same decay constant as in equation (3.20).

Spectral function To evaluate the spectral function defined as

$$A_{\alpha\beta}(\omega) = -\frac{1}{\pi} \Im\, G_{\alpha\beta}(\omega) \;, \tag{3.26}$$

we use the well-known identity

$$\frac{1}{x+i\eta} = \mathcal{P}\left(\frac{1}{x}\right) - i\pi\delta(x) \;, \tag{3.27}$$

where $\eta \to 0^+$ and $\mathcal{P}$ denotes the principal part. We apply it to equation (3.21), yielding

$$\begin{aligned} A_{\alpha\beta}(\omega) = & \sum_n \left\langle \psi_0 \middle| c_\alpha \middle| \psi_n^{N+1} \right\rangle \left\langle \psi_n^{N+1} \middle| c_\beta^\dagger \middle| \psi_0 \right\rangle \delta(\omega - (E_n^{N+1} - E_0)) \\ + & \sum_n \left\langle \psi_0 \middle| c_\alpha^\dagger \middle| \psi_n^{N-1} \right\rangle \left\langle \psi_n^{N-1} \middle| c_\beta \middle| \psi_0 \right\rangle \delta(\omega + (E_n^{N-1} - E_0)) \;, \end{aligned} \tag{3.28}$$

where α and β denote either sites i, j or momenta k, k'. The diagonal elements $A_{\alpha\alpha}$ are positive, which can easily be seen from,

$$\begin{aligned} A_{\alpha\alpha}(\omega) = & \sum_n \left| \left\langle \psi_0 \middle| c_\alpha \middle| \psi_n^{N+1} \right\rangle \right|^2 \delta(\omega - (E_n^{N+1} - E_0)) \\ + & \sum_n \left| \left\langle \psi_0 \middle| c_\alpha^\dagger \middle| \psi_n^{N-1} \right\rangle \right|^2 \delta(\omega + (E_n^{N-1} - E_0)) \;. \end{aligned} \tag{3.29}$$

Sum rules The anti-commutation relation for Fermions $\left\{c_\alpha, c_\beta^\dagger\right\} = \delta_{\alpha\beta}$ leads to a sum rule for the spectral weights. Expanding the anti-commutator and using the partition of unity yields

$$\delta_{\alpha\beta} = \left\langle \psi_0 \middle| \left\{c_\alpha, c_\beta^\dagger\right\} \middle| \psi_0 \right\rangle = \left\{ \begin{array}{ll} & \sum_n \left\langle \psi_0 \middle| c_\alpha \middle| \psi_n^{N+1} \right\rangle \left\langle \psi_n^{N+1} \middle| c_\beta^\dagger \middle| \psi_0 \right\rangle \\ + & \sum_n \left\langle \psi_0 \middle| c_\beta^\dagger \middle| \psi_n^{N-1} \right\rangle \left\langle \psi_n^{N-1} \middle| c_\beta \middle| \psi_0 \right\rangle \end{array} \right\} \;. \tag{3.30}$$

Comparing (3.30) to (3.28) shows that for the spectral function $A_{\alpha\beta}(\omega)$ the following formula holds,

$$\int_{-\infty}^{\infty} d\omega\; A_{\alpha\beta}(\omega) = \delta_{\alpha\beta} \;. \tag{3.31}$$

Note, that the photoemission spectral function

$$\int_{-\infty}^{\mu} d\omega\ A_{ij} = \int_{-\infty}^{\infty} d\omega\ A_{ij}^{\mathrm{PE}}(\omega) = \left\langle \psi_0 \middle| c_j^\dagger c_i \middle| \psi_0 \right\rangle = \rho_{ij}$$

yields the density matrix. Moreover, we can write for its diagonal elements

$$\int_{-\infty}^{\mu} d\omega\ A_{\alpha\alpha} = \int_{-\infty}^{\infty} d\omega\ A_{\alpha\alpha}^{\mathrm{PE}} = n_\alpha\ , \tag{3.32}$$

giving the occupation of state α. And similarly for the inverse photoemission spectral function

$$\int_{\mu}^{\infty} d\omega\ A_{\alpha\alpha} = \int_{-\infty}^{\infty} d\omega\ A_{\alpha\alpha}^{\mathrm{IPE}} = 1 - n_\alpha\ . \tag{3.33}$$

With

$$\sum_\alpha \int_{-\infty}^{\mu} d\omega\ A_{\alpha\alpha} = N\ , \tag{3.34}$$

where N is the number of particles, we have means to calculate μ.

n-particle Green's functions To generalize the concept of single particle Green's functions, we introduce higher-order or n-particle Green's functions. The n-body real-time Green's function is defined analogous to (3.15) as

$$\begin{aligned} &G^{(n)}(\alpha_1 t_1, \ldots, \alpha_n t_n | \alpha_1' t_1', \ldots, \alpha_n' t_n') = \\ &\quad (-i)^n \left\langle \psi_0 \middle| \mathcal{T}\left[c_{\alpha_1}(t_1) \ldots c_{\alpha_n}(t_n) c_{\alpha_n'}^\dagger(t_n') \ldots c_{\alpha_1'}^\dagger(t_1') \right] \middle| \psi_0 \right\rangle\ . \end{aligned} \tag{3.35}$$

Many-body expectation values and correlation functions can be expressed by choosing the corresponding order of the Green's function and suitable time arguments. The density-density correlation function in real-time can for example be written as,

$$\langle n_i(t) n_j(t') \rangle = \left\langle \psi_0 \middle| c_i^\dagger(t) c_i(t) c_j^\dagger(t') c_j(t') \middle| \psi_0 \right\rangle = G^{(2)}(it, jt' | it, jt')\ . \tag{3.36}$$

Moments of the distribution Let $\mathcal{O}$ be an arbitrary operator. The correlation function is then defined by,

$$I_{\mathcal{O}}(\omega) = -\frac{1}{\pi} \Im \left\langle \psi_0 \middle| \mathcal{O}^\dagger \frac{1}{\omega - (H - \mu) + i\eta} \mathcal{O} \middle| \psi_0 \right\rangle\ . \tag{3.37}$$

If $\mathcal{O}$, for example, denotes an annihilation operator c_i then the already discussed photoemission spectral function $A_{ii}^{\mathrm{PE}}(\omega)$ is obtained. $\mathcal{O}$ might also denote a two-body observable, for instance a density like $n_i = c_i^\dagger c_i$. $I_{\mathcal{O}}(\omega)$ would yield the imaginary part of the diagonal elements of (3.36) in the frequency domain.

The m^{th} moment of $I_{\mathcal{O}}(\omega)$ is defined by

$$\mu_m = \int_0^{\infty} d\omega\ \omega^m\, I_{\mathcal{O}}(\omega)\ . \tag{3.38}$$

$I_{\mathcal{O}}(\omega)$ is uniquely determined if all moments are given – at least in theory. The correlation function (3.37) can be rewritten in its spectral representation yielding

$$I_{\mathcal{O}}(\omega) = \sum_n \left| \left\langle \psi_n^X \middle| \mathcal{O} \middle| \psi_0 \right\rangle \right|^2 \delta\left(\omega - \left(E_n^X - E_0\right)\right) \ , \tag{3.39}$$

where $\left(E_n^X, \left|\psi_n^X\right\rangle\right)$ denotes an eigenpair in the subspace of Hamiltonian H where $\mathcal{O}\left|\psi_0\right\rangle$ belongs to. For instance if $\mathcal{O} = c_i$, X denotes the subspace with $N-1$ particles. With equation (3.39) and the definition of the moments (3.38) we directly obtain the moments of $I_{\mathcal{O}}(\omega)$

$$\mu_m^{\mathcal{O}} = \sum_n \left(E_n^X - E_0\right)^m \left| \left\langle \psi_n^X \middle| \mathcal{O} \middle| \psi_0 \right\rangle \right|^2 \ . \tag{3.40}$$

Galitskii-Migdal theorem The first moments of the spectral function $I_{c_{\mathbf{k}}}(\omega) = A^{\mathrm{PE}}_{\mathbf{kk}}$ are connected to the ground-state energy. This relation is known as the Galitskii-Migdal theorem. In this paragraph we will explicitly write out the spin-dependence, thus we define $I_{c_{\mathbf{k}}}$ as

$$I_{c_{\mathbf{k}}}(\omega) = \frac{1}{2}\left(I_{c_{\mathbf{k}\uparrow}}(\omega) + I_{c_{\mathbf{k}\downarrow}}(\omega)\right) \ .$$

We define the spin dependent first moment as

$$\mu_1^{\mathbf{k},\sigma} = \int d\omega\ \omega\, I_{c_{\mathbf{k},\sigma}}(\omega) = \sum_n \left(E_n^{N-1} - E_0\right) \left| \left\langle \psi_n^{N-1} \middle| c_{\mathbf{k},\sigma} \middle| \psi_0 \right\rangle \right|^2 \ , \tag{3.41}$$

which can, summing over the complete set of eigenstates, be formally written as

$$\mu_1^{\mathbf{k},\sigma} = \left\langle \psi_0 \middle| c_{\mathbf{k},\sigma}^\dagger \left[H, c_{\mathbf{k},\sigma}\right] \middle| \psi_0 \right\rangle \ . \tag{3.42}$$

We write the Hamiltonian H in k-space (2.34) which is given by

$$H = \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} + \frac{U}{N_e} \sum_{\mathbf{k},\mathbf{k}',\mathbf{q}} c_{\mathbf{k}\uparrow}^\dagger c_{\mathbf{k}-\mathbf{q},\uparrow} c_{\mathbf{k}'\downarrow}^\dagger c_{\mathbf{k}'+\mathbf{q},\downarrow} \ . \tag{3.43}$$

Evaluating $c_{\mathbf{k},\sigma}^\dagger \left[T, c_{\mathbf{k},\sigma}\right]$, where T denotes the kinetic part of the Hamiltonian, and using,

$$\left[A, BC\right] = \left\{A, B\right\} C - B \left\{A, C\right\} \ , \tag{3.44}$$

leads to

$$c_{\mathbf{k},\sigma}^\dagger \left[T, c_{\mathbf{k},\sigma}\right] = \sum_{\sigma\mathbf{k}'} \varepsilon_{\mathbf{k}} c_{\mathbf{k},\sigma}^\dagger \left[c_{\mathbf{k}'\sigma'}^\dagger c_{\mathbf{k}'\sigma'}, c_{\mathbf{k}\sigma}\right] = -\varepsilon_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} \ . \tag{3.45}$$

Similarly, $c_{\mathbf{k},\sigma}^\dagger \left[V, c_{\mathbf{k},\sigma}\right]$, where V denotes the Coulomb part of the Hamiltonian, yields,

$$c_{\mathbf{k},\sigma}^\dagger \left[V, c_{\mathbf{k},\sigma}\right] = -\frac{U}{N_e} \sum_{\mathbf{q},\mathbf{k}'} \begin{cases} c_{\mathbf{k}\uparrow}^\dagger c_{\mathbf{k}-\mathbf{q},\uparrow} c_{\mathbf{k}'\downarrow}^\dagger c_{\mathbf{k}'+\mathbf{q},\downarrow} & \sigma = \uparrow \\ c_{\mathbf{k}'\uparrow}^\dagger c_{\mathbf{k}'-\mathbf{q},\uparrow} c_{\mathbf{k}\downarrow}^\dagger c_{\mathbf{k}+\mathbf{q},\downarrow} & \sigma = \downarrow \end{cases} \ . \tag{3.46}$$

Summing $\mu_1^{\mathbf{k}}$ over all $\mathbf{k}$ we obtain,

$$\sum_{\mathbf{k}} \mu_1^{\mathbf{k}} = -\left(\frac{\langle T \rangle_0}{2} + \langle V \rangle_0 \right) , \tag{3.47}$$

where $\langle T \rangle_0$, $\langle V \rangle_0$ denote the expectation values of the kinetic, interaction energy respectively. Equation (3.47) provides the connection between the first moments of $I_{c_{\mathbf{k}}}$ and the ground-state energy. This is the Galitskii-Migdal theorem [27]. It is quite a remarkable result since with the kinetic energy expectation value and these first moments, which all can be derived from the single-particle Green's function, the expectation value of the two-particle observable, the Hamiltonian itself, is accessible.

3.4.2 Computing the spectral function: the third pass

We are interested in calculating dynamical correlation functions like (3.37). At first glance this seems to be very hard. Since there is a Hamiltonian H in the denominator its whole excitation spectrum contributes to the correlation function. Nevertheless the Lanczos method provides easy access to correlation functions because it rapidly finds the states which make the largest contribution. This is yet another great advantage of the method.

In order to evaluate equation (3.37) numerically it is practical to represent the Hamiltonian in the basis obtained by the Lanczos procedure. Instead of starting the Lanczos method with a random vector as in chapter 3.3, we now start with

$$|\phi_0\rangle = \frac{\mathcal{O}\,|\psi_0\rangle}{\sqrt{\langle\psi_0|\mathcal{O}^\dagger\mathcal{O}|\psi_0\rangle}} . \tag{3.48}$$

Thus the quantity (3.37) we are interested in is proportional to

$$\left\langle \phi_0 \middle| \frac{1}{z-H} \middle| \phi_0 \right\rangle , \tag{3.49}$$

where $z = \omega + E_0 + i\eta$. Following Fulde [17], we consider the identity

$$\sum_m (z-H)_{lm}(z-H)^{-1}_{mn} = \delta_{ln} \tag{3.50}$$

in the Lanczos basis given by equation (3.10).

$(z-H)$ thus reads,

$$(z-H) = \begin{pmatrix} z-a_0 & -b_1 & 0 & 0 & 0 & \dots \\ -b_1 & z-a_1 & -b_2 & 0 & 0 & \dots \\ 0 & -b_2 & z-a_2 & -b_3 & 0 & \dots \\ 0 & 0 & -b_3 & z-a_3 & -b_4 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{pmatrix} \tag{3.51}$$

and the term (3.49), we are interested in, is obviously equal to $(z-H)^{-1}_{00} =: x_0$.

Hence we can regard equation (3.50) with $n = 0$ as inhomogeneous system of linear equations with an unknown vector $x_m := (z-H)^{-1}_{m0}$,

$$\sum_m (z-H)_{lm} x_m = \delta_{l0} \; .$$

To solve this system for x_0 we use Cramer's rule, i.e.

$$x_0 = \frac{\det S_0}{\det (z-H)} \; ,$$

where S_0 is given by

$$S_0 = \begin{pmatrix} 1 & -b_1 & 0 & 0 & 0 & \dots \\ 0 & z-a_1 & -b_2 & 0 & 0 & \dots \\ 0 & -b_2 & z-a_2 & -b_3 & 0 & \dots \\ 0 & 0 & -b_3 & z-a_3 & -b_4 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{pmatrix} \; .$$

Using Laplace's formula for determinants one can easily show,

$$x_0 = \frac{1}{z - a_0 - b_1^2 \frac{\det D_2}{\det D_1}} \; ,$$

where D_i is defined similar to equation (3.51) but with the first i columns and rows removed. $\frac{\det D_2}{\det D_1}$ can be expanded likewise, leading to:

$$\frac{\det D_{i+1}}{\det D_i} = \frac{1}{z - a_i - b_{i+1}^2 \frac{\det D_{i+2}}{\det D_{i+1}}} \; ,$$

and thus

$$I_{\mathcal{O}}(\omega) = -\frac{1}{\pi} \Im \left\{ \frac{\langle \psi_0 | \mathcal{O}^\dagger \mathcal{O} | \psi_0 \rangle}{z - a_0 - \cfrac{b_1^2}{z - a_1 - \cfrac{b_2^2}{z - a_2 - \dots}}} \right\} \; . \tag{3.52}$$

To numerically evaluate this continued fraction we use the modified Lentz method (refer to chapter C or [28]).

Alternatively, we can compute the spectral functions as follows. Consider the spectral function (3.37) in its spectral representation

$$I_{\mathcal{O}}(\omega) = \sum_n |c_n|^2 \delta(\omega - (E_n^X - E_0)) \; , \tag{3.53}$$

where c_n is given by $c_n = \langle \psi_n^X | \mathcal{O} | \psi_0 \rangle$, E_n^X is the eigenvalue to the eigenvector $|\psi_n^X\rangle$ of H. The eigenvalues E_n^X are directly obtained from the Lanczos method as the

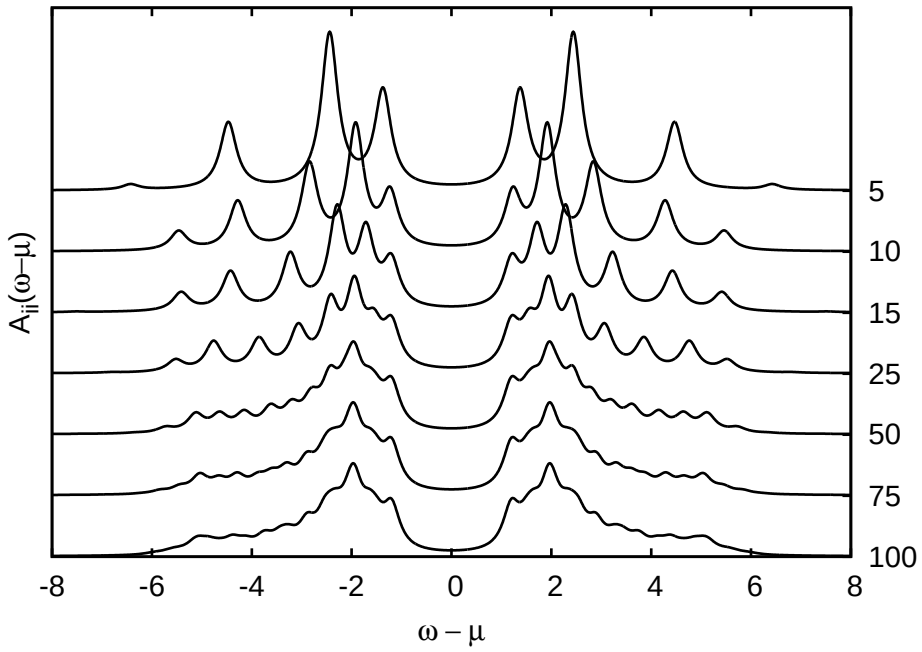


Figure 3.7: Convergence of $A_{ii}(\omega) = I_{a_i}(\omega) + I_{a_i^\dagger}(\omega)$ for different number of Lanczos iterations (5, 10, 15, 25, 50, 75, 100 steps) of a 14 sites system with half-filling and $U = 5t$. The number of iterations is equal to the maximal number of peaks. The peaks which have the largest contribution are there first.

eigenvalues of the tridiagonal matrix. The spectral weights can be easily calculated as well. The eigenvectors of the full Hamiltonian in the Krylov subspace are given by $\left|\psi_n^X\right\rangle = \sum_j U_{nj} \left|\phi_j\right\rangle$ where U_{ij} is the matrix, whose rows are the eigenvectors of the tridiagonal matrix, and $|\phi_i\rangle$ is the Lanczos basis with $|\phi_0\rangle$ being

$$|\phi_0\rangle = \frac{\mathcal{O}\,|\psi_0\rangle}{\sqrt{\left\langle\psi_0\middle|\mathcal{O}^\dagger\mathcal{O}\middle|\psi_0\right\rangle}}\,.$$

Now it is easy to see that

$$|c_n|^2 = \left|\left\langle\psi_n^X\middle|\mathcal{O}\middle|\psi_0\right\rangle\right|^2 = \left\langle\psi_0\middle|\mathcal{O}^\dagger\mathcal{O}\middle|\psi_0\right\rangle \left|\sum_j T_{nj}^* \left\langle\phi_j|\phi_0\right\rangle\right|^2 = \left\langle\psi_0\middle|\mathcal{O}^\dagger\mathcal{O}\middle|\psi_0\right\rangle |T_{n0}|^2\,. \tag{3.54}$$

Thus in order to calculate the dynamical response of a system we start the Lanczos method with a random vector to obtain the ground state (first two passes). Then we compute the start vector for the third pass with equation (3.48) and perform the third pass, storing all eigenvalues of the tridiagonal matrix E_n^X and all first elements of the eigenvectors $U_{n0} = c_n/\sqrt{\left\langle\psi_0\middle|\mathcal{O}^\dagger\mathcal{O}\middle|\psi_0\right\rangle}$. These are all ingredients needed to

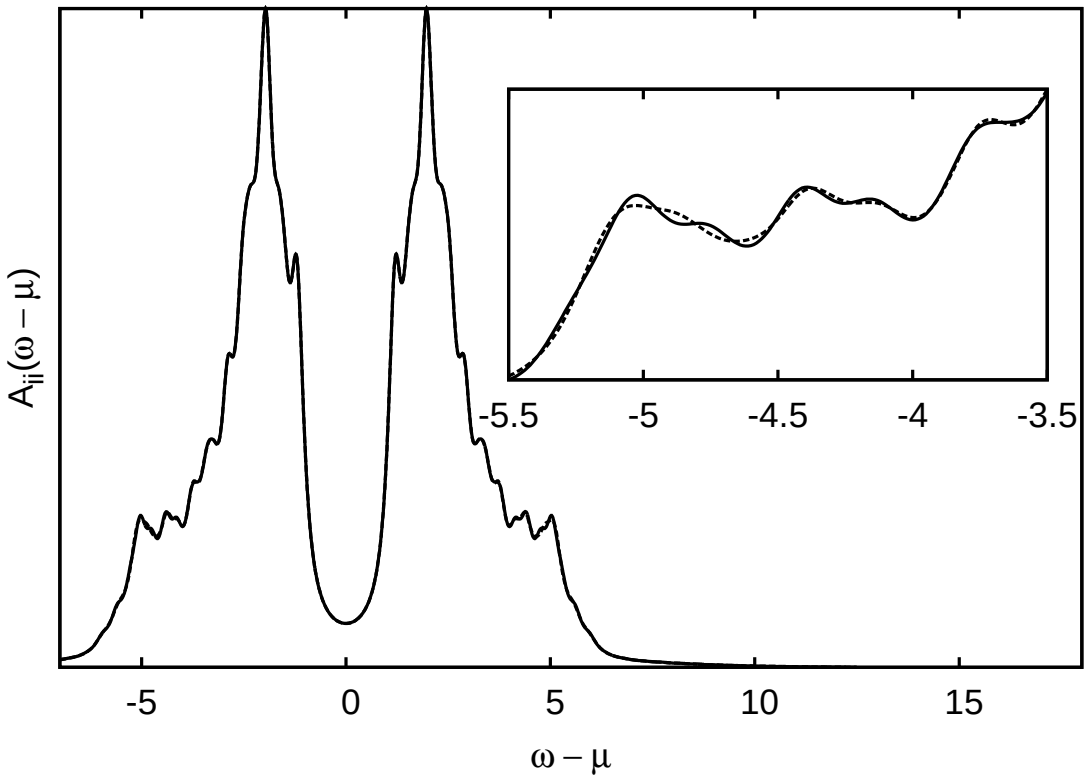


Figure 3.8: Comparison of spectral function with 200 iterations (dashed) and 100 iterations (solid) of a half-filled Hubbard model (14 sites) with $U = 5t$. The inset shows one of the two high energy parts where the function after 100 iterations is not yet converged.

compute (3.53).

Having these ingredients also enables us to calculate arbitrary moments of $I_{\mathcal{O}}(\omega)$, i.e. (3.38), using (3.53)

$$\mu_m^{\mathcal{O}} = \sum_n \left(E_n^X - E_0\right)^m |c_n|^2 \,. \tag{3.55}$$

How do we judge whether the correlation functions are converged? An easy method is to check by eye. Figure 3.7 shows how a spectral function A_{ii} changes with different numbers of third pass iterations for a half-filled one dimensional Hubbard model. We see that after only 5 iteration the main features are already visible, though the total spectral function is still coarse. This is because the number of Lanczos iterations determines the maximal number of peaks. We however observe, that the Lanczos method quickly finds the energies in the excitation spectrum which have the highest weight; or put another way, the energies whose eigenvalues have the largest overlap

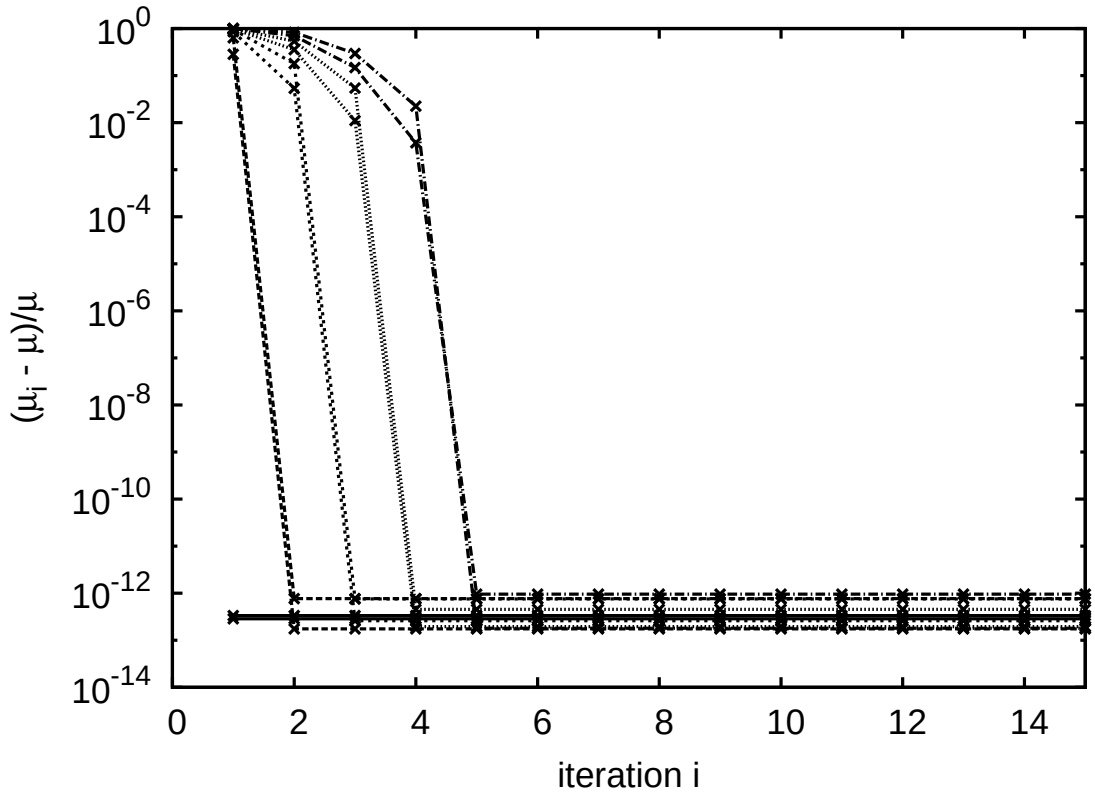


Figure 3.9: Logarithmic plot of the first 10 deviations of spectral moments against the number of iterations for a photoemission spectral function of a 10 sites Hubbard chain with 3 electrons of either spin, $U = 4t$, $t = 1$. In each iteration two moments converge which is yet another hint at the rapid convergence of the Lanczos method.

with the starting vector (see equation (3.48)). Figure 3.8 shows spectral functions for 100 and 200 iterations in the third Lanczos pass. We see that only the high energy parts are not converged yet. It shows in practice that normally 200 iterations suffice to have the spectral functions converged.

Note another heuristic hint at the fast convergence. Typically continuous fractions converge faster than power series and the Lanczos method's third pass provides us with all information needed to evaluate a continuous fraction which converges to the correlation function. Moreover we can also check how swiftly the moments converge. Momenta unambiguously determine the corresponding function and when the moments converge rapidly, so does the function. We see from figure 3.9, that the moments indeed converge rapidly. In each iteration two additional converge.

Until now we are only able to compute diagonal elements of the correlation functions. This is because we choose our start vector for the third pass as $\mathcal{O}_i |\psi_0\rangle$ and

from this the total spectral function

$$I_{\mathcal{O}_i}(\omega) = -\frac{1}{\pi}\Im \left\langle \psi_0 \middle| \mathcal{O}_i^\dagger \frac{1}{\omega - (H-\mu) + i\eta} \mathcal{O}_i \middle| \psi_0 \right\rangle \tag{3.56}$$

is obtained with the Lanczos method. However we often need to calculate off-diagonal elements, as well. For example the correlation function $I_{\tilde{\mathcal{O}}}^{\alpha\beta}$ of operator $\tilde{\mathcal{O}}_{\alpha\beta}$ for $\alpha \neq \beta$. We therefore start the third Lanczos pass with the initial vector (3.48), where $\mathcal{O}$ is given by

$$\mathcal{O}_\pm^{\alpha\beta} = \frac{1}{\sqrt{2}} \left(\tilde{\mathcal{O}}_\alpha \pm \tilde{\mathcal{O}}_\beta \right) . \tag{3.57}$$

According to equation (3.53) and (3.54) this leads to

$$I_{\mathcal{O}_\pm}^{\alpha\beta} = \sum_n \left\langle \psi_0 \middle| \left(\mathcal{O}_\pm^{\alpha\beta} \right)^\dagger \middle| \psi_n^X \right\rangle \left\langle \psi_n^X \middle| \left(\mathcal{O}_\pm^{\alpha\beta} \right) \middle| \psi_0 \right\rangle \delta(\omega - (E_n^X - E_0)) , \tag{3.58}$$

which can be expanded into

$$I_{\mathcal{O}_\pm}^{\alpha\beta} = \frac{1}{2} \sum_n \left\{ |c_n^\alpha|^2 + |c_n^\beta|^2 \pm 2\Re \left((c_n^\alpha)^* c_n^\beta \right) \right\} \delta(\omega - (E_n^X - E_0)) , \tag{3.59}$$

where $c_n^\alpha = \left\langle \psi_n^X \middle| \tilde{\mathcal{O}}_\alpha \middle| \psi_0 \right\rangle$. This result can be rewritten as

$$I_{\mathcal{O}_\pm}^{\alpha\beta} = \frac{1}{2} \left\{ I_{\tilde{\mathcal{O}}}^{\alpha\alpha} + I_{\tilde{\mathcal{O}}}^{\beta\beta} \right\} \pm I_{\tilde{\mathcal{O}}}^{\alpha\beta} . \tag{3.60}$$

Thus, we perform two third passes with each of the two starting vectors

$$|\phi_0, \pm\rangle = \frac{\mathcal{O}_\pm^{\alpha\beta} |\psi_0\rangle}{\sqrt{\left\langle \psi_0 \middle| \left(\mathcal{O}_\pm^{\alpha\beta} \right)^\dagger \mathcal{O}_\pm^{\alpha\beta} \middle| \psi_0 \right\rangle}} .$$

Then we can calculate the off-diagonal elements by

$$I_{\tilde{\mathcal{O}}}^{\alpha\beta} = \frac{1}{2} \left(I_{\mathcal{O}_+}^{\alpha\beta} - I_{\mathcal{O}_-}^{\alpha\beta} \right) . \tag{3.61}$$

If we were only interested in a single off-diagonal element this is the way to proceed. We only need two third Lanczos passes. But if we wanted to calculate the full matrix $\left(I_{\tilde{\mathcal{O}}}^{\alpha\beta} \right)$, we can choose a more efficient method. From equation (3.60) we see, that having the diagonal elements we can calculate all off-diagonal ones with a single additional third Lanczos pass, i.e.

$$I_{\tilde{\mathcal{O}}}^{\alpha\beta} = I_{\mathcal{O}_+}^{\alpha\beta} - \frac{1}{2} \left(I_{\mathcal{O}_\alpha} + I_{\mathcal{O}_\beta} \right) . \tag{3.62}$$

4 Implementation and Checking

In this chapter we discuss the actual implementation of our Lanczos code. The basic problem is the exponential growth of the Hilbert space. For a half-filled system of 14 sites we already need wave vectors of about 89 MB. This certainly is doable on ordinary personal computers. We often need larger systems, however, for instance in order to scale out finite-size effects or to gain a higher resolution for angular-resolved spectral functions. Another example are physical systems away from half-filling. For the organic metal TTF-TCNQ we will deal with in chapter 7 we need to treat systems with a filling of 0.6. Since we need an integer amount of electrons, this efficiently reduces the length of the chains to 10 or 20 sites. 20 sites with 6 electrons of either spin lead to wave vectors of 22 GB for the calculation of the spectral function. This cannot be handled on an ordinary personal computer anymore. Instead we have to resort to supercomputers.

The main topic of this chapter is, how to efficiently implement a Lanczos code on modern supercomputer architectures. At first an implementation on shared-memory systems is discussed. However, shared-memory systems are restricted to a relatively small number of processors. To treat large systems we write an implementation for distributed-memory systems and we show that we can take advantage of the latest massively parallel architectures like BlueGene/L supercomputers. The BlueGene/L system in Jülich, called JUBL, is at the time of writing on rank 8 of the top500 supercomputing list.

The last part of this chapter is devoted to tests that check whether the implementation is correct.

4.1 General considerations

Before one starts to optimize or parallelize code, it is important to know where the critical spots are, that is, in which parts of the code most of the time is spent? Knowing these spots we can judge, whether parallelization is a feasible option and focus our efforts. Parallelization makes sense if the ratio of the parallelizeable to the inherently sequential code fraction is large enough in terms of run-time. Amdahl's law, which is introduced in chapter A, states that the maximum achievable speed up is bounded by the inherently sequential code fraction.

What do we expect for the Lanczos method? Most of the time will certainly be spent in functions working on the huge many-body vectors. To verify this we use a profiler, a tool that measures the code's behavior at run-time. It gathers information like how often functions are called and how much time is spent in them.

4.1.1 Example profile of a serial Lanczos run

A profile of our code for a 12 sites Hubbard chain with half-filling and $U = 10t$ confirms our assumption. The call to `LancMult::pmassign()` actually performs the matrix-vector multiplication (line 10 in the pseudo code 3.1), which takes about 81% of the total execution time. The scalar products and norms of the Lanczos vectors are contained in the second item, the `lanczos()` function. All in all the code working on the Lanczos vectors takes more than 95% of the program's total execution time. Figure (4.1) illustrates this fact.

```
Each sample counts as 0.01 seconds.
  %   cumulative   self              self    total
 time   seconds   seconds    calls  s/call  s/call  name
 81.04    990.20   990.20      452    2.19    2.19  LancMult::pmassign(...)
 16.51   1191.93   201.73        4   50.43  297.55  void lanczos(...)
  1.06   1204.89    12.96 10267512    0.00    0.00  FBasis::getIndex(...)
  0.44   1210.24     5.35        1    5.35   11.82  void c\_k(...)
  0.40   1215.15     4.91        1    4.91   11.38  void cd\_k(...)
  0.18   1217.40     2.25        1    2.25    2.25  void RandomVec(...)
  ...
```

So these are obviously the two critical spots in our code on which optimization efforts should be focused. Moreover we will investigate whether they can be parallelized. For larger physical systems we expect the profile to become even sharper. Let $\mathcal{H}_\sigma$ denote the σ-electron Hilbert space with dimension $\dim(\mathcal{H}_\sigma)$. The setup of the corresponding spin-conserving hopping Hamiltonians scales linearly with this dimension, whereas the operations on the Lanczos vectors scale linearly with the dimension of the full Hilbert space $\dim(\mathcal{H}) = \dim(\mathcal{H}_\uparrow) \cdot \dim(\mathcal{H}_\downarrow)$. Thus, for larger systems the setup becomes more and more negligible since it scales with $\sqrt{\dim(\mathcal{H})}$ whereas the numerical part scales linearly.

4.1.2 Memory access

Modern CPUs are becoming increasingly fast, especially compared to the time spent on main memory access. To remedy this problem, caches were introduced. When a processor needs a memory element, it loads this element and some neighboring elements from main memory into a fast but small separate memory called cache. If the next element that the processor needs is sufficiently close to the first element in main memory, it is already cached and can thus be accessed rapidly. A programmer should strive to build his data structures in such a way that memory locality is ensured to exploit this effect (cf. chapter D1 by Goedecker in [12]).

4.1.3 Matrix-vector multiplication and cache-effects

It turns out that the matrix-vector multiplication, namely applying the Hamiltonian to the many-body wave function, will be the crucial problem for the efficient

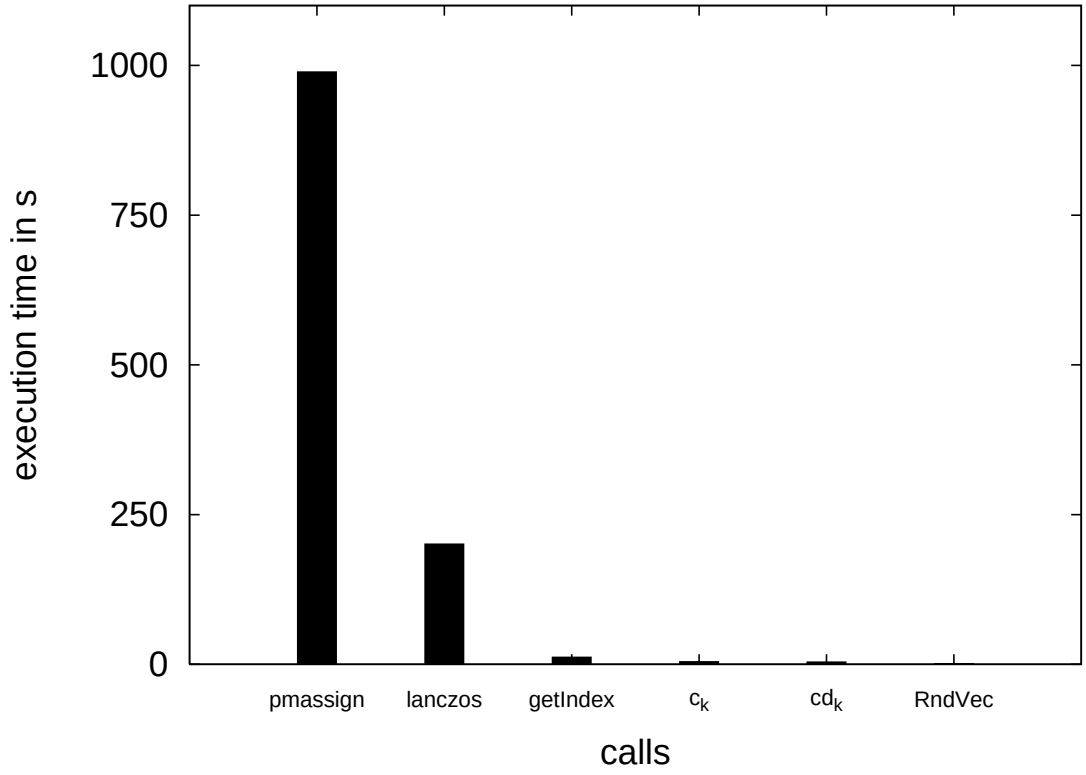


Figure 4.1: Profile of Lanczos code (serial run) for a 12 sites half-filled system. Such a sharp profile is typical of linear algebra programs. Optimization and parallelization efforts should focus on `LancMult::pmassign` and the `lanczos` function.

implementation, especially in case of distributed-memory systems.

The Hamiltonian H is a composition of two single spin-type Hamiltonians $T_\uparrow$ and $T_\downarrow$ and the diagonal Coulomb part V. In the real-space configuration basis the Hamiltonians are sparse. Off-diagonal elements denote the hopping of the electrons. The full Hamiltonian H is given by

$$H = (T_\downarrow \otimes 1_{\mathrm{dim}_\uparrow}) + (1_{\mathrm{dim}_\downarrow} \otimes T_\uparrow) + V \ .$$

The first term denotes the hopping of the down electrons in which we see that the up-electron configuration is fixed, while the down configuration changes; vice versa for the second term. The third term contains the Hubbard interaction term, which is diagonal.

From the basis construction in chapter 3.1 we know the order of the basis elements. For a fixed down electron all up-electron configurations follow contiguous in memory. Hence, hopping is local in memory for $1_{\mathrm{dim}_\downarrow} \otimes T_\uparrow$, but highly non-local for $T_\downarrow \otimes 1_{\mathrm{dim}_\uparrow}$. Therefore we expect cache effects to appear. In order to measure those effects we

time both hopping parts separately. Table 4.1 shows the results. As one can see

Table 4.1: Cache effects on hopping performance for half-filled Hubbard chain with different number of sites in PBC on the JUMP supercomputer (complex code). All times given in seconds.

sites	avg. up hop.	avg. down hop.	ratio
6	$0.090 \cdot 10^{-5}$	$0.092 \cdot 10^{-5}$	$\approx 2\%$
8	$1.21 \cdot 10^{-3}$	$1.24 \cdot 10^{-3}$	$\approx 2\%$
10	$1.68 \cdot 10^{-2}$	$1.79 \cdot 10^{-2}$	$\approx 6\%$
12	0.243	0.284	$\approx 14\%$
14	3.73	4.91	$\approx 24\%$
16	6.03	8.66	$\approx 30\%$

the up electron hopping takes less time compared to the down hopping. This is due to memory locality. For the down-electron hopping considerably more cache misses occur, and thus the data elements have to be fetched from main memory.

4.1.4 Is a parallelization possible?

For our code parallelization becomes feasible if we can parallelize the matrix-vector multiplication as well as the norms and scalar products. Therefore we have to check if the tasks can be decomposed into smaller tasks which can be carried out independently. This is clearly possible for scalar products which also yield the norms. We decompose the two vectors we want to multiply into n equal shares. A share can be regarded as a smaller vector on which the scalar product can be performed on these vectors, yielding n results. The sum of these intermediate results obviously gives the final result.

What about the matrix-vector multiplication? In order to calculate the i^{th} element of the result vector we need to calculate

$$q_i = \sum_j H_{ij} b_j \ .$$

We decompose the resulting vector q. Since the vector elements of b and the matrix elements H only need to be read and can be shared without locking mechanisms the iterations are independent. Each thread can calculate its share of the vector elements q_i. However, access to the vector elements b_i still is non-local, leading to the already discussed performance degradation for the down hopping. Moreover in case of distributed memory computing it will give rise to communication between the threads which has to be handled efficiently.

Parallelization indeed should be possible for the most time-consuming parts of the program.

4.2 Parallelization

4.2.1 Shared-memory parallelization with OpenMP

OpenMP is a standard for shared-memory multiprocessing, which means that each thread/processor has access to the full memory. It comprises a set of OpenMP directives (`#pragma`s in C/C++) and a run-time library which can be adjusted by environment variables at execution time. Often OpenMP parallelization requires only small changes to the source code. These changes mainly are `#pragma` directives, which are either interpreted by an OpenMP-capable compiler to generate parallel code or ignored by an ordinary sequential compiler. This obviously leads to a unified code base for parallel as well as sequential usage, which as a result bears the advantage of good maintainability.

The programming paradigm behind OpenMP is the so-called fork-join model. It means that only one master thread works its way through the serial code. In our case this is the setup of the basis and the Hamiltonian. When it reaches a section which can be parallelized, like the scalar products, it forks into multiple threads and the work is shared equally among them. When the parallel part is processed, the threads are joined again and the master continues with the serial code (cf. chapter D2 by Mohr in [12]).

To get a feeling for OpenMP we discuss the parallelization of a scalar product. In C++ the parallel scalar product may look like:

```
#pragma omp parallel for reduction(+:result)
for (long i = 0; i < a.size(); ++i)
    result += a[i]* b[i];
```

Here a and b denote the vectors which are to be multiplied. The `#pragma` statement in the first line actually is the OpenMP directive. It tells the compiler to parallelize the code in the way described above. The master forks into p threads. Each thread works on a share of $\dim(\mathcal{H})/p$ elements of the vector (in shared memory)[1] and performs the local scalar product, storing the result in a private variable `result`. Then all private `result` variables, yielding the intermediate results, are reduced by the OpenMP's `reduction(+:result)` directive. Hereafter the master holds the final result in its variable `result`.

The same can be done for the matrix-vector multiplication. The resulting vector elements can be calculated independently as seen above. We thus only need to parallelize the loop over the q_i. The matrix elements and the vector $\mathbf{b}$ are declared shared and thus no locking overhead is necessary. It all boils down to a single OpenMP directive. This so-called loop-level parallelization is very easy to employ. A shortened code fragment of our matrix-vector multiplication is shown in the following listing.

```
#pragma omp parallel for private(iup,i,nhop,hop)
```

[1] This is the simplest way of distributing the data, even though OpenMP implements more.

```
for(idn = 0; idn < dimdn; idn++)
    for(iup = 0; iup < dimup; iup++)
    {
        i=idn*dimup+iup;

        result[i]+=vec[i]*diag[i];

//      up electron hopping
        nhop = hup->numhop[iup];
        for (hop = 0; hop < nhop ; hop++)
            result[i]+=vec[uphopto(iup,hop)+idn*dimup]*upt(iup,hop);

//      dn electron hopping
        nhop = hdn->numhop[idn];
        for (hop = 0; hop < nhop; hop++)
            result[i]+=vec[iup+dnhopto(idn,hop)*dimup] * dnt(idn,hop);
    }
```

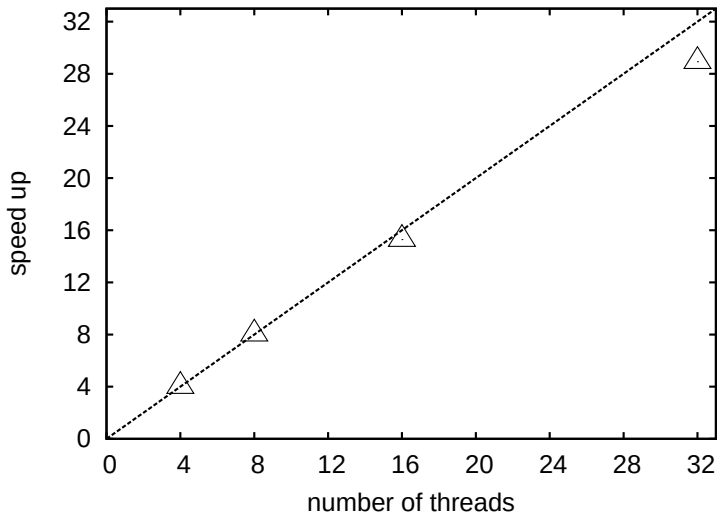


Figure 4.2: OpenMP speed up of Lanczos iterations for a half-filled 16 sites system with $U = 0$. The speed up is almost linear, thus using p processors reduces the execution time by a factor of about p.

Plot 4.2 shows the speed up of the OpenMP-parallelized Lanczos iterations on JUMP (see chapter B). We see that it benefits immensely from parallelization. The speed up is close to the ideal linear one. This means that there is only a negligible parallelization overhead and hardly any serial code (Amdahl's law). For our code shared-memory parallelization on JUMP scales very well. We thus have a means to efficiently calculate relatively large system like for example 20 sites systems with 0.6

filling and even the corresponding (photoemission) spectral functions. For complex boundary conditions and therefore complex wave vectors this amounts to a memory requirement of about 22 GB for each Lanczos vector.

Unfortunately shared-memory systems are restricted to relatively small number of processors. This is because all processors not only have to share the memory but also the bandwidth to access the memory, which quickly becomes a bottleneck. Moreover cache coherence has to be ensured. When a processor accesses a memory element it is copied to the processor's cache. If another processor does the same and changes this element, the change has to be synchronized with the other cache. The logic for maintaining this cache coherence is only feasible with relatively small processor counts.

A single symmetric multiprocessor (SMP) node of Jülich's IBM Regatta supercomputer JUMP comprises 32 processors and about 128 GB of RAM. This is a quite low processor count and for calculations of large systems we would like to use more processors and might need more RAM. In the example above of a 20 sites Hubbard chain calculating the angular-resolved (photoemission) spectral function for 11 k-values with complex boundary conditions takes about 20 h.

To gain a higher resolution in k-space we can employ the yet-to-be-discussed cluster perturbation theory. This would require the calculation of the Green's function matrices with 210 angular-resolved spectral functions, taking about 9 days with the real code. Moreover we would like to calculate the inverse photoemission spectral function as well, requiring $3 \cdot 44$ GB = 132 GB of RAM. Therefore we have to resort to another way of parallelization, the so-called distributed-memory parallelization. As the name suggests memory on those systems is not shared and thus the memory bandwidth in the previous sense is no issue anymore. Instead we have to handle communication due to remote memory access explicitly. We can probably even use thousands of processors which – if the parallelization works well – will significantly cut the execution time.

4.2.2 Distributed-memory parallelization on BlueGene

On distributed-memory systems message passing is the prevailing paradigm for parallelization. The de-facto standard is the *Message Passing Interface* (MPI). With MPI we have to explicitly distribute the data and take care for the calculations performed on each thread. Thus, we need to port the code to MPI and maintain coherence with the sequential/OpenMP one. Keeping a unified code base is more demanding compared to OpenMP.

In our case the data structures to be distributed are the Lanczos vectors. Each of the p threads holds a contiguous share of $\dim(\mathcal{H})/p$ elements in its local memory. This ensures load balancing, since each thread has to perform the same number of operations on average. Due to the sparsity of the Hamiltonians, they can still be stored locally on each thread. Porting of most mathematical operations to MPI is straightforward to do. The scalar product on the Lanczos vectors for example is performed on each thread with its own share of the vector. This is why the `for`-loop

in the following code fragment iterates up to `localdim`.

```
for (long i = 0; i < localdim; ++i)
    result += a[i]* b[i];
MPI_Allreduce(MPI_IN_PLACE,&result,1,
        MPI_DOUBLE,MPI_SUM,MPI_COMM_WORLD);
```

Similar to the `reduce` directive in OpenMP the global reduce operation (`MPI_Allreduce`) gathers all local `result` variables, adds them up and distributes the final result back to each thread. Thus, scalar products and norms are quite easily ported to MPI.

Whereas in the case of OpenMP shared-memory parallelization the matrix-vector multiplication is done by a straightforward loop parallelization, it is not as easy in the case of distributed-memory systems. The diagonal part of the multiplication can still be performed easily and yields no problems. But as we have discussed above the up-electron hopping is local in memory. The down-electron hopping, however, shows a very non-local and non-regular memory access. In case of a shared-memory computer system or a serial computer, this translates into cache misses and therefore to a performance degradation. For distributed memory systems we have an additional level in the memory hierarchy. A vector element can be residing in local or in a remote memory. Thus we need means to get this element, resulting in interprocess communication (IPC). We start with a straightforward MPI implementation similar to OpenMP.

Direct MPI-2 one-sided communication

MPI-2 one-sided communication extends the communicational means of MPI by enabling remote memory access (RMA). This means that each thread can declare a window, i.e. an area of its memory, which can be directly accessed by all other threads for read and write operations. It is not necessary for a thread to know, who accesses its memory. This is the reason that makes it preferable to the ordinary MPI two-sided communication. Because the matrix elements are scattered it would be difficult to write corresponding send and receive calls. Thus, MPI-2 one-sided communication can be regarded as a direct generalization of the OpenMP approach with a further hierarchy level of memory.

The algorithm then works as follows: At first all threads create a window containing their share of the vector. Hereafter the matrix-vector operation is performed as long as all the needed vector elements are local. Usually large parts of the up-electron hopping part can be carried out locally. If a vector element is missing, the thread calculates the owner's rank and the offset in its window. With `MPI_Get` it fetches the element.

This operation is however quite expensive. For each `MPI_Get` the latency, the time to set up the connection, is paid. Each thread needs vector elements from many other threads, leading to a huge amount of `MPI_Get` operations. Figure 4.3 shows the corresponding speed up. It actually does not deserve this name, since we observe a speed down, i.e. the run-time becomes longer with increasing numbers of processors.

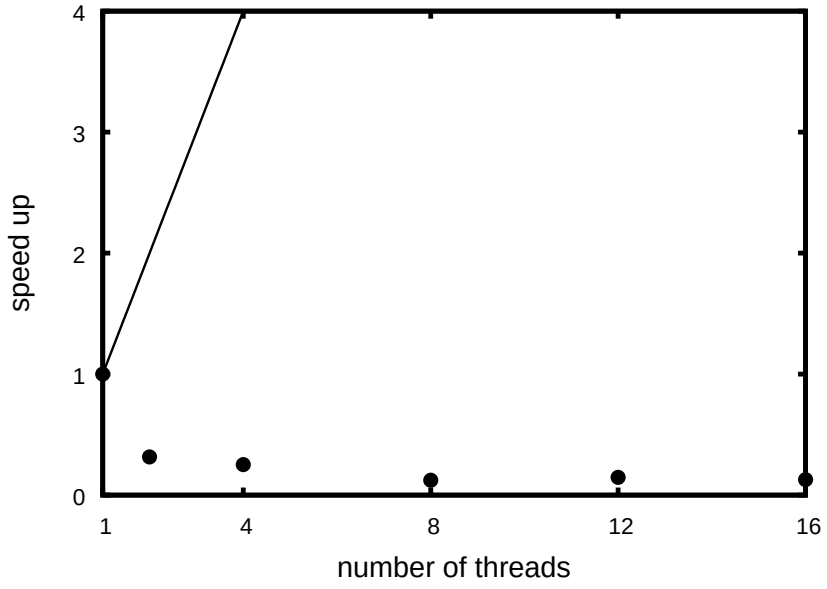


Figure 4.3: Speed up of RMA implementation (MPI-2) of the Lanczos code for a 14 sites one dimensional Hubbard chain. We actually observe a speed down. The line denotes the ideal speed up.

The implementation can be partly improved by local caching techniques, i.e. gathering sequential requests to the same remote window, and therefore cutting the number of get operations. However this does not show much promise. We clearly need better ideas.

Vector transposition operation

To obtain a more efficient code we need to find a way of restructuring the data such that all hopping can be performed locally. A simple yet practical observation provides us with the needed idea. We have seen in chapter 3.1 that we can unambiguously determine a state i by a tuple of labels $i \equiv (i_\uparrow, i_\downarrow)$, we know that the elements are ordered in memory in such a way that for a fixed down-electron configuration all up configurations are contiguous in memory, i.e. $i = i_\uparrow + \dim_\uparrow i_\downarrow$, where $\dim_\uparrow$ denotes the number of up-electron configurations. This is exactly how one addresses a matrix in a one-dimensional array. Hence we can consider a Lanczos vector as a matrix.

This situation is depicted in figure 4.4. To make the down hopping local in memory we exchange the indices $(i_\uparrow, i_\downarrow) \rightarrow (i_\downarrow, i_\uparrow)$. This corresponds just to transposing the Lanczos vector regarded as a matrix. If we store entire rows and columns, respectively, local in a thread's memory the following implementation of the matrix-vector product appears promising. Let q denote the result and b the vector to be multiplied in correspondence to listing 3.1.

- Perform the diagonal part and the up-electron hopping part locally.

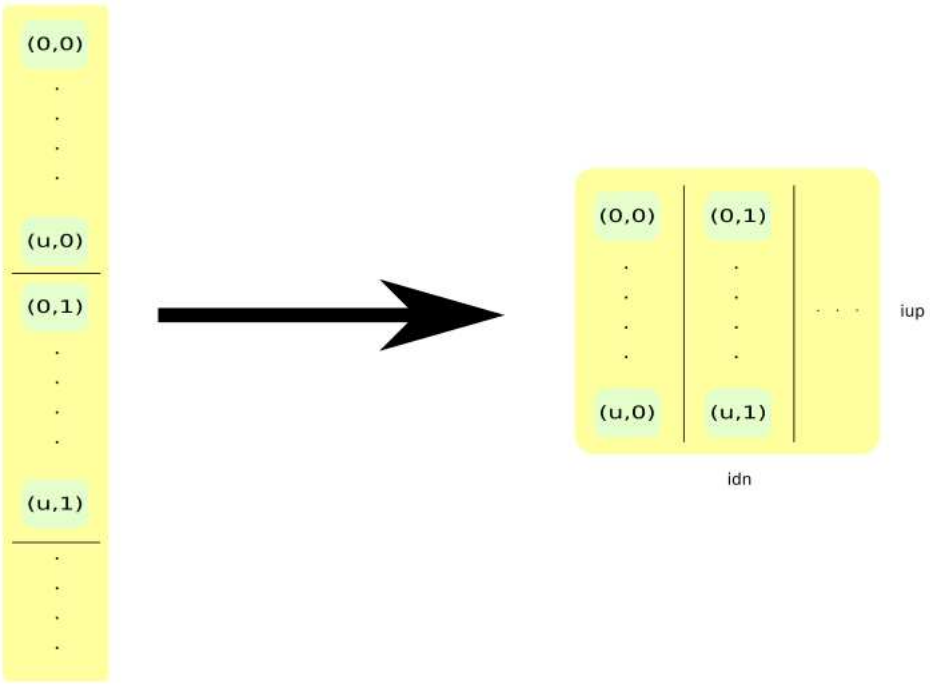


Figure 4.4: Considering a Lanczos vector as matrix. In the image $u = \dim_\uparrow$.

- Transpose vector b.
- Perform the down-electron part locally and store the result in a temporary vector.
- Transpose the temporary vector.
- Add transposed temporary vector to q.

It all boils down to an efficient matrix transposition routine for large dense distributed matrices. Those kind of transposition operations are, for instance, also needed for fast Fourier transforms (FFT). In such an operation, each thread has to communicate with all the others, apparently leading to an immense communicational effort. Thus a priori it is not clear whether this will work efficiently.

Matrix transpose routines

In the implementation we store entire columns of the matrix (vector) locally on each thread. The actual number of columns depends on the number of down-electron configurations and processors.

At first we will discuss the simplest case of $N_\uparrow = N_\downarrow$, or put differently, we regard Lanczos vectors which correspond to a square matrix, and $\dim_\uparrow = \dim_\downarrow$ is divisible by the number of threads.

MPI_Alltoall The preferred way of performing such all-to-all communication patterns is the use of MPI collective communication. For massively parallel systems like BlueGene these are said to be optimized and are recommended by IBM. For the matrix transposition operation we us the `MPI_Alltoall` call which is part of the MPI collective communication operations. Its C prototype looks like:

```
int MPI_Alltoall(void *sendbuf, int sendcount, MPI_Datatype sendtype,
                 void *recvbuf, int recvcount, MPI_Datatype recvtype,
                 MPI_Comm comm)
```

This call sends `sendcount` elements of `size(sendtype)` from position $j \cdot$ `(sendcount` $\cdot$ `sizeof(sendtype))` of thread i to thread j's position $i \cdot$ `(recvcount` $\cdot$ `sizeof(recvtype))` for all i,j. In principle this is exactly what we need. `MPI_Alltoall`, however, expects the data packages which will be sent to a given thread to be stored contiguously in memory. This does not apply to our case, since we would like to store the up-electron configurations sequentially in memory. Thus the matrix is stored column wise.

For `MPI_Alltoall` to work properly, we would have to bring the data elements in row-major order. This could be done by performing a local matrix transposition operation. The involved (sub-) matrices are, however, in general rectangular, leading to expensive local-copy and reordering operations. Fortunately this can be saved by calling `MPI_Alltoall` for each column separately. After calling `MPI_Alltoall` for each column with `sendcount` = `recvcount` = $\dim_\downarrow / \#p$ only a local (strided) transposition has to be performed to obtain the fully transposed matrix or Lanczos vector. These operations are depicted in scheme 4.5. The red arrows represent the `MPI_Alltoall` communication. They appear pairwise in parallel. The solid ones stem from the first call, the dashed ones from the second call. After the IPC the local transposition is performed in-place.

Figure 4.6 shows the speed up on JUMP and JUBL for a 16 sites half-filled Hubbard chain with periodic boundary conditions. The JUMP speed up of the code is represented by the filled triangles. Up until about 128 threads it features a very good speed up. 128 threads means we need as many processors and therefore 4 SMP nodes. The communication within a node is performed over the shared memory. But even between the nodes communication is quite efficient. JUBL has a significantly better speed up over the total range of processors shown in this plot (filled squares). Figure 4.8 shows that this speed up continues even for a much higher number of processors.

It is also interesting to compare the absolute timings. With our serial Lanczos version we compared the speed of a single processor of JUMP and JUBL. Execution on a single Power 4+ processor of JUMP is about two times faster than a single PowerPC 440 processor of JUBL. This is even better than the peak performance ratio from JUBL's perspective (about 40% of a single JUMP processor) suggests. The likely reason is that the fast Power4+ processors are limited more by the relatively slow memory access than the slower PowerPC. In [29] several JUBL timings of different codes are compared to corresponding JUMP ones. Ours is quite favourable compared to DMFT and laser-plasma calculations on JUBL, which achieved about

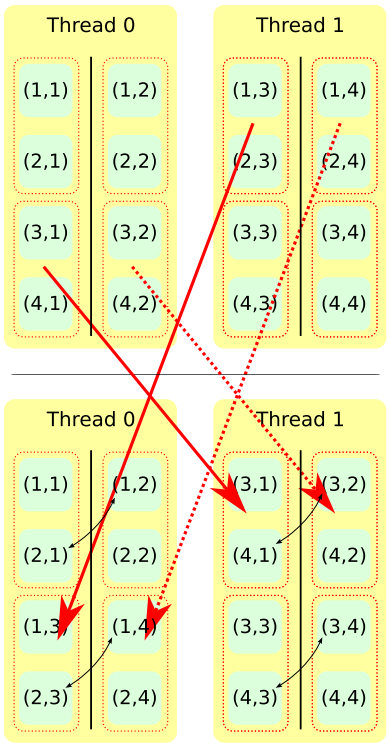


Figure 4.5: Scheme of the transpose operation for square matrices with the $\dim_{\uparrow}$ being divisible by the number of processors. The red arrows denote the `MPI_Alltoall` communication. The solid one represents the first, the dashed the second call. Finally the black ones denote the local inplace transpose.

33%, whereas structure optimizations with VASP about 22-43% compared to JUMP. QCD runs are about as fast as our code with 50% of Power4+ processor performance. For more information about the supercomputer systems refer to appendix B.

As promised by IBM, collective communication is indeed very efficient for BlueGene/L supercomputers like JUBL. A single node of a BlueGene/L system comprises two CPUs. In one mode, called coprocessor mode, one CPU handles the data processing while the other one takes care about communication and I/O. Thus a node essentially can be regarded as a single processor since it holds a single thread. The other mode is the virtual-node mode. Then each processor is indeed a single processor. It turns out that our implementation can make use of VN mode since `MPI_Alltoall` hardly needs an own processor for communication. This is because collective communication is blocking and all processors have to wait until the last pair exchanged their data packages. There is, however, an undocumented feature on BlueGene/L systems, the so called mixed mode. In this mode the two processors of each node are regarded as a small shared-memory system. We can then take

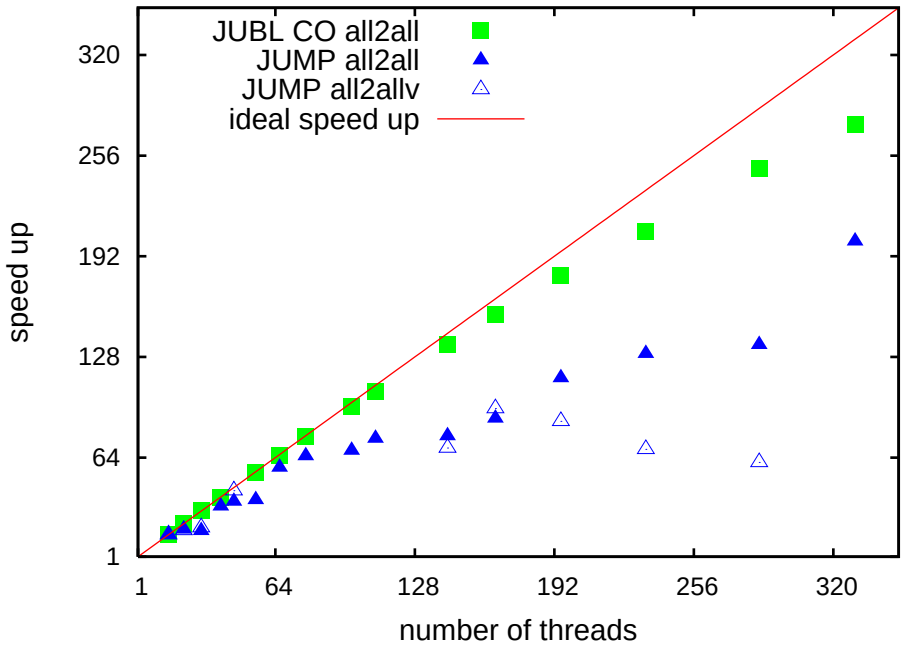


Figure 4.6: Speed up for a 16 site half-filled Hubbard chain. The green squares show the `MPI_Alltoall` implementation on JUBL in CO mode. On JUMP the filled triangles denote the `MPI_Alltoall`, whereas the other triangles represent `MPI_Alltoallv` one on JUMP. For runs with the smallest number of threads (15) the `MPI_Alltoallv` is about 1.17 times slower than the corresponding `MPI_Alltoall` implementation.

care for work distribution between them ourselves, namely performing the transpose operation on one processor, while the other processor works on the up-hopping part.

If this implementation had not worked this well, we could have tried another option for the matrix-vector transposition operation, namely the systolic matrix transposition described by Lippert [30]. The mechanism of a systeolic transpose is nicely illustrated in an interactive JAVA applet at
http://www.cs.rug.nl/ petkov/SPP/TRANSPOSE/transposition.html.

MPI_Alltoallv With the implementation discussed so far, the systems are restricted to $N_\uparrow = N_\downarrow$ and $\dim_\uparrow$ must be divisible by the number of threads. These constraints inhibit in particular the calculations of Green's functions. Therefore we need to generalize the matrix transposition. This can be achieved by the `MPI_Alltoallv` call. Its prototype looks like:

```
int MPI_Alltoallv(void *sendbuf, int *sendcnts, int *sdispls,
                  MPI_Datatype sendtype, void *recvbuf, int *recvcnts,
                  int *rdispls, MPI_Datatype recvtype, MPI_Comm comm)
```

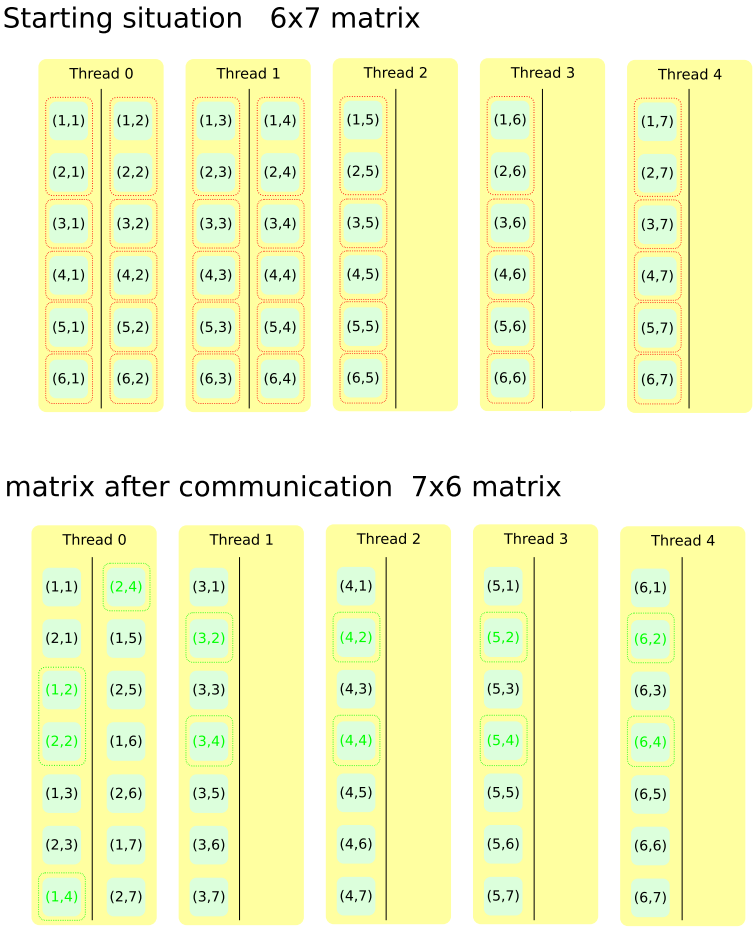


Figure 4.7: General matrix transposition scheme of the communication. Refer to text for details.

In principle this routine works in the same way as `MPI_Alltoall`. The difference is that the counts of elements sent and received and their displacements in the send and receive buffer can be specified by the integer arrays `*sendcnts, *sdispls and *rdispls, *recvcnts`.

In analogy to the above implementation, we call the `MPI_Alltoallv` for each columns on each thread. If some threads have one column less, they call the function nevertheless, without actually sending data.

The implementation is mainly a book-keeping problem. Thus let us look at an example of a 6×7 matrix, which is depicted in figure 4.7. It is decomposed and distributed on 5 threads, where the first two threads contain the two extra columns. Transposing leads to a 7×6 matrix, where an extra column resided in the memory of the first thread. Let col_i denote the number of columns on thread i after the transposition. We split all local columns in packages in such a way that the j^{th} package of thread i contains the data to be sent to thread j. More precisely the actual position on thread j is given by the i^{th} element of thread j's `rdispls` array.

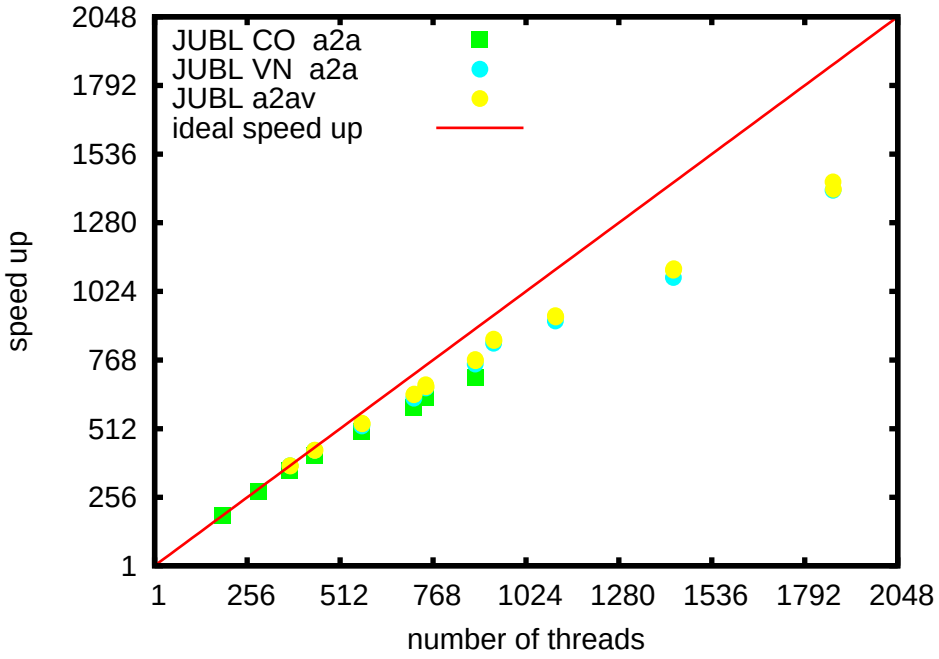


Figure 4.8: Speed up of Lanczos code on JUBL for a 18 sites half-filled Hubbard chain. The green squares denote the `MPI_Alltoall` implementation run in CO mode, the cyan circles the same implementation in VN mode. The `MPI_Alltoallv` implementation in VN mode is represented by yellow circles. The circles are actually pairs of circles. The second one denotes a system with a single-processor more, to gauge effects of load-imbalance.

The size of the package is col_j elements. This decomposition in packages is denoted in the scheme as the red boxes. The first thread will hold two columns after the transposition and thus all first packages comprise two elements.

On the receiving thread the displacement array for the incoming packages is set up in such a way, that there is enough room to store all the data of a thread i (over all calls of `MPI_Alltoallv`) contiguously in memory. In the figure this is denoted by the green boxes in the second matrix. After calling `MPI_Alltoallv` for the first column, the contents of the green boxes are undefined. The second call fills these boxes. Hence all data which originated from a thread i is stored contiguously.

After finishing all `MPI_Alltoallv` calls, it can be observed that the matrix elements are stored in row-major order in the columns. In our example this means we have a 2×7 matrix. Thus we need a local transposition operation of this rectangular matrix. As a result we obtain the globally transposed matrix we are interested in. The yellow circles in figure 4.8 show the speed up of this `MPI_Alltoallv` implementation.

We see that both transposition operations are very efficient and as a consequence this also applies to the matrix-vector multiplication. Thus the overall performance of the Lanczos implementation scales very well, even on massively parallel computers

with thousands of processors. Comparing `MPI_Alltoall` with `MPI_Alltoallv` we find that the more complicated `MPI_Alltoallv` code scales better; it takes, however, slightly more time to execute. We briefly analyze why.

`Alltoall` vs. `Alltoallv`

When comparing the time of all Lanczos iterations for matrix-vector products implemented by `MPI_Alltoall` and `MPI_Alltoallv`, we can see from table 4.2 that the latter code is about a factor of 1.14 slower for 1938 processors. This is due to two effects. First the call of `MPI_Alltoallv` takes longer to complete. This effect increases considerably when scaling to more processors. On the other hand we use an out-of-place matrix transpose which is more expensive than the strided in-place transpose used in the `MPI_Alltoall` implementation. This effect decreases with increasing number of processors for a fixed system, because the amount of data per thread becomes smaller.

How large are those two effects for the example system of 20 sites with 7 electrons of either spin? The timing difference per iteration for both implementations is about $0.7 - 0.8$ s on 1938 processors, where 5.9 and 6.7 are the absolute timing values for the `MPI_Alltoall` and `MPI_Alltoallv` implementation respectively.

0.2 s of this time is due to both `MPI_Alltoallv` calls. The setup of the book keeping, i.e. send, receive counts and displacements arrays, is negligible. Thus the rest of the time is spent in both out-of-place transposes.

Table 4.2: Run-time difference of Lanczos iterations for `MPI_Alltoall` and `MPI_Alltoallv` for a 20 sites system with 7 up and down electrons and 1938 threads. The more general code for rectangular matrices is slower by a factor of 1.14.

Implementation using	run-time in s
Alltoall	293
Alltoallv	333
ratio	1.14

What is the run-time impact of handling systems with $N_\downarrow \neq N_\uparrow$? These systems are particularly important for the calculation of Green's functions.

Table 4.3 shows the timings per Lanczos iteration for a 18 sites system with 9 electrons of either spin and a system with one up electron less. Since the run-time of the Lanczos iterations scales roughly linearly with the dimension of the Hilbert space we need to scale the time of the calculation for the large system in order to compare. The scaling factor is given by $\dim(\mathcal{H}_{8,9})/\dim(\mathcal{H}_{9,9}) = 0.9$. The table shows that there is no performance degradation for non-square matrices if the `MPI_Alltoallv` implementation is used. One might have expected an impact of load imbalance. But for 374 threads the system is not in a load imbalanced state. The dimension of both Hilbert spaces $\dim_\uparrow$ and $\dim_\downarrow$ can be divided by the number of threads without remainder. For 715 threads this is, however, not true. 20% of the threads have to

Table 4.3: Run-time difference of Lanczos iterations of `MPI_Alltoallv` implementation for a state vector which can be decomposed into a square and into a rectangular matrix. The Hilbert space of the 9 up and 8 down electron system is a factor of 0.9 smaller than the other one. With the assumption that the Lanczos code scales linearly with the dimension of the Hilbert space, the effective ratio is defined by $0.9 \cdot T_{9,9}/T_{9,8}$.

System with 18 sites and	#threads	run-time per iteration	effective ratio
9 up, down electrons	374 s	10.28 s	1.00
9 up, 8 down electrons	374 s	9.23 s	
9 up, down electrons	715 s	6.02 s	1.00
9 up, 8 down electrons	715 s	5.44 s	

work with one column extra. Since each thread has a minimum of 61 columns, the work for this extra column is negligible. Hence there is no problem of load imbalance as long as there are sufficiently many columns on each thread.

4.3 Checking the results

Checking is crucial when developing a complex code. We have to ensure that the results obtained by the complicated numerical calculations are correct. This can be done by comparing the results to analytically derived or otherwise known results.

4.3.1 Energy check

Band-limit check

In the band-limit case we consider the Hubbard model for $U = 0$ for which the solution is given by a Slater determinant. The Hamiltonian is diagonal in k-space and we obtain the well-known cosine band. The code uses the real-space configuration basis, clearly an unfavorable choice in this case. The computer performs complex calculations whose results we can obtain. It is however a good check for the setup of the hopping matrices, especially in order to spot errors in the handling of the Fermi-signs.

Table 4.4 shows the band-limit ground-state energies for half-filled one-dimensional Hubbard chains for reference purposes.

Lieb-Wu check

We thus have means to check the kinetic energy. To test whether the Coulomb interaction part works we can compare the numerical results to the analytical solution of Lieb and Wu [6] in one dimension.

Table 4.4: Band-limit ground-state energies for half-filled one-dimensional Hubbard chains with PBC

sites	energy in t	sites	energy in t
2	-4.0	4	-4.0
6	-8.0	8	-9.65685424949
10	-12.9442719100	12	-14.9282032303
14	-17.9758368297	16	-20.1093579685
18	-23.0350819326	20	-25.2550060587
22	-28.1066967333	24	-30.3830164509

They showed that such a system always is a Mott insulator except for $U = 0$ (refer to chapter 2.2.1). With the Bethe ansatz they calculated the ground-state energy per particle E_0/L as a function of U in an infinite half-filled system.

$$\frac{E_0(U)}{L} = -4 \int_0^\infty \frac{d\omega}{\omega} \frac{J_0(\omega) J_1(\omega)}{1 + \exp \frac{1}{2}\omega U} , \tag{4.1}$$

where $J_i(x)$ denote Bessel functions. Starting with this equation the following asymptotic forms can be derived (see [17]). For the band limit case,

$$\frac{E_{\text{band}}(U)}{L} \approx -\frac{4|t|}{\pi} + \frac{U}{4} - 0.017\frac{U^2}{|t|} , \tag{4.2}$$

and in the atomic limit,

$$\frac{E_{\text{atm}}(U)}{L} \approx -\frac{4t^2}{U} \ln 2 . \tag{4.3}$$

With these result we can check our implementation. Figure 4.9 shows the Lieb-Wu solution for an infinite system and the numerically obtained result for a 12 sites system with half-filling. For increasing values of U physics becomes more and more local and thus finite-size effects decrease. The results of the calculation for 12 sites quickly approach the analytical one for the infinite chain.

4.3.2 Green's function checks

How do we check whether the Green's function implementation yields the right result? We will discuss some techniques in this section.

Sum rules and momenta

A test which occurs directly is to check the sum rules, introduced in chapter 3.4. We can either integrate the spectral function yielding

$$\int_{-\infty}^{\infty} d\omega \; A_{\alpha\beta}(\omega) = \delta_{\alpha\beta}$$

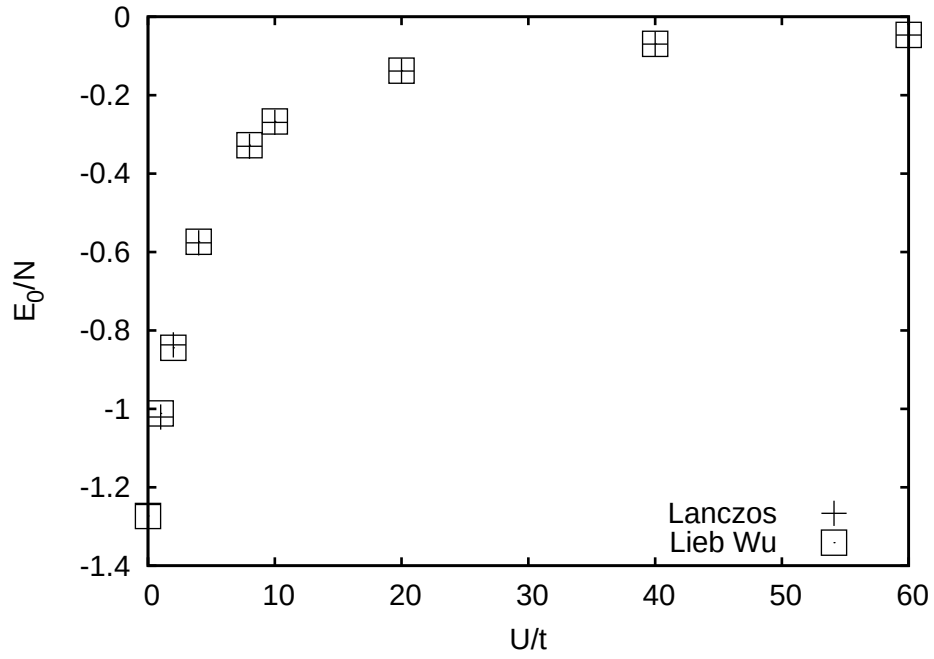


Figure 4.9: Comparison of the exact ground-state energy for an infinite half-filled system (squares) to Lanczos calculation for a 12 sites half-filled cluster with PBC. The approximation becomes better the larger U.

or equivalently sum up all spectral weights, since

$$\begin{aligned} A_{\alpha\beta}(\omega) = & \sum_n \left\langle \psi_0 \middle| c_\alpha \middle| \psi_n^{+1} \right\rangle \left\langle \psi_n^{+1} \middle| c_\beta^\dagger \middle| \psi_0 \right\rangle \delta(\omega - (E_n^{+1} - E_0)) \\ + & \sum_n \left\langle \psi_0 \middle| c_\alpha^\dagger \middle| \psi_n^{-1} \right\rangle \left\langle \psi_n^{-1} \middle| c_\beta \middle| \psi_0 \right\rangle \delta(\omega + (E_n^{-1} - E_0)) \end{aligned}$$

This is, however, not a proper check. The spectral weights are directly obtained from the diagonalization of a tridiagonal matrix. We discussed in chapter 3.4.2 that these weights are proportional to the first elements of its eigenvectors. So checking the zeroth moment merely checks the LAPACK routine used for the diagonalizing of the matrix.

Another try is to check, whether the Galitskii-Migdal theorem holds. As discussed in chapter 3.4.1 it connects the sum over the first moments of a spectral function to the ground-state energy

$$\sum_{\mathbf{k}} \mu_1^{\mathbf{k}} = -\left(\frac{\langle T \rangle_0}{2} + \langle V \rangle_0 \right) .$$

But we know, that the Lanczos method converges rapidly to the spectral function, which is especially reflected in the rapid convergence of the momenta. The

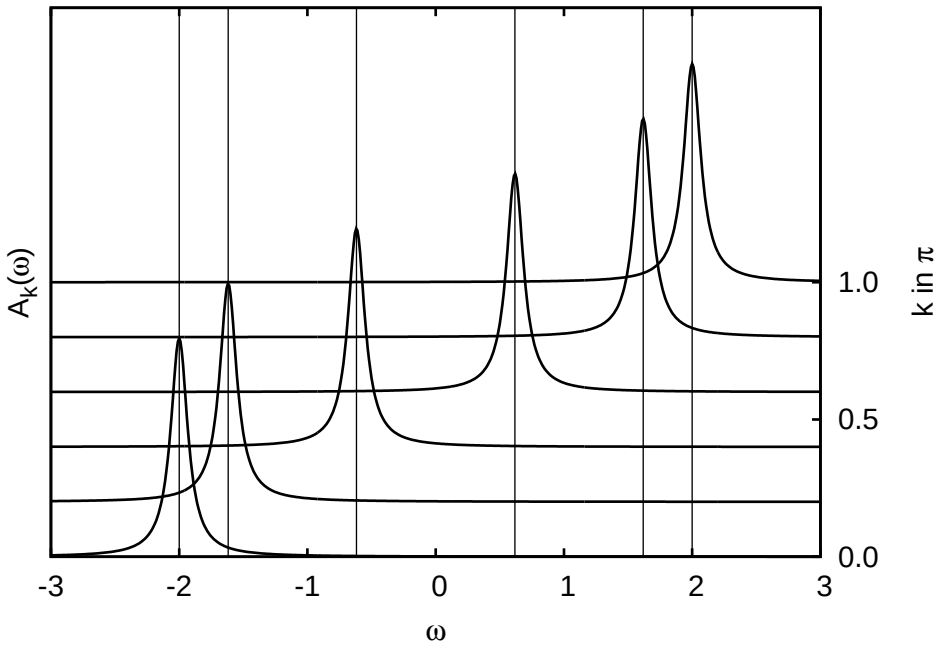


Figure 4.10: Band-limit check for a 10 sites Hubbard chain with half-filling and $t = 1$. The peaks are at the excitation energies of the corresponding single-particle system, denoted by vertical lines.

Galitzkii-Migdal test is passed after the first iteration. It therefore only gives a hint if the implementation works correctly. For judging whether the spectral function is converged it cannot be employed.

Band-limit check

The Galitzkii-Migdal check and the test of the sum rules do not thoroughly check the code for the discussed reasons. But we can again resort to the band-limit case. Here we know all results exactly and can test the computational results for the ground-state energy (see above) as well as the Green's functions. The angular-resolved Green's function is calculated for all possible values of k. We expect peaks with weight one to be at the excitation energies of a single-particle system. From the tight-binding approximation we know that these peaks for a one-dimensional system with L sites and periodic boundary conditions are at

$$\epsilon_i = -2t \cos\left(\frac{2\pi}{L} i\right) , \tag{4.4}$$

where i is an integer. Figure 4.10 shows the results for a 10 sites system. The vertical lines show the 5 energy eigenvalues for $k > 0$ calculated with equation (4.4). The numerically obtained spectral functions have their peaks at exactly these energies. Hence we see that the implementation is correct.

Atomic-limit check

In the atomic limit ($t = 0$), we can easily check the interaction part. Hubbard showed [1] that the angular integrated Green's function in this limit can be written as[2]

$$G_{ii}(\omega) = \left\{ \frac{1 - \frac{1}{2}n}{\omega} + \frac{\frac{1}{2}n}{\omega + U} \right\} ,$$

where n is the particle density per site.

Intuitively it is clear where the poles come from. Let us consider the case of a half-filled chain, i.e. $n = 1$. Each site is occupied with exactly one single electron. Thus, removing one electron does not change the energy of the system, leading to a peak at zero. The inverse process, i.e. adding an electron, leads to a double occupancy and thus to an energy increase of U. Hence we have a means to check the interaction part of the Hamiltonian. An example is shown in figure 4.11. We calculate the angular-integrated spectral function with the Lanczos method for $U = 8$ and obtain exactly the expected results.

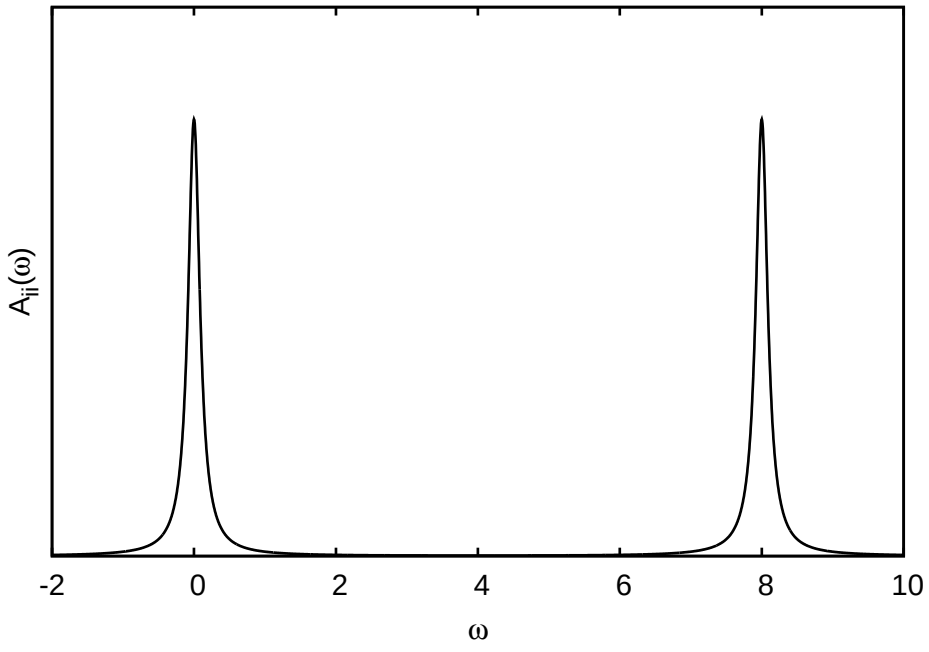


Figure 4.11: Atomic limit check for a 12 sites Hubbard chain with half-filling and $U = 8$. The photoemission peak is at $\omega = 0$ and the inverse photoemission peak at $\omega = U$.

[2]We will actually derive a slightly more general formula in chapter (5.2).

5 Angular-resolved spectral functions and CPT

In this chapter we discuss how to calculate angular-resolved spectral functions. Normally exact diagonalization enables us to efficiently calculate angular-resolved spectral functions, but only at a very low resolution in k-space. To improve the resolution we introduce complex boundary conditions. In principle we can then obtain a spectral function for each k value. In the case of non interacting particles this would yield the correct result. For interacting particles, however, we face severe finite-size effects.

Cluster perturbation theory (CPT) remedies this problem. The idea is to solve a finite cluster with open boundary conditions exactly and then treat hopping between those clusters perturbatively. This method gives a means to study the spectral function for each k value without treating different systems and quickly converges to the infinite-size limit.

5.1 Complex boundary conditions

In chapter 3.4.2 we have discussed how to efficiently calculate spectral functions with the Lanczos method. In the case of periodic boundary conditions, however, there are only $\lfloor L/2 \rfloor + 1$ independent angular-resolved spectral functions due to inversion symmetry. For an 8 sites system, for example, we only get a k-resolution of 5 spectral functions as shown in the left plot of figure 5.1. This resolution is much too poor to study details of spectra like for instance spin-charge separation. Hence, we need an improved technique. In chapter 2.1.3 we have already discussed, how to substitute a Hamiltonian of infinite dimension by finite dimensional but k-dependent replacement Hamiltonian. This was achieved by introducing complex boundary conditions. We proceed similarly in this case. Instead of periodic boundary conditions, i.e. $\psi(r_1, \ldots, r_i, \ldots, r_N) = \psi(r_1, \ldots, r_i + L, \ldots, r_N)$ where i denotes the site index we can also introduce general complex boundary conditions, i.e. $\psi(r_1, \ldots, r_i, \ldots, r_N) = e^{i\phi L}\psi(r_1, \ldots, r_i + L, \ldots, r_N)\ \forall i$. This requires us to use complex wave functions and to generalize the Lanczos code to Hermitian matrices, a task straightforward to do. If the development language supports templates (generics) like for instance C++ only minimal changes to the code are necessary. We can therewith maintain a unified code base.

Utilizing the complex phase shift ϕ we can sample the entire k-space. This is shown in the right plot of figure 5.1. However, we also observe that this method

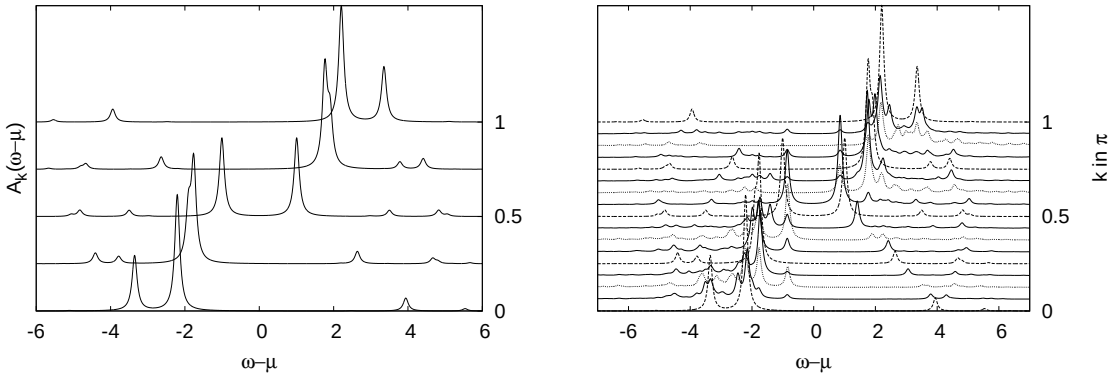


Figure 5.1: Angular-resolved spectral function for a 8 sites half-filled Hubbard chain with various periodic boundary conditions ($U/t = 4$). Left panel: periodic boundary conditions (PBC). Right: The dashed lines show the results from the left panel. The dotted lines represent the results for anti-PBC ($\phi = \pi$) and the solid ones for $\phi = \pi/2$.

does not seem to work well. There is a significant difference between the spectral function for $\phi = 0$ and other values of ϕ. This is due to the small size of the system. In chapter 2.1.3 the transformation to a k-dependent Hamiltonian was exact because we regarded the particles as independent. In chapter 2.1.4 we argued, why this does not hold anymore in the presence of correlations. Thus, for different boundary conditions we study essentially a different system which in particular can lead to shifts in the chemical potential. In chapter 6.2 we see that such a system is equivalent to a ring whose center is threaded by a magnetic flux. And thus it contains different physics. Figure 5.2 shows the mentioned shifts in the chemical potential. It can be observed, that some photoemission peaks are right of the Fermi level of the system with PBC, which clearly is an artifact, since electrons cannot be expelled from orbitals, which are not occupied.

Since we need reliable results as well as a high resolution we clearly need a more advanced method.

5.2 Cluster Perturbation Theory

5.2.1 Introduction

Cluster perturbation theory (CPT) is based on the strong coupling perturbation theory by Pairault, Sénéchal and Tremblay [31]. Using CPT we can calculate angular-resolved single-particle Green's functions with arbitrarily high resolution. Short-range correlation effects on the length scale of a cluster are treated exactly, whereas longer-range correlations are neglected. CPT works as follows:

1. Split the original infinite lattice into disconnected finite clusters as shown in

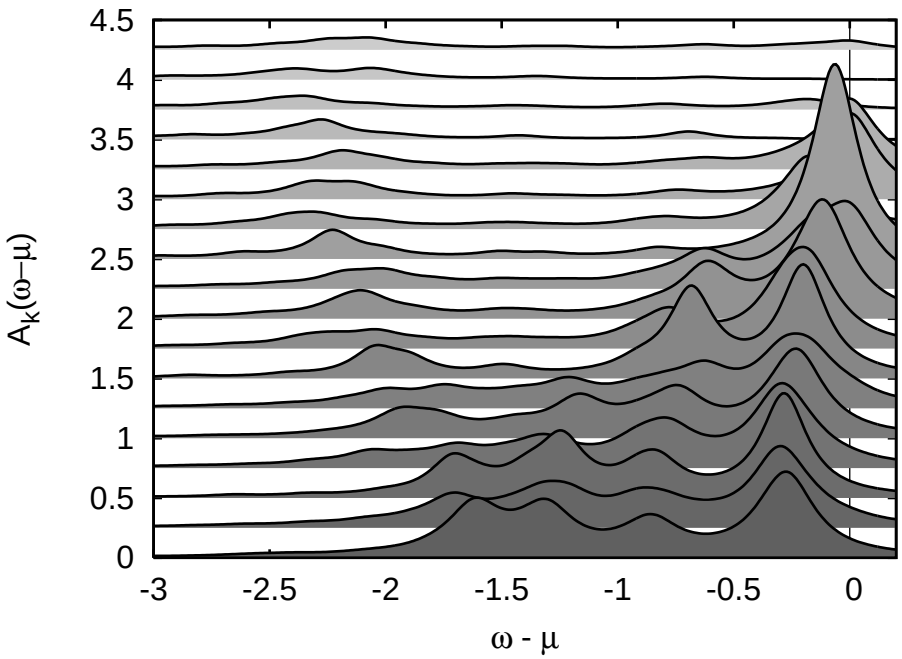


Figure 5.2: Photoemission spectral function $A_k^{PE}(\omega)$ of a 20 sites Hubbard chain with 6 electrons of either spin ($U/t = 10$) calculated with ED and complex boundary conditions.

figure 5.3

2. Calculate the full Green's matrix $G_{ij}^c(\omega)$ on such a cluster exactly (including full intra-cluster hopping)

3. Treat inter-cluster hopping perturbatively by strong-coupling perturbation theory to obtain the full system's Green's function

Let from now on γ denote the original lattice and Γ the superlattice of clusters. Formally a cluster is then given by γ/Γ.



Figure 5.3: Splitting the original infinite lattice (a) into identical finite clusters (b) (here with four sites). The dotted line denotes intra-cluster hopping, which is treated exactly in CPT. Inter-cluster hopping is treated perturbatively as sketched by arced arrows.

This method is applicable in any dimension, though we will only consider one dimensional systems here. In the first step we have the freedom to choose how many

sites a cluster consists of. Of course we expect finite-size effects to decrease with increasing number of sites. Experience shows, however, that even relatively small systems – as small as 8 sites for instance – yield good results. Especially for half-filled systems (with finite U) convergence is very rapid. This is because those systems are Mott insulators and their density matrix $\rho_{ij} = \rho_{0,|i-j|}$ decreases exponentially for $|i-j| \to \infty$ [8], leading to the so-called near-sightedness, i.e. local physics, which significantly reduces finite-size effects (cf. figure 5.4).

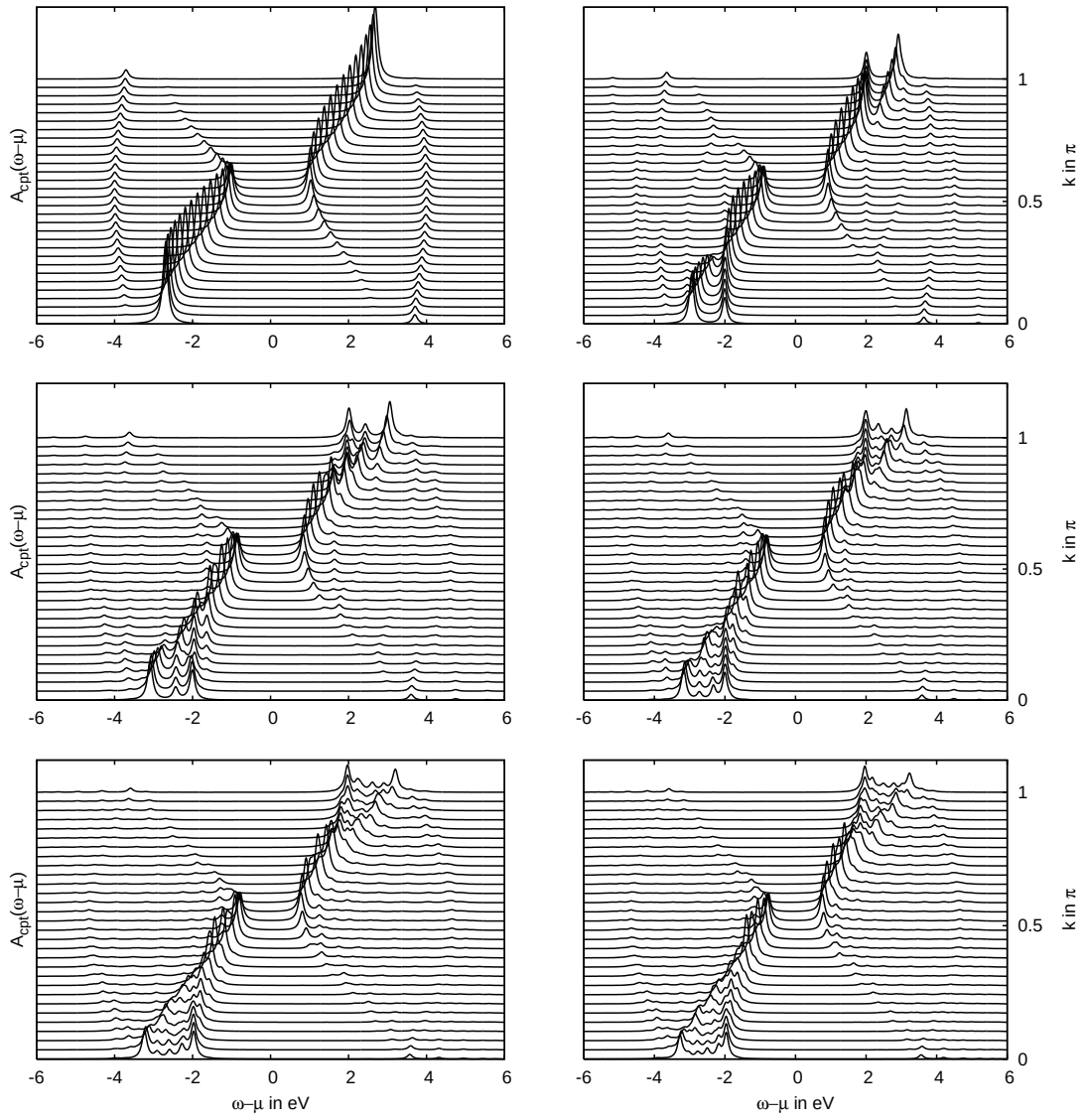


Figure 5.4: CPT spectral functions for (in row-wise order) 2, 4, 6, 8, 10, 12 sites at half-filling and $U = 4t$.

Moreover we have to choose boundary conditions for the finite clusters. Even though there are no a priori reasons for choosing open boundary conditions, aside

from seeming to be more natural, it shows that their convergence behavior is better.

As the second step we need to calculate the full Green's function matrix on the finite cluster. This can be pursued following the techniques introduced in chapter 3.4.2. The third step demands more explanation.

5.2.2 The CPT method

Following [32] and [33] we split the full lattice γ in an infinite number of finite clusters with L sites each. The clusters are labeled by an integer number m, their position is denoted by $R_m = m\,(La)$ with a the lattice constant being set to unity from now on. The sites within each cluster are labeled with the indices i, j where $i, j = 0, \ldots, L-1$. The position of site i within the cluster shall be denoted by $r_i = i$. Let $c_{m,i}$ denote the annihilation operator which removes an electron on site i in cluster m. Thus the Hamiltonian of the full system can be decomposed as

$$H = H^c + T \;, \tag{5.1}$$

where

$$H^c = \sum_m H^c_m, \qquad\qquad T = \sum_{mnij} T^{mn}_{ij} c^\dagger_{mi} c_{nj} \;, \tag{5.2}$$

with H^c_m being, in our case, the Hubbard model for cluster m, i.e.

$$H^c_m = -t \sum_{\langle i,j\rangle\sigma} \left(c^\dagger_{mi\sigma} c_{mj\sigma} + h.c. \right) + U \sum_i n_{mi\uparrow} n_{mi\downarrow} \;. \tag{5.3}$$

T denotes the nearest-neighbor hopping term between adjacent clusters with hopping matrix,

$$T^{mn}_{ij} = -t\left(\delta_{m,n-1}\delta_{i,L-1}\delta_{j,0} + \delta_{m,n+1}\delta_{i,0}\delta_{j,L-1}\right) \;, \tag{5.4}$$

which will be treated as perturbation.

Single-particle Green's function We are interested in the single-particle Green's function $G^{mn}_{ij}(\omega)$. By neglecting the perturbation term T in equation (5.1) the clusters decouple and the Green's function becomes diagonal in the cluster indices m, n, i.e.

$$G^{mn}_{ij}(\omega) = \delta_{mn} G^c_{ij}(\omega) \;, \tag{5.5}$$

where $G^c_{ij}(\omega)$ is calculated by the Lanczos method as described in chapter 3.4.2.

The hopping between the clusters is treated perturbatively. Pairault, Sénéchal and Tremblay showed in [31] and [34] that the Green's function of the full system in lowest order of strong-coupling perturbation theory is

$$\left(\hat{G}^{mn}\right)^{-1} = \left(\hat{G}^c\right)^{-1} - \hat{T}^{mn} \;, \tag{5.6}$$

with matrix notation $\hat{G}^{mn} = \left(G^{mn}_{ij}\right)$, $\hat{G}^c = (G_{ij})$ and $\hat{T}^{mn} = \left(T^{mn}_{ij}\right)$ or equivalently

$$\hat{G}^{mn} = \hat{G}^c \left(1 - \hat{T}^{mn}\hat{G}^c\right)^{-1} \;. \tag{5.7}$$

An alternative derivation of (5.6) via the self-energy was given already by Gros and Valentí in [35] .

Translation symmetry By splitting the translationally invariant lattice into finite clusters we obviously break the symmetry. It is, however, preserved on the superlattice Γ. This allows us to express the hopping term T of adjacent clusters as well as $\hat{G}^{mn}$ in terms of a wave vector Q, which belongs to the reduced Brillouin zone BZ_Γ of the superlattice. Fourier transformation of equation (5.4) yields

$$T_{ij}(Q) = -t\left(e^{iQL}\delta_{i,L-1}\delta_{j,0} + e^{-iQL}\delta_{i,0}\delta_{j,L-1}\right) \ . \tag{5.8}$$

Combining with equation (5.7) the Green's function reads

$$G_{ij}(Q,\omega) = \left(\hat{G}^c(\omega)\left(1 - \hat{T}(Q)\hat{G}^c(\omega)\right)^{-1}\right)_{ij} \ . \tag{5.9}$$

The Green's function is in a mixed representation: real space within and reciprocal space between the clusters. To obtain a full momentum-dependent Green's function we have to Fourier transform the full Green's function. Due to the missing translation symmetry the result will, however, depend on two momenta k, $k' \in \mathrm{BZ}_\gamma$ of the full Brillouin zone.

Fourier transforming the Green's functions yields

$$\begin{aligned} G_{kk'}(\omega) &= \frac{1}{NL}\sum_{\substack{mn\\ij}} G_{ij}^{mn}(\omega)e^{-ik(R_m+r_i)}e^{ik'(R_n+r_j)} \\ &= \frac{1}{N^2L}\sum_{\substack{mnQ\\ij}} G_{ij}(Q,\omega)e^{iQ(R_m-R_n)}e^{-ik(R_m+r_i)}e^{ik'(R_n+r_j)} \\ &= \frac{1}{L}\sum_{\substack{Q\\ij}} G_{ij}(Q,\omega)e^{ikr_i}e^{-ik'r_j}\delta(K-Q)\delta(K'-Q) \end{aligned} \tag{5.10}$$

In the last step the wave vector k was uniquely decomposed as $k = K + \zeta$, where $K \in \mathrm{BZ}_\Gamma$ and ζ belongs to the L reciprocal superlattice vectors. Thus, $e^{i\zeta R_i} = 1$. Therefore, $T(K) = T(k)$ and consequently $G(K) = G(k)$. Carrying out the summation over Q leads to

$$G_{kk'}(\omega) = \frac{1}{L}\sum_{ij} G_{ij}(k)e^{-ikr_i}e^{ik'r_j}\delta(K-K') \ , \tag{5.11}$$

where $\delta(K-K')$ can be written as $\delta(K-K') = \sum_i^L \delta(k-k'+\zeta_i)$ with ζ_i being one of the L reciprocal superlattice vectors.

Within the CPT approximation the full translational symmetry is restored by neglecting the off-diagonal elements or put another way by approximating $\delta(K-K') \approx \delta(k-k')$, neglecting all ζ_i but $\zeta_0 = 0$.

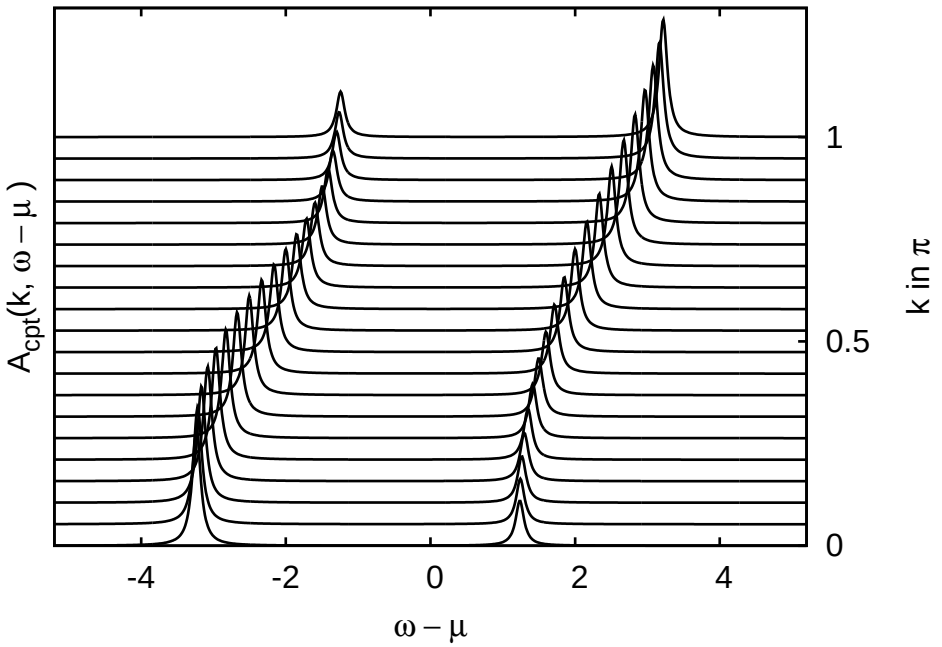


Figure 5.5: Spectral function $A(k, \omega - \mu) = -1/\pi \Im(G_{11})$ for the one dimensional Hubbard model with a single site ($U = 4t$).

Thus, the CPT approximation to obtain the single-particle Green's function is given by

$$G_{kk'}(\omega) = \frac{1}{L} \sum_{ij} G_{ij}(k) e^{-ik(r_i - r_j)} \,, \tag{5.12}$$

with $G_{ij}(k)$ given by equation (5.9).

5.2.3 Example calculation for a single site

As an example let us calculate the spectral function of a Hubbard chain in CPT for clusters consisting of a single site hosting a down electron. The Hilbert space is one dimensional and comprises the $|\downarrow\rangle$ state only. The cluster G^c therefore only contains the photoemission part of the down electron, since $|\langle \cdot | c_{1\downarrow} | \downarrow \rangle|^2 = 1$, and the inverse photoemission part of the up electron, since $|\langle \uparrow\downarrow | c^\dagger_{1\uparrow} | \downarrow \rangle|^2 = 1$. All the other matrix elements are zero, namely $|\langle \uparrow\downarrow | c^\dagger_{1\downarrow} | \downarrow \rangle|^2 = |\langle \cdot | c_{1\uparrow} | \downarrow \rangle|^2 = 0$. Hence the spin-averaged Green's function is given by

$$G^c = \frac{1}{2} \left(G^c_\uparrow + G^c_\downarrow \right) ,$$

with

$$G^c_\uparrow = \frac{1}{\omega - \mu - U} = \frac{1}{\omega - \frac{U}{2}} \quad \text{and} \quad G^c_\downarrow = \frac{1}{\omega - \mu} = \frac{1}{\omega + \frac{U}{2}}$$

and $\mu = -\frac{U}{2}$ due to half-filling's particle-hole symmetry. Thus,

$$G^c = \frac{1}{2}\left(\frac{1}{\omega - \frac{U}{2}} + \frac{1}{\omega + \frac{U}{2}}\right) = \frac{1}{\omega - \frac{U^2}{4\omega}} \; .$$

For $V_{11}(k)$ follows with equation (5.8),

$$V(k) = -t(\exp(ik) + \exp(-ik)) = -2t\cos(k) \; .$$

and therefore with equation (5.9),

$$G_{11}(k,\omega) = \frac{1}{\omega - \frac{U^2}{4\omega} + 2t\cos(k) + i\eta} \; .$$

This coincides with the Hubbard-I approximation for half-filling [1]. The corresponding spectral function is plotted in figure 5.5. Let us look at the asymptotic behavior for $U \to 0$ and $U \to \infty$. In the band-limit the ordinary dispersion relation of free particles is restored, i.e. the Green's function looks like

$$G_{11}(k,\omega) = \frac{1}{\omega + 2t\cos(k) + i\eta} \; .$$

In the atomic limit the energy gap E_g goes to infinity. At finite U the poles are situated at

$$\omega_{\pm} = -t\cos(k) \pm \frac{U}{2}\sqrt{1 + \frac{4t^2\cos^2(k)}{U^2}} \approx -t\cos(k) \pm \frac{U}{2} \; .$$

The Hubbard bands reside at $\pm U/2$ and the bandwidth is reduced by a factor of 2.

5.2.4 Limits of the CPT

In the case of infinite cluster length ($L \to \infty$) the CPT results obviously become exact. This also applies to the case of no inter-cluster hopping at all because if was considered as the perturbation in strong-coupling perturbation theory. With full translational symmetry, i.e. no hopping globally this is equivalent to the atomic limit.

It may seem surprising at first that CPT is also exact in the case of $U = 0$, i.e. the weak-coupling limit. This, however, becomes clear when looking at (5.6). In the case of $U = 0$ this equation becomes

$$\left(\hat{G}^{mn}\right)^{-1} = \left(\hat{G}_0^c\right)^{-1} - \hat{T}^{mn} \; , \tag{5.13}$$

where $\hat{G}_0^c$ is the Green's function of the cluster of the non-interacting system. Then, $\hat{T}^{mn}$ with the parameter t obviously is the exact self-energy. Hence, the fact that for $U = 0$ the band-limit results are restored can be used in practice to check the implementation.

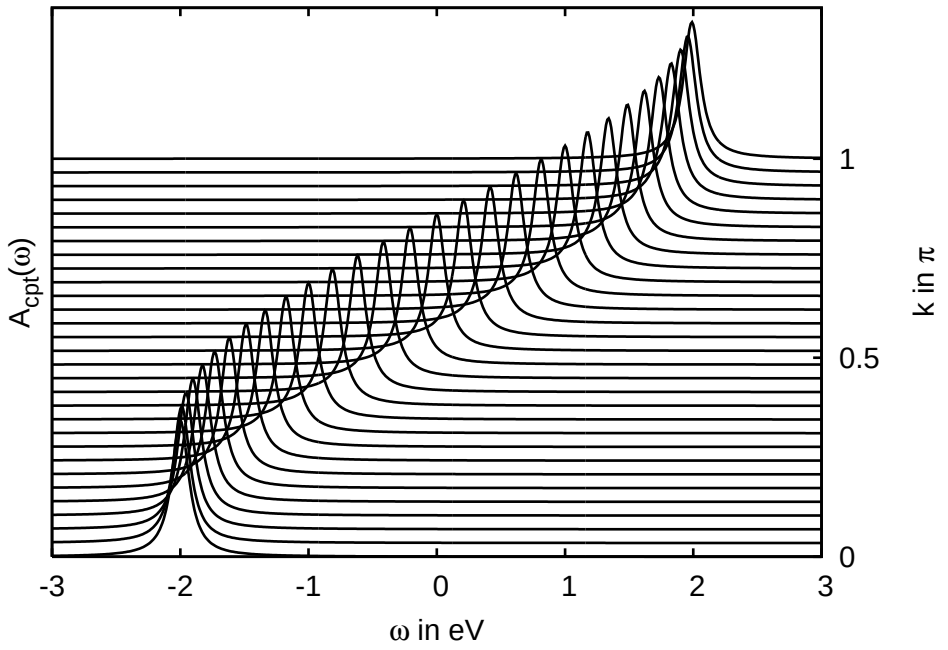


Figure 5.6: Weak-coupling limit check: CPT calculation $U = 0$ for a 6 sites system (band-limit). The tight-binding cosine band is restored.

5.2.5 Boundary conditions

We have already stated that open instead of periodic boundary conditions are used to compute the Green's function of the cluster. PBC would be favourable as far as computational resources are concerned. This is because PBC systems are translationally invariant and the Green's function only depends on $|i-j|$, i.e. $G_{ij} = G_{|i-j|}$. For OBC, on the other hand, the computational demand increases by a factor of L. Dahnken, Arrigoni and Hanke suggested [36] to use PBC anyway. In order to correct the wrong periodic hopping it has to be subtracted in strong-coupling perturbation theory. Therefore equation (5.8) changes to

$$T_{ij}(Q) = -t\left((e^{iQL}-1)\delta_{i,L-1}\delta_{j,0} + (e^{-iQL}-1)\delta_{i,0}\delta_{j,L-1}\right) . \tag{5.14}$$

The resulting spectral functions are inferior to the one calculated by OBC. This can be observed in figure 5.7. Inferior means that the spectral functions with OBC show quite strong finite size effects, whereas the PBC spectral functions are essentially converged (cf. figure 5.4). According to [33] this is because subtracting the periodic hopping leads to a long-range hopping in T for which CPT does not work well.

5.2.6 CPT vs. ordinary ED

With the ordinary symmetric Lanczos method we obtain only a very crude resolution in k-space as seen in the previous section. With the introduction of complex boundary

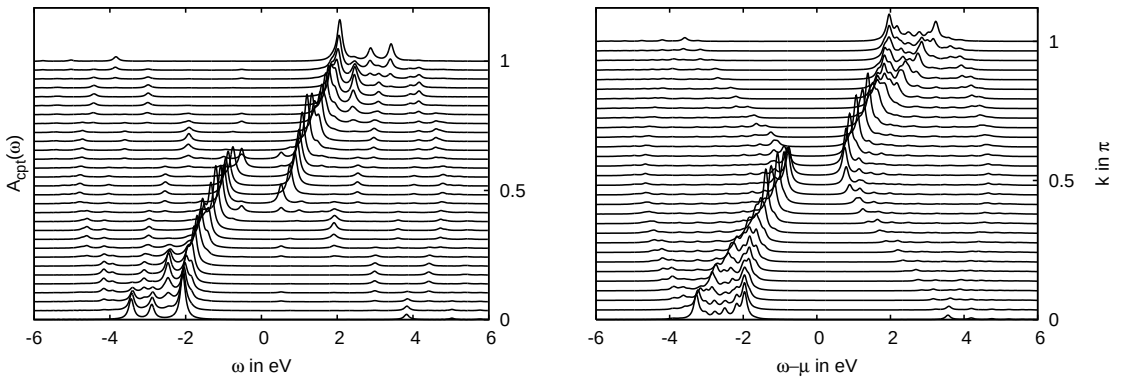


Figure 5.7: Comparison of a one-dimensional 12 sites half-filled Hubbard chain ($U = 4$) with PBC (left) and OBC (right). The OBC calculation yields considerably better convergence behavior, i.e. finite size scaling (cf. figure 5.4).

conditions we increase the resolution at the expense of accuracy. But finite-size problems arise mainly because of the change in the chemical potential. CPT is superior, as one can see in figure 5.8.

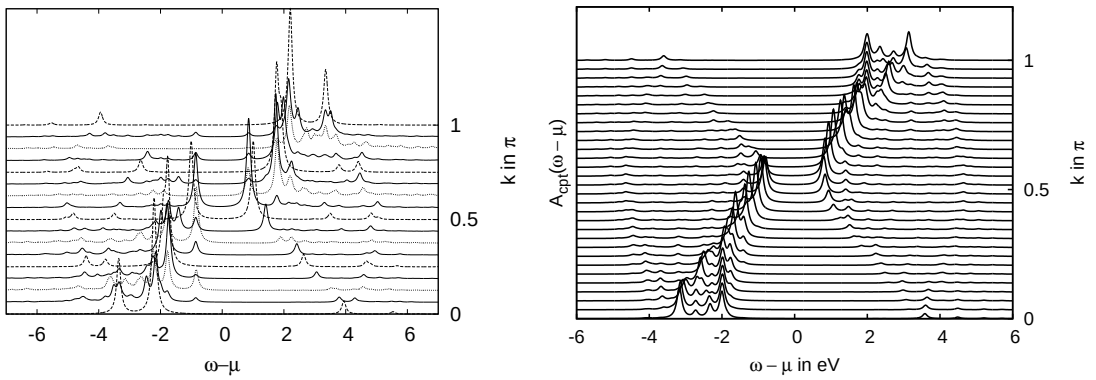


Figure 5.8: complexBC ED (left) vs. CPT (right) for a 8 sites half-filled Hubbard chain with $U/t = 4$. CPT is clearly superior.

Convergence behavior Aside from the resolution in k-space the CPT method has in addition a better convergence behavior with respect to the cluster size. From figure 5.4 we see that the CPT method converges very rapidly. For a six sites system the spectral function shows already qualitatively all the import features.

In order to directly compare the two methods we calculate the spectral functions for different numbers of sites at the Fermi wave vector. Figure 5.9 shows that indeed CPT converges much more rapidly compared to the ordinary Lanczos method. For

two sites the peaks with the largest weight are already at almost the correct energies. This is not the case for the ordinary Lanczos run where the peaks shift considerably.

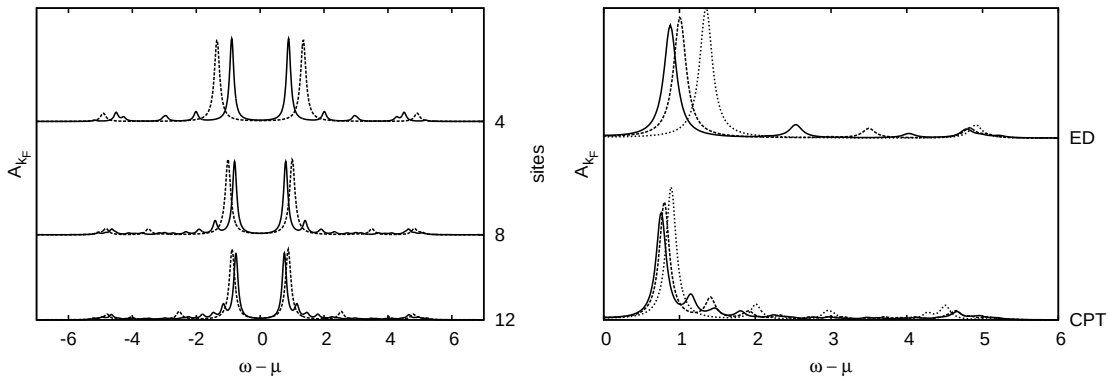


Figure 5.9: Direct comparison (left) of A_{k_F} for different number of sites (4, 8, 12) calculated with CPT (solid line) and ED with PBC (dashed line) for half-filled Hubbard chain ($U/t = 4$). The right plot shows the spectral functions of the same method in comparison for $\omega - \mu > 0$.

5.2.7 Next-neighbor interaction V

Including next-neighbor interactions into the CPT method yields a non-trivial problem. Whereas the Hubbard term is local to each site and therefore is not affected by the splitting of the cluster this is not the case for the next-neighbor interaction. It is not clear how to treat this term at both ends of a cluster. Aichhorn [37] used a mean field approach in his PhD thesis. We tested how the CPT with V behaves for two different treatments.

At first we neglect the next-neighbor interaction at both ends of the chain, i.e. we treat the next-neighbor interaction in open boundary conditions. The hopping stays untouched. Figure 5.10 shows the poor results. We observe strong artifacts which manifest themselves as many pronounced stripes.

It turns out that a better approach is to consider the next-neighbor interaction term in PBC. Figure 5.11 shows the results for two different numbers of sites. Comparing the results of our calculations with the result of Aichhorn (figure 5.4 in his thesis) shows a very good agreement. Both methods, however, share, the same problem, namely the finite-size stripes. Fortunately they are far less pronounced compared to the calculation with V in OBC. Moreover their number increases with L and their weight is reduced.

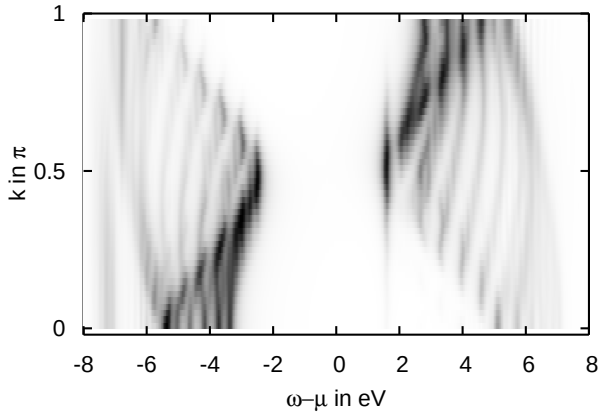


Figure 5.10: 12 sites half-filled Hubbard chain in CPT with $U/t = 8$ and $V/t = 2$, where V regarded in OBC. We see strong unphysical finite-size artifacts.

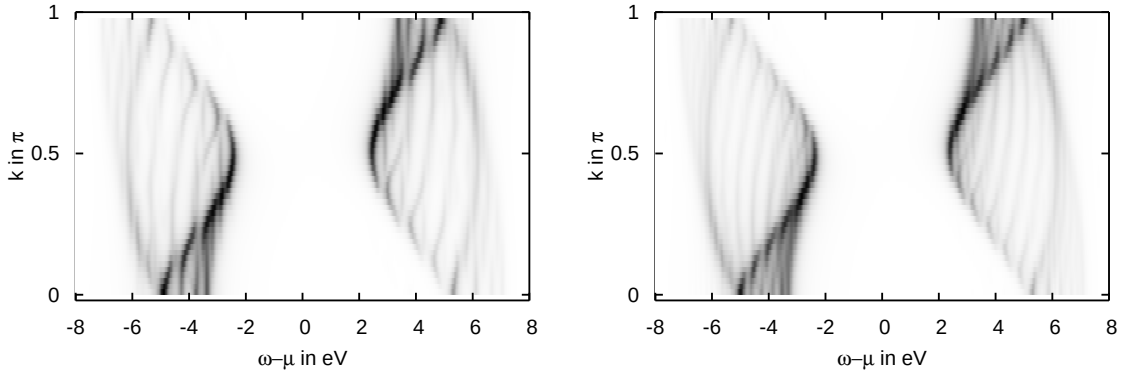


Figure 5.11: 8 (upper), 12 (lower) sites half-filled Hubbard chain in CPT with $U/t = 8$ and $V/t = 2$, where V considered in PBC. The stripes inside the Hubbard bands appear to be finite size effects, since their number increases and their weight is reduced with increasing L.

6 Metal-insulator transition

Experiments with strongly correlated materials like one-dimensional organic conductors and high-T_c materials show a strong deviation from the predictions of band theory. A prominent example are Mott insulators which are incorrectly predicted as metals.

In his paper "Theory of the insulating state" (1964) Kohn [38] developed a general way of determining whether a material is an insulator or metal. This characterization includes ordinary band insulators, predicted by band structure theory, as well as Mott insulators. It turns out that the corresponding calculations are relatively cheap: instead of looking at the spectrum the characterization is done by studying the response of the ground-state energy to changes in the boundary conditions which is a measure for the locality of the ground state.

In this chapter Kohn's method is used to study a system undergoing a Mott-band insulator transition. The studied system becomes a metal when the Mott to band transition occurs. In the final part we show how self-energies at the Fermi wave vector k_F look like for Mott and band insulators as well as metals.

6.1 Classical Drude conductivity

In order to describe the response of a metal subjected to a time-dependent electric field $\mathbf{E}(t) = \mathbf{E}(\omega)e^{i\omega t}$, let us look at the classical equation of motion of the electrons. It is given by

$$\frac{d\mathbf{p}}{dt} = -\frac{\mathbf{p}}{\tau} - e\mathbf{E} \,, \tag{6.1}$$

where $\mathbf{p}$ denotes the momentum of the electron and $1/\tau$ is its scattering rate. Using the Fourier representation

$$\mathbf{p}(t) = \mathbf{p}(\omega)e^{i\omega t} \,, \tag{6.2}$$

and inserting into equation (6.1) yields

$$-i\omega\mathbf{p}(\omega) = -\frac{\mathbf{p}(\omega)}{\tau} - e\mathbf{E} \,. \tag{6.3}$$

With the current density $\mathbf{j} = -ne\mathbf{p}/m$ where m is the mass of the electron and n their number this leads to

$$\mathbf{j}(\omega) = -\frac{ne}{m}\mathbf{p}(\omega) = \sigma(\omega)E(\omega) \,. \tag{6.4}$$

Here $\sigma(\omega)$ is the classical dynamical Drude conductivity defined as

$$\sigma(\omega) = \frac{\sigma_0}{1 - i\omega\tau} , \tag{6.5}$$

where $\sigma_0 = ne^2\tau/m$ is the corresponding static Drude conductivity.

For $1/\tau \to 0$ the static Drude conductivity obviously is retained. The electrons are independent particles and free. This leads to the Drude peak

$$\sigma(\omega) = \sigma(0)\delta(\omega) , \tag{6.6}$$

which is a result of freely accelerating electrons. It can be observed in spectra whenever quasi-free particles are present. In the following part we calculate a quantity, the so called stiffness or Drude weight, which is directly connected to $\sigma(0)$ by

$$D = \frac{\sigma(0)}{2\pi e^2} . \tag{6.7}$$

6.2 Optical conductivity in the Hubbard model

How do we determine whether the ground state of a given Hamiltonian is metallic or insulating? Or equivalently is $\Re\left(\sigma(\omega \to 0)\right) = 0$? Ordinary techniques to answer this question need the full spectrum of the Hamiltonian. This is not really satisfactory since whether a system is insulating or not is a property of the ground state. Indeed, Walter Kohn [38] showed in 1964 that it is possible to extract information about $\sigma(\omega)$ with the help of ground-state properties only. In the following we will consider a one-dimensional Hubbard chain which is connected to a ring with varying boundary conditions.

Physically, changing the boundary conditions is equivalent to a magnetic flux Φ through the center of the ring of L sites. This translates into a so called Peierls phase factor which a particle gathers when hopping from site to site. Adjusting the hopping matrix elements of the kinetic energy operator $T = -t \sum c^\dagger_{\sigma i+1} c_{\sigma i} + h.c.$, results in

$$T_\Phi = -t \sum \left(e^{-i\Phi/L} c^\dagger_{\sigma i+1} c_{\sigma i} + e^{+i\Phi/L} c^\dagger_{\sigma i} c_{\sigma i+1} \right) , \tag{6.8}$$

where Φ can be written in terms of a vector potential A as $\Phi = LAe$. The equivalence of the effects of a magnetic flux leading to a phase factor when hopping from site to site and complex boundary conditions has already been observed and used in chapter 2.1.3, when the infinite dimensional Hamiltonian was reduced to a finite one.

Assuming that the system is an insulator the ground state wave function can be decomposed into a sum of localized wave functions which have essentially vanishing overlap. Changing the boundary conditions slightly we expect that the response of the ground-state eigenvalue is very small, since it is only a surface effect. In case of a metal the wave functions overlap and the ground-state eigenvalues will probably change significantly.

Expanding the exponential function in the perturbed Hamiltonian up to second order in Φ yields

$$H_\Phi = H_0 - \frac{j}{eL}\Phi - \frac{1}{2}\frac{1}{L^2}\Phi^2 \,, \tag{6.9}$$

where H_0 is the Hamiltonian with periodic boundary conditions $\Phi = 0$ and j the paramagnetic current operator defined by

$$j = ite \sum \left(c^\dagger_{\sigma i} c_{\sigma i+1} - c^\dagger_{\sigma i+1} c_{\sigma i} \right) \,. \tag{6.10}$$

The ground-state energy shift is given in second-order perturbation theory as,

$$E_0(\Phi) - E_0 = D\frac{\Phi^2}{\text{Ł}} + \mathcal{O}(\Phi^4) \,, \tag{6.11}$$

with $E_0 = E_0(0)$ the ground-state energy of the unperturbed Hamiltonian and D denotes the so called stiffness constant given by

$$D = \frac{1}{L}\left(\frac{\langle -T \rangle}{2} - \sum_{p \neq 0} \frac{|\langle \psi_0 | j | \psi_p \rangle|^2}{E_p - E_0} \right) \,. \tag{6.12}$$

In the ground state the expectation value of the current is zero, i.e. $\langle \Psi_0 | j | \Psi_0 \rangle = 0$, since otherwise it would break the symmetry of the system and favour the direction of the current.

Being interested in the optical conductivity we specialize the vector potential $A \to A^0 e^{-i\omega t}$ such that it gives rise to an electric field $F = -i\omega A$, where the name F is chosen in order to avoid confusion with the energy E. The linear-response formalism, the Kubo equation, gives us the current conductivity σ by using equation $j(\omega) = \sigma(\omega)E(\omega)$. The actual derivation is performed in [39]. The results specifically for this model are presented in [40] and [17]. For small ω the real part of the conductivity yields

$$\sigma(\omega) = 2\pi e^2 \left(D\delta(\omega) + \frac{1}{L}\sum_{p \neq 0} |\langle \psi_0 | j | \psi_p \rangle|^2 \cdot \delta\left((E_p - E_0)^2 - \omega^2 \right) \right) \,. \tag{6.13}$$

The Drude weight shown here is indeed analogous to the Drude weight D introduced in equation (6.7) in the case of non-interacting particles. The delta-function implies quasi-free acceleration or put another way infinite static conductivity. This is reasonable since there are no dissipative mechanisms in the model.

The criteria From equation (6.11) we observe that D can be calculated from $E_0(\Phi)$ by

$$D = \frac{L}{2} \left. \frac{\partial^2 E_0(\Phi)}{\partial \Phi^2} \right|_{\Phi=0} \,. \tag{6.14}$$

Now we have means to calculate the Drude weight D and can therefore decide whether a many-particle Hamiltonian is insulating, i.e. , $D = 0$, or metallic, $D \neq 0$ [41] by performing ground-state calculations for different boundary conditions.

6.3 In practice

In practice only the ground-state energies for three different complex boundary conditions are needed to numerically evaluate the second derivative. Thus such computations are relatively cheap, in particular compared to the techniques discussed earlier, namely computing the spectral function.

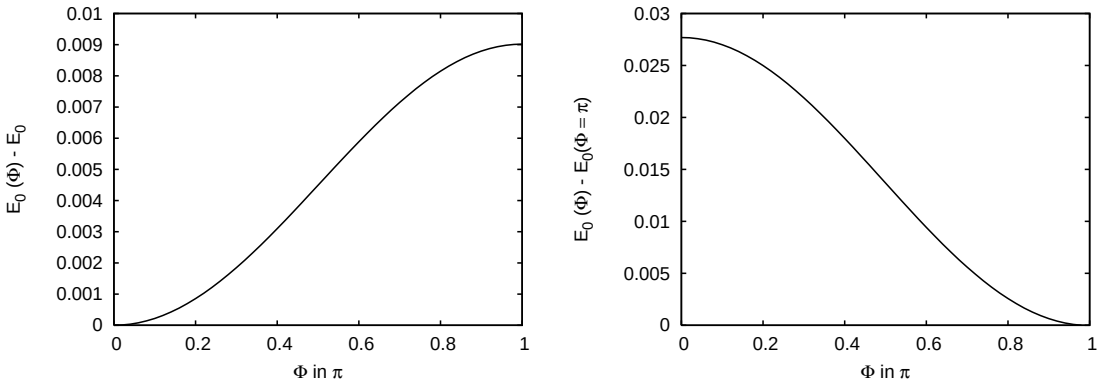


Figure 6.1: Comparison of $E(\Phi)$ for 10 (left) and 8 (right) sites systems with half-filling and $U = 6$. Whereas the 10 sites system reaches its minimum in energy for $\Phi = 0$ and therefore has a positive curvature the 8 sites system has its minimum at $\Phi = \pi$.

Degenerate ground state In half-filled systems where the number of sites is a multiple of four the calculated D, being the curvature, is negative, whereas in systems with $4n + 2$ sites, where i is an arbitrary integer, D is positive. This behavior can be found in strongly and weakly interacting systems. It is depicted in figure 6.1. Stafford, Millis and Shastry argued in [42] where the sign comes from. We will briefly discuss the small-U limit. In case of $4n$ sites and half-filling the system is an open-shell system and the ground state is degenerate. Hence there is a level crossing at $\Phi = 0$. Any small perturbation lifts the degeneracy and gives rise to a negative curvature of $E(\Phi)$ in the resulting lower band. Note that for $4n + 2$ sites systems with $\Phi = 0$ and $4n$ sites systems with $\Phi = \pi$ level crossings are sufficiently far away so they can be neglected. For $d > 1$ dimensional systems refer to [41].

This problem can be avoided by introducing anti-periodic boundary conditions for systems having $4n$ sites. Then the second derivative of $E(\Phi)$ has to be evaluated at $\Phi = \pi$.

Finite-size effects Consider Kohn's insulating ring, where the ground-state wave function can be decomposed into localized wave functions with vanishing overlap. The change of the boundary condition at an arbitrary place in the ring changes the neighboring localized wave functions. This perturbation decays swiftly because of the

vanishing overlap between the wave functions. In insulators this effect is a surface effect (c.f. chapter 2.1.4). Treating relatively small systems the surface effects can have an impact on the result and may lead to a misjudgement of the phase.

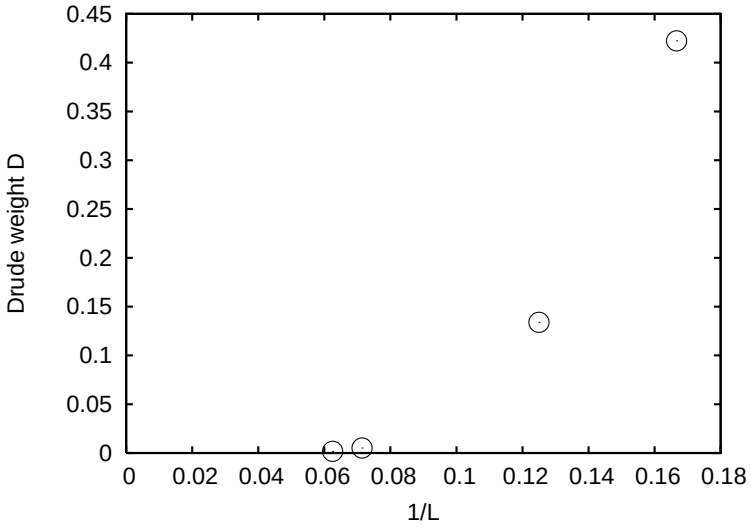


Figure 6.2: Finite-size scaling for half-filled Hubbard model with $U/t = 6$ and $\Delta = 0$. For too small system sizes the criterium misjudges the phase and tells us, that it is a metal. The finite-size effects are of higher than linear order and thus vanish quickly when increasing the system size.

The value of the Hubbard-U adjusts the "locality" of the system, the higher the ratio U/t the more the system becomes local and finite-size effects disappear. If we have a small system with relatively small ratio U/t, the perturbation does not decay quickly enough. Hence, the system we study pretends to be a metal though it is an insulator. An example is shown in the lower left plot of figure 6.4. With the Bethe ansatz it has been proven [6] that a half-filled Hubbard chain with finite $U > 0$ is a Mott insulator. For too small system sizes the criterium tells us that it is a metal. Increasing the number of sites, however, directly reveals, that this misjudgement was due to finite size effects. Figure 6.2 emphasises this fact. The finite-size effects are suppressed more than linearly.

6.4 Mott-band insulator transition

This section revisits the model studied with an effective single-particle theory in chapter 2.1.3 including correlation effects. The system is depicted in figure 6.3.

Half-filling For $\Delta = 0$ the half-filled Hubbard system for finite U/t is a Mott insulator. The electrons are distributed as uniformly as possible in the system and the occupancy $n_i \ \forall i$ of the sites is equal to one without any charge fluctuation, i.e.



Figure 6.3: Hubbard chain with next-neighbor hopping. Each second site is shifted in energy by an energy offset of Δ.

$\langle n_i^2 \rangle - \langle n_i \rangle^2 = 0$. Since the system is periodic we only need to take a unit cell with two sites $i = 0, 1$ into consideration.

Increasing Δ breaks the uniform electron distribution. As Δ rises, the electrons tend to doubly occupy the lower energy states since the higher states become too costly in energy. This can be seen in the first row plots of figure 6.4. It also can be observed that for higher values of U/t this redistribution starts at higher values of Δ/t.

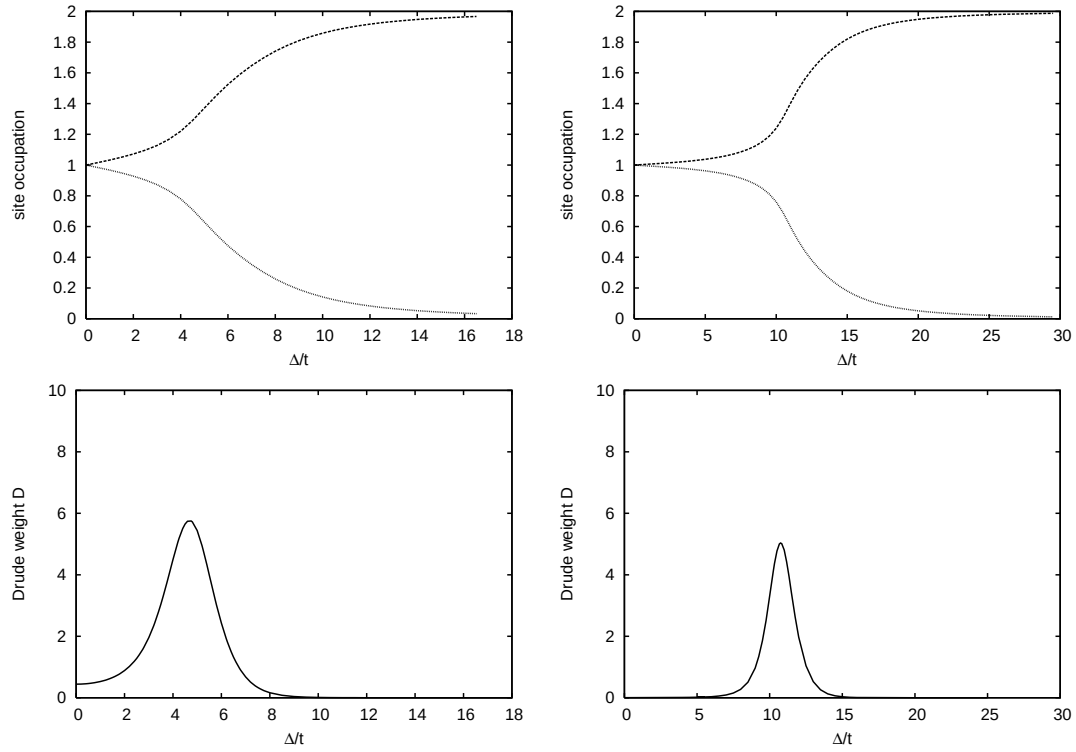


Figure 6.4: The first column shows a 10 sites system with $U/t = 6$ for different values of Δ/t, whereas the second column describes a similar system with $U/t = 12$. The first row shows the occupancy of the upper (lower line) and lower energy (upper line) orbitals and the second the corresponding Drude weights. The lower left plot shows for $\Delta \approx 0$ a finite size effect (see figure 6.2).

In the opposite limit, i.e. $\Delta/t \gg U/t$ the electrons prefer to doubly occupy the low energy orbitals since an occupancy of a higher orbital is too expensive. Then the half-filled system becomes a band insulator.

Most interesting is of course the domain where $\Delta/t \approx U/t$. Kohn's criterium shows that when the system changes its phase from a Mott to a band insulator it becomes a metal. This can be understood heuristically. Assuming a two site system the first electron put into the system will surely occupy the lower orbital. If $U = \Delta$ it does not matter for a second electron whether it occupies the higher or the lower state doubly. Therefore it can move freely and the system is a metal. The expected occupation of 1.5 on the lower orbital and consequently an occupation of 0.5 on the higher orbital can be seen in the leftmost plot of figure 6.4.

Because of the finite size of the systems no conclusions can be drawn about the width of the metal phase regime as a function of U, Δ. Resta and Sorella showed in 1999 [43] that the ground state at the transition point is indeed metallic. However, it is a singular point, though with our technique we cannot directly identify it as such.

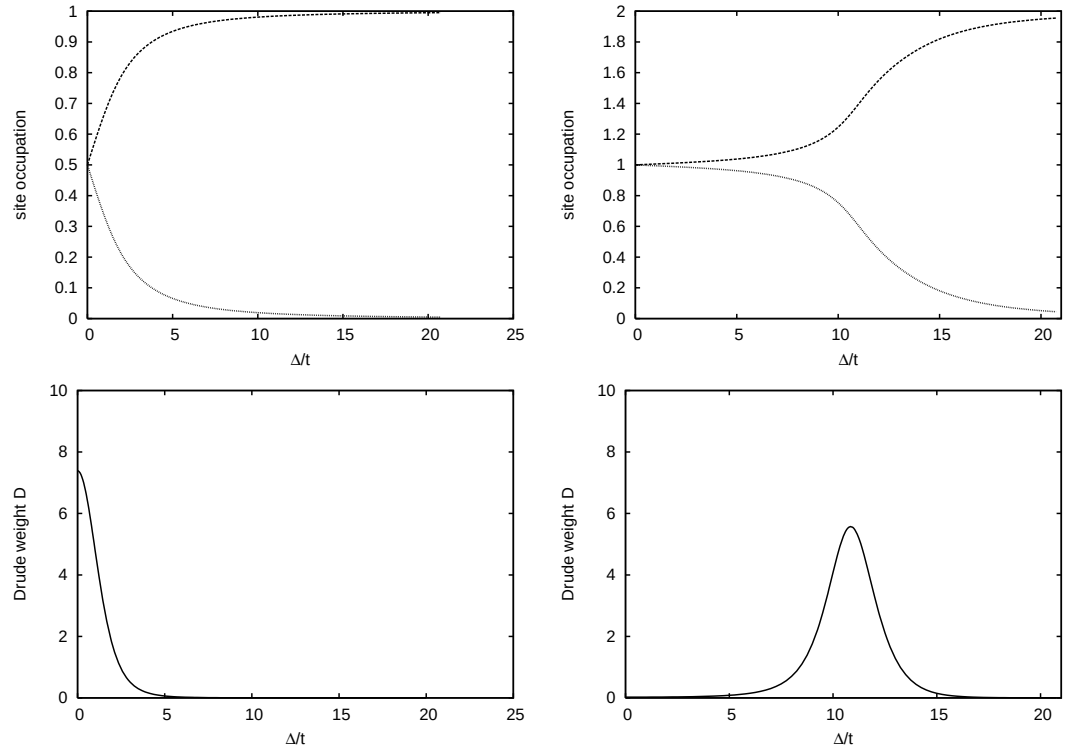


Figure 6.5: First column: Occupancy of the two orbitals of a quarter filled 8 sites Hubbard system with $U/t = 12$. Second column is a similar system with half-filling. Because this is a $4n = 8$ system, the Drude weight is evaluated at $\Phi = \pi$.

Quarter filling In case of quarter filling finite U/t and $\Delta/t = 0$ the system is a metal. The criterium confirms this as shown in figure 6.5. With increasing Δ/t the electrons distribute themselves more and more uniformly on the lower orbitals in such a way that each lower orbital is singly occupied. Neither Δ nor U is paid energetically and the system becomes a Mott insulator.

6.5 Self-energy

Green's functions of non-interacting and interacting systems can be written in a formally equivalent way by using the self-energy. The non-interacting Green's function looks like

$$G^0(k,\omega) = \frac{1}{\omega - \epsilon(k)} \,, \tag{6.15}$$

and those of interacting particles

$$G(k,\omega) = \frac{1}{\omega - \epsilon(k) - \Sigma(k,\omega)} \,, \tag{6.16}$$

where all many-body physics is contained in the self-energy $\Sigma(k,\omega)$. It can be advantageous to study $\Sigma(k,\omega)$ directly. Once G has been obtained we can calculate $\Sigma(k,\omega)$ with the help of equations (6.15) and (6.15) in the following way

$$\Sigma(k,\omega) = \omega - \epsilon(k) - \frac{1}{G(k,\omega)} = \frac{1}{G^0(k,\omega)} - \frac{1}{G(k,\omega)} \,. \tag{6.17}$$

Usually the self-energy $\Sigma(k,\omega)$ is a complex function of (k,ω).

Equation (6.15) shows that the poles of this Green's function denote the excitation energies of the non-interacting system. We can regard them as particles in a certain energy eigenstate and hence they have an infinite life time. In case of interaction the poles are at the energies for which the denominator of (6.16) is zero, i.e.

$$\omega = \epsilon(k) + \Sigma(k,\omega) \,. \tag{6.18}$$

$\Sigma(k,\omega)$ is most usually a complicated function and therefore equation (6.18) often is evaluated numerically or graphically. There might be multiple solutions for fixed values of k. This can be interpreted as the decay of a particle in the non-interacting system into several excitations in the interacting system, called quasi-particles. Their spectral weights, however, have to sum up to 1 because of particle number conservation. Whereas the life time of a particle in an energy eigenstate is infinite in the non-interacting system, the life times of the quasi-particles are inversely-proportional to the imaginary part of the self-energy function.

Having the self-energy we can tell how the spectral function looks like. Graphically each intersections of ω with $\epsilon(k) + \Sigma(k,\omega)$ for a fixed value of k yields a peak in the spectral function, whose width is proportional to the inverse imaginary part of Σ. This can be seen in figures 6.6 and 6.7.

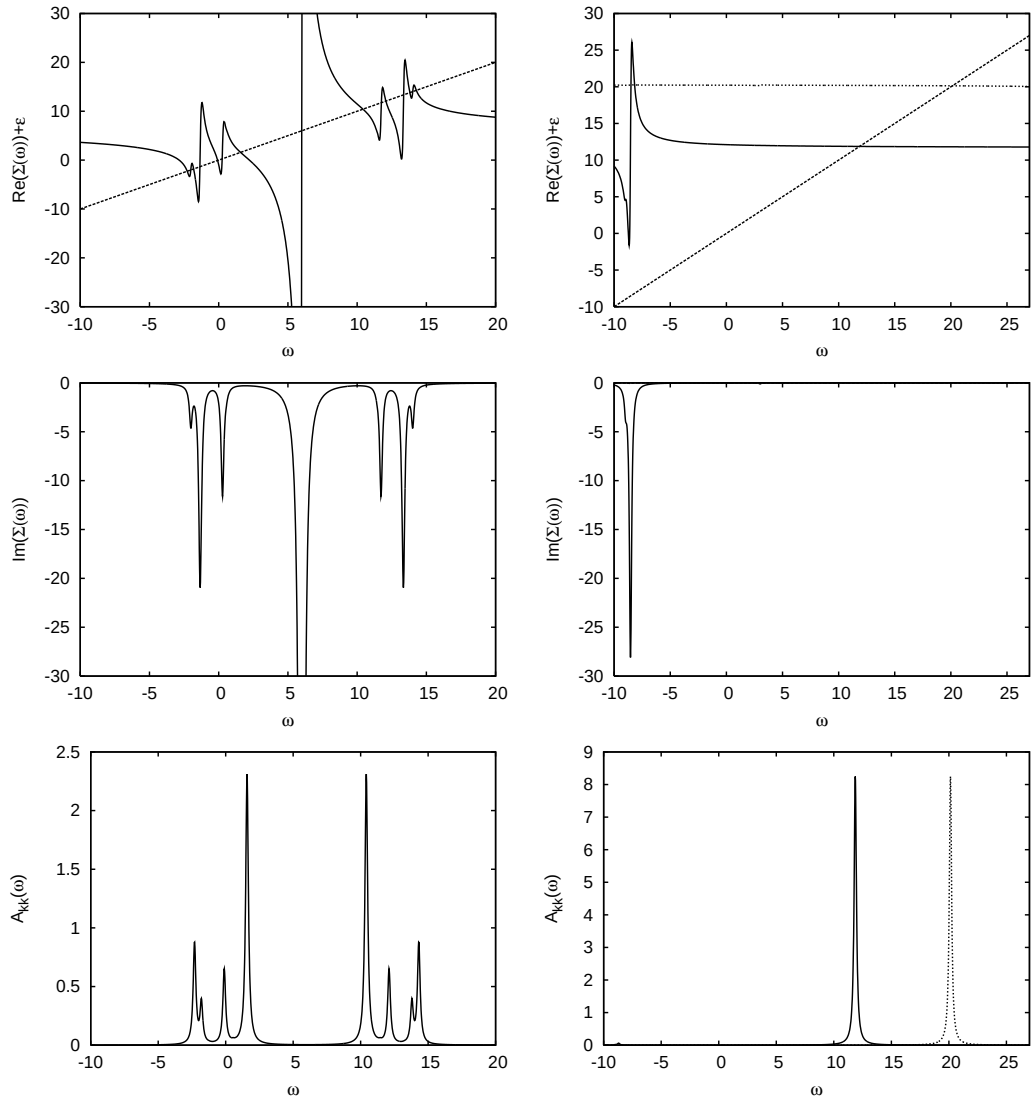


Figure 6.6: Self-energy $\Sigma(k_F, \omega)$ of a Mott insulator (left) and band insulator (right). The divergence of $\Im\ \Sigma$ at the chemical potential is typical of Mott insulators at k_F. The plots of the third row show the spectral functions A_{k_F}. Calculations are done for a half-filled 10 sites Hubbard chain with $U/t = 12$ and $\Delta/t = 0$ (left), $\Delta/t = 20$ (right). Dashed curves denote the higher, solid the lower orbitals.

6.5.1 Mott and band insulator

Let us take a look at the self-energy of the model studied in this chapter. We are particularly interested in the self-energy at the Fermi wave vector k_F for each of the three occurring phases.

The first column of figure 6.6 presents the self-energy of the Mott insulating phase.

The real part of Σ shows that the single-energy state of the non-interacting system splits into two bands. The band-width of the non-interacting system $W/t = 4$ is retained in both bands. The width of the peaks is connected to the imaginary part of Σ. The most striking feature of this imaginary part is the divergence at the Fermi level, characteristic of Mott insulators. Furthermore the energy gap around the Fermi level makes the system an insulator.

The self-energy of the band insulator (second column) yields hardly any features compared to the one of the Mott insulator. There are two peaks in the spectral function. The first one at $\omega = U$ shows the photoemission peak of the lower orbital and the second one at $\omega = \Delta$ the inverse photoemission peak for the higher orbital. The imaginary part of the self-energy is zero for those two peaks, hence the quasi-particles have a long life time.

The inverse/direct photoemission peaks have no weight respectively because the lower orbitals are already doubly occupied and the higher ones are unoccupied.

6.5.2 Metal

Figure 6.7 shows the self-energy for the Fermi wave vector k_F of a metal. Note the long living ($\Im\Sigma(k_F, \omega_F) = 0$) quasi-particles at the Fermi level. The solid line denotes one which is removed by photoemission from a lower orbital yielding the energy gain of $U = 12$. The dashed line represents correspondingly the inverse photoemission of an additional quasi-particle in a higher orbital.

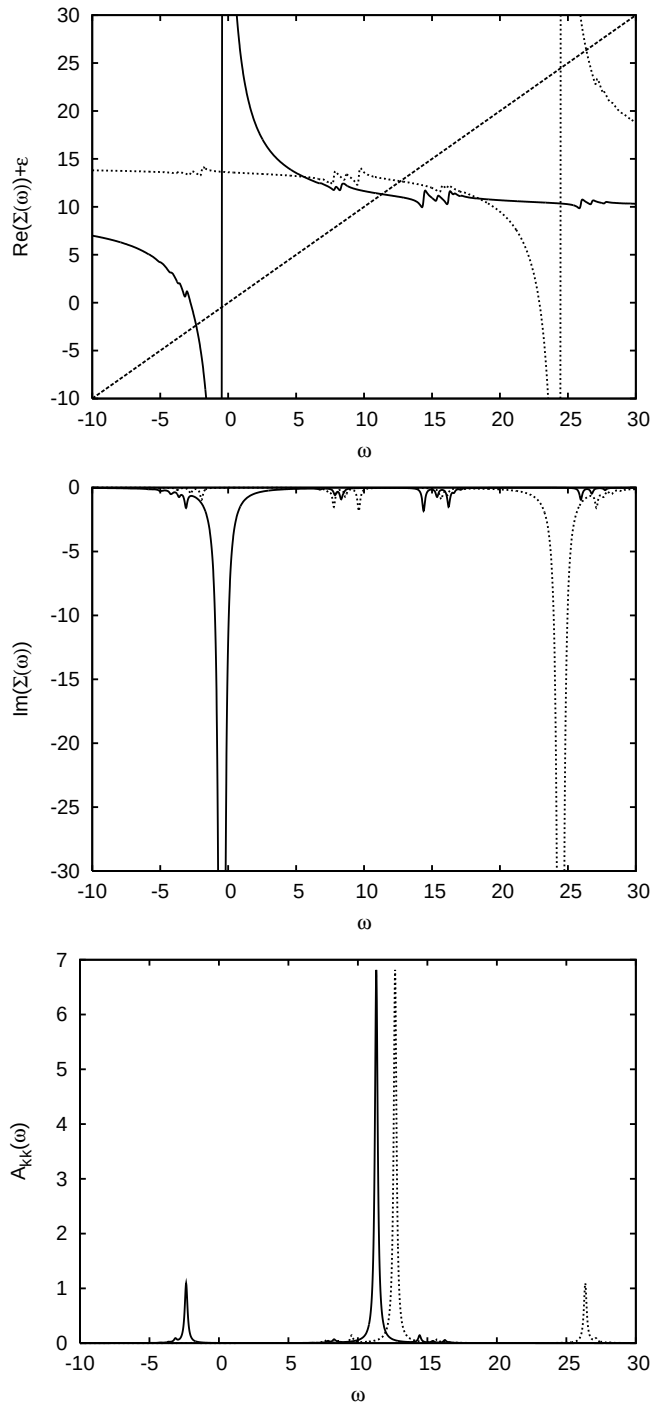


Figure 6.7: Self-energy $\Sigma(k_F, \omega)$ of a metal. For a 10 sites half-filled Hubbard chain with $\Delta = U = 12$. The dashed curves denote the higher, the solid the lower orbitals.

7 Organics

In 1973 Alan Heeger and colleagues at the University of Pennsylvania prepared an ionic salt consisting of two organic compounds, TTF and TCNQ. Containing only carbon, hydrogen, sulfur and nitrogen the salt made out of those two compounds at $-220°$ C has a conductivity comparable to the one of copper at room temperature. It is a metal that contains no metal atoms — a metal-free metal! Moreover it turns out that electron hopping in those materials happens along stacks of like molecules giving rise to nearly one-dimensional bands. This low dimensionality in tandem with strong Coulomb repulsion compared to the kinetic energy leads to many-body effects which can be observed in angular-resolved photoemission spectroscopy (ARPES) experiments. TTF-TCNQ is one of the very few systems where exotic physics like spin-charge separation has been clearly seen experimentally.

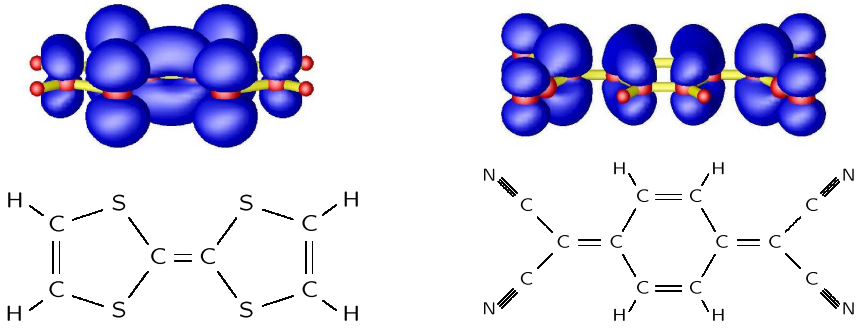


Figure 7.1: The figures show the highest occupied (upper left) and lowest unoccupied (upper right) molecular orbital of TTF and TCNQ respectively. The second row presents the corresponding molecular structures.

With new realistic Coulomb parameters obtained from density-functional theory [44] we will study the spectral function of the molecular metal TTF-TCNQ with the Lanczos technique. Until now only t-U models were employed to understand this metal. Since DFT calculations of the Coulomb parameters give a value of the next-neighbor interaction parameter V which is approximately $U/2$ the effects of V cannot be neglected. We thus study its impact in terms of a t-U-V model on the spectral function. We find that V broadens the bands similar to an enlarged effective hopping. This effect of the next-neighbor interaction V seems to resolve a long-standing puzzle in the theoretical interpretation of ARPES data.

7.1 Charge-transfer salt TTF-TCNQ

Isolated TTF (tetrathiofulvalene) and TCNQ (7,7,8,8-tetracyano-p-quinodimethane) molecules are stable. This is because they are closed shell molecules. The single-particle energy levels are plotted in figure 7.2. We see that the highest occupied molecular orbital (HOMO) of TTF is significantly higher in energy compared to the lowest unoccupied molecular orbital (LUMO) of TCNQ. Both orbitals are shown in figure 7.1. Thus, when making a crystal out of these molecules a charge transfer

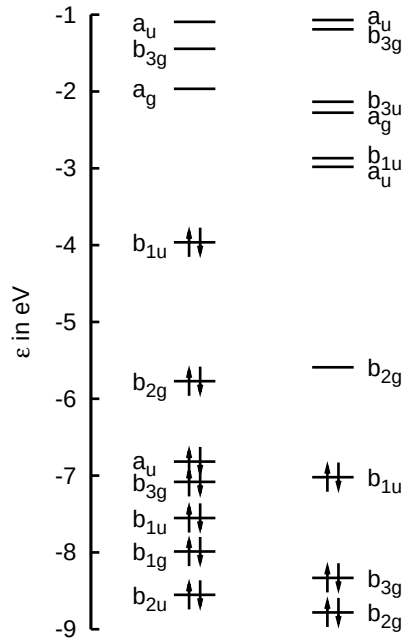


Figure 7.2: Single-particle energy levels of TTF (left) and TCNQ (right) calculated with all-electron DFT for a single molecule.

takes place. 0.6 electrons are, on average, transferred from the TTF-HOMO to the TCNQ-LUMO. If this were not the case the crystal would be a band-insulator since closed shells lead to completely filled bands. The charge transfer, however, leads to partially filled bands and thus metallic behavior.

In figure 7.3 the crystal structure of the salt is shown. The two-dimensional planar molecules of the same type are stacked on top of each other. π-molecular orbitals on adjacent molecules overlap and therefore give rise to the one-dimensional partially filled band. The overlap, however, is relatively weak. It does not lead to a strong covalent bonding but rather to a weak van der Waals-like bonding. Thus, a description of the hopping in terms of a tight-binding approximation seems reasonable.

ARPES experiments [45], however, do not only show ordinary tight-binding bands. Close to the Γ-point the band splits into two branches which is a signature typical of spin-charge separation. This suggests that effects of strong correlations become

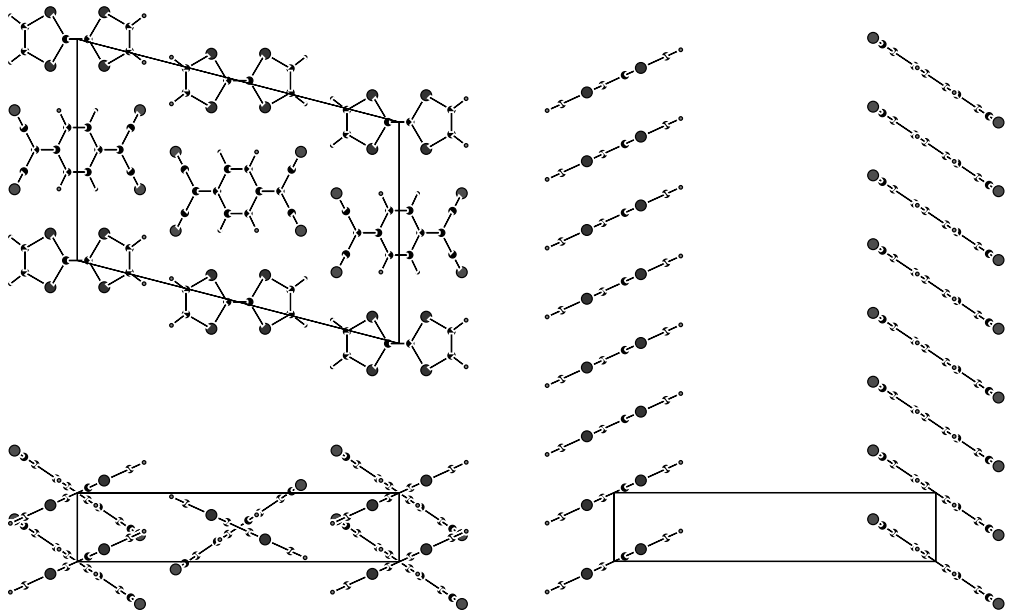


Figure 7.3: Crystal structure of TTF-TCNQ from different perspectives. The molecules are flat and like molecules are stacked on top of each other (right picture). Perpendicular π-molecular orbitals on adjacent molecules overlap and form a one-dimensional band.

important.

7.2 Realistic parameters

Having a kinetic energy part that can be described by the tight-binding approximation the Hubbard model suggests itself for treating the correlation effects. All we need for our calculations are the parameters t, U and possibly some values V_l for the near neighbor interaction. For TTF and TCNQ these were calculated in [44]. The energy levels depicted in figure 7.2 are the results of an all-electron DFT calculation for isolated molecules. In this way also the charge densities $\rho(\mathbf{r})$ for the LUMO and the HOMO are accessible. Since these molecular orbitals (MO) do not change considerably when two molecules come close to each other we can calculate the bare Coulomb matrix elements

$$V_{\text{bare}}(l) = \int\int d^3\mathbf{r} d^3\mathbf{r}' \, \frac{\rho_0(\mathbf{r})\rho_l(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}$$

between two HOMOs (LUMOs) of TTF (TCNQ) a distance l apart from each other from these charge densities. For $l = 0$, V^0_{bare} yields the Hubbard-U_{bare}. This parameter would have to be used if we wanted to employ a Hubbard model for all electrons in all orbitals. Because of the exponential growth of the Hilbert space this is, however, infeasible. We only treat the "conduction" electrons explicitly in our effective

Table 7.1: Realistic Hubbard parameters for TTF and TCNQ from [44]. U, V, V', V'' denote on-site, nearest-neighbor, ... interaction. All energies are in eV.

Stack	U	V	V'	V''
TTF	2.0	1.0	0.55	0.4
TNCQ	1.7	0.9	0.4	0.3

Hubbard model. The effect of the other electrons, which are not explicitly taken into account, is a renormalization of the parameters of the model. The Coulomb parameters for example are decreased due to screening processes. The actual derivation of the parameters is described in [44].

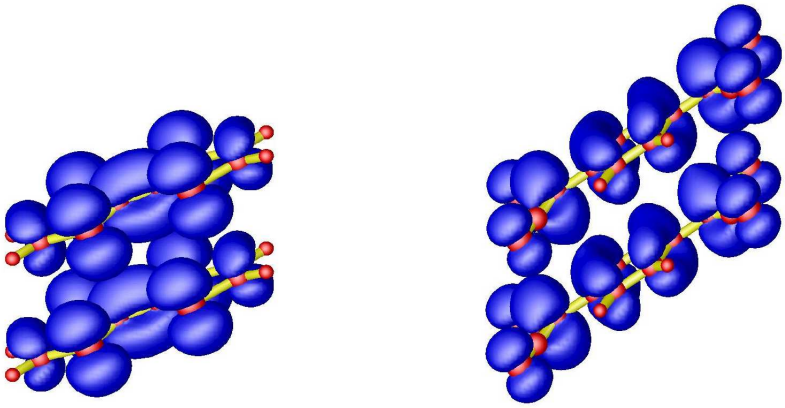


Figure 7.4: Two TTF and TCNQ molecules next to each other. With a DFT calculation for this setup, the hopping matrix element t can be derived from the bonding-antibonding splitting of the molecular orbitals.

With DFT calculations for two molecules (cf. figure 7.4), we get the hopping parameters t from the bonding-antibonding splitting of neighboring molecular orbitals. Hopping occurs along the stack of like molecules. These parameters are, $t_{\mathrm{TCNQ}} \approx 0.18\,\mathrm{eV}$ and $t_{\mathrm{TTF}} \approx -0.15\,\mathrm{eV}$. The realistic parameters for the Hubbard model obtained by this method are shown in table 7.1.

7.3 TTF-TCNQ in the t-U model

Up until now the TTF-TCNQ was mainly described by the ordinary Hubbard t-U-model [45],[46], i.e.

$$H = -\sum_{i \neq j,\nu\sigma} t_{ij,\nu} c^{\dagger}_{i\nu\sigma} c_{j\nu\sigma} + U \sum_{i} n_{i\uparrow} n_{i\downarrow} \,,$$

using parameters based on rough estimates from experiments [47] and theory [48], [49]. Figure 7.5 shows the result of such a calculation for a TCNQ stack. It was

obtained by the CPT technique described in chapter 5.2 for a 20 sites system with 6 electrons of either spin and the parameters $U = 1.96$ eV and $t = 0.4$ eV.

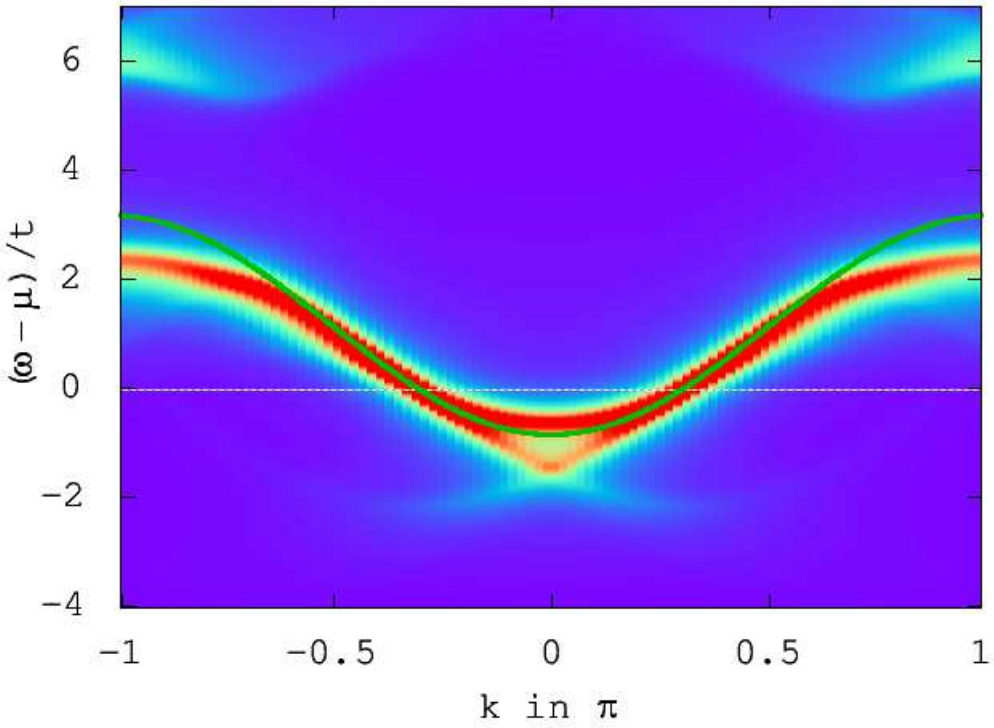


Figure 7.5: Angular-resolved spectral function obtained by CPT for a 20 sites TCNQ system with 6 electrons of either spin ($U = 1.96$ eV, $t = 0.4$ eV). The white line shows the chemical potential, the green cosine represents the independent-particle band. Signatures of spin-charge separation can be observed around the Γ-point. For a discussion refer to text.

As the main feature we observe that the tight-binding band is retained. The dispersion is a bit narrower compared to the cosine curve describing independent-particles. But the Coulomb interaction leads to striking changes due to correlation effects. In the interval $-k_F < k < k_F$, with $k_F/\pi = 0.3$ we observe three dispersing features. Figure 7.6 shows a magnification of this interesting area and since $A_{-k}(\omega) = A_k(\omega)$ we only show the spectral function for $k_F > 0$. Close to the Fermi level which is denoted by the white line at $\omega - \mu = 0$ there are peaks with high weight ranging from $\omega \approx -0.5t$ at $k = 0$ to $\omega \approx 0t$ at $k = k_F$, thus showing a rather narrow

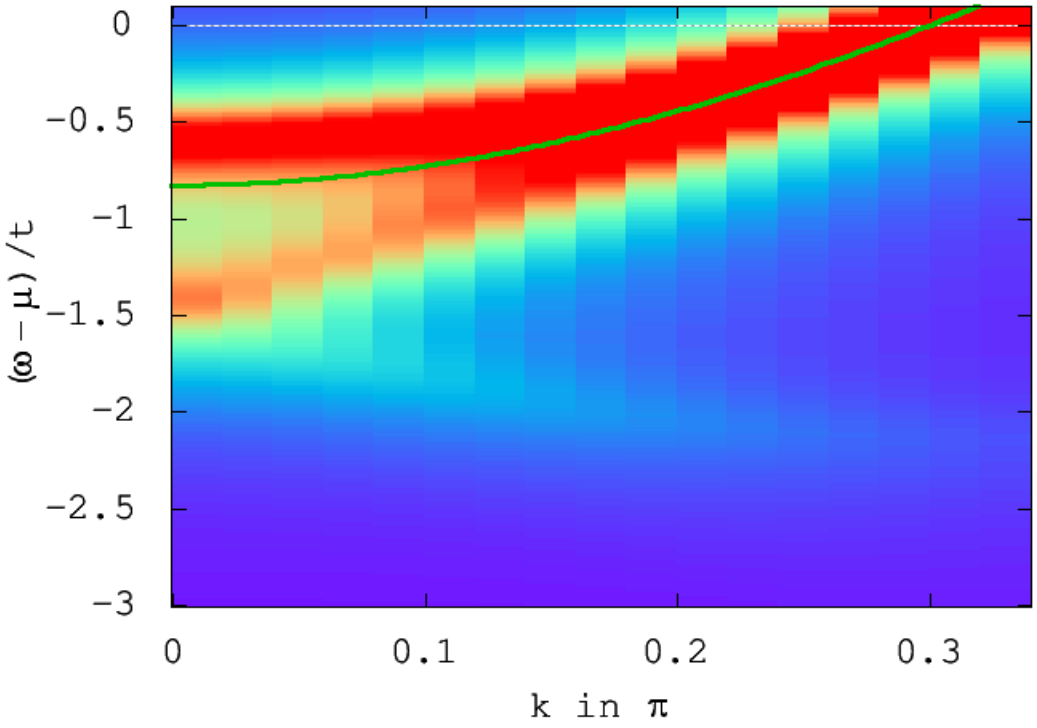


Figure 7.6: Magnification of $A_k(\omega)$ in interval $0 < k < k_F$ of the TCNQ calculation shown in figure 7.5.

dispersion. According to Luttinger liquid theory these spectral weights correspond to the spinon branch. In figure 7.5 we see that there is only a single spinon branch, since the partial branches for $k < 0$ and $k > 0$ join in $k = 0$ with the same slope. At the Γ-point and higher binding energies, i.e. lower values of ω, there seem to start two branches. Figure 7.5 suggests, and from the Bethe ansatz solution [50] we know, however, that the lower so called shadow branch is the continuation of the upper branch for $k < 0$. It runs from $k = 0$ and $\omega \approx -1.5t$ to $k = k_F$ from $\omega \approx -2.2t$ and quickly loses weight with increasing $|k|$. The actual holon band extends from $k = 0$ and $\omega \approx -1.5t$ to $\omega \approx -0t$ at $k = k_F$. It seems to join the spinon branch at k_F. This is, however, because of the finite broadening for plotting since we know from the Bethe-ansatz solution that they intersect at the Fermi level but do not join. The holon and spinon branches have almost constant weight in this region with the spinon branch being considerably more pronounced.

Comparing these results to the results of a DDMRG calculation done by Benthien, Gebhard and Jeckelmann (figure 1 of [46]) shows very good agreement. Moreover it agrees very well to the experimental data. However, in the previous section we calculated the parameter $t = 0.18$ eV. In order to fit the numerical results to the experimental data we have to double the value of t. This has already been taken into account in the above calculation since we used $t = 0.4$ eV.

This ad-hoc change is, however, unsatisfactory. How does this come about? Up until here the next-neighbor interaction term V was neglected. But from [44] we know that $V \approx U/2$, and thus should not be neglected. We will study its effects in the following section.

7.4 TTF-TCNQ in the t-U-V model

7.4.1 Hubbard-Wigner approach

The realistic parameters (cf. table 7.1) show that U and V are considerably larger than the bandwidth $W = 4t$. As an approximation to this case Hubbard suggested in 1978 [48] to use the atomic limit – he calls it zero bandwidth limit – and additionally regard U in a first approximation as infinitely large, since $U > V$. In the following we will mainly look into the TCNQ molecules. Similar results can be obtained for TTF when regarding holes instead of electrons (see particle-hole transformation in chapter D).

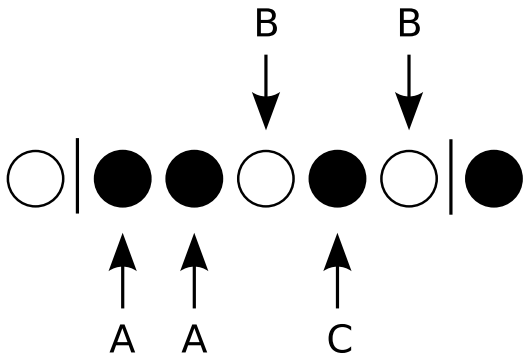


Figure 7.7: Hubbard-Wigner crystal for TCNQ. Empty sites are denoted by an open, singly occupied sites by a filled circle. The crystal has a periodicity of 5 sites. Occupied sites with a single occupied neighbor are named A, occupied sites without occupied neighbor C and free sites B.

Since the electron density in the TCNQ-LUMO is $\rho = 0.6$ there are no double occupancies for $U \to \infty$. In this case the spin of the electron does not play a role anymore and we regard sites only as occupied or unoccupied.

Then the effect of the next-neighbor interaction is that the electrons try not to occupy neighboring sites. Hubbard showed for this case that the many-body

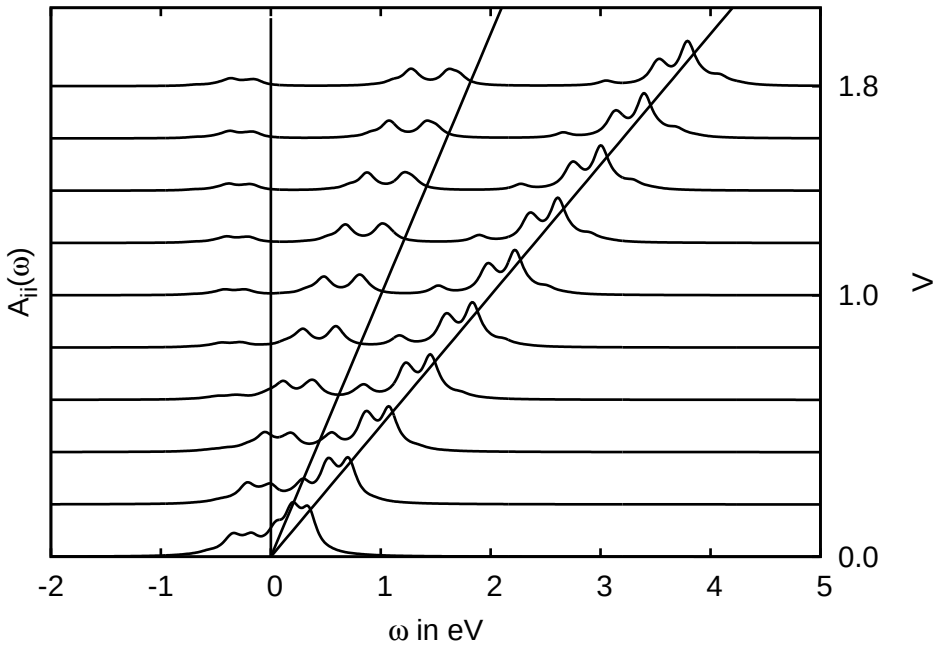


Figure 7.8: Angular-integrated spectral function $A_{ii}(\omega)$ for a 10 sites system with 3 spins of either type $U = \infty$, $t = 0.18$ and varying values of V. The dotted lines show the branches, created by the V interaction. From left to right the energy shifts are proportional to $0, V, 2V$.

ground state can be described in terms of a generalized Wigner lattice [51]. For the case of TCNQ [48] the electrons arrange themselves in a lattice that looks like $\cdots \bullet \circ | \bullet \bullet \circ \bullet \circ | \bullet \bullet \circ \bullet \circ | \bullet \bullet \circ \bullet \circ | \bullet \bullet \circ \cdots$, where $\bullet$ or $\circ$ denote an occupied or unoccupied orbital, respectively. Such a lattice is called *generalized Wigner lattice* with periodicity of 5 sites. The lattice is depicted in figure 7.7

Effect of hopping parameter t Of course t cannot be neglected in our system. So what does actually change when we introduce a finite value of t but keep $U \to \infty$ and V considerably larger than t?

To study such a system we use the Lanczos method. For $U \to \infty$, $t = 0.18$ and varying values of V we calculate the angular-integrated spectral function $A_{ii}(\omega)$. To actually perform the Lanczos calculation for $U = \infty$ we set all amplitudes which have double occupancies to zero. What kinds of processes do we expect? In photoemission-like processes an electron can be removed from an A-site (cf. figure 7.7), leading to a gain in energy of about $-V$. Alternatively, an electron from a C-site can be expelled with hardly any change in the energy at all. Thus, we expect photoemission peaks to be at $\omega = 0$ and $\omega \approx V$, broadened by the hopping bandwidth $W = 4t$. Inverse photoemission processes add one electron to the cell. Since double occupancies are suppressed, the electron can only occupy a B-site, resulting in an energy increase of

about $2V$.

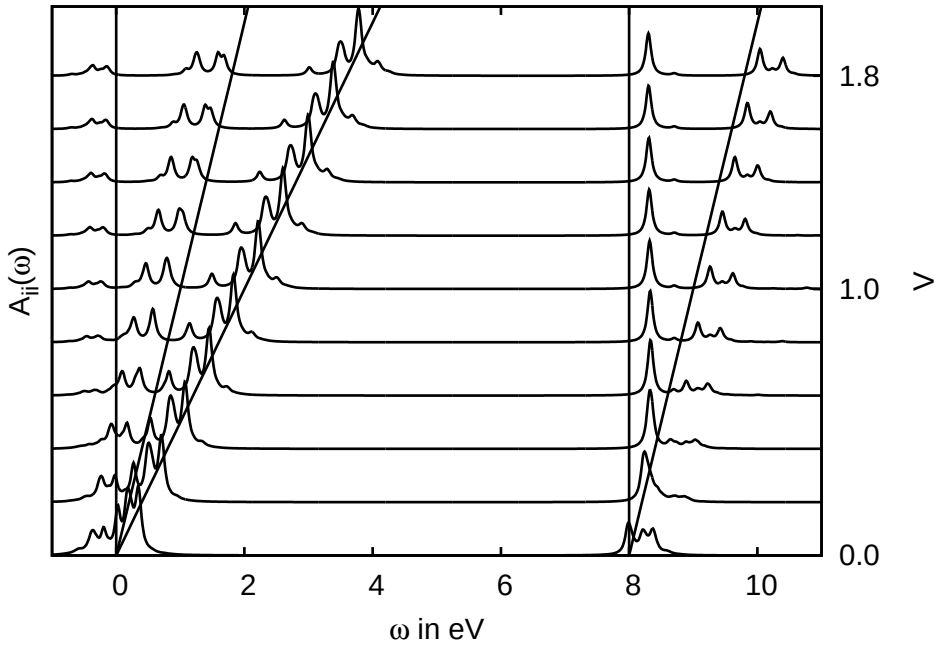


Figure 7.9: Angular-integrated spectral function $A_{ii}(\omega)$ for a 10 sites system with 3 spins of either type, $U = 8$, $t = 0.18$ and varying values of V. The solid lines show the branches, created by the V interaction.

This can be seen from figure 7.8. These processes are denoted by straight lines. The outer left two lines represent the shift in the photoemission peaks and the third one the shift in the inverse photoemission peaks.

Due to $U \to \infty$ the higher Hubbard band cannot be observed. For TCNQ, however, U is finite. Can its physics still be described in the Hubbard-Wigner image for finite but large U? To answer this question we redid the calculation now for a finite value of $U = 8$. The resulting spectral function is shown in figure 7.9. The low-energy features essentially stay the same. The new feature is the upper Hubbard band around U which can still be understood in the Hubbard-Wigner image. Since now double occupancies are allowed electrons can be put into an already occupied site leading to an energy offset of U. If an A-site gets the new electron then additionally V has to be spent, leading to the $U + V$ branch. If a C-site is doubly occupied only U is paid.

Thus, we fully understand this spectrum and system. Does this Hubbard-Wigner description still work for the real parameters that describe TCNQ?

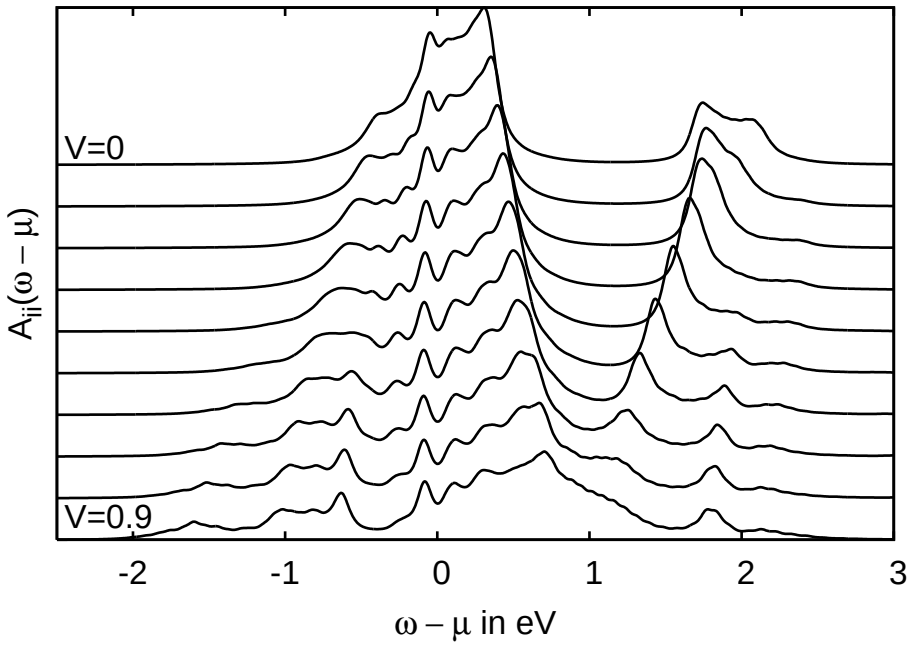


Figure 7.10: Angular-integrated Spectral function for TCNQ in t-U-V model of 20 sites ($U = 1.7$, $t = 0.18$) with varying values of V.

7.4.2 Realistic t-U-V model

Now we perform exactly the same calculation with the proper parameters, namely $U = 1.7$ eV, $t = 0.18$ eV. We vary V to see how the system responds to the next-neighbor interaction. The resulting angular-integrated spectral function can be seen in figure 7.10. It looks qualitatively differently. Introducing V seems to simply broaden the spectrum. This can no longer be understood in terms of the Hubbard-Wigner description.

A key assumption in the Hubbard-Wigner approximation was that U is large enough to completely suppress double occupancies. The question is if this assumption is still valid. In an uncorrelated system the probability of double occupations is given by $d = n_\uparrow \cdot n_\downarrow$. Thus it is $d = 0.09$ in case of TCNQ filling.

For the t-U model discussed above with $U = 1.7$ eV and $t = 0.18$ eV (proper parameters except for $V = 0$) we find that d is still about 10% of the uncorrelated value. Increasing V from 0 to its actual value $V = 0.9$ eV increases the double occupancy d further. This is shown in figure 7.11, where the squares denote the probability of a site being doubly occupied as a function of V. We can understand this behaviour intuitively. Increasing V leads to a decrease of the weight of configurations where neighboring sites are occupied. As a result the weight of the other configurations increases. Since there is not enough space for the electrons to distribute themselves such that the double occupations as well as the next-neighbor occupation are prevented the weight of configurations with double occupancies grows. For

$V = 0.9$ eV the double occupancy has increased to $d = 0.027$. All this suggests that the Hubbard-Wigner approximation is no longer applicable.

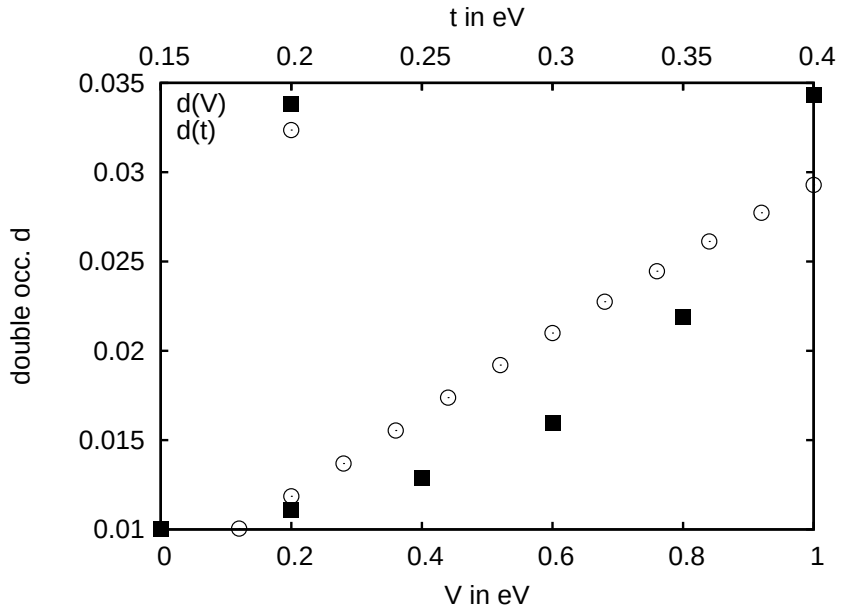


Figure 7.11: Probability of double occupation d as function of V (squares) for fixed $t = 0.18$ and as function of t (circles) for fixed $V = 0$ in a 10 sites system with 3 electrons of either spin ($U = 1.7$).

Figure 7.11 shows d (circles) as a function of the hopping parameter t in the t-U model for the same value of U. We find that d depends almost linearly on the hopping parameter t. This is also intuitively clear since for increasing ratio of kinetic to Coulomb energy the electrons become more and more free particles. It is more important, however, that d looks roughly similar as a function of t and V. To get the same value of $d = 0.027$ of the t-U-V model ($V = 0.9$ eV) in the t-U model we need to double t from $t = 0.18$ eV to about $t = 0.37$ eV. This is probably why t has to be doubled in order to fit experimental results to the t-U-model calculations. The effect of increasing V thus seems to be to encourage hopping which leads to a broadening of the spectrum. An intuitive argument for this behaviour is as follows: consider two electrons, which want to pass each other. At first they have to be neighbors, paying an energy of V. When occupying the same orbital, they pay U but gain V. After passing each other they again pay V. Thus, this process needs the energy $U - V$ to happen. This suggests the introduction of an effective hopping parameter t_{eff} given by

$$t_{\text{eff}} = \frac{U}{U - V} t \; .$$

Figure 7.12 shows d as a function of the ordinary t and effective hopping parameter t_{eff}. Again we observe that for our parameters of TCNQ we need an effective t_{eff} of twice the original value and the dependence of d on V and t is quite similar.

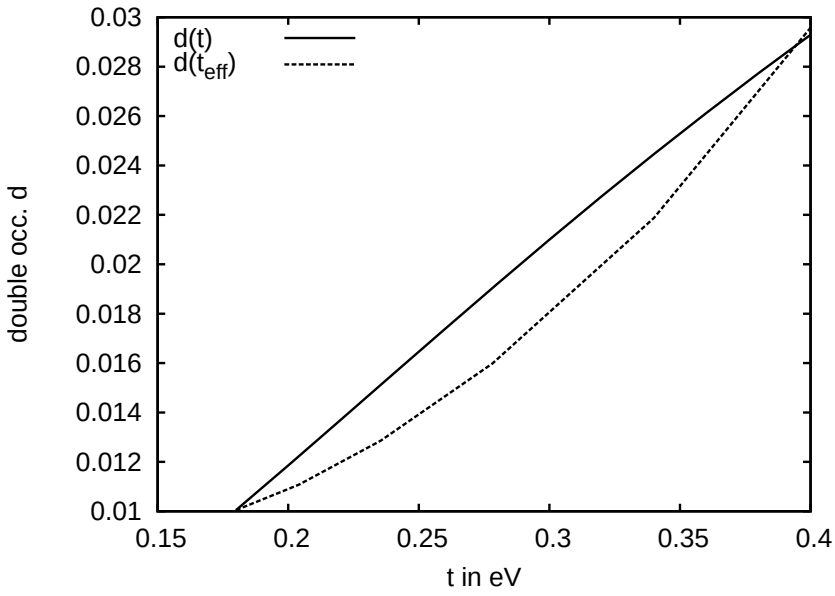


Figure 7.12: Probability of double occupation d as function of effective $t_{\text{eff}} = tU/(U - V)$ and ordinary t hopping parameter.

Plotting the spectral function for different t shows indeed that in this regime the effects of larger t are comparable to the effects of increased V (see figure 7.13). Thus we understand why former calculation in the t-U model yielded good results when t was doubled.

Perturbation theory To further study the effect of V we replot the data of figure 7.10 without centering around the Fermi level. From the resulting figure 7.14 we see that the effect of V is to shift spectral peaks linearly in V. The further the peaks are away from the chemical potential the stronger this shift is, leading to the already described broadening of the spectrum. This is the case in a regime from $V = 0$ until about $V \approx 1.0$ eV, where the chemical potential reaches the upper Hubbard band. For larger V the spectra change qualitatively. This linear behavior can also be observed in the next-neighbor interaction energy part of the ground state. Figure 7.15 shows how the different energy parts of the ground-state energy change when V is increased. And again, up until $V \approx 1$ eV the energy increases linearly in V. Above, the situation changes completely and for $V \approx 1.2$ eV all energy scales become comparable.

The linearity in V suggests a perturbative treatment of the next-neighbor interaction V, i.e. we consider

$$H_V = V \sum_{<ij>} n_i n_j$$

as a perturbation to the t-U model.

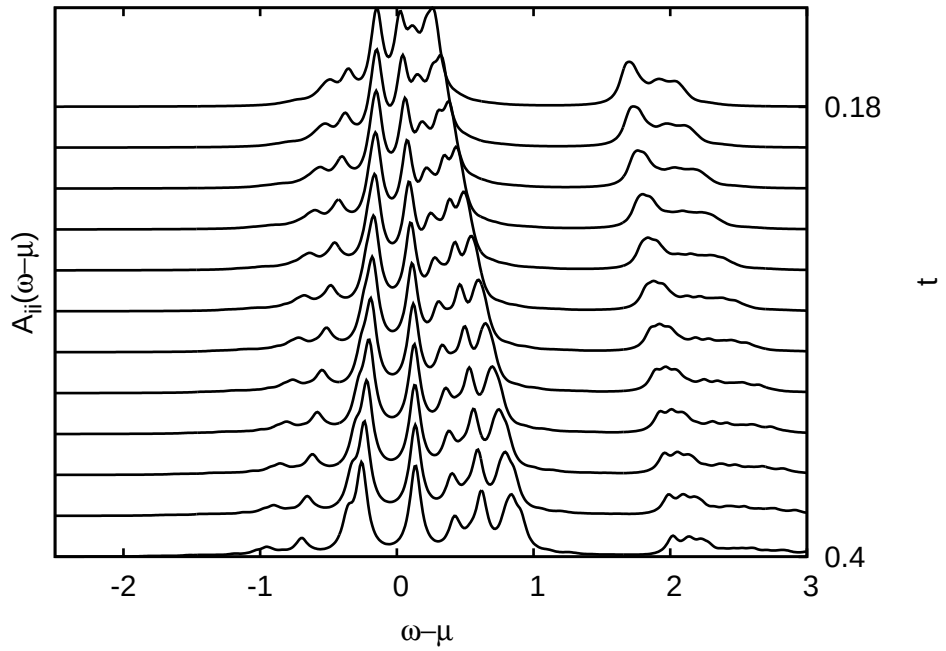


Figure 7.13: Spectral function $A_{ii}(\omega)$ in the t-U model for different values of $t = \{0.18, 0.3, \ldots 0.40\}$ ($U = 1.7$, $V = 0$) and 10 sites with 3 electrons of either spin. The effects of increasing t is a broadening of the band similar to increasing V. Compare to figure 7.9. The system seems to become an insulator, which is not possible, however, in an t-U model away from half-filling. Indeed, it is an effect of finite size (10 sites).

We start with the spectral function in the Lehmann representation (cf. equation (3.29))

$$A_{ii} = \sum_j |c_j|^2 \, \delta\left(\omega - (E_j^{N\pm1} - E_0)\right) \ ,$$

where $c_j = \left\langle \psi_j^{N\pm1} \middle| c_i \middle| \psi_0 \right\rangle$ and $\left|\psi_j^{N\pm1}\right\rangle$ denotes the eigenvector of the Hamiltonian with energy $E_j^{N\pm1}$ in the Hilbert space with one electron added or removed, respectively. To understand how the energies $E_n^{N\pm1}$ change if an infinitesimal V is in effect we calculate the next-neighbor occupation

$$v_n^{N\pm1} = \left\langle \psi_n^{N\pm1} \middle| \sum_{\langle ij \rangle} n_i n_j \middle| \psi_n^{N\pm1} \right\rangle \ .$$

In first order perturbation theory the energy $E_n^{N\pm1}$ is then shifted by $V v_n^{N\pm1}$, while the wave functions remains unchanged.

For relatively small systems (10 sites in this case) the required eigenvectors and eigenenergies can be directly computed using ARPACK. 100 eigenpairs of the spaces

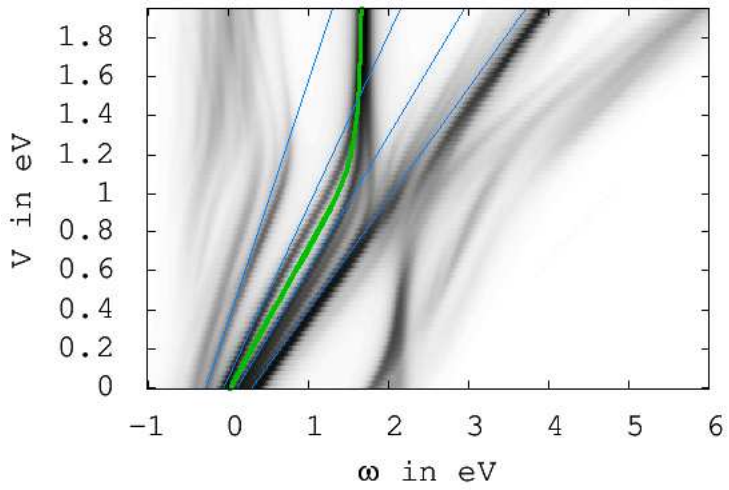


Figure 7.14: Density plot of spectral function for TCNQ ($U = 1.7$, $t = 0.18$) as a function of V of a 10 sites chain, with 3 electrons of either spin type. The green curve denotes the chemical potential. The four blue lines show the shift in the peaks with the largest spectral weight in first-order perturbation theory.

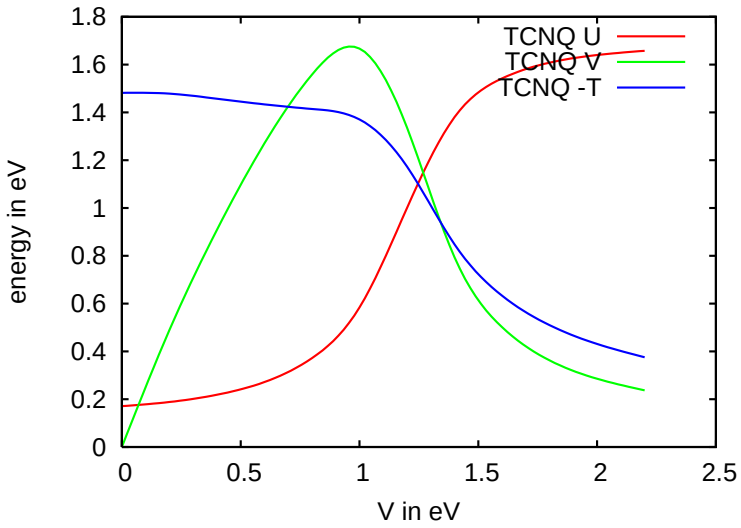


Figure 7.15: Kinetic, local and next-neighbor Coulomb energy for a 10 sites system ($U = 1.8$, $t = 0.18$, $V = 0.9$) for various values of V.

with $N \pm 1$ particles are calculated to obtain $v_n^{N\pm1}$, yielding the energy shifts in this regime. For the largest four spectral weights $|c_j|^2$ these shifts are plotted as blue lines in figure 7.14. The further the peaks are from the Fermi level, the more the slope increases, leading to the broadening. This is because $v_n^{N\pm1}$ tends to increase as a function of the energy in the low-energy regime. Up until $V \approx 1$ eV the broadening can thus be understood in first-order perturbation theory.

7.4.3 Angular-resolved spectral function with CPT and V

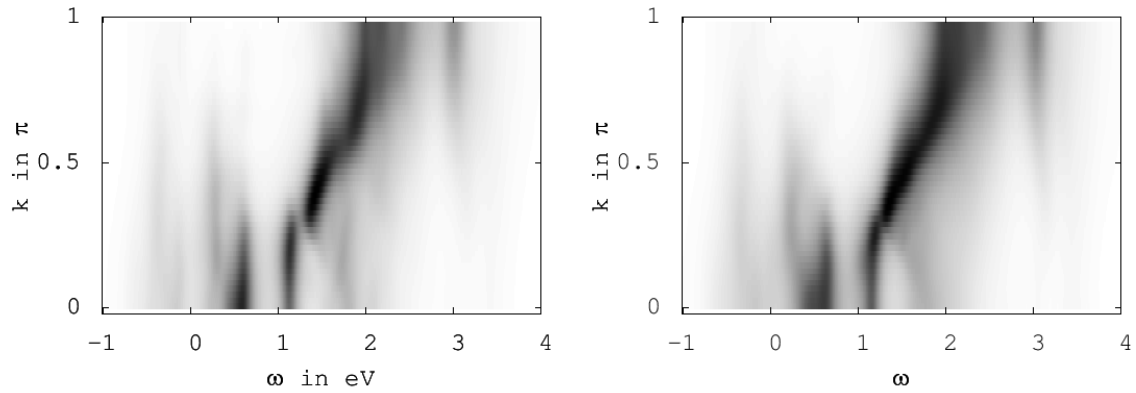


Figure 7.16: CPT calculation of spectral function $A_k(\omega)$ with next-neighbor interaction $V = 0.9$ taken into account in PBC. System parameters are 10/20 sites (left/right) with $t = 0.18$, $U = 1.7$.

It would definitely be useful to have angular-resolved spectral functions of these compounds. With CPT we have means to calculate these functions. But as we have already seen in chapter 5.2.7 we face considerable finite-size effects for the t-U-V model. In case of half-filling they could be effectively scaled out with increasing system size. Here however, we consider systems away from half-filling. Figure 7.16 shows, that finite-size effects remain strong even for 20 sites.

7.4.4 TTF and particle-hole symmetry

The spectral functions of TTF and TCNQ are intimately connected. The charge-transfer puts 0.6 electrons on average to the TCNQ LUMO, creating in turn 0.6 holes in the TTF HOMO. Since the parameter $t = -0.15$, $U = 2.0$ and $V = 1.0$ are quite similar compared to the ones of TCNQ we can obtain approximately the spectral function of TTF having the one of TCNQ (see chapter D).

Figure 7.17 shows the actual spectral function of the TTF system in the t-U model. The photoemission part ($\omega < \mu$) of the system looks as if it is uncorrelated, i.e. we

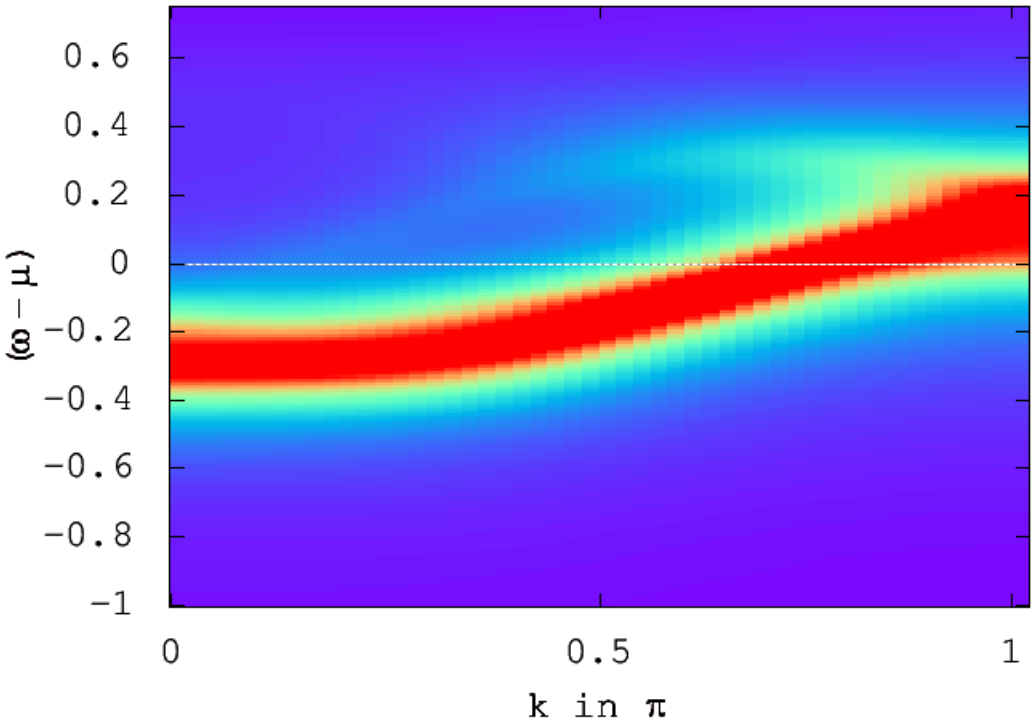


Figure 7.17: Angular-resolved spectral function of TTF for 10 sites with $U = 2.0$ eV and $t = 0.15$ eV.

observe a slightly narrowed tight-binding band. This coincides with experiments. The correlation effects are obvious only in the inverse photoemission part of the spectral function, which is unfortunately not accessible by inverse photoemission spectroscopy, since the electrons impinging on the material would essentially destroy it.

In theory, however, it can be calculated and we see that because of the larger values of U and V as well as the smaller t, TTF is even slightly more correlated than TCNQ.

7.5 Accuracy of the results

In this section we will briefly study the impact of finite-size effects. Figure 7.18 compares the angular-integrated spectral functions for the realistic parameters of

TCNQ for 10 and 20 sites. We see that all major features are at the correct position even in the 10 sites calculation. For larger values of V, however, finite size effects become more pronounced.

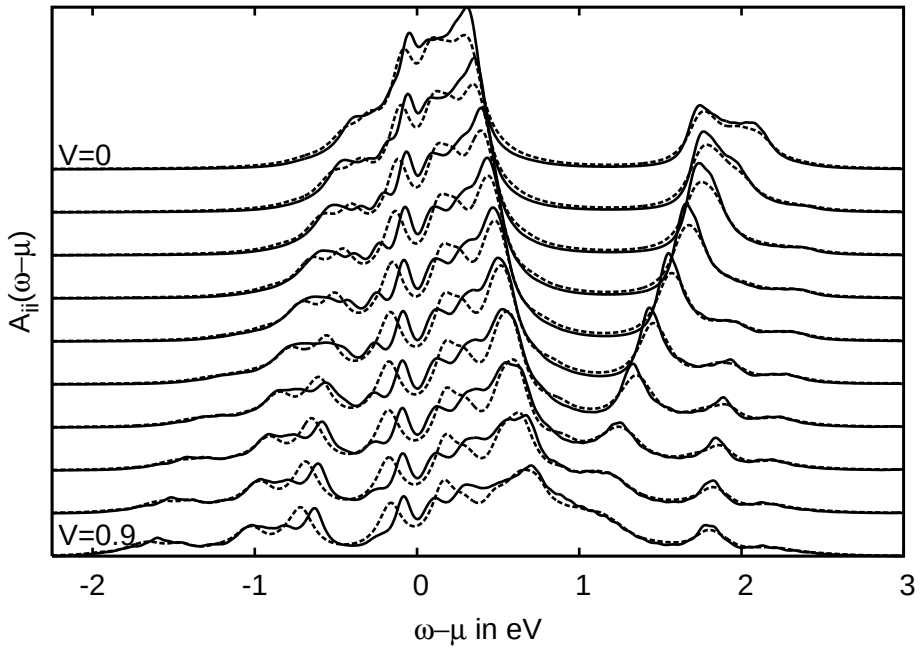


Figure 7.18: Finite size effects: Spectral function of TCNQ for $U = 1.7$, $t = 0.18$ and various values of $V = \{0, 0.1, \ldots, 0.9\}$ on a 10 (dashed) and 20 (solid) sites chain, with 3 and 6 electrons of either spin, respectively.

Figure 7.19 shows how the results depend on the choice of boundary conditions. Periodic and anti-periodic boundary conditions are pictured. Hardly any deviations can be noticed.

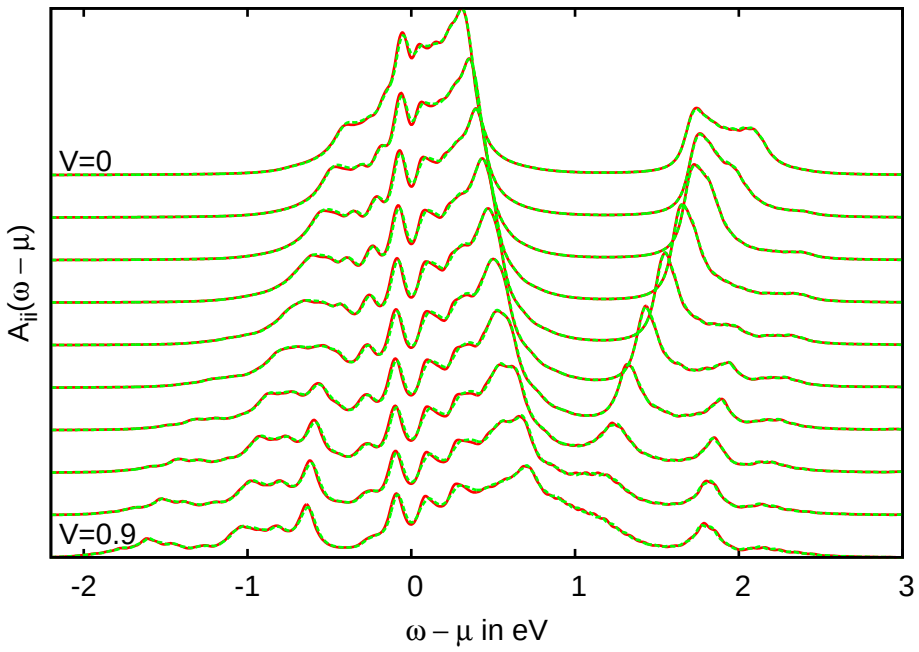


Figure 7.19: Angular-integrated spectral function for TCNQ on a 20 sites system with periodic (full) and anti-periodic boundary (dashed) conditions ($U = 1.7$, $t = 0.18$) and $V = \{0, 0.1, \ldots, 0.9\}$ (upper to lower).

8 Summary

The field of strongly correlated electrons yields fascinating physics. This is due the interplay between Coulomb repulsion and kinetic energy. The prototype to study systems of strongly correlated Fermions is the Hubbard model. Though looking formally quite simple it describes spectacular physics. In one-dimension it properly describes Luttinger liquid effects like spin-charge separation whose signatures can also be found in the metal-free metal TTF-TCNQ.

Traditional techniques of solid state theory cease to work in such systems since they fail to incorporate the Coulomb interaction perturbatively as a pair interaction. Therefore non-perturbative methods are needed. Among those techniques is the Lanczos method, which solves these kind of systems exactly. This method is fast and gives access to arbitrary many-body expectation values and dynamical response functions. However, due to the exponential increase in the Hilbert space for growing physical system sizes it is limited to relatively small systems. In this work we studied extended Hubbard models. On ordinary computers – without exploiting further symmetries except spin- and charge- conservation – half-filled systems up to 14 sites are accessible. For many applications, however, these systems are too small, either because of filling constraints or because of finite-size scaling. We thus developed a C++ code from scratch which is not only as efficient as a comparable FORTRAN shared-memory Lanczos code but also more flexible. Due to template meta-programming techniques we have a unified code for complex and real Hamiltonians and wave functions in single and double precision, making systems with complex boundary conditions treatable. Moreover this code can exploit the latest supercomputer architectures of shared and distributed memory environments within a unified programming interface.

Most of our calculations were performed on the massively parallel BlueGene/L system in Jülich, called JUBL, comprising 16 384 processors. Due to the distributed memory of such systems the matrix-vector multiplication posed the main challenge for the efficient implementation, since it requires non-local memory access. The realisation that the Lanczos vectors can be decomposed into a matrix whose indices denote the up/down electron configurations lead to the idea of performing a matrix transposition [52] on these vectors in order to get the needed elements local in memory. It turned out that such an approach is indeed very efficient and offers a means to access large systems with reasonable wall-clock times. Thus, we can efficiently compute expectation values of arbitrary observables as well as dynamical response functions.

With this code we treated systems with two orbitals per unit cell which show an interesting Mott-band insulator transitions. Kohn's criterium which determines

the phase of a system by ground-state properties alone, namely the response of the ground-state energies to a vector potential, was employed. This vector potential can be equivalently regarded as a change in boundary conditions, requiring complex wave vectors. We found that half-filled systems become metallic when going over from a band to a Mott insulating phase.

The systems that can be treated with the Lanczos method are finite. To extrapolate to infinite system sizes and getting access to an arbitrarily high resolution in k-space, for instance to compute angular-resolved spectral functions, we use cluster perturbation theory (CPT). It works by diagonalizing a finite cluster exactly and then treating the hopping between identical cluster perturbatively in strong coupling perturbation theory. This leads to an effectively infinite chain. Due to our efficient Lanczos implementation it is possible to calculate relatively large clusters, minimizing finite-size effects.

With our code and the CPT technique we investigated the organic metal TTF-TCNQ and were able to solve a long-standing problem in the interpretation of its experimental spectra. Because it is a one-dimensional metal TTF-TCNQ has been studied thoroughly for more than 30 years. Theoretical estimates as well as experimental measurements suggest values for the parameters of the Hubbard model t and U. With these values Hubbard model calculations show qualitatively the same features as APRES experiments, namely the signatures of spin-charge separation at the Γ-point. But in order to fit the experimental data to the numerical calculations t has to be doubled, which translates into a broadening of the spectra. This ad-hoc adjustment is not only unsatisfactory but also the temperature dependence of the spectral function is incorrectly described. We resolved this problem [44] by properly including nearest-neighbor Coulomb interactions.

Recent DFT calculations show that indeed the commonly used values of t and U are correct. But Coulomb interactions between electrons on neighboring sites are also significant and must not be neglected. Thus we studied the effect of the next-neighbor interaction parameter V having a value of about $U/2$. Including V in our Lanczos calculations for obtaining the angular-integrated spectral function shows that its effect is to broaden the spectra, comparable to increasing the value of t. Moreover we found that doubling t indeed mimics the effect of $V \approx U/2$. Surprisingly, the broadening of the spectrum as a result of V can be understood in first-order Rayleigh-Schrödinger perturbation theory, as we showed with calculations of our code.

A Speed up and Amdahl's law

Why multiprocessing? We expect programs to run faster, i.e. have less total run time, if several processors work in parallel on the same problem.

Naïvely one would expect that doubling the number of processors cuts the execution time by a factor of two. This is, however, usually not the case because of (a) communication and management overhead and – more importantly – (b) inherently sequential code. Inherently sequential code is code that cannot be parallelized. In order to judge if an algorithm scales well on more than a single processor we define the speed up $S(p)$ as

$$S(p) = \frac{T_1}{T_p} \ , \tag{A.1}$$

where T_1 denotes the execution time on a single and T_p on p processors.

A speed up of $S(p) = p$ is called linear or ideal speed up and corresponds to the naïve expectation.

Let us assume a code with an inherently sequential part of γ and with a parallel part $1 - \gamma$ that can be fully parallelized. Increasing the number of processors cuts the run-time in the parallel part but the sequential part does not profit at all. Then, for an infinite number of processors the wall clock time of the parallel part vanishes and only the sequential part remains. Thus, the maximum achievable speed up is bounded by $1/\gamma$. This is known as Amdahl's law [53].

Formally this law can be derived in the following way. Let the total run-time (wall clock) on p processors be T_p^{tot}. Let T_p^{par} and T^{seq} denote the run-time for the parallel and sequential part respectively. The total run-time can be decomposed as

$$T_p^{\text{tot}} = T_p^{\text{par}} + T^{\text{seq}} = \frac{T_1^{\text{par}}}{S_p^{\text{par}}} + T^{\text{seq}} \geq \frac{T_1^{\text{par}}}{p} + T^{\text{seq}} \ .$$

With

$$\gamma = \frac{T^{\text{seq}}}{T_1^{\text{tot}}} \ \in [0, 1]$$

and

$$T_1^{\text{par}} = (1 - \gamma) T_1^{\text{tot}} \ ,$$

this leads to Amdahl's law,

$$S_p \leq \frac{T_1^{\text{tot}}}{T_1^{\text{tot}}(\frac{1-\gamma}{p} + \gamma)} = \frac{1}{(1-\gamma)/p + \gamma}, \tag{A.2}$$

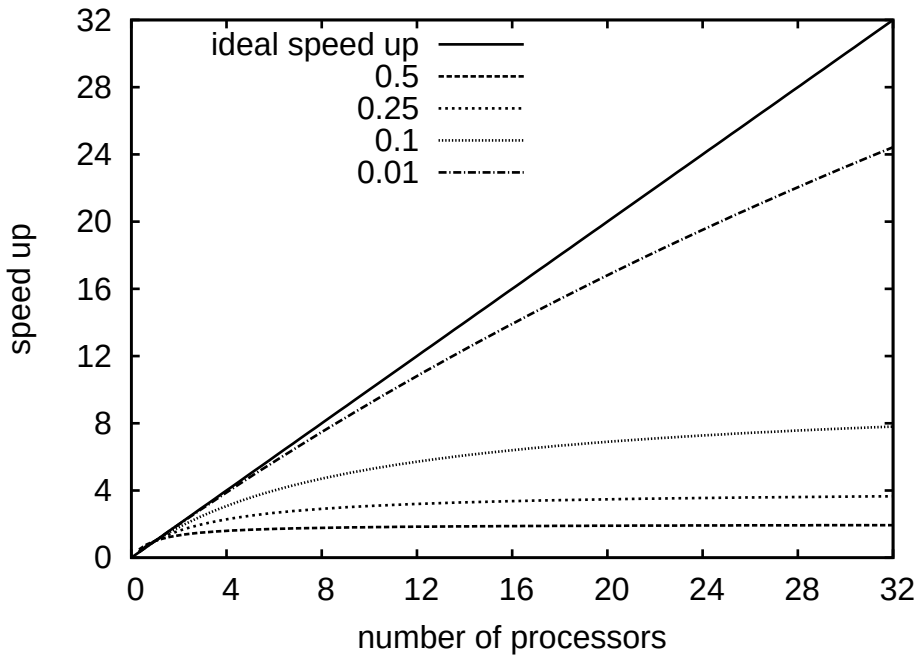


Figure A.1: Typical "look" of Amdahl's law. The plot shows the function (A.2) for $\gamma = \{0.01, 0.1, 0.25, 0.5\}$ and processors from 1 to 32.

which is presented for some values of γ in figure A.1. In the limit of infinitely many processors we retain

$$S_\infty \leq \frac{1}{\gamma} \, .$$

For instance a code with 1% serial code will never gain a speed up of more than 100 irrespective of the number of processors. To see, how efficiently the processors are used, the efficiency is defined as

$$E_p = \frac{S_p}{p} \, .$$

Hence, using this code with 1000 processors would yield an efficiency below 0.1, which would be waste of computational resources.

In our code the ratio of the sequential setup time to the time of the parallelizable Lanczos calculations decreases with the system size, i.e. γ becomes smaller the larger the system. Thus solving large systems many processors can be used efficiently.

B Supercomputers at the Forschungszentrum Jülich

B.1 JUMP

JUMP is a general purpose system. 32 processors on each node share their memory and can be efficiently used with OpenMP. As we have seen in chapter 4 OpenMP is a very convenient way of parallelizing a given serial code. Computing on JUMP is, however, relatively expensive [54]. One wall-clock hour with p processors costs

$$C = p \cdot 1.90\,€ \quad . \tag{B.1}$$

For instance, a CPT calculation to obtain the angular-resolved spectral function $A_k^{\mathrm{PE}}(\omega)$ (photoemission only!!) of TCNQ (20 sites) would take almost 9 days on 32 processors and thus cost about 13 000 €.

B.1.1 Architecture

The JUMP cluster consists of 41 SMP (Symmetric MultiProcessing) nodes, called IBM p690 frame, with 32 processors respectively (peak performance: 218 GFLOPS). This amounts to 1312 processors in total (overall peak performance: 8.9 TFLOPS). Each processor is a Power4+ CPU clocked at 1.7 GHz and features a L1 cache with 64 KB for instructions and 32 KB for data. The L2 cache (1.5 MB, 10-12 cycles) is shared between two processors and the L3 cache of 512 MB (92-100 cycles) is shared frame-wide. The total shared main memory of a frame is 128 GB and therefore the aggregated total main memory of the whole cluster is 5.2 TB. For more information refer to `jumpdoc.fz-juelich.de`. The interframe communication is managed by IBM's High Performance Switch (HPS). It has a MPI bandwidth of about 1.6 GB/s per link and a latency of about 5.6 μs. For more information concerning the interconnect refer to [55] and [56].

B.2 JUBL

JUBL is the new massively parallel supercomputer in Juelich, which belongs to the BlueGene family. Target for the development of the BlueGene system was high performance/price and performance/power consumption ratios. To put this into practice IBM took a new approach in supercomputing. The main idea is to build the

system out of a very large number of nodes, which taken for themselves have only a relative modest clock rate and performance, leading to lower power consumption and low cost. Therefore the nodes (2 processors) can be packed very densely. Instead of having only 32 processors in a rack, a single BlueGene rack comprises 2048 processors. Using relatively slow processors also implies a better balance between CPU speed and the considerably slower memory and network access. Moreover the memory-bus bottleneck of shared memory systems plays no role in two processor nodes.

On JUBL accounting works differently compared to JUMP. One has to reserve a partition and is charged for the reservation time not for the actual run-time of the program. A reserved hour of a partition comprising p processors on JUBL costs

$$C = p \cdot 0.03 \text{ €} . \tag{B.2}$$

On JUBL the job of calculating a 20 sites TCNQ system to obtain the angular-resolved spectral function (photoemission only) takes 16 hours and costs about 1000 Euro which is considerably faster and cheaper than on JUMP. The calculation of the full spectral function is feasible and takes slightly over two days.

B.2.1 Architecture

JUBL consists of 8 racks hosting each (2 x 16) x 32 compute nodes, which actually contain dual processors, thus leading to 16 384 processors. Those processors are 32-bit PowerPC 440 cores at 700 MHz, leading to an overall theoretical peak performance of 45.8 Teraflops. Each compute node has 512 MB of RAM which is shared between the two processors (aggregated 4.1 TB). The compute nodes can be used in CO (coprocessor) mode, which means that only one processor performs the computations, while the other one performs communication and I/O. A second mode is the VN (virtual node) mode, meaning that both processors compute, communicate and perform I/O operations. In this case each processor is a single virtual node. A third mode is the so called mixed mode. It is unfortunately not yet documented. In this mode, a single node can be regarded as a small shared memory dual-processor system.

B.3 Run-time restriction

JUMP and JUBL, however, share the restriction that jobs cannot run longer than 24 h. Therefore one has to carefully estimate the execution time and split the jobs accordingly. Alternatively one needs check-pointing. In our case check-pointing is relatively simple. For the Lanczos passes we just need to dump the current Lanczos vector. When the calculation is resumed we can restart the iterations with the dumped vector.

C Evaluation of continued fractions

In chapter 3.4.2 we discussed how to compute spectral functions by using the Lanczos method. The intermediate results of the third Lanczos pass, namely the tridiagonal matrix elements a_i, b_i, define a continued fraction, which rapidly converges to the spectral function. Here, we briefly discuss how to evaluate such fractions.

A continued fraction is a fraction of the form

$$f = b_0 + \frac{a_1}{b_1 + \frac{a_2}{b_2 + \frac{a_3}{b_3 + \ldots}}} \; . \tag{C.1}$$

An equivalent representation often found in print is

$$f = b_0 + \frac{a_1}{b_1 +} \; \frac{a_2}{b_2 +} \; \frac{a_3}{b_3 +} \; \ldots \; .$$

Continued fractions are often superior to series, since they typically converge more rapidly. They are, however, harder to evaluate. Computations of power series are usually stopped, when the next contribution is sufficiently small. A naive approach to evaluate a continued fraction is to guess a starting point, i.e. some a_n and b_n, and start the evaluation from there to the left. If the result is not sufficiently well converged a new evaluation has to be started from an a_n and b_n further right. This apparently is not an efficient solution. In 1655 J. Wallis developed a method which allows the evaluation from left to right. Let f_n denote the approximation to f by computing the fraction through to coefficients a_n and b_n. We can write f_n as

$$f_n = \frac{A_n}{B_n}$$

with the following recurrence relations for A_n and B_n and $j \in [1, n]$

$$\begin{aligned} A_{-1} &= 1 && \text{(C.2)} \\ B_{-1} &= 0 && \text{(C.3)} \\ A_j &= b_j A_{j-1} + a_j A_{j-2} && \text{(C.4)} \\ B_j &= b_j B_{j-1} + a_j B_{j-2} \; , && \text{(C.5)} \end{aligned}$$

which can be easily proven by induction.

This method has, however, some numerical problems. It often generates very large and very small numbers which lead to over- and/or underflows. To remedy this problem, renormalization methods can be used. According to [28] the best general method is the modified Lentz's method. Listing (C.1) shows an example implementation of the modified Lentz method is Python.

```
def calc_continued_fraction(a,b):
    if abs(b[0])==0:
        f = [1e-25,]
    else:
        f = [b[0],]
    C = [f[0],]
    D = [0.,]
    delta=[0.,]
    for j in range(1,len(a)):
        D.append(b[j]+a[j]*D[j-1])
        if abs(D[j]) == 0:
            D[j]=1e-25
        C.append(b[j]+a[j]/C[j-1])
        if abs(C[j]) == 0:
            C[j]=1e-25
        D[j]=1./D[j]
        delta.append(C[j]*D[j])
        f.append(f[j-1]*delta[j])
    return f[-1]
```

Listing C.1: Python implementation of the modified Lentz method to evaluate continued fractions. It evaluates a continued fraction: $b_0 + (a_0/b_1+)(a_1/b_2+)$ ...

D Particle - hole symmetry

Bipartite lattices, i.e. lattices that can be decomposed into two identical sub-lattices A and B such that the nearest-neighbor always belongs to the other lattice, give rise to an additional symmetry, namely the particle-hole symmetry or, in our, case electron-hole symmetry. Loosely following [57], the particle-hole transformation is given by

$$c_{i,\sigma} = \begin{cases} +d^\dagger_{\sigma,i} & : \quad i \in A \\ -d^\dagger_{\sigma,i} & : \quad i \in B \end{cases} \quad . \tag{D.1}$$

Under this transformation the Hamiltonian H^{el} becomes, as one can easily verify,

$$\begin{aligned} H^{\mathrm{el}} &= -\sum_{\langle i,j\rangle,\sigma} t_{ij} c^\dagger_{i,\sigma} c_{j,\sigma} + U \sum_i n_{i,\uparrow} n_{i,\downarrow} + V \sum_i n_i n_{i+1} && \text{(D.2)} \\ &= H^{\mathrm{hole}} + (U + 4L)(L - N) \; , && \text{(D.3)} \end{aligned}$$

where $N = N_\uparrow + N_\downarrow$ is the total number of particles of either spin and L, as usual, the number of sites. For half-filling, i.e. $L = N$, we find $H^{\mathrm{el}} = H^{\mathrm{hole}}$.

For identical parameter t, U, V we can thus map spectra of systems with $N_\uparrow$, $N_\downarrow$ and $N^h_\uparrow = L - N_\uparrow$, $N^h_\downarrow = L - N_\downarrow$ onto each other. This can also be done for the spectral functions. The energies have to be shifted according to (D.3) and inverse photoemission part has to be exchanged with the photoemission part, because ejecting an electron is equivalent to creating a hole and vice versa.

Acknowledgements

Every beginning is hard, so was writing the first scientific work, my diploma thesis. Thanks to the assistance of many helping brains and hands, it is finally accomplished.
First of all I would like to thank my supervisor Dr. Erik Koch for the introduction into the field of strongly correlated systems and who never got out of patience when answering my innumerable questions. Moreover his proofreading of my thesis as well as his constructive comments, suggestions and corrections were helpful. So was the pointing out of apt references.
I am grateful to Prof. Dr. Stefan Blügel who offered me the chance of writing my thesis in the research center Jülich and who took over the "Erstgutachten" of my thesis.
Big thanks go to Dipl. Phys. Manfred Niesert, who strenuously scrutinized my thesis and helped to eliminate typos as well as linguistic weaknesses. Furthermore he proved to be a great support.
My thanks also go to Swantje Heers, Gerrit Bette and Katharina Peter for proofreading.
Finally I'd like to thank the members of the institute "Theory 1" in the research center Jülich, above all my room mates Andreas Gierlich, Swantje Heers and Manfred Niesert, and our nearest neighbor Dr. Christoph Friedrich for creating the excellent working atmosphere and helpful environment. You made even the work enjoyable.
I would like to acknowledge research grant JIFF22 under which all supercomputer calculations were performed.

Bibliography

[1] J. Hubbard.
Electron Correlations in Narrow Energy Bands.
Royal Society of London Proceedings Series A, 276:238–257, November 1963.

[2] J. Hubbard.
Electron Correlations in Narrow Energy Bands. II. The Degenerate Band Case.
Royal Society of London Proceedings Series A, 277:237–259, January 1964.

[3] J. Hubbard.
Electron Correlations in Narrow Energy Bands. III. An Improved Solution.
Royal Society of London Proceedings Series A, 281:401–419, September 1964.

[4] Martin C. Gutzwiller.
Effect of Correlation on the Ferromagnetism of Transition Metals.
Phys. Rev. Lett., 10(5):159–162, Mar 1963.

[5] Junjiro Kanamori.
Electron Correlation and Ferromagnetism of Transition Metals.
Prog. Theor. Phys., 30:275, 1963.

[6] E. H. Lieb and F. Y. Wu.
Absence of Mott Transition in an Exact Solution of the Short-Range, One-Band Model in One Dimension.
Physical Review Letters, 20:1445–1448, June 1968.

[7] N. W. Ashcroft and N. D. Mermin.
Solid State Physics.
ITPS Thomson Learning, 1988.

[8] Erik Koch and Stefan Goedecker.
Locality properties and Wannier functions for interacting systems.
Solid State Communications, 119:105, 2001.

[9] G.Nenciu.
Existence of the exponentially localised Wannier functions.
Commun. Math. Phys., 91:81–85, 1983.

[10] Jacques Des Cloizeaux.
Energy Bands and Projection Operators in a Crystal: Analytic and Asymptotic Properties.
Phys. Rev., 135(3A):A685–A697, Aug 1964.

[11] Nicola Marzari and David Vanderbilt.

Maximally localized generalized Wannier functions for composite energy bands. *Phys. Rev. B*, 56(20):12847–12865, Nov 1997.

[12] Stefan Bluegel, Gerhard Gompper, Erik Koch, Heiner Müller-Krumbhaar, Robert Spatschek, and Roland Winkler, editors.
Computational Condensed Matter Physics.
Number 37. Schriften des FZ Juelich, 2006.

[13] W. Kohn and L. J. Sham.
Self-Consistent Equations Including Exchange and Correlation Effects.
Phys. Rev., 140(4A):A1133–A1138, Nov 1965.

[14] Assa Auerbach.
Interacting Electrons and Quantum Magnetism.
Springer-Verlag, 1994.

[15] A. J. Millis and S. N. Coppersmith.
Variational wave functions and the Mott transition.
Phys. Rev. B, 43(16):13770–13773, Jun 1991.

[16] W. F. Brinkman and T. M. Rice.
Application of Gutzwiller's Variational Method to the Metal-Insulator Transition.
Phys. Rev. B, 2(10):4302–4304, Nov 1970.

[17] Peter Fulde.
Electron Correlations in Molecules and Solids.
Springer, 1995.

[18] J. Kleinberg.
Authoritative sources in a hyperlinked environment.
Proc. Ninth Ann. ACM-SIAM Symp. Discrete Algorithms, pages 668–677, 1998.

[19] Sergey Brin and Lawrence Page.
The anatomy of a large-scale hypertextual Web search engine.
Proceedings of the seventh international conference on World Wide Web 7, pages 107–117, 1998.

[20] H. Q. Lin and J. E. Gubernatis.
Exact diagonalization methods for quantum systems.
Computers in Physics, 7(4):400, 1993.

[21] G.W. Stewart.
Afternotes Goes to Graduate School: Lectures on Advanced Numerical Analysis.
SIAM, 1998.

[22] E. Dagotto.
Correlated electrons in high-temperature superconductors.
Reviews of Modern Physics, 66:763–840, July 1994.

[23] C. C. Paige.
Error analysis of the Lanczos algorithm for tridiagonalizing a symmetric matrix.

J. Inst. Maths Applics, 18:341–349, 1976.

[24] Willoughby, Ralph and Cullum, Jane K.
Lanczos Algorithms for Large Symmetric Eigenvalue Problems, Volume 2.
Birkhauser-Boston, 1984.

[25] Golub, G. H. and Van Loan, C. F.
Matrix Computations.
Johns Hopkins University Press, Baltimore and London, 3^{nd} edition, 1996.

[26] B. Parlett and D. Scott.
The Lanczos algorithm with selective orthogonalization.
Math. Comp., 33:217–238, 1979.

[27] V.M. Galitskii and A.B. Migdal.
Zh. Éksp. Teor. Fiz. 34 [Sov. Phys. JETP 139, 96 (1958)], 139, 1958.

[28] William H. Press Saul A. Teukolsky William T. Vetterling Brian P. Flannery.
Numerical Recipes in C.
Cambridge University Press, 1992.

[29] N. Attig.
IBM Blue Gene/L in Jülich: A First Step to Petascale Computing.
Inside, 2, 2005.
http://inside.hlrs.de/pdfs/inSiDE_autumn2005.pdf.

[30] Thomas Lippert, Klaus Schilling, F. Toschi, S. Trentmann, and R. Tripiccione.
Transpose Algorithm for FFT on APE/Quadrics.
pages 439–448, 1998.

[31] Stéphane Pairault, David Sénéchal, and A.-M. S. Tremblay.
Strong-Coupling Expansion for the Hubbard Model.
Phys. Rev. Lett., 80(24):5389–5392, Jun 1998.

[32] D. Sénéchal, D. Perez, and M. Pioro-Ladrière.
Spectral Weight of the Hubbard Model through Cluster Perturbation Theory.
Phys. Rev. Lett., 84(3):522–525, Jan 2000.

[33] David Sénéchal, Danny Perez, and Dany Plouffe.
Cluster perturbation theory for Hubbard models.
Phys. Rev. B, 66(7):075129, Aug 2002.

[34] Stéphane Pairault, David Sénéchal, and A.-M. S. Tremblay.
Strong-Coupling Expansion for the Hubbard Model.
Eur. Phys. J. B, 85(16), 2000.

[35] Claudius Gros and Roser Valentí.
Cluster expansion for the self-energy: A simple many-body method for interpreting the photoemission spectra of correlated Fermi systems.
Phys. Rev. B, 48(1):418–425, Jul 1993.

[36] E. Arrigoni C. Dahnken and W. Hanke.

Spectral properties of high-Tc cuprates via a Cluster-Perturbation Approach.
J. Low Temp. Phys., 126:949–959, 2002.

[37] Markus Aichhorn.
Ordering phenomena in strongly-correlated systems: Cluster perturbation theory approaches.
PhD thesis, TU Graz, 2004.

[38] Walter Kohn.
Theory of the Insulating State.
Phys. Rev., 133(1A):A171–A181, Jan 1964.

[39] Wolfgang Nolting.
Grundkurs Theoretische Physik 7 - Viel-Teilchen-Theorie.
Springer, 2005.

[40] B. S. Shastry and B. Sutherland.
Twisted boundary conditions and effective mass in Heisenberg-Ising and Hubbard rings.
Physical Review Letters, 65:243–246, July 1990.

[41] D. J. Scalapino, S. R. White, and S. Zhang.
Insulator, metal, or superconductor: The criteria.
Physical Review B, 47:7995–8007, April 1993.

[42] C. A. Stafford, A. J. Millis, and B. S. Shastry.
Finite-size effects on the optical conductivity of a half-filled Hubbard ring.
Physical Review B, 43:13660–13663, June 1991.

[43] Raffaele Resta and Sandro Sorella.
Electron localization in the insulating state.
Phys. Rev. Lett., 82(2):370–373, Jan 1999.

[44] Laura Cano-Cortes, Andreas Dolfen, Jaime Merino, Jorg Behler, Bernard Delley, Karsten Reuter, and Erik Koch.
Coulomb parameters and photoemission for the molecular metal TTF-TCNQ.
submitted to Phys. Rev. Lett., cond-mat/0609416, 2006.

[45] R. Claessen, M. Sing, U. Schwingenschlögl, P. Blaha, M. Dressel, and C. S. Jacobsen.
Spectroscopic Signatures of Spin-Charge Separation in the Quasi-One-Dimensional Organic Conductor TTF-TCNQ.
Phys. Rev. Lett., 88(9):096402, Feb 2002.

[46] H. Benthien, F. Gebhard, and E. Jeckelmann.
Spectral function of the one-dimensional Hubbard model away from half filling.
Physical Review Letters, 92:256401, 2004.

[47] J. B. Torrance, Y. Tomkiewicz, and B. D. Silverman.
Enhancement of the magnetic susceptibility of TTF-TCNQ (tetrathiafulvalene-tetracyanoquinodimethane) by Coulomb correlations.

Phys. Rev. B, 15(10):4738–4749, May 1977.

[48] J. Hubbard.
Generalized Wigner lattices in one dimension and some applications to tetracyanoquinodimethane (TCNQ) salts.
Phys. Rev. B, 17:494505, 1978.

[49] S. Mazumdar and A. N. Bloch.
Systematic Trends in Short-Range Coulomb Effects among Nearly One-Dimensional Organic Conductors.
Phys. Rev. Lett., 50(3):207–211, Jan 1983.

[50] H. J. Schulz.
Interacting fermions in one dimension: from weak to strong correlation.
ArXiv Condensed Matter e-prints, February 1993.

[51] E. Wigner.
On the Interaction of Electrons in Metals.
Phys. Rev., 46(11):1002–1011, Dec 1934.

[52] Andreas Dolfen, Eva Pavarini, and Erik Koch.
New Horizons for the Realistic Description of Materials with Strong Correlations.
Inside, 1, 2006.
http://inside.hlrs.de/htm/Edition_01_06/article_05.htm.

[53] G. M. Amdahl.
Validity of the single processor approach to achieving large scale computing capabilities.
Proceedings AFIPS, pages 483–485, 1967.

[54] Meier and Gürich.
Abrechnungs- und Kontingentierungsverfahren für die zentralen Rechnersysteme und Dienste (TKI0015).
2006.
http://www.fz-juelich.de/zam/files/docs/tki/tki-0015.pdf.

[55] Octavian Lascu.
An Introduction to the New IBM eServer pSeries High Performance Switch.
IBM, 2004.
http://www.redbooks.ibm.com/Redbooks.nsf/RedbookAbstracts/SG246978.html.

[56] Frank Johnston.
High Performance Switch - Performance White Paper.
IBM, 2005.
http://jumpdoc.fz-juelich.de/doc_pdf/hw/pseries_hps_perf.pdf.

[57] E.Fradkin.
Field Theories of Condensed Matter Systems.
AddisonWesley Publishing Company - Frontiers in Physics, 1991.

List of Figures

1.1 Band vs. atomic limit 4

2.1 Band structure of 1d/2d HM 10
2.2 Energy surface of 2d HM 11
2.3 Density of states in low dimensions 12
2.4 Two orbitals per cell, separated by Δ in energy. 12
2.5 Band structure of two band model 14
2.6 Boundary conditions in one dimension 16
2.7 Finite size scaling 17
2.8 Band vs. atomic limit 20
2.9 Mott insulation 22
2.10 DOS in half-filled system 23
2.11 Bandwidth of lowest band in a half-filled system 24
2.12 Gutzwiller wave function variation 25
2.13 Gutzwiller wave function variation 26

3.1 Development of energy with iterations 36
3.2 Large Lanczos run 38
3.3 Convergence of residual norm 40
3.4 Duplicate eigenvalue problem 41
3.5 Density matrix 44
3.6 Momentum distribution with Coulomb interaction at $T = 0$ 45
3.7 Spectral function for different number of iterations 51
3.8 High energy convergence 52
3.9 Error of first 10 momenta 53

4.1 Sharp profile of simulations program 57
4.2 OpenMP speed up of a half-filled 16 site system 60
4.3 Speed up of direct MPI-2 one-sided fetch 63
4.4 Considering a LancVec as matrix 64
4.5 Scheme of alltoall-transpose 66
4.6 Speed up of Alltoall(v) implementation for 16 67
4.7 General matrix transposition operation 68
4.8 Speed up of Alltoall(v) implementation for 18 sites 69
4.9 Exact and Lanczos ground state energy 73
4.10 Band-limit check 74
4.11 Atomic limit check 75

5.1 (complex) ED's angular resolved spectral function 78
5.2 20 sites photoemission spectral function 79
5.3 Lattice to super lattice for CPT 79
5.4 Comparison of CPT spectral functions for different number of sites . 80
5.5 Single site CPT calculation 83
5.6 $U = 0$ CPT check 85
5.7 CPT calculation for clusters in PBC and OBC 86
5.8 CPT vs. complex Lanczos 86
5.9 Comparison CPT,ED with respect to L 87
5.10 CPT for 12 sites with V term in OBC 88
5.11 CPT for 8 and 12 sites with V in PBC 88

6.1 Qualitative different behavior of $E(\Phi)$ 92
6.2 Finite-size effects (2) 93
6.3 Sketch of two-band system 94
6.4 Occupation and Drude weight in half-filled Hubbard chain 94
6.5 Occupation and Drude weight with quarter-filling 95
6.6 Self-energy of insulators 97
6.7 Self-energy of a metal 99

7.1 TTF and TCNQ 101
7.2 Single-particle energy levels 102
7.3 Charge-transfer salt TTF-TCNQ 103
7.4 TTF and TCNQ two molecules 104
7.5 $A_k(\omega)$ for TCNQ in t-U model 105
7.6 Magnification of $A_k(\omega)$ 106
7.7 Hubbard-Wigner crystal 107
7.8 Wigner crystal effect 108
7.9 Wigner crystal with V 109
7.10 TCNQ spectral function in t-U-V model 110
7.11 Double occupation d as function of V and t 111
7.12 Double occupation d as function of t_{eff} 112
7.13 Spectral function for different values of t 113
7.14 TCNQ density plot with PT 114
7.15 Kinetic and potential energy of TCNQ 114
7.16 CPT with V 115
7.17 TTF angular-resolved spectral function 116
7.18 Gauge finite size effects for TCNQ 117
7.19 Periodic and anti-periodic bc 118

A.1 Typical "look" of Amdahl's law 122

Selbständigkeitserklärung

Hiermit versichere ich, die vorliegende Arbeit selbständig und nur unter Zuhilfenahme der angegebenen Quellen und Hilfsmittel angefertigt zu haben.

Jülich im Oktober 2006

Andreas Dolfen