

Analytic Continuation of Electronic Green's Functions from Imaginary to Real Time using Maximum Entropy

Master's Thesis at the Department of Applied Physics

JOHAN SCHÖTT

Department of Applied Physics
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2013

Analytic Continuation of Electronic Green's
Functions from Imaginary to Real Time using
Maximum Entropy

Johan Schött

July 22, 2013

Abstract

This thesis is mainly a computational work studying the analytical continuation of Green's functions using the Maximum Entropy method. A strongly correlated electron system is described with the single-band Hubbard model and paramagnetic solutions are studied using Dynamic Mean Field Theory on a Bethe lattice. Continuous Time Quantum Monte Carlo is used as Impurity solver, for the infinite Anderson model at a finite temperature, to obtain the Matsubara single-particle Green's function propagator. Both metallic and insulating spectral functions are obtained using the Maximum Entropy Method. General properties of the Maximum Entropy Method as an analytic continuation method from imaginary to real time are also discussed.

Contents

1	Introduction	5
1.1	The Electronic Structure Problem in Solids	6
1.2	Measurements	7
1.2.1	Schrödinger Picture	7
1.2.2	Heisenberg Picture	8
1.2.3	Imaginary Time	8
1.2.4	Finite Temperature	8
2	Green's Function Propagators	11
2.1	Two Operator Propagator	11
2.2	Normal Single-Particle Propagator	14
2.3	Non-interacting Green's Function	15
2.4	Dyson Equation	16
2.5	Local Green's Function	17
3	Hubbard Model	19
3.1	Bethe Lattice	22
4	Dynamic Mean Field Theory (DMFT)	25
4.1	Single Impurity Anderson Model (SIAM)	26
4.2	Self-Consistent Equation	28
4.2.1	DMFT on a Bethe Lattice	28
4.3	Continuous Time Quantum Monte Carlo (CTQMC)	30
4.3.1	Legendre Polynomials	31
4.4	Simulations	32
5	Maximum Entropy Method (MEM)	41
5.1	Bayesian Inference	42
5.2	Likelihood function	45
5.2.1	Correlations	46
5.3	Entropic Prior	48

5.4	Q Functional Minimization	50
5.4.1	Discretization Approach	53
5.4.2	Functional Approach	54
5.5	Hyper-parameter α	55
5.6	Batch Averaging	57
5.7	Stochastic Sampling of the Likelihood Function	57
5.8	Analytic Continuation Kernels	58
5.9	Simulations	58
5.9.1	Tests of MEM	59
5.9.2	Spectral Functions for the Hubbard model	62
6	Conclusion	69
	Appendices	75
A	Second Quantization	77

Chapter 1

Introduction

In Condensed matter physics electron correlations are sometimes so strong that they totally change the system's thermodynamic properties. In such materials Density Functional Theory does not work since it is based on the approximation of weakly interacting electrons. The fully interacting electronic structure problem in solids is impossible to solve. Effective models including local interaction, e.g. the Hubbard model, are therefore used to understand the underlying physics. Examples of phenomena due to strong correlations are high temperature superconductivity and Mott metal-insulator transitions (MIT). By turning on the Hubbard interaction parameter U , the system undergoes a MIT at half-filling.

The thesis is organized as follows. First the electronic structure problem is introduced. Single-particle *Green's functions* and *spectral functions*, objects to describe the many-body problem, are introduced in chapter 2. Chapter 3, about the Hubbard model, define our system. The Hubbard model is solved in chapter 4 by Dynamic Mean Field Theory (DMFT). Using Maximum Entropy Method (MEM), an analytic continuation of the obtained Green's functions, shown in chapter 5, finally gives us the wanted spectral functions.

In this work the half-filled paramagnetic single-band Hubbard model is investigated on a Bethe lattice for a finite temperature and various interaction strengths. It can be solved by DMFT in a mean-field fashion using an approximation which is exact for infinite number of nearest neighbors. DMFT maps the many-body lattice problem to a local many-body impurity problem where a single interacting site is coupled to a non-interacting environment. The impurity problem can be solved by different techniques, in this thesis Quantum Monte Carlo is used. It allows calculations of the Green's functions with imaginary time argument, expressed using a Legendre basis, see section 4.3.

However the physical information of interest is for real argument. Viewing

the Green's function in complex frequency plane, the transformation from the imaginary to the real frequency axis is actually an analytic continuation. The statistical noise from the Quantum Monte Carlo (QMC) sampling and the inverse nature of the analytical continuation makes this transformation ill-conditioned. A regularization is needed, in MEM it is to maximize the information entropy of the spectral function. The most probable spectral function is calculated using Bayesian probability theory, see section 5.

The spectral function, describing the effective single-particle Density of States (DOS), is defined in chapter 2. For a non-interacting system the spectral function is the DOS of the non-interacting eigenenergies. Momentum dependent spectral functions contain information about the quasiparticle bandstructure and can be measured experimentally with Angle-Resolved Photoemission Spectroscopy (ARPES) [1].

1.1 The Electronic Structure Problem in Solids

The system of interacting electrons has, within the Born-Oppenheimer approximation, a Hamiltonian of the form

$$H = \underbrace{\sum_i \left(\frac{\mathbf{p}_i^2}{2m} + V_{\text{ext}}(\mathbf{r}_i) \right)}_{H_0} + \frac{1}{2} \underbrace{\sum_{i,j} V_I(\mathbf{r}_i, \mathbf{r}_j)}_{H_I} \quad (1.1)$$

where $V_{\text{ext}}(\mathbf{r})$ is a external potential and $V_I(\mathbf{r}_i, \mathbf{r}_j) = \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|}$ is the electron-electron Coulomb potential. By using a basis-set of orbital single-particle states $\{f_i\}$ the Hamiltonian in second quantization takes the form

$$H = \underbrace{\sum_{ij\sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma}}_{H_0} + \frac{1}{2} \underbrace{\sum_{ijkl\sigma\sigma'} V_{ijkl} c_{i\sigma}^\dagger c_{k\sigma'}^\dagger c_{l\sigma'} c_{j\sigma}}_{H_I} \quad (1.2)$$

with the Hamiltonian matrix elements

$$t_{ij} = \int d^3r f_i^*(\mathbf{r}) \left(\frac{\mathbf{p}^2}{2m} + V_{\text{ext}}(\mathbf{r}) \right) f_j(\mathbf{r}) \quad (1.3)$$

$$V_{ijkl} = \int d^3r \int d^3r' f_i^*(\mathbf{r}) f_j(\mathbf{r}) V_{e-e}(\mathbf{r}, \mathbf{r}') f_k^*(\mathbf{r}') f_l(\mathbf{r}'). \quad (1.4)$$

The equation 1.2 needs to be simplified in order to be solvable. In materials where electron–electron interaction is strong the electron eigen states are

usually localized in space [2]. This is in contrast to non-interacting Bloch-states which are extended over the entire crystal. Due to the localized nature of the orbitals, a set of localized functions is a good choice of basis functions. The localization enables approximations of only nearest-neighbor hopping and only intra-orbital correlations. Hence an effective Hamiltonian called the Hubbard model is obtained, see chapter 3.

With Bloch-single particle eigenstates $\{\phi_{n,\mathbf{k},\sigma}\}$, thus $i = \{n, \mathbf{k}\}$, with bandindex n , lattice momentum \mathbf{k} and spin σ , as basis set, the hopping matrix t_{ij} is diagonal,

$$t_{ij} = t_{n\mathbf{k},n'\mathbf{k}'} = t_{n\mathbf{k},n'\mathbf{k}'} \langle n\mathbf{k} | n'\mathbf{k}' \rangle = \epsilon_{n\mathbf{k}} \delta_{nn'} \delta_{\mathbf{k}\mathbf{k}'} \quad (1.5)$$

so

$$H_0 = \sum_{n\mathbf{k}\sigma} \epsilon_{n\mathbf{k}} c_{n\mathbf{k}\sigma}^\dagger c_{n\mathbf{k}\sigma} \quad (1.6)$$

where $\epsilon_{n\mathbf{k}}$ is the dispersion relation for the non-interacting system for different bands n .

1.2 Measurements

There exist different, but mathematically equivalent, formulations of quantum mechanical dynamics, called pictures. The expectation value of operator \mathcal{O} in state Ψ is

$$\bar{\mathcal{O}} = \langle \Psi | \mathcal{O} | \Psi \rangle. \quad (1.7)$$

Its time evolution can be calculated by viewing the time dependence of the state Ψ , or through the time dependence of the operator \mathcal{O} or by a combination of both.

1.2.1 Schrödinger Picture

In this picture, all the time-dependence is captured by the quantum state $\Psi(t)$. Its time dependence is determined by the time dependent Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\Psi(t)\rangle = H |\Psi(t)\rangle. \quad (1.8)$$

We will work with $\hbar = 1$ in the rest of the thesis. This equation can formally be solved for a time independent Hamiltonian

$$|\Psi(t)\rangle = U(t) |\Psi(t=0)\rangle \quad (1.9)$$

where

$$U(t) = e^{-itH} \quad (1.10)$$

is the unitary time evolution operator.

1.2.2 Heisenberg Picture

In the Heisenberg picture, the time dependence is determined by the operator $\mathcal{O}(t)$. Ψ is here time independent. $\mathcal{O}(t)$ is related to the Schrödinger picture operator \mathcal{O} through

$$\mathcal{O}(t) = e^{+itH} \mathcal{O} e^{-itH}. \quad (1.11)$$

The expectation values are the same in both pictures which is easily seen by

$$\bar{\mathcal{O}}(t) = \langle \Psi(t) | \mathcal{O} | \Psi(t) \rangle = \langle \Psi | U^\dagger(t) \mathcal{O} U(t) | \Psi \rangle = \langle \Psi | U(-t) \mathcal{O} U(t) | \Psi \rangle = \langle \Psi | \mathcal{O}(t) | \Psi \rangle. \quad (1.12)$$

Using the grand canonical ensemble, Heisenberg time-dependent operators are using $\bar{H} = H - \mu N$ instead of H , which just shifts the single particle energies with μ ,

$$H_0 - \mu N = \sum_{n\mathbf{k}\sigma} (\epsilon_{n\mathbf{k}} - \mu) c_{n\mathbf{k}\sigma}^\dagger c_{n\mathbf{k}\sigma}. \quad (1.13)$$

1.2.3 Imaginary Time

It is usually computationally easier to calculate thermal averages using imaginary time τ instead of the physical time. Let $t = -i\tau$. The time evolution operator $U(t)$, in eq 1.10, is exponential for imaginary times in contrast to oscillating behavior for real times,

$$U(t = -i\tau) = e^{-\tau H}. \quad (1.14)$$

In the Heisenberg picture the time-dependent operator is therefore given by

$$\mathcal{O}(t = -i\tau) = e^{+\tau H} \mathcal{O} e^{-\tau H}. \quad (1.15)$$

If one use the variable name τ , imaginary time is implicitly indicated. In the literature

$$\mathcal{O}(\tau) = e^{+\tau H} \mathcal{O} e^{-\tau H} \quad (1.16)$$

is a common way of writing the operator \mathcal{O} , when dependent on the imaginary time.

1.2.4 Finite Temperature

Let's consider a thermodynamic system given by a Hamiltonian H at a finite temperature T , volume V and chemical potential μ . The particle number

$N = \sum_j c_j^\dagger c_j$ is an observable operator and the many body eigenstate $|\Psi_i\rangle$ obey

$$H|\Psi_i\rangle = E_i|\Psi_i\rangle \quad (1.17)$$

$$N|\Psi_i\rangle = N_i|\Psi_i\rangle. \quad (1.18)$$

The operator $\bar{H} = H - \mu N$, with eigenvalues $\bar{E}_i = E_i - \mu N_i$, will make the formulas for the grand canonical ensemble look the same as for the canonical ensemble.

If assuming that the eigenstates are known, measurements can be done by averaging over the eigenstates $|\Psi_i\rangle$ according to the grand canonical ensemble with probability weights

$$\mathcal{P}_i = \frac{\exp(-\beta\bar{E}_i)}{\sum_n \exp(-\beta\bar{E}_n)} \quad (1.19)$$

where $\beta = \frac{1}{k_B T}$ is the inverse temperature. By introducing the density matrix operator

$$\rho = \exp(-\beta\bar{H}) \quad (1.20)$$

get

$$\rho_i = \langle \Psi_i | \rho | \Psi_i \rangle \quad (1.21)$$

$$\mathcal{Z} = \sum_i \rho_i = \sum_i \langle \Psi_i | \rho | \Psi_i \rangle = \text{Tr } \rho \quad (1.22)$$

$$\mathcal{P}_i = \frac{\rho_i}{\mathcal{Z}} \quad (1.23)$$

where \mathcal{Z} is the grand partition function and Tr is an abbreviation for taking the trace of an operator. This notation simplifies thermodynamic measurements of observables,

$$\begin{aligned} \langle \mathcal{O} \rangle &= \sum_i \mathcal{P}_i \langle \Psi_i | \mathcal{O} | \Psi_i \rangle = \sum_i \frac{\rho_i}{\mathcal{Z}} \langle \Psi_i | \mathcal{O} | \Psi_i \rangle = \frac{1}{\mathcal{Z}} \sum_i \langle \Psi_i | \mathcal{O} \rho_i | \Psi_i \rangle \\ &= \frac{1}{\mathcal{Z}} \sum_i \langle \Psi_i | \mathcal{O} \rho | \Psi_i \rangle = \frac{1}{\mathcal{Z}} \text{Tr } \mathcal{O} \rho. \end{aligned} \quad (1.24)$$

Measurements are thus calculated by the ratio of two traces. Since the trace is invariant of the choice of basis representation one can use any complete basis set. This is advantageous since the basis with eigenstates of the full Hamiltonian is usually unknown.

Chapter 2

Green's Function Propagators

Green's functions are time dependent expectation values of products of operators. First, properties for a general two-operator Green's function is presented, where the two operators A and B do not need to be hermitian. The normal Green's function propagator with A and B equal to the annihilation and creation operator respectively is analyzed in detail.

The analytic continuation of Green's functions via the Hilbert transformation, from imaginary to real time, is investigated. In this chapter we also see the simple connection between the Green's function and the spectral function.

For interacting systems, which are studied in this thesis, the Dyson equation describes how the self-energy in an elegant way relates the non-interacting and interacting Green's function. This will be applied intensely in chapter 4 about DMFT.

2.1 Two Operator Propagator

The temperature Green's function for operators A and B for imaginary time within the Heisenberg picture is defined as

$$G(\tau) = G_{A,B}(\tau) = -\langle T A(\tau) B(0) \rangle \quad (2.1)$$

where $\langle \rangle$ is the thermal average. A time independent Hamiltonian is assumed so the system is time invariant. This makes the Green's function only dependent on the relative time difference between two events $\tau = \tau_2 - \tau_1$. A, B are Heisenberg operators with $A(\tau) = e^{\tau \bar{H}} A e^{-\tau \bar{H}}$ and $\bar{H} = H - \mu N$. T is the time-ordering operator. T switches A and B when τ is negative. If both A, B are even functions of fermionic operators, T also change the sign of the expression when $\tau < 0$. It will be true in this thesis.

The imaginary time propagator is defined for $\tau \in (-\beta, \beta]$. Because the trace is invariant under cyclic permutations, the Green's function is anti-periodic. This is seen by letting $\tau < 0$ and use the definition of the thermal average,

$$\begin{aligned} G(\tau) &= \langle B(0)A(\tau) \rangle = \frac{1}{Z} \text{Tr} e^{-\beta\bar{H}} B(0)A(\tau) \\ &= \frac{1}{Z} \text{Tr} A(\tau)e^{-\beta\bar{H}} B(0) = \frac{1}{Z} \text{Tr} e^{-\beta\bar{H}} \underbrace{e^{\beta\bar{H}} A(\tau)e^{-\beta\bar{H}}}_{A(\tau+\beta)} B(0) \\ &= -G(\tau + \beta). \end{aligned} \quad (2.2)$$

By periodically repeating $G(\tau)$ from $\tau \in (-\beta, \beta]$, the Fourier representation is given by

$$G(\tau) = \frac{1}{\beta} \sum_{n=-\infty}^{\infty} e^{-i\nu_n\tau} G_n \quad (2.3)$$

with

$$G_n = \frac{1}{2} \int_{-\beta}^{\beta} d\tau e^{i\nu_n\tau} G(\tau) \quad (2.4)$$

and $\nu_n = \frac{\pi n}{\beta}$. By using the anti-periodicity in eq 2.2, only odd n are allowed,

$$\nu_n = \frac{(2n-1)\pi}{\beta} \quad (2.5)$$

ν_n are called the Matsubara frequencies. Eq 2.4 simplifies to

$$G_n = \int_0^{\beta} d\tau e^{i\nu_n\tau} G(\tau). \quad (2.6)$$

A unique *spectral function* $\rho(\omega)$ correspond to a complete set of G_n values [3],

$$G_n = \int_{-\infty}^{\infty} d\omega \frac{\rho(\omega)}{i\nu_n - \omega}. \quad (2.7)$$

This can be seen as a special case of

$$G(z) = \int_{-\infty}^{\infty} d\omega \frac{\rho(\omega)}{z - \omega} \quad (2.8)$$

for function $G(z)$ when $z = i\nu_n$. $G(z)$ is analytic in the lower and upper half plane with a cut at the real axis according to the Sokhotski-Plemelj theorem. Because the spectral function is unique, the *analytical continuation* from $G_n = G(i\nu_n)$ to $G(z)$ is also unique, see figure 2.1. By inserting the spectral

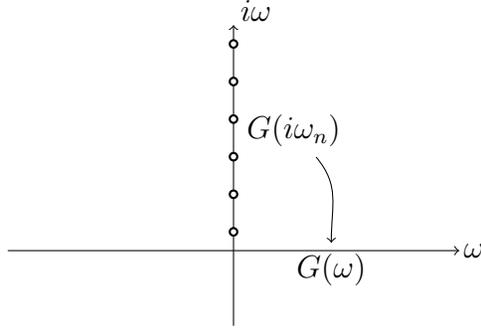


Figure 2.1: Analytic continuation of $G(z)$ from Matsubara points to real axis.

representation of the Matsubara values from eq 2.7 into eq 2.3 we get

$$\begin{aligned}
 G(\tau) &= \frac{1}{\beta} \sum_{n=-\infty}^{\infty} e^{-i\nu_n \tau} G_n = \int_{-\infty}^{\infty} d\omega \rho(\omega) \frac{1}{\beta} \sum_{n=-\infty}^{\infty} \frac{e^{-i\nu_n \tau}}{i\nu_n - \omega} \\
 &= \begin{cases} \int_{-\infty}^{\infty} d\omega \rho(\omega) \underbrace{\frac{-e^{-\omega \tau}}{e^{-\beta \omega} + 1}}_{k_{\tau}(\omega)}, & \text{if } 0 < \tau < \beta, \\ -G(\tau + \beta), & \text{if } -\beta < \tau < 0. \end{cases} \quad (2.9)
 \end{aligned}$$

where $k_{\tau}(\omega) \in \mathbb{R}_-$ is an important integral kernel.

If all the eigenstates in our system would be known, the spectral function could be written as a sum over them. This formulation is called the Lehmann representation and is given by [4]

$$\rho(\omega) = (1 + e^{-\beta \omega}) \frac{1}{Z} \sum_{n,m} e^{-\beta E_m} \langle m|A|n\rangle \langle n|B|m\rangle \delta(\omega - E_n + E_m). \quad (2.10)$$

For a system with infinite number of eigenstates, the discrete sum forms a continuous function. Green's functions with $A = B^{\dagger}$ are real (and non-negative),

$$\rho(\omega) = (1 + e^{-\beta \omega}) \frac{1}{Z} \sum_{n,m} e^{-\beta E_m} |\langle n|B|m\rangle|^2 \delta(\omega - E_n + E_m). \quad (2.11)$$

A real spectral function is related to the imaginary part of the Green's func-

tion just above the real axis. For a positive ϵ , equation 2.8 becomes

$$\begin{aligned}
-\frac{1}{\pi} \text{Im}[G(\omega' + i\epsilon)] &= -\frac{1}{\pi} \int_{-\infty}^{\infty} d\omega \text{Im}\left[\frac{\rho(\omega)}{\omega' + i\epsilon - \omega}\right] \\
&= -\frac{1}{\pi} \int_{-\infty}^{\infty} d\omega \text{Im}\left[\frac{\omega' - \omega - i\epsilon}{(\omega' - \omega)^2 + \epsilon^2}\right] \rho(\omega) \\
&= \int_{-\infty}^{\infty} d\omega \underbrace{\frac{1}{\pi} \frac{\epsilon}{(\omega' - \omega)^2 + \epsilon^2}}_{L_\epsilon(\omega' - \omega)} \rho(\omega)
\end{aligned} \tag{2.12}$$

where $L_\epsilon(\omega' - \omega)$ is a Lorentzian of width ϵ and centered at $\omega = \omega'$. By using $L_\epsilon(\omega) \xrightarrow{\epsilon \rightarrow 0^+} \delta(\omega)$ we get

$$\rho(\omega) = -\frac{1}{\pi} \text{Im}[G(\omega + i0^+)]. \tag{2.13}$$

Eq 2.8 can be seen as a straightforward analytic continuation of the Green's function from the real axis to any point z in the complex plane as a consequence of this. Unfortunately the problem is usually inverse with $G(z = i\nu_n)$ known and $\rho(\omega)$ unknown. This requires solving the Hilbert transform in eq 2.8 as a set of integral equations, making the analytic continuation hard.

For a Green's function $G(\tau) \in \mathbb{R}$, $G(-i\nu_n) = G(i\nu_n)^*$ and $G(\tau)$ can be represented by G_n for only positive Matsubara frequencies,

$$G(\tau) = \frac{2}{\beta} \text{Re}\left[\sum_{n=1}^{\infty} G_n e^{-i\nu_n \tau}\right]. \tag{2.14}$$

2.2 Normal Single-Particle Propagator

The normal single-particle Green's function propagator for imaginary time is obtained by letting $A = c_{\mathbf{k}\sigma}$ and $B = c_{\mathbf{k}\sigma}^\dagger$,

$$G_{\mathbf{k}\sigma}(\tau) = -\langle T c_{\mathbf{k}\sigma}(\tau) c_{\mathbf{k}\sigma}^\dagger(0) \rangle, \tag{2.15}$$

where $c_{\mathbf{k}\sigma}$, $c_{\mathbf{k}\sigma}^\dagger$ are the annihilation and creation operators for an electron with momentum \mathbf{k} and spin σ . We will only study spin invariant systems so the spin-index $\sigma \in \{\uparrow, \downarrow\}$ is usually omitted. The normal Green's function for real time $G_{\mathbf{k}\sigma}(t)$ describes how the system reacts if add an electron at time 0 with momentum \mathbf{k} and spin σ and annihilate an electron at time t with same momentum and spin. Since $B = c_{\mathbf{k}\sigma}^\dagger = A^\dagger$, the spectral function is real. By using eq 2.11, its Lehmann representation is

$$\rho_{\mathbf{k}\sigma}(\omega) = (1 + e^{-\beta\omega}) \frac{1}{Z} \sum_{n,m} e^{-\beta E_m} \left| \langle n | c_{\mathbf{k}\sigma}^\dagger | m \rangle \right|^2 \delta(\omega - E_n + E_m). \tag{2.16}$$

The normal spectral function $\rho_{\mathbf{k}\sigma}(\omega)$ describe the probability density for an electron with momentum \mathbf{k} and spin σ of having the energy ω relative to the chemical potential μ .

2.3 Non-interacting Green's Function

If the system Hamiltonian can be written as sum of a non-interacting and a interacting Hamiltonian,

$$H = H_0 + H_I, \quad (2.17)$$

a non-interacting Green's function $G_{0,\mathbf{k}\sigma}(\omega)$, also called the bare propagator, is useful. By non-interacting we mean that interaction between electrons is neglected but external potentials and the potential from the lattice is included. For a periodic potential, the single-particle eigenstates of H_0 are Bloch-states. $G_{0,\mathbf{k}\sigma}(\omega)$ is defined like the interacting Green's function but instead of using the full Hamiltonian H one use the non-interacting one,

$$H_0 = \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} \quad (2.18)$$

where $\epsilon_{\mathbf{k}}$ is the electron energy dispersion for the non-interacting system. Tracing over the non-interacting eigenstates gives

$$\begin{aligned} G_{0,\mathbf{k}\sigma}(\tau) &= -\langle T c_{\mathbf{k}\sigma}(\tau) c_{\mathbf{k}\sigma}^\dagger(0) \rangle_0 \\ &= e^{-\tau(\epsilon_{\mathbf{k}} - \mu)} \begin{cases} -\langle 1 - c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} \rangle_0, & \text{if } 0 < \tau < \beta, \\ \langle c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} \rangle_0, & \text{if } -\beta < \tau < 0. \end{cases} \end{aligned} \quad (2.19)$$

where

$$\langle c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} \rangle_0 = \frac{1}{e^{\beta(\epsilon_{\mathbf{k}} - \mu)} + 1} \quad (2.20)$$

is the Fermi-Dirac occupation number. An interesting observation is that $G_{0,\mathbf{k}\sigma}(\tau = 0_-) = \langle c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} \rangle_0$. By Fourier transforming eq 2.19 according to eq 2.6, the non-interacting Matsubara Green's function becomes

$$G_{0,\mathbf{k}\sigma}(i\nu_n) = \frac{1}{i\nu_n - \epsilon_{\mathbf{k}} + \mu}. \quad (2.21)$$

The analytic continuation to the complex plane is simply done by replacing $i\nu_n$ with z

$$G_{0,\mathbf{k}\sigma}(z) = \frac{1}{z - \epsilon_{\mathbf{k}} + \mu}. \quad (2.22)$$

The non-interacting spectral function is therefore

$$\begin{aligned}
\rho_{0,\mathbf{k}\sigma}(\omega) &= \lim_{\delta \rightarrow 0^+} -\frac{1}{\pi} \text{Im}[G_{0,\mathbf{k}\sigma}(\omega + i\delta)] \\
&= \lim_{\delta \rightarrow 0^+} -\frac{1}{\pi} \text{Im}\left[\frac{1}{\omega - \epsilon_{\mathbf{k}} + \mu + i\delta}\right] \\
&= \lim_{\delta \rightarrow 0^+} L_{\delta}(\omega - \epsilon_{\mathbf{k}} + \mu) \\
&= \delta(\omega - \epsilon_{\mathbf{k}} + \mu),
\end{aligned} \tag{2.23}$$

so for a given \mathbf{k} , the electron can only have energy $\epsilon_{\mathbf{k}}$. $\rho_{\mathbf{k}}(\omega)$ is thus given by the bandstructure.

2.4 Dyson Equation

The interacting Green's function, also called the dressed propagator, can be written as an infinite sum of Feynman diagrams [4]. These sums can formally be calculated, resulting in the Dyson equation

$$G_{\mathbf{k}\sigma}(z) = \frac{1}{G_{0,\mathbf{k}\sigma}(z)^{-1} - \Sigma_{\mathbf{k}}(z)} \tag{2.24}$$

connecting the bare and the dressed Green's function propagator. All interactions are captured by the self-energy $\Sigma_{\mathbf{k}}(z) \in \mathbb{C}$. According to eq 2.22 the dressed Green's function reads

$$G_{\mathbf{k}\sigma}(z) = \frac{1}{z - \epsilon_{\mathbf{k}} + \mu - \Sigma_{\mathbf{k}}(z)}. \tag{2.25}$$

By using eq 2.13, the self-energy modify the spectral function to a Lorentzian

$$\begin{aligned}
\rho_{\mathbf{k}}(\omega) &= \frac{1}{\pi} \frac{-\text{Im}[\Sigma_{\mathbf{k}}(\omega)]}{(\omega - \epsilon_{\mathbf{k}} + \mu - \text{Re}[\Sigma_{\mathbf{k}}(\omega)])^2 + (-\text{Im}[\Sigma_{\mathbf{k}}(\omega)])^2} \\
&= L_{-\text{Im}[\Sigma_{\mathbf{k}}(\omega)]}(\omega - \epsilon_{\mathbf{k}} + \mu - \text{Re}[\Sigma_{\mathbf{k}}(\omega)]).
\end{aligned} \tag{2.26}$$

Since the spectral function is non-negative, the imaginary part of the self-energy needs to be non-positive. For sufficiently small self-energy the system is a Fermi liquid with quasiparticles, where the real part of the self-energy acts as a energy shift and the imaginary part broadens the peak in the spectral function. The energy shift effectively renormalize the quasiparticle mass and the peak-width is the inverse quasiparticle lifetime. For strong interaction is the Fermi-liquid peak replaced by collective spectral weight.

The self-energy itself is an infinite sum of all irreducible self-energy diagrams. A self-energy diagram is a Feynman diagram which can be inserted into a particle line without having outgoing lines. If a self-energy diagram can be divided into two unconnected self-energy parts it is called reducible, if not it is irreducible. Once $\Sigma_{\mathbf{k}}(z)$ is known, the Green's function is known. Approximations of only including a certain irreducible self-energy diagrams in $\Sigma_{\mathbf{k}}$ gives e.g. Hartree-Fock, Random Phase Approximation or the Ladder Approximation [4]. In DMFT the self-energy is assumed to be momentum independent. This approximation is exact for lattices with infinite number of nearest lattice neighbors. More about this is found in chapter 4.

2.5 Local Green's Function

By summing over all possible \mathbf{k} values, a local Green's function is obtained. More precise, the local Green's function $G_{\sigma}(z)$ is defined as the spatial Fourier transform of $G_{\mathbf{k}\sigma}(\omega)$ for $\mathbf{r} = \mathbf{0}$,

$$G_{\sigma}(z) = G_{\sigma, \mathbf{r}=\mathbf{0}}(z) = \frac{1}{\Lambda} \sum_{\mathbf{k}} G_{\mathbf{k}\sigma}(z) e^{i\mathbf{k}\cdot\mathbf{0}} = \frac{1}{\Lambda} \sum_{\mathbf{k}} \frac{1}{z - \epsilon_{\mathbf{k}} + \mu - \Sigma_{\mathbf{k}}(z)} \quad (2.27)$$

where Λ is a normalization factor. This sum of single particle Green's functions is itself a Green's function. Its spectral function is the effective local DOS

$$\rho(\omega) = -\frac{1}{\pi} \text{Im}[G(\omega + i0^+)] = \frac{1}{\Lambda} \sum_{\mathbf{k}} -\frac{1}{\pi} \text{Im}[G_{\mathbf{k}}(\omega + i0^+)] = \frac{1}{\Lambda} \sum_{\mathbf{k}} \rho_{\mathbf{k}}(\omega). \quad (2.28)$$

We choose $\Lambda = \#\mathbf{k}$, with $\#\mathbf{k}$ being the number of possible momentums, to get a probability density interpretation of the local spectral function. The non-interacting system, described by letting $\Sigma_{\mathbf{k}}(z) = 0$, thus has local Green's function and corresponding spectral density of the form

$$G_0(z) = \frac{1}{\Lambda} \sum_{\mathbf{k}} \frac{1}{z - \epsilon_{\mathbf{k}} + \mu} \quad (2.29)$$

$$\rho_0(\omega) = \frac{1}{\Lambda} \sum_{\mathbf{k}} \rho_{0, \mathbf{k}}(\omega) = \frac{1}{\Lambda} \sum_{\mathbf{k}} \delta(\omega - \epsilon_{\mathbf{k}} + \mu). \quad (2.30)$$

$\rho_0(\omega)$ is the DOS for the non-interacting system. Since the only momentum dependence in eq 2.29 enters through $\epsilon_{\mathbf{k}}$, the momentum sum can be

expressed as an integral over energy with weight $\rho_0(\omega)$,

$$G_0(z) = \frac{1}{\Lambda} \sum_{\mathbf{k}} \frac{1}{z - \epsilon_{\mathbf{k}} + \mu} = \frac{1}{\Lambda} \sum_{\mathbf{k}} \frac{1}{z - \underbrace{(\epsilon_{\mathbf{k}} - \mu)}_{\xi_{\mathbf{k}}}} = \int_{-\infty}^{\infty} d\xi \frac{\rho_0(\xi)}{z - \xi}. \quad (2.31)$$

This is actually just the Hilbert transform between a spectral function and its corresponding Green's function, see eq 2.8.

Chapter 3

Hubbard Model

To describe interacting condensed matter systems, effective models which simplify the full electronic structure problem but still capture the underlying physics are used. Solving the Hamiltonian in eq 1.2 without any approximations is impossible. For strongly interacting condensed matter systems, the electrons are localized in space around lattice sites \mathbf{R}_i . This localization allows us to approximate the matrix elements t_{ij} , V_{ijkl} in eq 1.3,1.4. We choose a basis set with one orbital per lattice site. If we neglect all off-diagonal correlation terms (inter-orbital interactions) so only consider the often dominant intra-orbital repulsion, we get

$$V_{ijkl} \approx U\delta_{ijkl} \quad (3.1)$$

so the interacting Hamiltonian reduces to

$$H_I = U \sum_i n_{i\uparrow} n_{i\downarrow} \quad (3.2)$$

where $n_{i\sigma} = c_{i\sigma}^\dagger c_{i\sigma}$ is the occupation number operator. Double occupation of one site therefore costs U in energy due to the simplified electron-electron interaction. H_I is diagonal in real space and hence tends to make the eigenstates of the full Hamiltonian local in real space.

The hopping overlap element t_{ij} will be small for orbitals far away from each other so it is restricted to only nearest neighbors,

$$t_{ij} \approx -t\delta_{\langle ij \rangle} \quad (3.3)$$

so

$$H_0 = -t \sum_{\langle ij \rangle, \sigma} (c_{i\sigma}^\dagger c_{j\sigma} + h.c.). \quad (3.4)$$

This kinetic energy part is diagonal in momentum space and hence tries to make the eigenstates local in momentum space. This implies extended states in real space.

The total Hamiltonian is thus

$$H = H_0 + H_1 = -t \sum_{\langle ij \rangle, \sigma} (c_{i\sigma}^\dagger c_{j\sigma} + h.c.) + U \sum_i n_{i\uparrow} n_{i\downarrow}. \quad (3.5)$$

This is the *Hubbard Model* which was proposed in 1963 to study interacting electrons in narrow bands [5]. $\frac{U}{t}$ is the only free parameter, determining the localization tendency of the electrons. For $U = 0$, the electrons are free to hop between neighboring sites and the Hamiltonian reduce to the Tight-Binding model which is diagonalized by Fourier transformation,

$$H_0 = -t \sum_{\langle ij \rangle, \sigma} (c_{i\sigma}^\dagger c_{j\sigma} + h.c.) = \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} \quad (3.6)$$

where the dispersion of the non-interacting electrons is

$$\epsilon_{\mathbf{k}} = -t \sum_{\langle 0,j \rangle}^d e^{-i\mathbf{k}\cdot\mathbf{R}_j} \quad (3.7)$$

and d is called the coordination number and represent the number of nearest neighbors. Given $\epsilon_{\mathbf{k}}$, the non-interacting spectral function ρ_0 is easily calculated from eq 2.30. If $t \rightarrow 0$ the dispersion becomes flat, i.e. momentum independent. The Hamiltonian is diagonal in real space with the eigenvalues separated by U .

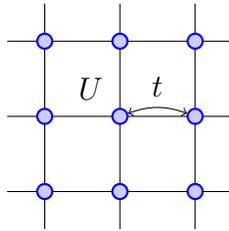


Figure 3.1: The Hubbard model on a square lattice with hopping parameter t and local interaction parameter U .

In this thesis half-filling is considered. For a finite U the ground-state is an anti-ferromagnetic insulator and for $U = 0$ a metal. However this is often a pathology of the rough approximations of only nearest neighbor

hopping. Therefore a paramagnetic solution is required in the calculations. Then a first order transition exists at low temperatures between a metallic (for low U) and an insulating (for high U) phase [6]. For weak interaction, the system is a Fermi liquid metal with dominating quasiparticle excitations. In the insulating regime collective magnetic excitations are present. Their lifetimes are increasing when lowering the interaction strength U [7]. The first non-temperature driven Metal to Insulator Transition (MIT) was proposed by Nevill F. Mott in 1945 and is therefore called the Mott-transition [8]. Half-filling (i.e. one electron per site on average) is obtained by letting the chemical potential

$$\mu = \frac{U}{2}. \quad (3.8)$$

For the non-interacting case the chemical potential is thus zero but for any finite interaction μ is finite. The Dyson equation, relating the non-interacting and interacting propagator, in eq 2.24 is for a constant μ . Hence it needs to be modified since μ change for the half-filled Hubbard model

$$G_{\mathbf{k}\sigma}(z) = \frac{1}{G_{0,\mathbf{k}\sigma}(z)^{-1} + \mu - \Sigma_{\mathbf{k}}(z)}. \quad (3.9)$$

The local Green's functions and corresponding spectral functions for the paramagnetic half-filled Hubbard model thus reads

$$G_0(z) = \frac{1}{\Lambda} \sum_{\mathbf{k}} \frac{1}{z - \epsilon_{\mathbf{k}}} \quad (3.10)$$

$$G(z) = \frac{1}{\Lambda} \sum_{\mathbf{k}} \frac{1}{z - \epsilon_{\mathbf{k}} + \mu - \Sigma_{\mathbf{k}}(z)} \quad (3.11)$$

$$\rho_0(\omega) = -\frac{1}{\pi} \text{Im}[G_0(\omega + i0^+)] = \frac{1}{\Lambda} \sum_{\mathbf{k}} \delta(\omega - \epsilon_{\mathbf{k}}) \quad (3.12)$$

$$\rho(\omega) = -\frac{1}{\pi} \text{Im}[G(\omega + i0^+)]. \quad (3.13)$$

By having the chemical potential $\mu = \frac{U}{2}$ the system has electron-hole symmetry. This implies that the spectral function $\rho(\omega)$ will be an even function. A consequence is that the Matsubara Green's function is purely imaginary since eq 2.7 can be rewritten as

$$G(i\nu_n) = \int_{-\infty}^{\infty} d\omega \underbrace{\frac{-\omega}{\omega^2 + \nu_n^2} \rho(\omega)}_{\text{odd}} + i \int_{-\infty}^{\infty} d\omega \underbrace{\frac{-\nu_n}{\omega^2 + \nu_n^2} \rho(\omega)}_{\text{even}}. \quad (3.14)$$

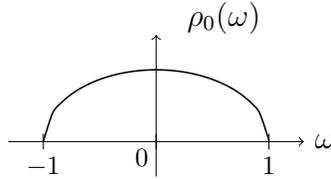


Figure 3.2: Non-interacting spectral function for a Bethe lattice.

Analytic continuation to the real axis can thus be done by matching with just $\text{Im}[G_n]$ or an extended version of the spectral function,

$$\rho(i\nu_n) = -\frac{1}{\pi} \text{Im}[G(i\nu_n)] = \int_{-\infty}^{\infty} d\omega L_{\nu_n}(\omega) \rho(\omega) \quad (3.15)$$

where $L_{\nu_n}(\omega)$ is Lorentzian centered at $\omega = 0$ with characteristic width ν_n .

In principle, one way of solving the Hubbard model is to use a local basis: $\{|\uparrow\rangle, |\downarrow\rangle, |\uparrow\downarrow\rangle, |0\rangle\}$ and diagonalize the hamiltonian. Due to the exponential growth of the number of basis states, 4^N , as function of lattice site N , this can only be done when N is small. Since the thermodynamic limit $N \rightarrow \infty$ is of interest, DMFT is a better choice. It becomes exact when $d \rightarrow \infty$.

3.1 Bethe Lattice

The non-interacting system's eigenvalues, given by eq 3.7, depends on the lattice structure. The DOS for a hyper-cubic lattice becomes gaussian when $d \rightarrow \infty$ [6]. Another structure is the Bethe lattice (Cayley tree). It is not really a lattice but a graph where all sites have d neighbors. If one site is removed the graph is divided into d graphs. It can be divided into two groups so that all neighboring sites are from the other group. The Bethe lattice is commonly used because of the properties of its non-interacting DOS $\rho_0(\omega)$. $\rho_0(\omega)$ has a finite bandwidth when $d \rightarrow \infty$. Low dimensional lattices usually have a finite bandwidth so the Bethe lattice can be seen as a approximation of such a lattice. It also allows the local, momentum independent, Green's function, in eq 3.11, to be calculated in a closed form, which simplifies the DMFT equations. For infinite d , bandwidth $W = 4t$ and a renormalized hopping $t \rightarrow \frac{t}{\sqrt{d}}$, the DOS is [9]

$$\rho_0(\omega) = \begin{cases} \frac{2}{\pi} \sqrt{1 - \left(\frac{2\omega}{W}\right)^2}, & \text{if } |\omega| < \frac{W}{2}, \\ 0, & \text{otherwise} \end{cases} \quad (3.16)$$

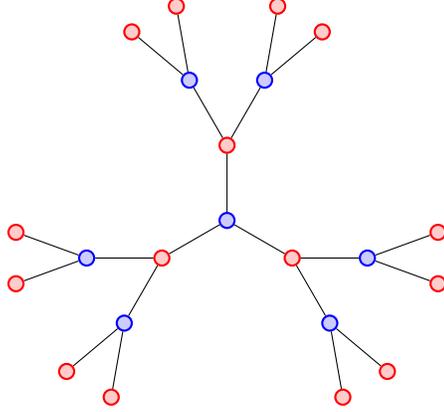


Figure 3.3: Bethe lattice for coordination number $d = 3$.

This semi-circular DOS, see figure 3.2, is thus the solution to the Hubbard model on a Bethe lattice for $U = 0$. The corresponding Green's function $G_0(z)$ becomes according to the Hilbert transform in eq 2.8

$$G_0(z) = \frac{8}{W^2} (z - \text{sgn}(\text{Im}[z])) \sqrt{z^2 - \frac{W^2}{4}}. \quad (3.17)$$

Chapter 4

Dynamic Mean Field Theory (DMFT)

DMFT is a many-body method that maps the original lattice problem to an impurity problem which is easier to solve. This mapping is exact for lattices with infinite coordination number. The impurity model parameters a priori are not known, but get a self-consistent equation that normally is solved by fixpoint iteration. The converged Green's function is the wanted lattice Green's function. This chapter uses the Hubbard Hamiltonian, eq 3.5, on a Bethe lattice. DMFT can be used for other models and lattices, e.g. multi-orbital and multi-impurity sites.

As mentioned earlier, the hopping term is appropriately scaled $t \rightarrow \frac{t}{\sqrt{d}}$ when taking the limit of $d \rightarrow \infty$. For Hamiltonians with only local interaction, only *local* dressed skeleton diagrams contribute in the standard diagrammatic expansion with this scaling. This causes the total self-energy Σ to be momentum independent [7, 10],

$$\Sigma_{\mathbf{k}}(z) \xrightarrow{d \rightarrow \infty} \Sigma(z), \quad (4.1)$$

which in real space mean that the self-energy is completely local $\Sigma_{ij}(z) = \Sigma(z)\delta_{ij}$. For a finite d this is an approximation, neglecting the spatial fluctuations. But, as in classical MFT, becomes more accurate as the number of neighbors increase. For a lattice with low d , the DMFT approximation is thus questionable in principle.

The expression of the interacting local lattice Green's function in eq 3.11

simplifies when the self-energy becomes momentum independent,

$$\begin{aligned}
G(z) &= \frac{1}{\Lambda} \sum_{\mathbf{k}} \frac{1}{z - \epsilon_{\mathbf{k}} + \mu - \Sigma(z)} = \frac{1}{\Lambda} \sum_{\mathbf{k}} \frac{1}{\underbrace{(z + \mu - \Sigma(z))}_x - \epsilon_{\mathbf{k}}} \\
&= \int_{-\infty}^{\infty} d\epsilon \frac{\rho_0(\epsilon)}{x - \epsilon}.
\end{aligned} \tag{4.2}$$

Given the bare spectral function $\rho_0(\omega)$ and the self-energy $\Sigma(z)$, the interacting Green's function can thus be calculated. The mapping of the lattice to the impurity problem is described next.

4.1 Single Impurity Anderson Model (SIAM)

The SIAM describe one impurity site coupled to a set of conducting fermionic bath sites [11]. One representation of the Hamiltonian is

$$H = \underbrace{U n_{\uparrow} n_{\downarrow}}_{\text{local}} + \underbrace{\sum_{i,\sigma} \epsilon_i n_{i\sigma}}_{\text{bath}} + \underbrace{\sum_{i,\sigma} V_i (c_{i\sigma}^{\dagger} c_{\sigma} + h.c.)}_{\text{mix}} \tag{4.3}$$

with ϵ_i bath site i 's energy, U local electron repulsion, c_{σ} (n_{σ}) annihilation (number) operator for the impurity site, $c_{i\sigma}$ ($n_{i\sigma}$) annihilation (number) operator for bath site i and V_i coupling parameter between impurity and bath site i . The chemical potential μ is still $\mu = \frac{U}{2}$ but acts only on the impurity site.

The hybridization function $\Delta(z)$ captures the effect of the bath states coupling to the impurity. It is given by

$$\Delta(z) = \sum_i \frac{V_i^2}{z - \epsilon_i} \tag{4.4}$$

and acts like a self-energy for the local non-interacting Green's function due to the bath coupling

$$g_0(z) = \frac{1}{z - \Delta(z)}. \tag{4.5}$$

For the impurity problem, the interaction is only on one site. The dressed propagator from the impurity site to the impurity site, g , is expressed through g_0 and the impurity self-energy, σ , in the impurity Dyson equation

$$g(z) = \frac{1}{g_0^{-1}(z) + \mu - \sigma(z)}. \tag{4.6}$$

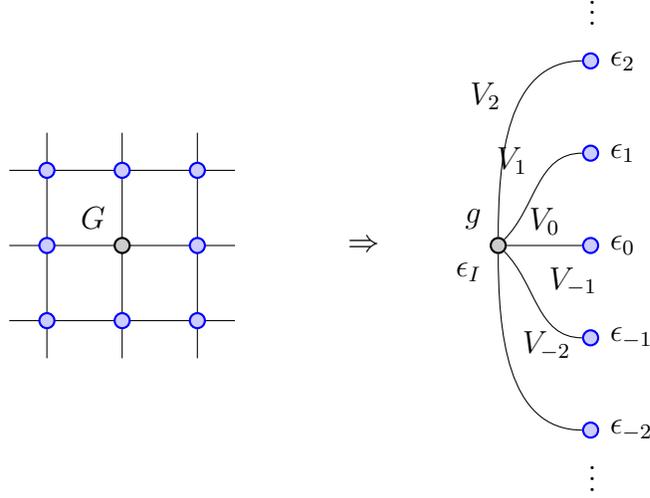


Figure 4.1: The lattice problem and the associated SIAM.

Inserting the hybridization expression of g_0 from eq 4.5 gives

$$g(z) = \frac{1}{z + \mu - \Delta(z) - \sigma(z)}. \quad (4.7)$$

There exists many methods that approximately solve a SIAM, so g is obtained, for a given hybridization function Δ and local impurity parameters U and μ . Methods solving a SIAM usually are called impurity solvers. Many different impurity solvers exist. Diagonalizing the Hamiltonian and use the Lehmann representation to calculate $g(z)$ is called Exact Diagonalization (ED). Due to the exponential scaling of the number of sites, only systems with moderate number of bath sites can be solved. But $g(z)$ becomes known for all z in the upper half-plane so the spectral function can directly be calculated without any analytic continuation. Rewriting the Hamiltonian to a linear chain form, with exponentially decreasing coupling parameters, allows a iterative treatment which is called Numerical Renormalization Group (NRG). Other commonly used impurity solvers are QMC algorithms. They require finite temperature. SIAMs with arbitrary many bath sites can be solved because the hybridization function, not the bath-site parameters V_i, ϵ_i , is needed as input data. But a big drawback is that the Green's function can only efficiently be calculated in the imaginary domain. Therefore an analytic continuation from $G(\tau)$ is necessary to get $\rho(\omega)$. In this thesis Continuous Time Quantum Monte Carlo (CTQMC) is used, see section 4.3.

4.2 Self-Consistent Equation

The mapping of the lattice problem onto a SIAM is the core of DMFT [9], see figure 4.1. There is one SIAM such that the dressed impurity site's Green's function $g(z)$ and local self-energy $\sigma(z)$ obey

$$G(z) = g(z) \tag{4.8}$$

$$\Sigma(z) = \sigma(z). \tag{4.9}$$

This can be shown either by observing that exactly the same local dressed diagrams occur in the two systems or by a cavity construction of the action [6, 7]. The bare local Green's function, $g_0(z)$, is however unknown ($g_0(z) \neq G_0(z)$).

The two Dyson equations 4.2,4.6 together with the mapping equations 4.8,4.9 and the impurity solver forms a set of coupled equations that can be used to obtain the 5 unknown functions $G, g, g_0, \Sigma, \sigma$. For a fix ρ_0, T and U , the 5 coupled equations are usually solved self-consistently in a iterative fashion, see figure 4.2,:

1. Make an initial guess of the non-interacting impurity Green's function g_0 .
2. Solve the SIAM using e.g. ED or CTQMC.
3. Use the obtained impurity self-energy σ as lattice self-energy, $\sigma \rightarrow \Sigma$, eq 4.9.
4. Σ inserted in the lattice Dyson equation 4.2 gives the dressed lattice Green's function G .
5. The mapping equation 4.8 is then used $G \rightarrow g$
6. A new version of g_0 is now calculated using the impurity Dyson equation 4.6.

Steps 2-6 are repeated until convergence is reached.

4.2.1 DMFT on a Bethe Lattice

For some special bare lattice DOS, $\rho_0(\omega)$, the self-consistent iteration scheme can be simplified. We will now present the simplified DMFT iteration scheme for a semielliptical ρ_0 , eq 3.16, corresponding to the Bethe lattice with $z \rightarrow \infty$. For this lattice, neglecting the momentum dependence for the self-energy

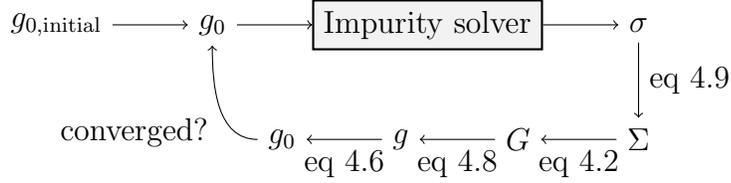


Figure 4.2: DMFT iteration cycle for a general lattice.

is not an approximation. One can also view our system as a translational invariant lattice with a semielliptical bare DOS. If this lattice has $d \neq \infty$, the locality of the self-energy is an approximation. For this particular bare DOS the integral in eq 4.2 is evaluated to

$$G(z) = \frac{8}{W^2} \left(x(z) - \text{sgn}(\text{Im}[x(z)]) \sqrt{x(z)^2 - \frac{W^2}{4}} \right) \quad (4.10)$$

where $x(z) = z + \mu - \Sigma(z)$. The bandwidth W work as an energy reference and without loss of generality we set $W = 2$, such that

$$G(z) = 2 \left(x(z) - \text{sgn}(\text{Im}[x(z)]) \sqrt{x(z)^2 - 1} \right). \quad (4.11)$$

By letting $x = z$ one obtain the non-interacting Green's function in eq 3.17. Expression 4.11 obeys

$$\frac{1}{4}G + \frac{1}{G} = x. \quad (4.12)$$

This expression combined with the impurity Dyson equation 4.6, $G = g$ and $\Sigma = \sigma$ gives a direct coupling between the interacting and non-interacting impurity Green's function,

$$g_0(z) = \frac{1}{z - \frac{1}{4}g(z)}. \quad (4.13)$$

This simplifies the DMFT loop, see figure 4.3, to:

1. Make an initial guess of the non-interacting impurity Green's function g_0 .
2. Solve the SIAM using e.g. ED or CTQMC.
3. Insert the obtained Green's function g into the self-consistency equation 4.13 to get a new version of g_0 .

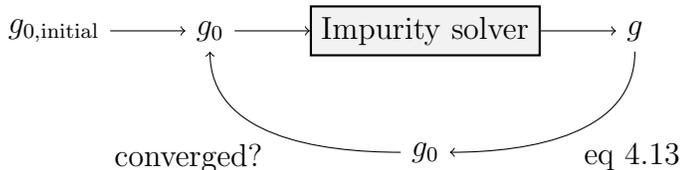


Figure 4.3: DMFT iteration cycle when using the Bethe lattice.

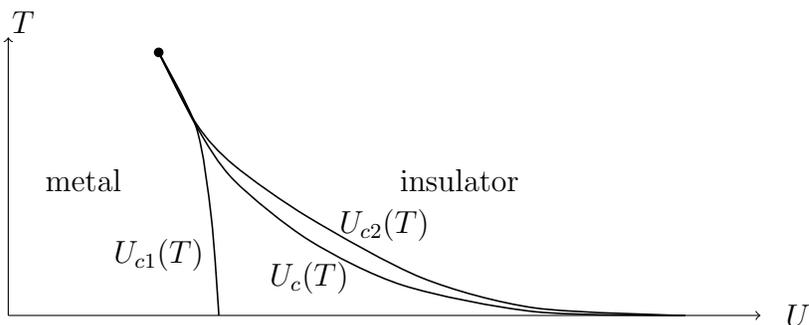


Figure 4.4: Sketch of the paramagnetic Hubbard model’s phase diagram [6]. $U_c(T)$ shows the thermodynamic first-order transition line. $U_{c1}(T)$ ($U_{c2}(T)$) indicate border where insulating (metallic) DMFT solutions can be found.

Steps 2-3 are repeated until convergence is reached. In order to suppress the anti-ferromagnetic phase in the Hubbard model, spin averaging is done between step 2 and 3 in each iteration $g = \frac{1}{2}(g_{\uparrow} + g_{\downarrow})$. In DMFT get a hysteresis region in the (U, T) -phase diagram where both a metallic and an insulating solution exist, see figure 4.4. The solution with the lowest free energy is the equilibrium phase.

4.3 Continuous Time Quantum Monte Carlo (CTQMC)

To calculate the interacting Green’s function g of a SIAM different CTQMC algorithms can be used. We will here just briefly sketch the ideas of how to obtain the Green’s function in the imaginary domain, i.e. $g(\tau), g(i\nu_n)$ or legendre coefficients G_l . The impurity model is specified by the partition function [12]

$$Z = \int \mathcal{D}(c^{\dagger}, c) e^{-S_{\text{eff}}} \quad (4.14)$$

where the effective action is

$$S_{\text{eff}} = - \int_0^\beta \int_0^\beta d\tau d\tau' \sum_{\sigma} c_{\sigma}^{\dagger}(\tau) g_0^{-1}(\tau - \tau') c_{\sigma}(\tau') + \int_0^\beta d\tau U n_{\uparrow}(\tau) n_{\downarrow}(\tau) \quad (4.15)$$

and $g_0(\tau)$ is the Fourier transform of the Matsubara Green's function

$$g_0(i\nu_n) = \frac{1}{i\nu_n - \Delta(i\nu_n)}. \quad (4.16)$$

In the Continuous Time Hybridization (CT-HYB) algorithm, the hybridization function Δ is used as a perturbation parameter to expand the partition sum Z . This algorithm approximates the hybridization but treats the interactions exact, making it useful for strongly interacting systems. The configurations to be summed over are $\mathcal{C} = (\{\sigma_1, \tau_1, \tau'_1\}, \{\sigma_2, \tau_2, \tau'_2\}, \dots, \{\sigma_n, \tau_n, \tau'_n\})$, where n is the (variable) number of segments of occupied fermion states and $\sigma_j, \tau_j, \tau'_j$ denotes the spin, endtime and starttime for segment j . The probability weight for a configuration \mathcal{C} , used in the QMC Markov chain, is $|w(\mathcal{C})|$ where $w(\mathcal{C})$ essentially is the trace of the configuration, $\text{Tr}(\prod_{i=1}^n c_{\sigma_i}^{\dagger}(\tau_i) c_{\sigma_i}(\tau'_i))$, times the determinant of the hybridization matrix $\Delta_{\sigma_i \sigma_j}(\tau_i - \tau'_j)$. The partition function and the average of an arbitrary function $f(\mathcal{C})$ are

$$Z = \sum_{\mathcal{C}} w(\mathcal{C}) \quad (4.17)$$

$$\langle f(\mathcal{C}) \rangle = \frac{1}{Z} \sum_{\mathcal{C}} f(\mathcal{C}) w(\mathcal{C}). \quad (4.18)$$

The Green's function is the functional derivative of Z with respect of the hybridization function

$$G(\tau) = -\frac{1}{Z} \frac{\partial Z}{\partial \Delta(\tau)}. \quad (4.19)$$

4.3.1 Legendre Polynomials

The MC sampling of the Green's function in imaginary time can be done using a finite Legendre polynomial basis. Instead of measuring the values of the Green's function at discrete imaginary times, Legendre polynomial coefficients are calculated. It is a compact representation since the main features are captured with a few Legendre polynomials. By using a finite number of polynomials, high frequency features are suppressed, and therefore it works as an efficient statistical noise filter [12]. Expanding $G(\tau)$ for $\tau \in$

$[0, \beta]$ in Legendre polynomials $P_l(x)$, where $x \in [-1, 1]$, gives

$$G(\tau) = \sum_{l=0}^{\infty} G_l \frac{\sqrt{2l+1}}{\beta} P_l\left(2\frac{\tau}{\beta} - 1\right) \quad (4.20)$$

$$G_l = \sqrt{2l+1} \int_0^{\beta} d\tau P_l\left(2\frac{\tau}{\beta} - 1\right) G(\tau). \quad (4.21)$$

Setting this equal to the Hilbert transformation expression in eq 2.9, then multiplying with P_l and integrating over $[0, \beta]$ gives

$$G_l = \int_{-\infty}^{\infty} d\omega \rho(\omega) \underbrace{\int_0^{\beta} d\tau \frac{-\sqrt{2l+1} P_l\left(2\frac{\tau}{\beta} - 1\right) e^{-\tau\omega}}{1 + e^{-\beta\omega}}}_{k_l(\omega)}. \quad (4.22)$$

By using the fact that $\int_{-1}^1 dx P_l(x) e^{-ax} = 2i^l j_l(ia)$ where j_l is the spherical Bessel function, the kernel $k_l(\omega)$ is evaluated to

$$k_l(\omega) = \frac{-\sqrt{2l+1} \beta i^l e^{-\frac{\beta\omega}{2}} j_l\left(i\frac{\beta\omega}{2}\right)}{1 + e^{-\beta\omega}}. \quad (4.23)$$

k_l is real for all l . For even l , k_l is even and non-positive and for odd l , k_l is odd. Hence for an even spectral function, $G_l = 0$ for odd l and non-positive for even l .

4.4 Simulations

The DMFT calculations for the Bethe lattice, sampling Legendre coefficients, were performed with the TRIQS project [13] according to figure 4.3. TRIQS implements the theory in the articles [12, 14, 15]. Since $G(i\nu_n)$ and G_l are coefficients in a basis expansion of $G(\tau)$, changing from one basis representation to the other is straightforward. Different U is considered to obtain a MIT. It is done for $\beta = 150$ and $G(i\nu_n)$ is plotted. But first we analyze calculations for $U = 4$ and $\beta = 100$.

The DMFT convergence is shown in figure 4.5 for $U = 4$ and $\beta = 100$. As starting point the interacting Green's function G is chosen to be equal to the non-interacting Green's function G_0 . The norm of the change of G_l , with G_l seen as a vector with l_{\max} elements, after one DMFT iteration is measured as a function of the iteration number. $l_{\max} = 80$ and 16 parallel processes, each with $N_c = 10^6$ cycles and 200 QMC-steps within one cycle, are used in the QMC impurity solver. One measurement per cycle is done and for burn-in

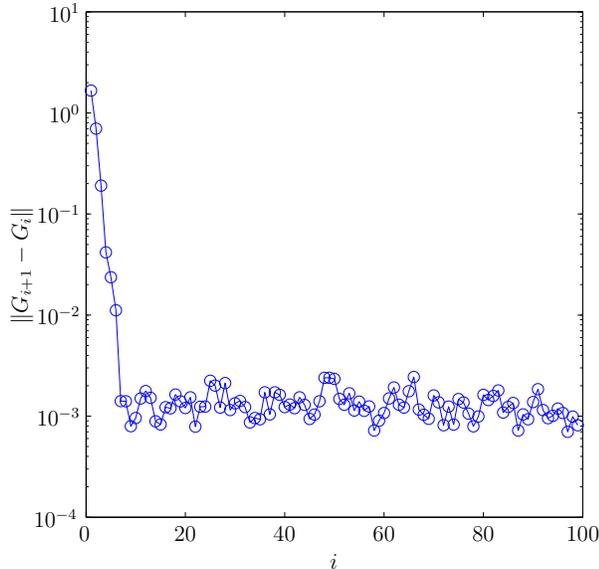


Figure 4.5: Convergence of legendre coefficients as function of DMFT-iteration number i .

$\frac{N_c}{10}$ cycles extra are performed. The DMFT self-consistent equations converge fast to a plateau. Due to statistical noise and a finite number of Legendre coefficients total convergence is not reached. If the norm of the change of G_l reaches 10^{-3} is the DMFT fixpoint iteration considered to be converged.

A finite number of Legendre coefficients, l_{\max} , are used since they are expected to decrease faster than any power of $\frac{1}{l}$ [12]. The number l_{\max} shall be sufficiently big so $G(\tau)$ can be well represented but low enough to suppress statistical noise from rapidly oscillating Legendre functions. For a given bare Green's function propagator, the QMC measurements of the different Legendre coefficients are independent of each other. So a given coefficient G_l is independent of the total number of Legendre coefficients l_{\max} . But in DMFT the calculated Green's function is used to calculate a new bare Green's function. This makes different G_l values dependent of l_{\max} . Simulations with the same settings as above but for different l_{\max} are done. The l_{\max} dependence for three different Legendre coefficients is shown in figure 4.6. The error-bars are standard deviations of converged DMFT-data. The coefficient values for $l_{\max} = 110$ is removed to be able to visualize the three G_l -values in the same figure. For $l_{\max} \geq 50$, G_l is saturated within the statistical noise.

Variations of G_l , for even l , between converged DMFT-iterations are seen in figure 4.7. We notice that the standard deviations are independent of l and

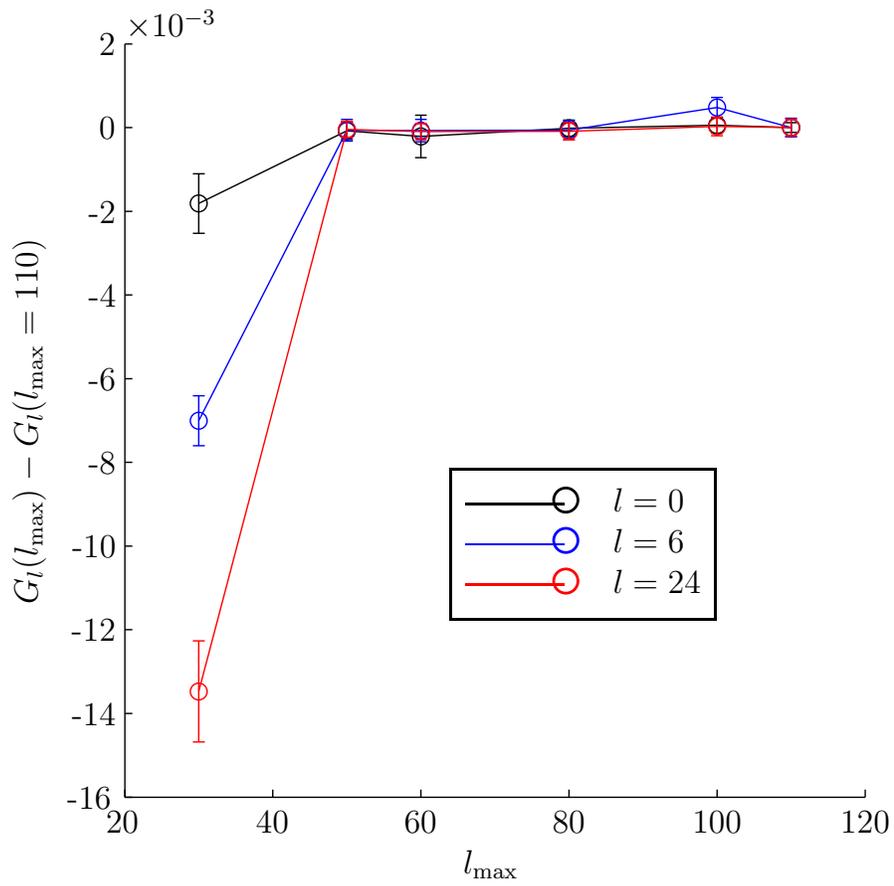


Figure 4.6: Convergence of G_l as function of l_{\max} is reached for $l_{\max} = 50$.

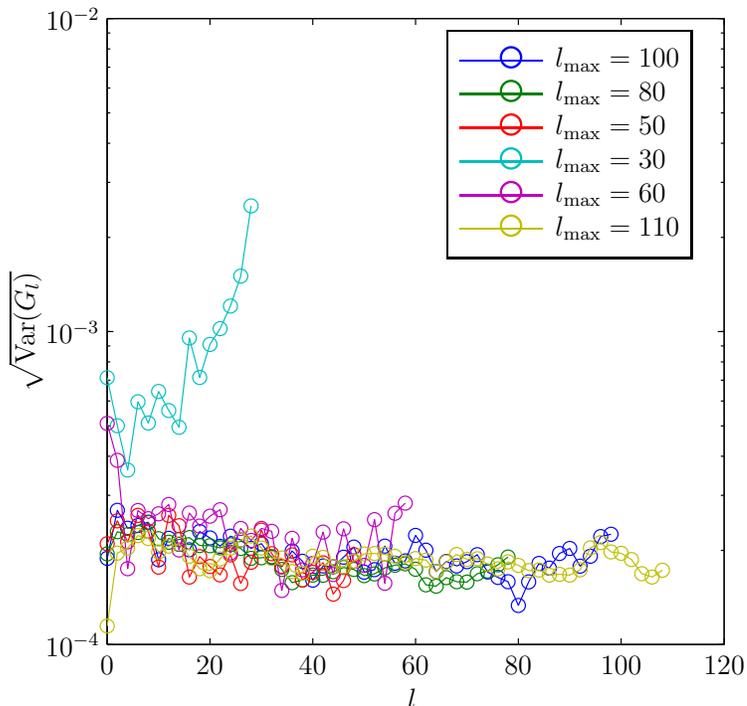


Figure 4.7: Standard deviations of the DMFT solutions for even l .

l_{\max} except for the calculation using $l_{\max} = 30$. Hence, for $50 < l_{\max} < 110$ do we expect to get good results and variations of the same order of magnitude.

The correlation length, see section 5.2.1, is calculated to investigate if G_l from different DMFT iterations are correlated with each other. Since the correlation length τ is lower than one, see figure 4.8, G_l from different DMFT iterations are independent. G_l for odd l are not shown since they are not used in the analytic continuation.

$U = 1, 2, 2.5, 4$ for $\beta = 150$ are also studied. The temperature is low enough so the system is within the coexistence region of the (U, T) phase diagram for $U = 2.5$. The DMFT settings are the same as above except for $U = 2.5$ where 160 instead of 16 parallel process are used. The number of Legendre coefficients is $l_{\max} = 80$. After DMFT fixpoint convergence, $N_i > 20$ additional DMFT iterations are performed. Hence Legendre coefficients $G_{l,i}$ and Fourier coefficients $G_{n,i}$, where $i \in \{1, 2, \dots, N_i\}$, are obtained. This allow us to estimate the standard deviations of G_l and G_n , which are used in MEM. The Legendre coefficients for even l are seen in figure 4.9. The corresponding Matsubara Green's function $G(\tau)$ and its Fourier coefficients G_n are seen in figures 4.10 and 4.10. The known non-interacting systems

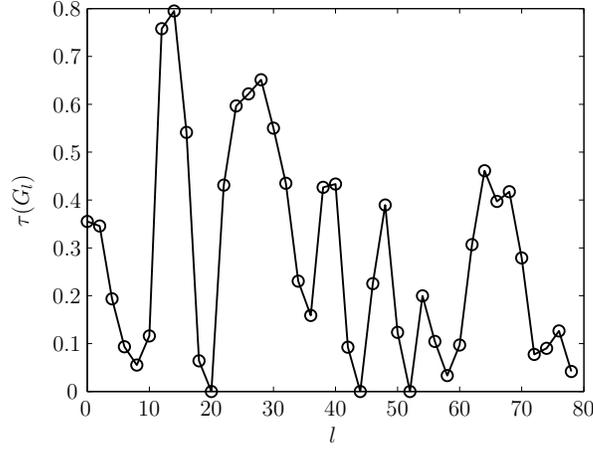


Figure 4.8: Correlation between different DMFT-iterations for G_l .

Green's function, see section 3, is added in the figures for comparison. Due to anti-periodicity and $G_\sigma(\tau = 0^-) = \#$ electrons on average with spin σ per site, half-filling is verified.

Since $\rho(0) = -\frac{1}{\pi} \text{Im} G(i0^+)$, the Fourier values tell us if a system is metallic or insulating. $U = 0, 1, 2, 2.5$ thus give us a metal and $U = 2.5, 4$ give us an insulator within DMFT.

An improvement, which was not implemented due to lack of time, would be to put $G_l = 0$ for odd l in every DMFT iteration step. This would improve the accuracy in the step where the new bare Greens' function is calculated.

Next chapter will be about how to perform the analytic continuation of the CTQMC calculated Green's functions in figures 4.9 and 4.10.

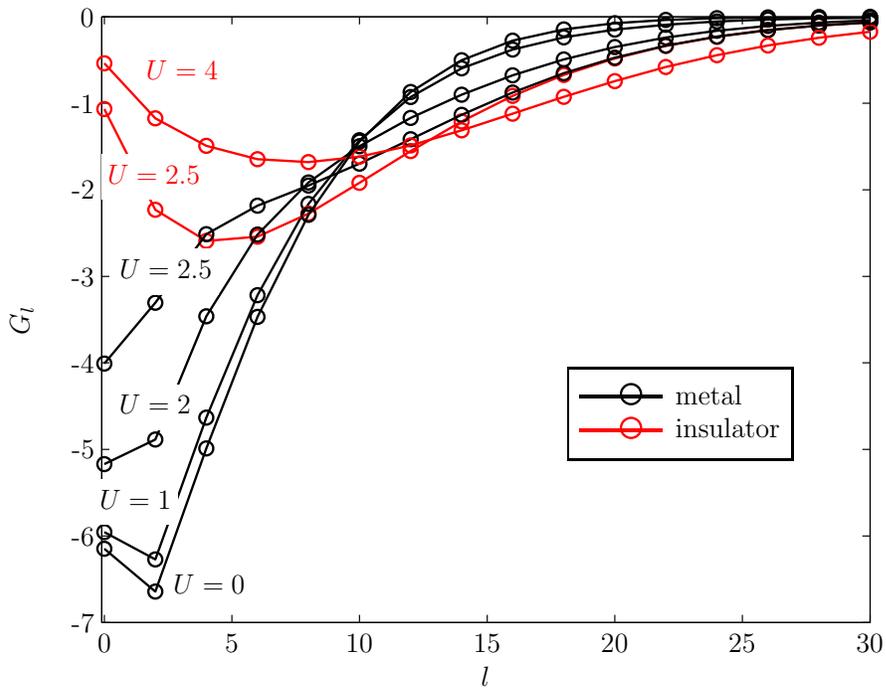


Figure 4.9: Legendre coefficients for $\beta = 150$.

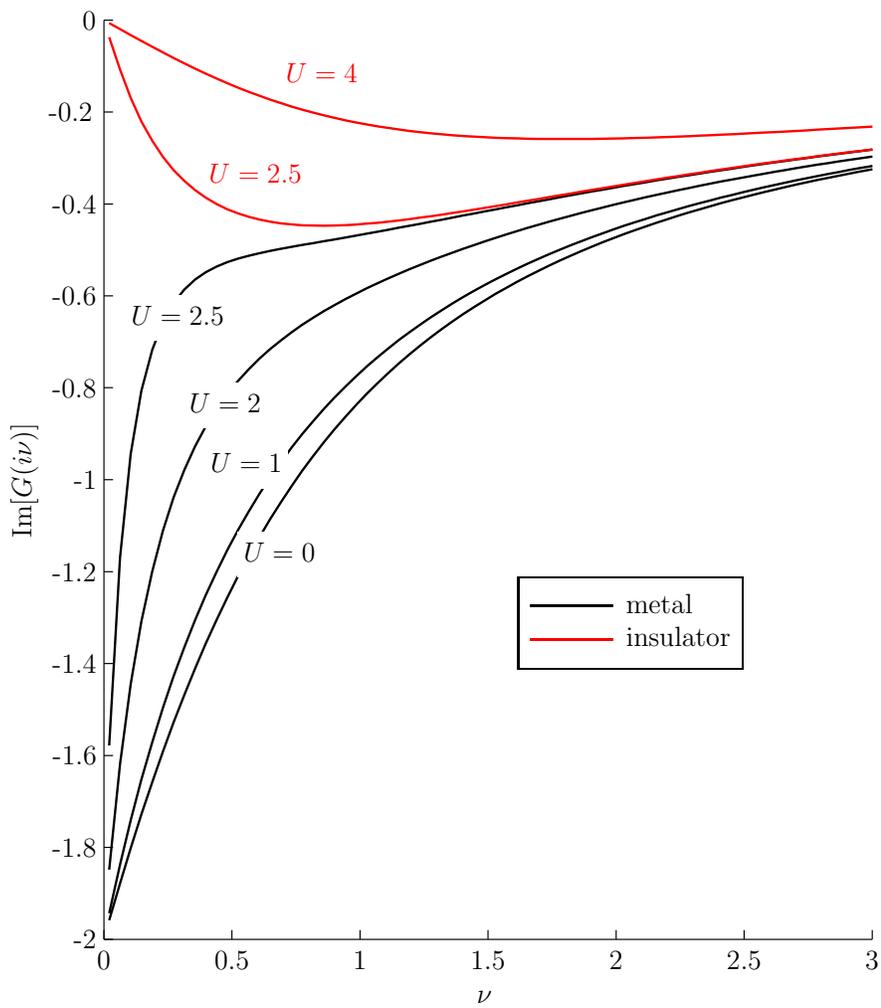


Figure 4.10: Green's function $G(i\nu_n)$ for $\beta = 150$.

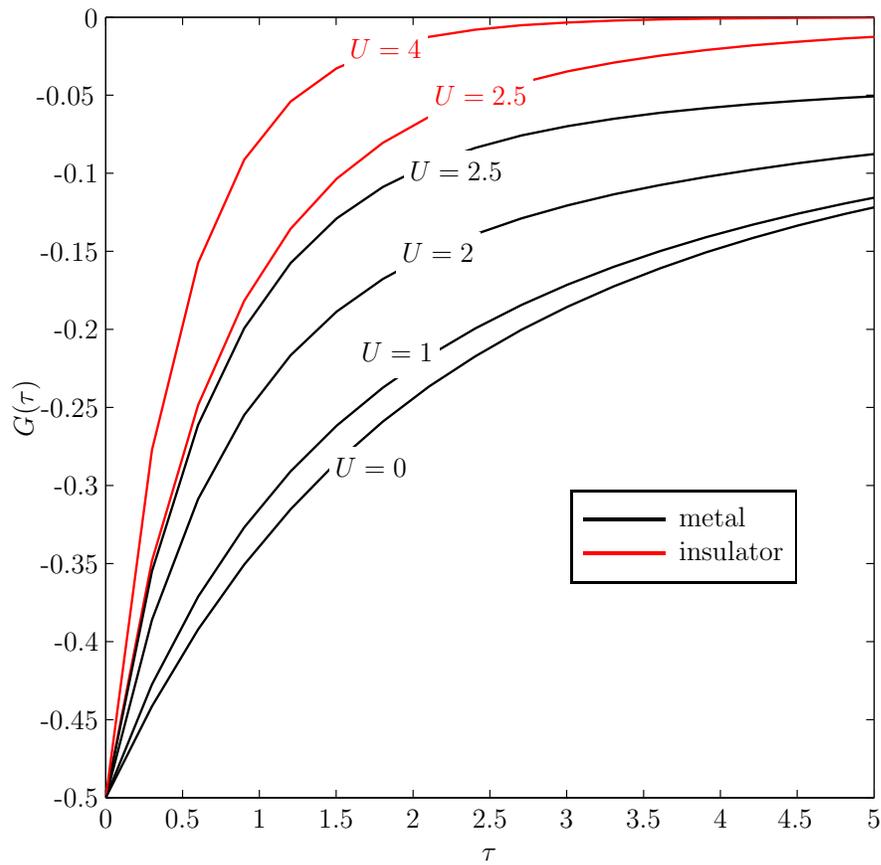


Figure 4.11: $G(\tau)$, for small positive τ -values where the U dependence is most prominent.

Chapter 5

Maximum Entropy Method (MEM)

MEM is a general method for solving inverse problems. Until section 5.8, MEM is therefore discussed for a general inverse integral problem. The analytic continuation, using the Hilbert transformation in eq 2.7, is an inverse integral problem since the Green's function is easy to calculate on the imaginary axis given it's values on the real axis, but not the other way around. In section 5.8 the specific kernels used for the analytic continuation are summarized. The obtained spectral functions, using some of these kernels, are presented in section 5.9.

An integral inverse problem is of the form

$$G_i = G_i[\rho] = \int d\omega k_i(\omega)\rho(\omega) \quad , \text{for } i \in \{1, 2, \dots, N_G\} \quad (5.1)$$

where ρ is the wanted unknown function, G_i is a set of N_G known values and $k_i(\omega)$ is a known kernel. The brackets around ρ indicate that G_i is a functional of the function ρ .

Unfortunately the kernels for the analytic continuation, like many inverse problems, are extremely ill-posed. This mean that small changes in G correspond to big changes in ρ . Another way of formulating the difficulties is that there exist many ρ configurations with the same G . Actually an infinite-dimensional manifold of solutions that satisfies eq 5.1 exists if N_G is finite [16]. We are thus required to somehow select one solution from this manifold. There are many ways to do that. In MEM it is done by maximizing the information entropy for the spectral function [17].

Another issue which makes the inverse problem even harder is that one usually does not have infinite accuracy of the input data G . This is the case for the analytical continuation problem where the Green's function for

imaginary argument is measured by a Monte Carlo average and therefore has statistical noise [14]. Let's call this measured data \tilde{G} .

MEM view the spectral function ρ as a probability distribution. A probability density functional is constructed, saying the likeliness for every spectral function. This functional consists of two parts. The first include how much G_i , in eq 5.1, deviate from the measured input data \tilde{G} , where

$$\Delta G_i = \tilde{G}_i - G_i[\rho]. \quad (5.2)$$

The second part is what makes MEM special. This factor of the functional takes into account the probability of every spectral function from an entropic point of view. The construction of the functional is done using Bayesian logic.

5.1 Bayesian Inference

In the construction of finding the most likely spectral function, Bayes' theorem is used. For two stochastic variables A and B it says [18]

$$P[A = a, B = b] = P[A = a|B = b] P[B = b] = P[B = b|A = a] P[A = a] \quad (5.3)$$

where $P[A = a, B = b]$ is the joint probability for $A = a$ and $B = b$ and $P[A = a|B = b]$ is the conditional probability of $A = a$ given $B = b$. Let's us use the short-form notation $P[a]$ for $P[A = a]$ in the rest of the thesis. By simply integrating out one of the variables in eq 5.3 one gets

$$P[a] = \int db P[a, b]. \quad (5.4)$$

For normalized distributions,

$$1 = \int da P[a] \quad (5.5)$$

$$1 = \int da P[a|b]. \quad (5.6)$$

An initial guess of how to solve an inverse problem might be to get an expression for a probability functional $P[\rho|\tilde{G}]$ and then average over all possible ρ . It is actually what you do in MEM but with the functional

$$P[\rho|m, \tilde{G}] \quad (5.7)$$

where $m = m(\omega)$ is a Lebesgue measure of the distribution-space for ρ [17]. $m(\omega)$ is often called the default model because it is the MEM solution in

the absence of input data. The choice of the default model is hence an a priori guess of the spectral function. Effectively it works as a way of implementing prior knowledge about the spectral function, e.g. normalization, characteristic features. This is explained more in section 5.3.

For a given default model m and a measured input data \tilde{G} the MEM solution to the inverse problem is given by

$$\langle \rho \rangle(m, \tilde{G}) = \int \mathcal{D}\rho P[\rho|m, \tilde{G}] \rho. \quad (5.8)$$

It is important to keep in mind that the function $\langle \rho \rangle(m, \tilde{G})$ is ω -dependent even if not expressed explicitly in the above formula.

A positive hyper-parameter α is introduced in MEM to balance between maximizing the entropy and making the difference in eq 5.2 small. By using first eq 5.4 and then eq 5.3, $P[\rho|m, \tilde{G}]$ can be expanded into

$$P[\rho|m, \tilde{G}] = \int_0^\infty d\alpha P[\rho, \alpha|m, \tilde{G}] = \int_0^\infty d\alpha P[\alpha|m, \tilde{G}] P[\rho|\alpha, m, \tilde{G}]. \quad (5.9)$$

Inserting eq 5.9 into eq 5.8 gives

$$\langle \rho \rangle(m, \tilde{G}) = \int_0^\infty d\alpha P[\alpha|m, \tilde{G}] \underbrace{\int \mathcal{D}\rho P[\rho|\alpha, m, \tilde{G}] \rho}_{\langle \rho \rangle(\alpha, m, \tilde{G})}. \quad (5.10)$$

Notice that $\langle \rho \rangle(\alpha, m, \tilde{G})$ is the average distribution for ρ given α , m and \tilde{G} . This function is then just averaged over α with the probability weight $P[\alpha|m, \tilde{G}]$.

But the posteriori distributions with respect to \tilde{G} , $P[\alpha|m, \tilde{G}]$ and $P[\rho|\alpha, m, \tilde{G}]$ are still unknown. The rest of this section describes how they can be determined using Bayesian rules. We start with the joint probability distribution for ρ , α , m and \tilde{G} and apply the Bayesian rule in eq 5.3 repeatedly

$$\begin{aligned} P[\rho, \alpha, m, \tilde{G}] &= P[\tilde{G}|\rho, \alpha, m] P[\rho, \alpha, m] \\ &= P[\tilde{G}|\rho, \alpha, m] P[\rho|\alpha, m] P[\alpha, m]. \end{aligned} \quad (5.11)$$

We note that $P[\rho|\alpha, m]$ is the a priori distribution for ρ . It is very important and in MEM determined by the entropy, see section 5.3.

Another very important observation is that $P[\tilde{G}|\rho, \alpha, m]$ has the inverse conditioning to the original distribution, in eq 5.7, for ρ and \tilde{G} . It is thus a likelihood function, saying the probability of \tilde{G} given ρ , depending on the differences in eq 5.2. α and m does not influence the likelihood function so $P[\tilde{G}|\rho, \alpha, m] = P[\tilde{G}|\rho]$. The likelihood function is investigated in section 5.2.

Let's now assume the entropic prior $P[\rho|\alpha, m]$ and the likelihood function $P[\tilde{G}|\rho]$ are known. Expanding the joint distribution $P[\rho, \alpha, m, \tilde{G}]$ again, using the Bayesian rule, gives

$$\begin{aligned} P[\rho, \alpha, m, \tilde{G}] &= P[\rho, \alpha, m|\tilde{G}] P[\tilde{G}] = P[\rho, \alpha|m, \tilde{G}] P[m|\tilde{G}] P[\tilde{G}] \\ &= P[\rho|\alpha, m, \tilde{G}] P[\alpha|m, \tilde{G}] P[m|\tilde{G}] P[\tilde{G}]. \end{aligned} \quad (5.12)$$

Combining now eq 5.11 and 5.12 gives

$$P[\rho|\alpha, m, \tilde{G}] = \frac{1}{Z_\rho} P[\tilde{G}|\rho] P[\rho|\alpha, m] \quad (5.13)$$

where the likelihood function and the entropic functional are the last two factors. $Z_\rho = Z_\rho(\alpha, m, \tilde{G}) = P[\alpha|m, \tilde{G}] P[m|\tilde{G}] P[\tilde{G}] / P[\alpha, m]$ is just a normalization factor and does not depend on ρ . By using that $P[\rho|\alpha, m, \tilde{G}]$ is a normalized Probability Density Function (PDF) and integrating eq 5.13 over ρ gives the partition sum

$$Z_\rho = \int \mathcal{D}\rho P[\tilde{G}|\rho] P[\rho|\alpha, m]. \quad (5.14)$$

By combining eq 5.13 and 5.14 we get a useful expression for calculating $\langle \rho \rangle(\alpha, m, \tilde{G})$ defined in eq 5.10

$$\langle \rho \rangle(\alpha, m, \tilde{G}) = \frac{1}{Z_\rho} \int \mathcal{D}\rho P[\tilde{G}|\rho] P[\rho|\alpha, m] \rho. \quad (5.15)$$

What is left in MEM now is just an expression for $P[\alpha|m, \tilde{G}]$ so eq 5.10 can be evaluated. Using the Bayesian rule again: $P[\rho, \alpha, m, \tilde{G}] = P[\rho, \alpha, \tilde{G}|m] P[m]$. Combine it with eq 5.11 gives

$$P[\rho, \alpha, \tilde{G}|m] = \frac{1}{P[m]} P[\tilde{G}|\rho] P[\rho|\alpha, m] P[\alpha, m]. \quad (5.16)$$

The posteriori distribution for α thus reads

$$P[\alpha|m, \tilde{G}] = \int \mathcal{D}\rho P[\rho, \alpha|m, \tilde{G}] = \int \mathcal{D}\rho \frac{P[\rho, \alpha, \tilde{G}|m]}{P[\tilde{G}|m]} = \frac{P[\alpha, m] Z_\rho}{P[m] P[\tilde{G}|m]}. \quad (5.17)$$

α and m are a priori uncorrelated: $P[\alpha, m] = P[\alpha] P[m]$ so eq 5.17 simplifies to

$$P[\alpha|m, \tilde{G}] = \frac{1}{P[\tilde{G}|m]} P[\alpha] Z_\rho \quad (5.18)$$

where $P[\tilde{G}|m]$ works just as a normalization parameter when $P[\alpha|m, \tilde{G}]$ is used in eq 5.10. Both $P[\alpha]$ and Z_ρ depend on α and are described in detail in section 5.5.

One can also check how likely a certain default model is by calculating $P[m|\tilde{G}]$. It is done by integrating eq 5.18 over α ;

$$P[\tilde{G}|m] = \int_0^\infty d\alpha P[\alpha] Z_\rho \quad (5.19)$$

and using eq 5.3, resulting in

$$P[m|\tilde{G}] = \frac{P[m]}{P[\tilde{G}]} \int_0^\infty d\alpha P[\alpha] Z_\rho. \quad (5.20)$$

Since the normalization evidence $P[\tilde{G}]$ is unknown and can't be calculated, eq 5.20 can only be used for comparison of different default models [19].

5.2 Likelihood function

If the input data \tilde{G}_i , $i \in \{1, 2, \dots, N_G\}$ are assumed to be averages of N_s measurements,

$$\tilde{G}_i = \frac{1}{N_s} \sum_{k=1}^{N_s} \tilde{G}_{i,k}, \quad (5.21)$$

\tilde{G}_i will due to the central limit theorem become gaussian distributed as $N_s \rightarrow \infty$. The likelihood function then reads

$$P[\tilde{G}|\rho] = \frac{1}{Z_L} e^{-\frac{\chi^2}{2}} \quad (5.22)$$

with the normalization factor

$$Z_L = \int \prod_{i=1}^{N_G} d\tilde{G}_i e^{-\frac{\chi^2}{2}} \quad (5.23)$$

where χ^2 depends on the kernels $k_i(\omega)$ and if $\tilde{G}_{i,k}$ and $\tilde{G}_{i',k'}$ are correlated.

Let's at first assume no correlations between $\tilde{G}_{i,k}$ and $\tilde{G}_{i,k'}$ for $k \neq k'$. For a MC measurement, e.g. the one described in section 4.3, this means that the correlation length is zero. This is never the case for a MC Markov chain but can be achieved by e.g. binning analysis, see section 5.2.1 [2]. We also assume no correlations between different observables at the same step in the Markov chain, thus $\tilde{G}_{i,k}$ and $\tilde{G}_{i',k}$ for $i \neq i'$ are uncorrelated. Actually the

collected data presented in section 4.4 are not from a single Markov chain but from many chains put together forming a DMFT-chain. This is why the correlation length in figure 4.8 are almost zero. If \tilde{G}_i is real, the kernel $k_i(\omega)$ is real and χ^2 is given by

$$\chi^2 = \sum_{i=1}^{N_G} \frac{(\Delta G_i)^2}{\sigma_i^2} \quad (5.24)$$

where ΔG_i is the difference between \tilde{G} and G and σ_i is the standard deviation for \tilde{G}_i . The likelihood is then just a multivariate gaussian distribution [19] with normalization

$$Z_L = \prod_{i=1}^{N_G} \int_{-\infty}^{\infty} d\tilde{G}_i e^{-\frac{1}{2} \frac{\Delta G_i^2}{\sigma_i^2}} = \prod_{i=1}^{N_G} \sqrt{2\pi} \sigma_i = (2\pi)^{\frac{N_G}{2}} \prod_{i=1}^{N_G} \sigma_i. \quad (5.25)$$

If $k_i(\omega)$ is generalized to being complex, $\tilde{G}_i \in \mathbb{C}$ and χ^2 is instead given by

$$\chi^2 = \sum_{i=1}^{N_G} \frac{|\Delta G_i|^2}{\sigma_i^2}. \quad (5.26)$$

The likelihood is still just a multivariate gaussian distribution but for twice as many variables. Its normalization is

$$Z_L = (2\pi)^{N_G} \prod_{i=1}^{N_G} \sigma_i^2. \quad (5.27)$$

For correlated $\tilde{G}_{i,k}$, methods for creating uncorrelated observables are needed, see section 5.2.1.

5.2.1 Correlations

There are two possible types of correlations of the data $\tilde{G}_{i,k}$. Chain correlations can effectively be eliminated by binning analysis. The other issue is that different observables can be correlated (within the same chain step). It is solved by diagonalizing the covariance matrix.

Binning Analysis

Given a set of eventually correlated samples $O_i^{(0)}$, with $i \in \{1, 2, \dots, N\}$, from a distribution, the auto correlation length τ can be estimated by iteratively creating new sets of data according to

$$O_i^{(l)} = \frac{1}{2} \left(O_{2i-1}^{(l-1)} + O_{2i}^{(l-1)} \right), \text{ for } i \in \left\{ 1, 2, \dots, N_l \equiv \frac{N}{2^l} \right\}. \quad (5.28)$$

These bin values becomes less correlated as l increases, even if the mean value remains the same. The correlation length is [2]

$$\tau = \frac{1}{2} \left(\frac{\sigma^2}{\frac{\text{Var}[O_i^0]}{N}} - 1 \right) \quad (5.29)$$

where σ is the standard deviation of the mean value of O_i^0 , which can be approximated by the binning values

$$\sigma^2 = \lim_{l \rightarrow \infty} \frac{\text{Var}[O_i^{(l)}]}{N_l}. \quad (5.30)$$

Diagonalization Analysis

Diagonalization analysis generalizes χ^2 in eq 5.24 to

$$\chi^2 = \sum_{i=1}^{N_G} \sum_{j=1}^{N_G} \Delta G_i C_{i,j}^{-1} \Delta G_j \quad (5.31)$$

where C is the hermitian covariance matrix, measuring the variance of each data point \tilde{G}_i individually but also how they are correlated with each other. The element $C_{i,j}$ is an estimator for the covariance of the mean values \tilde{G}_i and \tilde{G}_j

$$C_{i,j} = \frac{1}{N_s(N_s - 1)} \sum_{k=1}^{N_s} \left(\tilde{G}_i - \tilde{G}_{i,k} \right) \left(\tilde{G}_j - \tilde{G}_{j,k} \right). \quad (5.32)$$

The diagonal terms $C_{i,i} = \sigma_i^2$ contains the uncorrelated part. If we ignore the off-diagonal elements, eq 5.31 reduces to eq 5.24. The generalized expression for the likelihood normalization factor Z_L reads [19]

$$Z_L = (2\pi)^{\frac{N_G}{2}} \sqrt{\det C}. \quad (5.33)$$

By diagonalizing C , in eq 5.32, the correlated χ^2 can be written in the form of eq 5.24. Start by diagonalize C ,

$$U^{-1} C U = \begin{pmatrix} \sigma_1'^2 & 0 & \dots & 0 \\ 0 & \sigma_2'^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{N_G}'^2 \end{pmatrix}, \quad (5.34)$$

where U is an unitary matrix. By inserting

$$C_{i,j}^{-1} = \sum_{a=1}^{N_G} \sum_{b=1}^{N_G} U_{i,a} \frac{1}{\sigma_a'^2} \delta_{a,b} U_{b,j}^{-1} \quad (5.35)$$

into eq 5.31, get

$$\chi^2 = \sum_{a=1}^{N_G} \left(\sum_{i=1}^{N_G} U_{a,i}^{-1} \Delta G_i \right) \frac{1}{\sigma_a'^2} \left(\sum_{j=1}^{N_G} U_{a,j}^{-1} \Delta G_j \right). \quad (5.36)$$

By defining the rotated variables

$$\Delta G'_i = \sum_{j=1}^{N_G} U_{i,j}^{-1} \Delta G_j \quad (5.37)$$

$$\tilde{G}'_i = \sum_{j=1}^{N_G} U_{i,j}^{-1} \tilde{G}_j \quad (5.38)$$

$$G'_i = \sum_{j=1}^{N_G} U_{i,j}^{-1} G_j, \quad (5.39)$$

χ^2 is finally expressed as

$$\chi^2 = \sum_{i=1}^{N_G} \frac{(\Delta G'_i)^2}{\sigma_i'^2} \quad (5.40)$$

with $\Delta G'_i = \tilde{G}'_i - G'_i$.

5.3 Entropic Prior

The MEM is characterized by the choice of the prior PDF. It shall incorporate all prior knowledge of the spectral function. Since the spectral function is viewed as a PDF it should be normalized and non-negative. There are different statistical inference arguments [20], [21] that the prior should have the form

$$P[\rho|\alpha, m] = \frac{1}{Z_S} e^{\alpha S[\rho]} \quad (5.41)$$

with S the information theory entropy, defined by

$$S[\rho(\omega)] = \int_{-\infty}^{\infty} d\omega \left(\rho(\omega) - m(\omega) - \rho(\omega) \ln\left(\frac{\rho(\omega)}{m(\omega)}\right) \right). \quad (5.42)$$

The normalization factor, Z_S , reads

$$Z_S(\alpha) = \int \mathcal{D}\rho. \quad (5.43)$$

$m(\omega)$ is the default model incorporating all prior knowledge. The entropy is non-positive and maximized for $\rho = m$. The regularization parameter α determines the relative importance between minimizing the least square fit χ^2 and maximizing the entropy S . For $\alpha = 0$ we have no entropy, so $P[\rho|\alpha, m, \tilde{G}] \propto P[\tilde{G}|\rho]$, thus only fitting to input data. If $\alpha \rightarrow \infty$, so $P[\rho|\alpha, m, \tilde{G}] \propto P[\rho|\alpha, m]$, the optimal spectral function is m . See section 5.5 for the choice of α .

If we have no prior knowledge, a flat $m(\omega)$ is a good choice. But if we have an idea of how the solution may look like, this should be implemented by setting m equal to the guess. The choice will bias the spectral function towards m . With a default model of more structure than there actually exist, this bias may lead to wrong solutions. One can also choose a flat default and use the MEM solution as a new default model. By iterating, this will eventually converge to the right answer. This sounds promising but the risk is that the default model develops non-existing features which will be enforced by the MEM [22].

Other prior PDFs, which also take into account that $\rho(\omega)$ should be a smooth function, has been proposed, which can be achieved by a factor

$$\int_{-\infty}^{\infty} \left(\frac{\partial \rho(\omega)}{\partial \omega} \right)^2 d\omega. \quad (5.44)$$

In order to calculate the normalization factor Z_S in eq 5.43, discretization of ω -space is needed. Let's use N_ω points ω_i and approximate the entropy as

$$S \approx \sum_{i=1}^{N_\omega} f_i \left(\rho_i - m_i - \rho_i \log \frac{\rho_i}{m_i} \right) \quad (5.45)$$

where f_i are the integral weights. We choose the same metric as in [19], thus

$$\mathcal{D}\rho = \prod_{i=1}^{N_\omega} \frac{1}{\sqrt{\rho_i/f_i}} d\rho_i. \quad (5.46)$$

A convenient variable substitution is

$$\frac{\partial \rho_i}{\partial x_j} = \sqrt{\rho_i/f_i} \delta_{ij} \quad (5.47)$$

such that

$$\mathcal{D}\rho = \prod_{i=1}^{N_\omega} \frac{1}{\sqrt{\rho_i/f_i}} d\rho_i = \prod_{i=1}^{N_\omega} dx_i. \quad (5.48)$$

Hence, the normalization factor Z_S in 5.43 becomes

$$Z_S(\alpha) = \int \mathcal{D}\rho e^{\alpha S[\rho]} = \int \prod_{i=1}^{N_\omega} dx_i e^{\alpha S(x)} = \prod_{i=1}^{N_\omega} \int_0^\infty dx_i e^{\alpha f_i(\rho_i - m_i - \rho_i \log \frac{\rho_i}{m_i})}. \quad (5.49)$$

The integrals are approximated by Taylor expanding the exponent in x_i around x'_i , where x'_i correspond to ρ_i equal to m_i ,

$$\alpha f_i \left(\rho_i - m_i - \rho_i \log \frac{\rho_i}{m_i} \right) \approx -\alpha \frac{1}{2} (x_i - x'_i)^2. \quad (5.50)$$

This is a good approximation if $\alpha > 1$, so that only values of x_i close to x'_i contribute in eq 5.45. To get a simple form we assume $e^{-\frac{\alpha}{2} x_i'^2} \ll 1$, negative values can then also safely be integrated. The gaussian integrals give

$$Z_S(\alpha) \approx \prod_{i=1}^{N_\omega} \int_{-\infty}^{\infty} dx_i e^{-\alpha \frac{1}{2} (x_i - x'_i)^2} = \left(\frac{2\pi}{\alpha} \right)^{\frac{N_\omega}{2}}. \quad (5.51)$$

5.4 Q Functional Minimization

Now the likelihood function and the entropic functional are known so $\langle \rho \rangle(\alpha, m, \tilde{G})$, in Eq 5.15, can be studied in detail. We define the functional

$$Q[\rho] = \frac{\chi^2[\rho]}{2} - \alpha S[\rho] \quad (5.52)$$

with χ^2 equal to eq 5.24 if $k_i(\omega)$ is real and eq 5.26 if $k_i(\omega)$ is complex. S is defined in eq 5.42. We have

$$P[\tilde{G}|\rho] P[\rho|\alpha, m] = \frac{1}{Z_L Z_S} e^{-Q[\rho]} \quad (5.53)$$

and

$$Z_\rho = \frac{1}{Z_L Z_S} \int \mathcal{D}\rho e^{-Q[\rho]}. \quad (5.54)$$

Hence we can write

$$\langle \rho \rangle(\alpha, m, \tilde{G}) = \frac{1}{\int \mathcal{D}\rho e^{-Q[\rho]}} \int \mathcal{D}\rho e^{-Q[\rho]} \rho. \quad (5.55)$$

Here $e^{-Q[\rho]}$ is expected to have a sharp maximum, and in MEM one therefore approximates

$$e^{-Q[\rho]} \propto \delta(\hat{\rho}_\alpha - \rho) \quad (5.56)$$

where $\min_{\rho} Q[\rho] = Q[\hat{\rho}_{\alpha}]$. This reduces the problem of averaging over all ρ to just finding the one that minimizes Q , thus

$$\langle \rho \rangle(\alpha, m, \tilde{G}) \approx \hat{\rho}_{\alpha}. \quad (5.57)$$

The MEM spectral function solution, see eq 5.10, to the inverse problem is approximately

$$\langle \rho \rangle(m, \tilde{G}) = \int_0^{\infty} d\alpha P[\alpha|m, \tilde{G}] \hat{\rho}_{\alpha} \quad (5.58)$$

where $P[\alpha|m, \tilde{G}]$ is described in detail in section 5.5.

To justify the approximation in eq 5.56 we have to discretize the ω -space as was done in section 5.3. Using a Taylor expansion of Q around $\hat{\rho}_{\alpha}$, let \hat{x}_{α} correspond to $\hat{\rho}_{\alpha}$ and $\Delta x_i = x_i - \hat{x}_{\alpha,i}$, thus

$$Q(x) \approx Q(x = \hat{x}_{\alpha}) + \sum_{i=1}^{N_{\omega}} \Delta x_i \frac{\partial Q(x)}{\partial x_i} \Big|_{x=\hat{x}_{\alpha}} + \frac{1}{2} \sum_{i=1}^{N_{\omega}} \sum_{j=1}^{N_{\omega}} \Delta x_i \Delta x_j \frac{\partial^2 Q(x)}{\partial x_i \partial x_j} \Big|_{x=\hat{x}_{\alpha}}. \quad (5.59)$$

The derivatives are

$$\frac{\partial Q(x)}{\partial x_i} = \frac{\partial Q(\rho)}{\partial \rho_i} \frac{\partial \rho_i}{\partial x_i} \quad (5.60)$$

$$\frac{\partial^2 Q(x)}{\partial x_i \partial x_j} = \frac{\partial^2 Q(\rho)}{\partial \rho_i \partial \rho_j} \frac{\partial \rho_i}{\partial x_i} \frac{\partial \rho_j}{\partial x_j} + \frac{\partial Q(\rho)}{\partial \rho_i} \frac{\partial^2 \rho_i}{\partial x_i \partial x_j} \quad (5.61)$$

Since $Q(\rho)$ is maximal at $\rho = \hat{\rho}_{\alpha}$ is $\frac{\partial Q(\rho)}{\partial \rho_i} \Big|_{\rho=\hat{\rho}_{\alpha}} = 0$. By using this and eq 5.47 get

$$Q(\rho) \approx Q(\hat{\rho}_{\alpha}) + \frac{1}{2} \sum_{i=1}^{N_{\omega}} \sum_{j=1}^{N_{\omega}} \Delta x_i \Delta x_j \sqrt{\hat{\rho}_{\alpha i} / f_i} \sqrt{\hat{\rho}_{\alpha j} / f_j} \frac{\partial^2 Q}{\partial \rho_i \partial \rho_j} \Big|_{\rho=\hat{\rho}_{\alpha}}. \quad (5.62)$$

To get an explicit expression of the second derivative of Q , we write ΔG_i in its discretized version,

$$\Delta G_i = \tilde{G}_i - \sum_{n=1}^{N_{\omega}} K_{in} \rho_n, \quad (5.63)$$

where $K_{in} = k_{in} f_n = k_i(\omega_n) f_n$. K is a matrix of dimension $N_G \times N_{\omega}$. If $k_i(\omega) \in \mathbb{R}$, the second order derivative of Q is equal to

$$\frac{\partial^2 Q}{\partial \rho_i \partial \rho_j} = \alpha \frac{f_i}{\rho_i} \delta_{ij} + \sum_{n=1}^{N_G} \frac{K_{ni} K_{nj}}{\sigma_n^2}. \quad (5.64)$$

For the general case with $k_i(\omega) \in \mathbb{C}$, instead get

$$\frac{\partial^2 Q}{\partial \rho_i \partial \rho_j} = \alpha \frac{f_i}{\rho_i} \delta_{ij} + \operatorname{Re} \left[\sum_{n=1}^{N_G} \frac{K_{ni}^* K_{nj}}{\sigma_n^2} \right]. \quad (5.65)$$

Inserting the real version of the second derivative into the Taylor expansion in eq 5.62 gives

$$Q(\rho) \approx Q(\hat{\rho}_\alpha) + \frac{1}{2} \sum_{i=1}^{N_\omega} \sum_{j=1}^{N_\omega} \Delta x_i \Delta x_j \underbrace{\left(\alpha \delta_{ij} + \sqrt{\hat{\rho}_{\alpha i} / f_i} \sqrt{\hat{\rho}_{\alpha j} / f_j} \sum_{n=1}^{N_G} \frac{K_{ni} K_{nj}}{\sigma_n^2} \right)}_{\hat{T}_{ij}} \quad (5.66)$$

and the general complex version is

$$Q(\rho) \approx Q(\hat{\rho}_\alpha) + \frac{1}{2} \sum_{i=1}^{N_\omega} \sum_{j=1}^{N_\omega} \Delta x_i \Delta x_j \underbrace{\left(\alpha \delta_{ij} + \sqrt{\hat{\rho}_{\alpha i} / f_i} \sqrt{\hat{\rho}_{\alpha j} / f_j} \operatorname{Re} \left[\sum_{n=1}^{N_G} \frac{K_{ni}^* K_{nj}}{\sigma_n^2} \right] \right)}_{\hat{T}_{ij}}. \quad (5.67)$$

This may be conveniently expressed in matrix notation

$$Q(\rho) \approx Q(\hat{\rho}_\alpha) + \frac{1}{2} \Delta x^T \hat{T} \Delta x. \quad (5.68)$$

The validity of eq 5.56 is hence determined by the eigenvalues of \hat{T} . Big eigenvalues indicate a sharp minimum and makes the approximation of only considering the minimum value of Q reasonable. It is also computationally easier to find a well-defined minimum.

Before discussing how to minimize Q in order to solve the inverse problem, the problematic nature of the inverse problem is mentioned. By setting the left hand side of equation 5.63 to zero, for a well behaved matrix K one could just invert the matrix to get ρ_i . This is what MEM will try to do in order to minimize Q if $\alpha = 0$. Unfortunately this is not possible since the condition number of K for a typical analytic continuation problem is huge. The difficulties of the analytic continuation can also be seen from the fact that the rank (number of independent columns) of K is lower than N_ω . Regularizations are usually used for ill-conditioned problems, introducing assumptions of the solution to obtain a unique solution and prevent overfitting. MEM can thus be viewed as a regularization technique where the default model is used to introduce known information of the solution.

5.4.1 Discretization Approach

MEM is about minimizing $Q[\rho]$. We can do that approximately by using the same ω -space discretization as earlier. This reduces the problem from finding a continuous function to finding a finite number of unknown variables. Q becomes a function of N_ω variables instead of a functional. The minimum is found by setting the gradient of $Q(\rho_i)$ to zero.

For $k_i(\omega) \in \mathbb{R}$ the equations to solve are

$$0 = \frac{\partial Q}{\partial \rho_i} = - \sum_{n=1}^{N_G} \frac{\Delta G_n}{\sigma_n^2} K_{ni} + \alpha f_i \log \frac{\rho_i}{m_i}, \text{ for } i \in \{1, 2, 3, \dots, N_\omega\} \quad (5.69)$$

and for $k_i(\omega) \in \mathbb{C}$ they are

$$0 = \frac{\partial Q}{\partial \rho_i} = - \operatorname{Re} \left[\sum_{n=1}^{N_G} \frac{\Delta G_n}{\sigma_n^2} K_{ni}^* \right] + \alpha f_i \log \frac{\rho_i}{m_i}, \text{ for } i \in \{1, 2, 3, \dots, N_\omega\} \quad (5.70)$$

The number of degrees of freedom is independent of the number of matching values so that one can have more variables than input data. This is not possible even for a well-posed problem unless regularization (i.e. MEM) is introduced. More ω_i points gives better approximation of the continuation integral in 5.1. But as N_ω/N_G increases, the importance of the entropy increases. It can therefore be useful to have a more dense ω spacing just in certain important ω -intervals, e.g. where the spectral function is expected to have sharp features. A finite ω interval has to be used, so one needs to estimate ω_{\max} where $\rho(|\omega| > \omega_{\max}) \approx 0$. The discretization of ω points and the type of discrete integration method is somewhat arbitrary. None of this is present with the functional approach, see section 5.4.2

Integral Discretization Error

The discretization error of

$$\int_{-\infty}^{\infty} d\omega k_i(\omega) \rho(\omega) \approx \sum_{j=1}^{N_\omega} k_{ij} \rho_j f_j \quad (5.71)$$

depends on the shape of the unknown function $\rho(\omega)$, the kernels $k_i(\omega)$, the number of ω -points and the numerical integration method f_j . For rapidly oscillating functions a dense ω grid is required in order to suppress this error. There are many different numerical integration techniques. In this thesis the Trapezoidal method is used. It is exact if the integrand is a piecewise linear function. Call the discretization error $\sigma_{i,\text{disc}}$. To avoid overfitting, due to

neglection of this error, we here propose that the standard deviations in the likelihood function should be modified to

$$\sigma_i^2 = \sigma_{i,\text{MC}}^2 + \sigma_{i,\text{disc}}^2 \quad (5.72)$$

The statistical MC noise and the discretization error are assumed to be independent. $\sigma_{i,\text{disc}}$ depends on the unknown function $\rho(\omega)$. If we use MEM for $\sigma_i = \sigma_{i,\text{MC}}$ and use an analytic function, similar to the solution, one can estimate the discretization error. This factor becomes relevant when $\sigma_{i,\text{disc}} \geq \sigma_{i,\text{MC}}$. For accurate input data $\sigma_{i,\text{MC}}$ is the dominant error term.

5.4.2 Functional Approach

In this section, the functional Q is minimized by setting its functional derivative to zero [23]. This leads to a set of N_G non-linear equations to solve. For simplicity, the kernel k is first assumed to be real and later the general case is given.

For $k_i(\omega) \in \mathbb{R}$, the minimum of Q is found by

$$0 = \frac{\delta Q}{\delta \rho} = \frac{1}{2} \frac{\delta \chi^2}{\delta \rho} - \alpha \frac{\delta S}{\delta \rho} = - \sum_{i=1}^{N_G} \frac{\Delta G_i}{\sigma_i^2} k_i(\omega) + \alpha \log \frac{\rho(\omega)}{m(\omega)} \quad (5.73)$$

which gives

$$\rho(\omega) = m(\omega) \exp \left(\underbrace{\sum_{i=1}^{N_G} \frac{1}{\alpha} \frac{\Delta G_i}{\sigma_i^2} k_i(\omega)}_{\lambda_i} \right) \quad (5.74)$$

where λ_i is dependent of the unknown $\rho(\omega)$. One gets a self-consistent set of equations for λ_i by using the definition of ΔG_i

$$\begin{aligned} \lambda_i &= \frac{1}{\alpha} \frac{\Delta G_i}{\sigma_i^2} = \frac{1}{\alpha \sigma_i^2} \left(\tilde{G}_i - \int_{-\infty}^{\infty} d\omega k_i(\omega) \rho(\omega) \right) \\ &= \frac{1}{\alpha \sigma_i^2} \left(\tilde{G}_i - \int_{-\infty}^{\infty} d\omega k_i(\omega) m(\omega) \exp \left(\sum_{i=1}^{N_G} \lambda_i k_i(\omega) \right) \right). \end{aligned} \quad (5.75)$$

These N_G equations can be solved by a standard non-linear solver. The solution,

$$\rho(\omega) = m(\omega) \exp \left(\sum_{i=1}^{N_G} \lambda_i k_i(\omega) \right), \quad (5.76)$$

is a continuous positive function. If $\alpha \rightarrow \infty$ all $\lambda_i \rightarrow 0$. This means that by increasing the importance of the entropy the solution converges to the default

model. On the other hand $\alpha \rightarrow 0$ or $\sigma_i \rightarrow 0$ turns eq 5.75 into $\Delta G_i = 0$. If one would use eq 5.75 as an ansatz, insert it in Q , and minimize by λ_i we would obtain the same set of equations to solve.

If exact input data is given, $\Delta G_i = 0$ is desired. Another functional Q' , called the *Least Biased* [23], maximizes the entropy and implements the requirements through Lagrange multipliers λ_i .

$$Q' = S + \sum_{i=1}^{N_G} \lambda_i \Delta G_i \quad (5.77)$$

Maximizing Q' gives the same ansatz as eq 5.76. The Lagrange multipliers λ_i are determined by requiring $\Delta G_i = 0$. So this gives the same result as ordinary functional approach MEM if either all $\sigma_i \rightarrow 0$ or $\alpha \rightarrow 0$. The *Least Biased* method is thus a special case of MEM.

If $k_i(\omega) \in \mathbb{C}$, the procedure is the same as for $k_i(\omega) \in \mathbb{R}$,

$$0 = \frac{\delta Q}{\delta \rho} = \frac{1}{2} \frac{\delta \chi^2}{\delta \rho} - \alpha \frac{\delta S}{\delta \rho} = - \sum_{i=1}^{N_G} \frac{\text{Re}[\Delta G_i^* k_i(\omega)]}{\sigma_i^2} + \alpha \log \frac{\rho(\omega)}{m(\omega)} \quad (5.78)$$

which gives

$$\rho(\omega) = m(\omega) \exp\left(\sum_{i=1}^{N_G} \underbrace{\text{Re}\left[\frac{1}{\alpha} \frac{\Delta G_i^*}{\sigma_i^2} k_i(\omega)\right]}_{\lambda_i}\right). \quad (5.79)$$

λ_i is now complex and obey

$$\begin{aligned} \lambda_i &= \frac{1}{\alpha} \frac{\Delta G_i^*}{\sigma_i^2} = \frac{1}{\alpha \sigma_i^2} \left(\tilde{G}_i^* - \int_{-\infty}^{\infty} d\omega k_i^*(\omega) \rho(\omega) \right) \\ &= \frac{1}{\alpha \sigma_i^2} \left(\tilde{G}_i^* - \int_{-\infty}^{\infty} d\omega k_i^*(\omega) m(\omega) \exp\left(\sum_{i=1}^{N_G} \text{Re}[\lambda_i k_i(\omega)]\right) \right). \end{aligned} \quad (5.80)$$

5.5 Hyper-parameter α

To complete the MEM the PDF $P[\alpha|m, \tilde{G}]$ for the hyper-parameter α is needed. By combining eq 5.18 with 5.54 and insert the Taylor expansion of Q from eq 5.68 we get

$$\begin{aligned} P[\alpha|m, \tilde{G}] &= \frac{P[\alpha]}{P[\tilde{G}|m]} \frac{1}{Z_L Z_S} \int \mathcal{D}\rho e^{-(Q(\hat{\rho}_\alpha) + \frac{1}{2} \Delta x^T \hat{T} \Delta x)} \\ &= \frac{P[\alpha]}{P[\tilde{G}|m]} \frac{e^{-Q(\hat{\rho}_\alpha)}}{Z_L Z_S} \int \prod_{i=1}^{N_\omega} dx_i e^{-\frac{1}{2} \Delta x^T \hat{T} \Delta x}. \end{aligned} \quad (5.81)$$

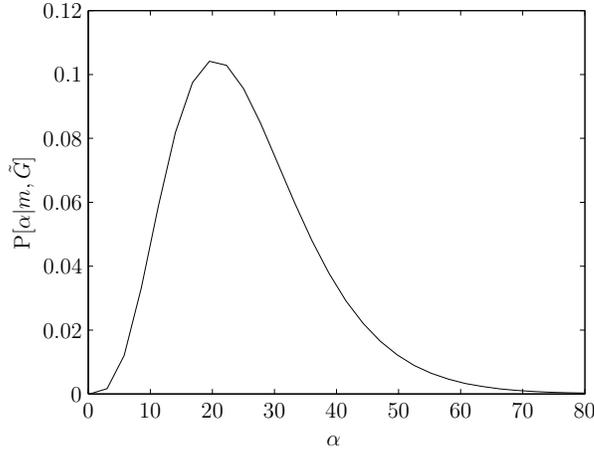


Figure 5.1: A typical probability distribution of the entropy parameter α .

For the kernels defined in section 5.8, \hat{T} is a positive-definite matrix [19]. The inverse to a positive-definite matrix is also positive-definite [24]. The integrand is therefore a well-known multivariate gaussian so

$$P[\alpha|m, \tilde{G}] = \frac{P[\alpha]}{P[\tilde{G}|m]} \frac{e^{-Q(\hat{\rho}_\alpha)} (2\pi)^{\frac{N_G}{2}}}{Z_L Z_S \sqrt{\det \hat{T}}}. \quad (5.82)$$

$P[\alpha]$ is usually taken to be a constant. $P[\alpha] \propto \frac{1}{\alpha}$ is called the Joffreys prior [17]. In this master-thesis project the two priors gave almost the same MEM result. $P[\tilde{G}|m]$ is unknown but is constant since \tilde{G} and m are fixed. In practice we let $P[\tilde{G}|m] = 1$. The α integral is discretized and the finite number of $P[\alpha|m, \tilde{G}]$ values are normalized to one. The distribution is almost always skewed, an example is seen in figure 5.1.

There exist several versions of MEM and the one presented is usually called the *Bryan's method* [25]. The early versions of MEM picks one α -value by some rules instead of average using eq 5.82. In what is called *Historic Maximum Entropy*, α is chosen so that $\chi^2 = \frac{1}{2}N_G$ [19]. That makes the deviation ΔG_i of the same order of magnitude as the noise σ_i . It is a good approach to avoid overfitting with the data. Overfitting creates spurious features in the spectral function. The *Classic Maximum Entropy* method uses the α which maximize the posteriori distribution $P[\alpha|m, \tilde{G}]$. This usually gives a closer fit to the input data than Historic Maximum Entropy [19]. If $P[\alpha|m, \tilde{G}]$ is sharply peaked at some α , this method gives the same result as the Bryan's method. But if $P[\alpha|m, \tilde{G}]$ has a broad peak, Bryan's method will average over many $\hat{\rho}_\alpha$ solutions instead of only picking

one. Unphysical oscillations is an artifact in all MEM [26]. They are less present if a solution average, as in Bryan's method, is used.

5.6 Batch Averaging

To remove unphysical oscillations in the calculated $\rho(\omega)$, batch averaging can be used [26]. Normally the input data $\tilde{G}_{i,j}$, $i \in \{1, 2, \dots, N_G\}$ and $j \in \{1, 2, \dots, N_s\}$, is averaged over all j so get

$$\tilde{G}_i = \frac{1}{N_s} \sum_{j=1}^{N_s} \tilde{G}_{i,j}. \quad (5.83)$$

\tilde{G}_i is then used as input to the inverse integral problem. Instead $\tilde{G}_{i,j}$ can be divided into N_b batches. Averaging is performed within each batch, yielding $\tilde{G}_{i,\text{batch}_b}$ for $b \in \{1, 2, \dots, N_b\}$. MEM is used for each batch average, and the obtained $\rho_{\text{batch}_b}(\omega)$ are finally averaged

$$\rho(\omega) = \frac{1}{N_b} \sum_{b=1}^{N_b} \rho_{\text{batch}_b}(\omega). \quad (5.84)$$

5.7 Stochastic Sampling of the Likelihood Function

In this section we briefly mention a method similar to MEM. It uses Bayesian inference but with another a priori probability distribution for ρ . The trick of introducing an entropic prior makes the posterior distribution sharply peaked around some spectral function. That choice favors spectrums with high entropy. If one instead uses the likelihood function and as a priori knowledge only that the spectral function must be non-negative, no bias towards some spectrums is introduced.

$$P(\rho|G) = \frac{P(G|\rho) \cdot P(\rho)}{P(G)} \propto \begin{cases} P(G|\rho) \propto \exp(-\frac{\chi^2}{2}) & \text{if } \rho(\omega) \geq 0 \forall \omega \\ 0 & \text{otherwise} \end{cases} \quad (5.85)$$

However this distribution contains many different local maximums with $\chi^2 \approx 0$. With statistical noise on the input data (in our case the Matsubara Green's function) these solutions are usually not smooth. Instead one can calculate the mean spectrum by sample spectrums from Eq. 5.85 and then the average of them [27], [22]. With a discretized representation of the spectral function

the sampling over the high-dimensional space is advantageously done by MC. In order not to get trapped in some local maximum simulated-annealing can be applied.

5.8 Analytic Continuation Kernels

So far MEM has been described for *any* inverse integral problem. For the analytic continuation the integral kernels are, from eq 2.7,2.9,4.23,

$$\begin{aligned}
k_{\nu_n}(\omega) &= \frac{1}{i\nu_n - \omega} \quad , \text{for } \nu_n = (2n - 1)\frac{\pi}{\beta} \quad \text{with } n \in \mathbb{Z} \\
k_\tau(\omega) &= \frac{-e^{-\omega\tau}}{e^{-\beta\omega} + 1} \quad , \text{for } \tau \in (0, \beta) \\
k_l(\omega) &= \frac{-\sqrt{2l+1}\beta i^l e^{-\frac{\beta\omega}{2}} j_l(i\frac{\beta\omega}{2})}{1 + e^{-\beta\omega}} \quad , \text{for } l \in \mathbb{N}_0.
\end{aligned} \tag{5.86}$$

Since the half-filled Hubbard model has particle-hole symmetry, thus $\rho(-\omega) = \rho(\omega)$, $G(i\nu_n)$ is purely imaginary according to eq 3.14. Continuation from $\text{Im}[G(i\nu_n)]$ can therefore be done with the kernel

$$k_{\nu_n}(\omega) = \frac{-\nu_n}{\nu_n^2 + \omega^2} \quad , \text{for } \nu_n = (2n - 1)\frac{\pi}{\beta} \quad \text{with } n \in \mathbb{N}. \tag{5.87}$$

If using the Legendre coefficients G_l , the particle-hole symmetry reduces the continuation conditions from $l \in \mathbb{N}_0$ to even l . Due to the particle-hole symmetry, calculating $\rho(\omega)$ for $\omega \geq 0$ is sufficient, halving the degrees of freedom in the discretization approach.

Once a kernel is chosen, the number of input values, N_G , for the continuation has to be determined. The analytic continuation from imaginary time has a problem that τ is not discrete. Arbitrariness is hence introduced when picking N_G discrete times. This kind of choice is not present when $G(\tau)$ is expressed in its Fourier coefficients $G(i\nu_n)$ or Legendre coefficients G_l . But still a cut-off of the number of coefficients are needed. The Fourier coefficients $G(i\nu_n)$ decay slow, $G(i\nu_n) \propto \frac{1}{\nu_n}$ for large n , so the cut-off is somewhat arbitrary. The Legendre coefficients G_l are expected to decay exponentially fast and $N_G = 50$ usually seem to be sufficient.

5.9 Simulations

This section is divided into three parts. First, implementation issues for all simulations are first presented. Secondly, we investigate the performance

of MEM as an analytic continuation method for Green's functions. The obtained spectral functions using MEM, corresponding to the DMFT calculated Matsubara Green's functions are finally presented.

When minimizing Q , both with the discretization and functional approach, MEM boils down to solving a set of (non-linear) equations. These equations are solved with MATLAB's built-in methods *fmincon* and *fsolve*. In all presented simulations the discretized version of MEM is used. The primary reason is because it was computationally easier to solve the discretization equations than those from the functional approach.

The norm of ρ is indirectly influenced by the default model and the input data. A strict normalization condition could be introduced in the a priori distribution, but it was found to be unnecessary. $\rho_i \geq 0 \forall i \in \{1, 2, \dots, N_\omega\}$ are constraints in the Q minimization to ensure that the spectral function is non-negative.

A finite ω -window is used and ω_{\max} is chosen so that $\rho(\omega_{\max}) \approx 0$. The MEM solutions was weakly dependent of the ω -mesh. In this report a logarithmic mesh defined by N_ω , ω_{\max} and a distribution parameter γ is used [26].

A flat default model is used except for when different default models are compared, see figure 5.5.

5.9.1 Tests of MEM

To test MEM four cases are studied. A spectral function is chosen and by the Hilbert transformation the Matsubara Green's function is obtained. MEM tries to retrieve the original spectral function from the exact Matsubara values $G(i\nu_n)$. σ is chosen to be the same for all ν_n but change between different simulations. An average of spectral functions, not plotted but obtained from Matsubara data with added gaussian noise of strength σ , and the spectral function obtained using exact data and the same σ were very similar. The chosen spectral functions are normalized to one and described only for $\omega > 0$ since we study only even spectral functions.

High Energy Independence

All the Matsubara time and frequency kernels decay as a function of ω . This mean that the high energy parts of the spectral function contributes less than for small ω . For large ω , the imaginary time kernels $k_\tau(\omega) \propto \exp(-\tau\omega)$ and are exponentially decaying since $\tau \geq 0$. To verify these arguments a simple spectral function is created and Matsubara data is used in the analytical continuation, see figure 5.2. As expected the lowest energy peak is found and the high energy peak gets smeared out by MEM. The exact spectral

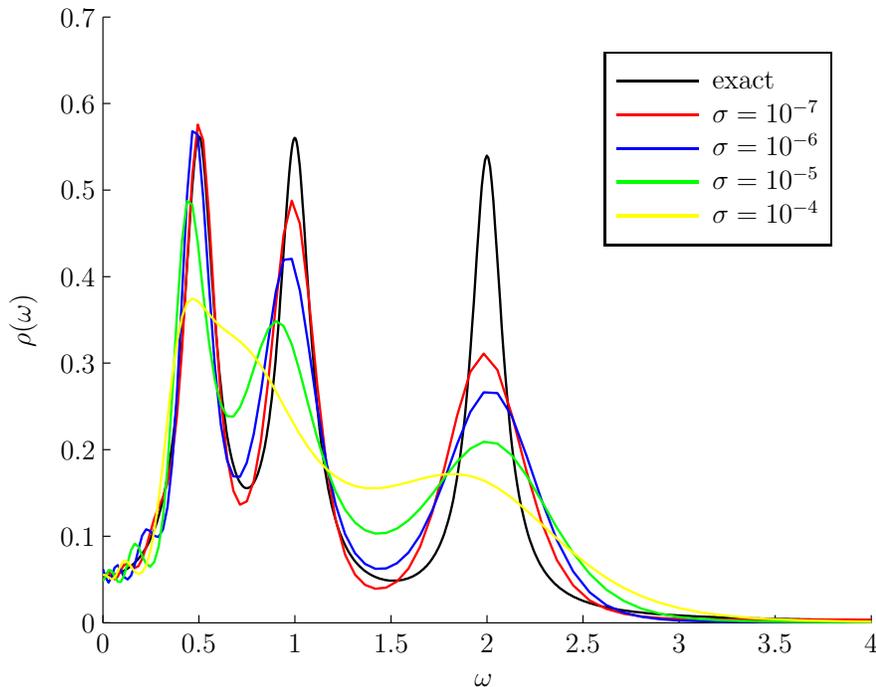


Figure 5.2: Spectral functions obtained for different σ .

function was chosen to be a sum of three Lorentzians positioned around 0.5, 1 and 2 each with width 0.1. We use a flat default model, a logarithmic grid with $N_\omega = 121$, $\omega_{\max} = 15$ and $\gamma = 0.5$, $\beta = 100$ and $\nu_{\max} = 10$ in this simulation. The slow decay of Lorentzians, $\propto \frac{1}{\omega^2}$, obliges a high ω_{\max} . Using high σ -values, the entropic prior will dominate over the likelihood function so spectral functions, biased towards the flat default model, are obtained.

Featureless Spectral Function

The analytic continuation works well, as expected, for smooth and featureless spectral functions i.e. a gaussian, see figure 5.3. In this and the following two tests, $\omega_{\max} = 4$. Otherwise the same ω -mesh, β and ν_{\max} as above are used.

Sharp Hubbard Peak

In order to see if our MEM algorithm is expected to be able to extract an eventual inner side peak to the Hubbard band [7, 28], a known trial spectral function is used. It is chosen with a gaussian centered at $\omega = 1$ with width

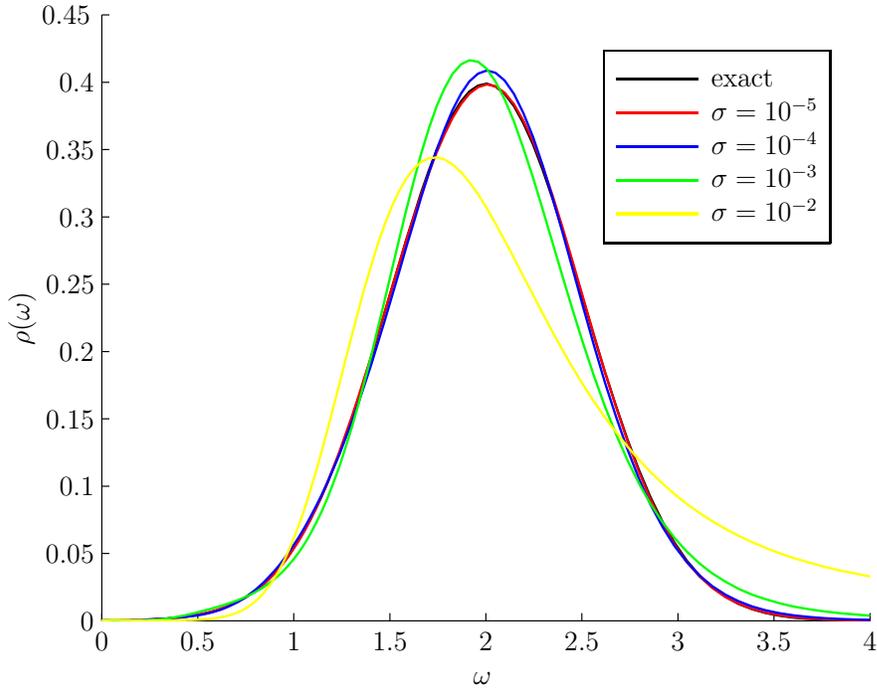


Figure 5.3: Spectral functions obtained for different σ .

0.06 and another gaussian centered at $\omega = 2$ with width 0.5. Using MEM for several values of σ , get smeared out spectral functions, see figure 5.4. No peak is obtained for $\sigma > 10^{-5}$, indicating high accuracy is required to resolve sharp peaks.

Default Model

No unphysical peaks were observed in spectral function when a flat default model was used in the test above. We understand this since MEM can be seen to maximize the entropy under constrain of matching the input data. In order to resolve sharp peaks with MEM, prior knowledge is needed i.e. another default model should be used. Figure 5.5 shows an example of how a good default model increases the accuracy of the MEM solver, using the same spectral function as above. The two non-flat default models both have gaussians with same centers as the exact ρ , but the sharp peak is broadened by a factor 1.5 and 3 respectively. The probability weights $P[m|\tilde{G}]$, from eq 5.20, are normalized to one to simplify comparison between the three default models. The default model most similar to the exact spectral function got the highest probability weight.

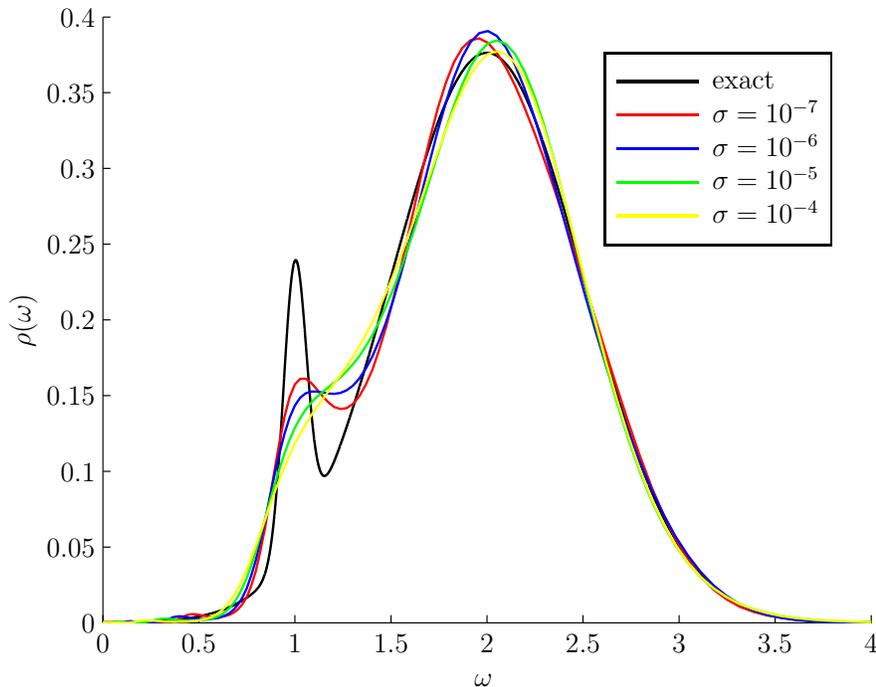


Figure 5.4: Spectral functions obtained for different σ .

5.9.2 Spectral Functions for the Hubbard model

We use MEM to obtain ρ from the DMFT calculated Matsubara Green's functions for $U = 1, 2, 2.5, 4$ and $\beta = 150$. Continuation from even Legendre coefficients $\tilde{G}_{l,i}$ and the imaginary part of the Fourier coefficients $\text{Im}[\tilde{G}_{n,i}]$, where i is DMFT iteration number, are considered. To estimate the covariance between $\tilde{G}_{l,i}$ and $\tilde{G}_{l',i}$ for $l \neq l'$, the number of converged DMFT-iterations N_i should be larger than the number of observables N_G [19]. But even for $N_i > 2 \times N_G$ some eigenvalues of the covariance matrix C became zero, making MEM unusable. The same results were obtained for the Fourier coefficients $\tilde{G}_{n,i}$. Only σ_l and σ_n are therefore considered in the MEM algorithm, and not the off-diagonal terms of the covariance matrices.

The same ω -mesh and ν_{\max} as above are used. No cut-off in the number of Legendre coefficients are used, G_l for $l \in \{0, 2, 4, \dots, l_{\max}\}$.

For $U = 1$, an analytic continuation of 10 converged DMFT iterations using Matsubara and Legendre kernel are seen in figure 5.6. The spectral functions obtained from the Fourier coefficients vary more than those from the Legendre representation, even though the mean spectrums are similar. An underestimation of the variations of G_n , due to neglect of the full

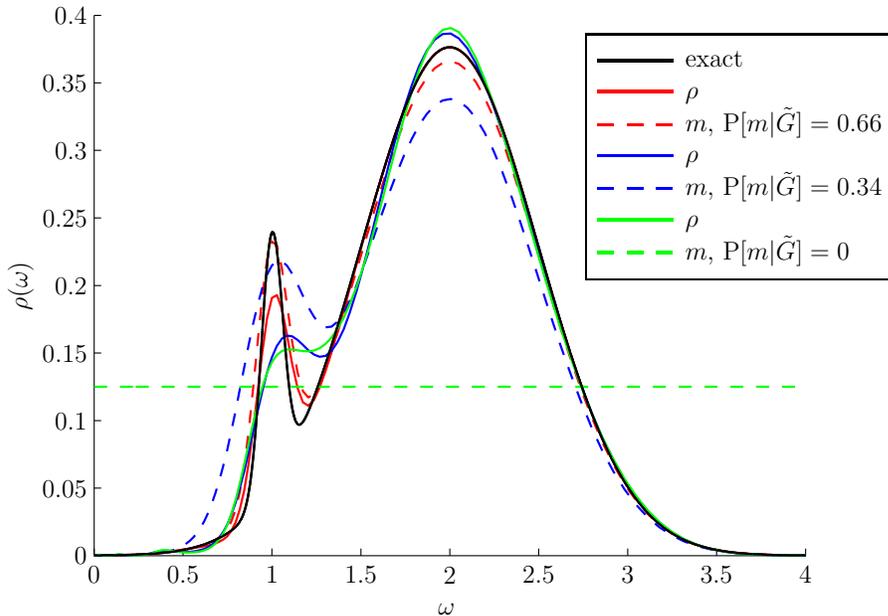


Figure 5.5: Obtained spectral functions together with corresponding default models. $\sigma = 10^{-6}$ is used.

covariance matrix, might be an explanation. In the rest of this section continuation is done only from Legendre coefficients. For each U , Legendre coefficients from 10 converged DMFT iterations are separately continued. The obtained spectral functions are averaged and the mean values are seen in figure 5.7. When the interaction is increased from $U = 0$, the quasiparticle peak around $\omega = 0$ becomes sharper and finally disappears, illustrating an interaction driven MIT. In Fermi liquid theory, el-el interaction enhances the electron mass so get a smaller bandwidth i.e. sharper spectral density. The insulator in some sense indicates a breakdown of the Fermi liquid. Worth notice, $\rho(0)$ is almost constant in the metallic phase. The small drop at $U = 2.5$ might be due to finite temperature. Figures 5.8, 5.9, 5.10 and 5.11 compare the obtained ρ with zero temperature Dynamic-Density Matrix Renormalization Group (D-DMRG) calculations [7]. For low interaction ($U = 1$ and $U = 2$) the two methods give similar results. For the metallic solution at $U = 2.5$, the sharp inner peak using D-DMRG is not obtained by MEM. The insulating spectrums are similar but the outer Hubbard band is less sharp using MEM. For $U = 4$ the obtained Hubbard band has a two-peak structure instead of a single peak. The conservative choice of a flat default model should prevent unphysical features but the two peak structure might be an artifact from underestimation of the QMC statistical noise. It could also be

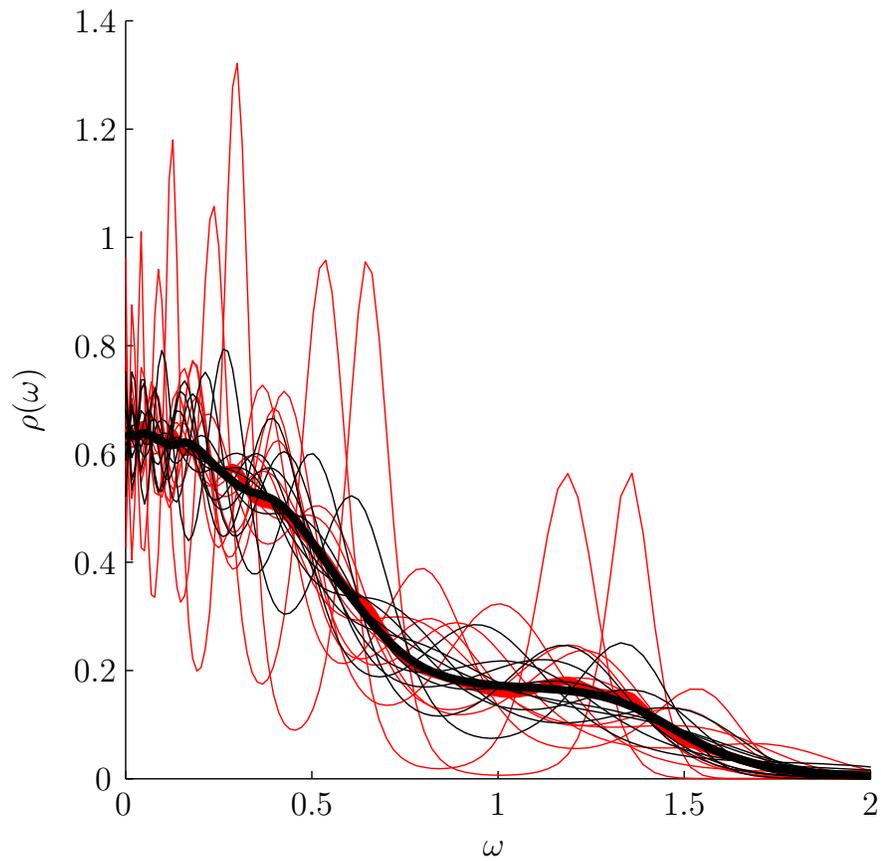


Figure 5.6: Spectral functions obtained from Matsubara (red) and Legendre (black) representations of imaginary Green's function for 10 different set of data. The thick lines are mean values.

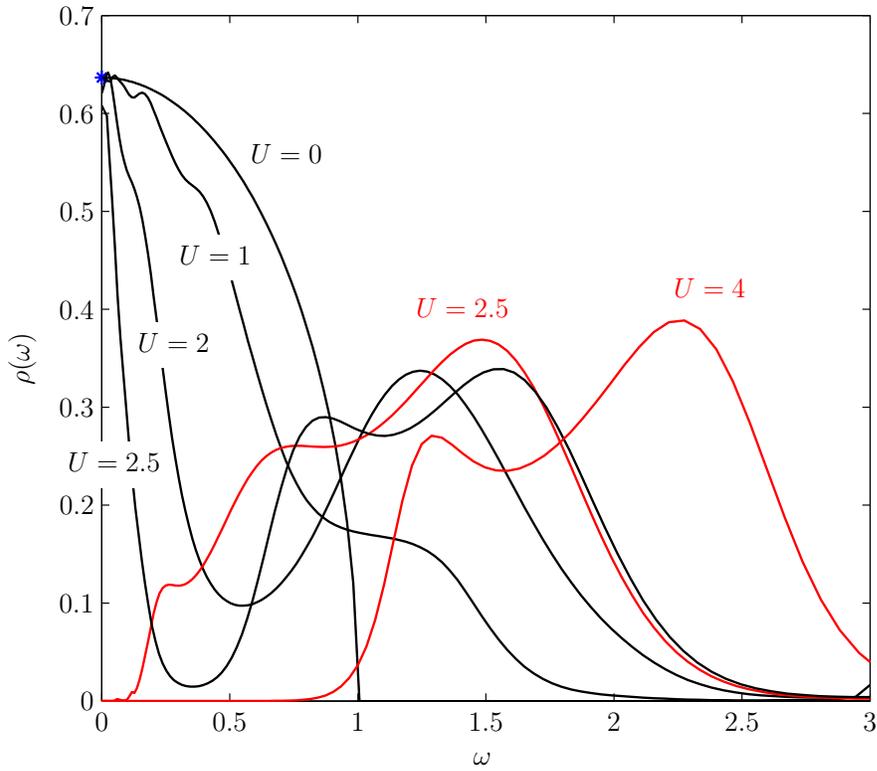


Figure 5.7: Spectral functions for different U and $\beta = 150$ obtained using CTQMC and MEM. Black (red) spectrums are metals (insulators) to illustrate the MIT.

a MEM changeover artifact due to very sharp band edges.

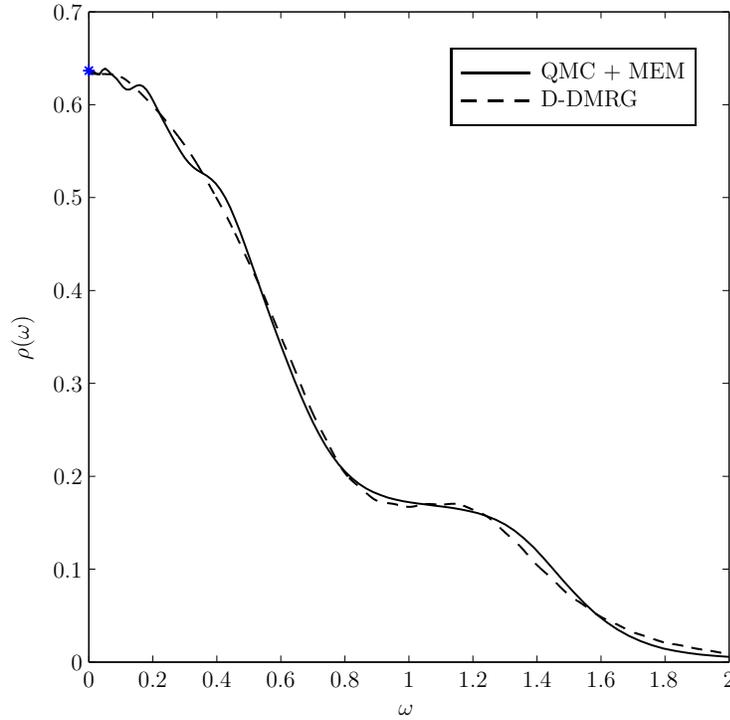


Figure 5.8: Spectral functions for interaction strength $U = 1$.

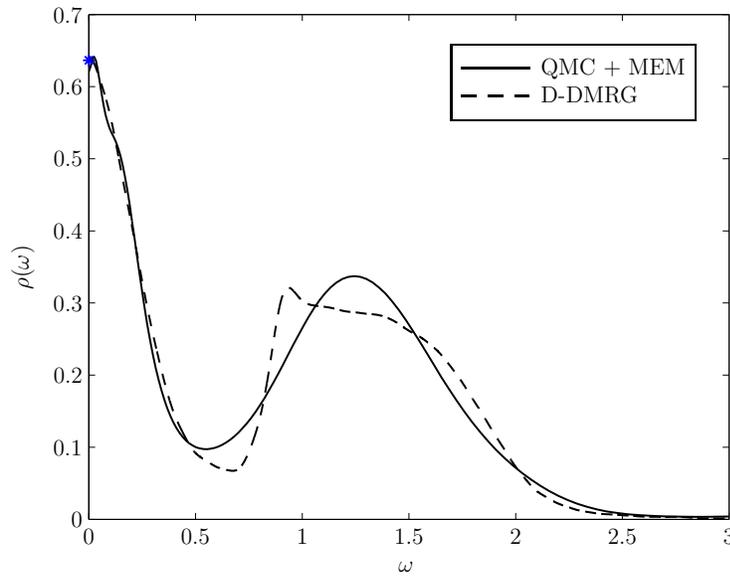


Figure 5.9: Spectral functions for interaction strength $U = 2$.

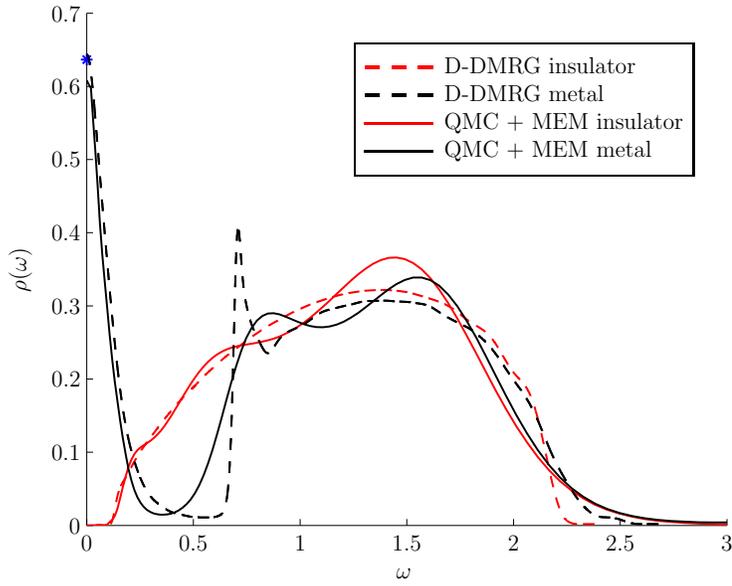


Figure 5.10: Spectral functions for interaction strength $U = 2.5$.

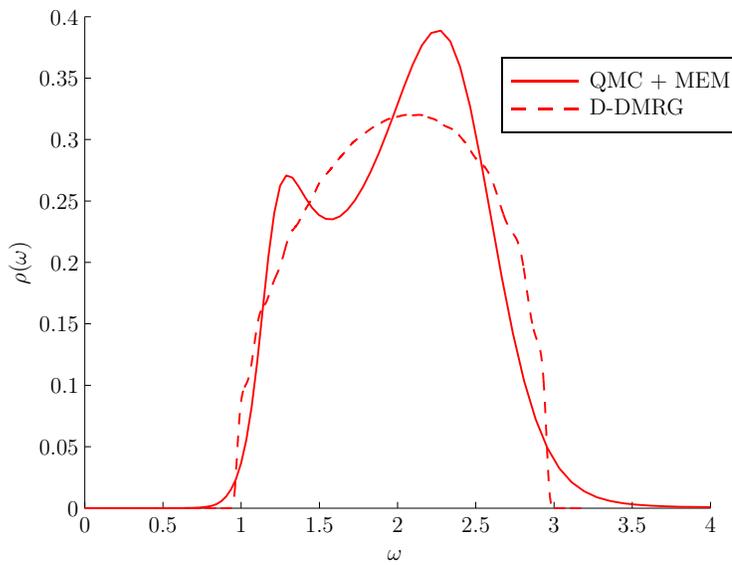


Figure 5.11: Spectral functions for interaction strength $U = 4$.

Chapter 6

Conclusion

We studied the Hubbard model using DMFT. The analytic continuation of the Matsubara single-particle Green's function was done with MEM. Parts of the well-known DMFT two-solution region in the phase diagram for the paramagnetic Hubbard model was observed. At least 50 Legendre polynomials were needed to converge the Legendre coefficients when using the CT-HYB algorithm.

The focus of this thesis was to investigate MEM. Test spectral functions showed us that MEM usually produces smeared out spectral functions compared to the exact one. This smearing was extra pronounced at high energies, which is expected due to the analytic continuation kernels. Featureless spectral functions were easy to obtain, even for low accuracy. For moderate Matsubara accuracy, a very good default model was needed to improve the spectral function obtained using a flat default model.

The analytic continuation of the DMFT Green's functions was done directly from the sampled Legendre coefficients but also from the Fourier coefficients. Highly oscillating spectral functions were obtained from continuation of Matsubara Fourier coefficients, due to underestimation of the statistical error. The qualitative quasiparticle peak and the Hubbard bands were obtained.

To easier obtain sharp features, PADE methods together with a maximum entropy constrain might be useful. It would also be interesting to study other models, i.e. the multi-band Hubbard model, or real materials using LDA+DMFT approaches.

Acknowledgements

I would like to express my gratitude to my supervisor Mats Granath for the many hours of explanation, guidance and discussion along the learning process of this thesis. A big thanks also goes to Hugo Strand, especially for helping me to use Triqs simulations on SNIC resources. I also would like to thank my family and friends for their invaluable support.

Bibliography

- [1] S. Hufner, “Photoelectron spectroscopy,” 2003.
- [2] M. Troyer, “lecture notes from course: Computational quantum physics, ETH Zurich, spring semester 2012,”
- [3] C. Titchmarsh, “The theory of functions,” 1939.
- [4] R. Mattuck, “A guide to feynman diagrams in the many-body problem,” 1992.
- [5] J. Hubbard, “Electron correlations in narrow energy bands,” *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, vol. 276, no. 1365, pp. 238–257, 1963.
- [6] H. Strand, “Critical properties of the mott-hubbard metal-insulator transition,” 2011.
- [7] M. Karski and G. U. C. Raas, “Single-particle dynamics in the vicinity of the mott-hubbard metal-to-insulator transition,” vol. 77, no. 7, p. 075116, 2008.
- [8] N. F. Mott, “The basis of the electron theory of metals, with special reference to the transition metals,” *Proceedings of the Physical Society. Section A*, vol. 62, no. 7, p. 416, 1949.
- [9] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, “Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions,” *Rev. Mod. Phys.*, vol. 68, pp. 13–125, Jan 1996.
- [10] W. Metzner and D. Vollhardt, “Correlated lattice fermions in $d = \infty$ dimensions,” *Phys. Rev. Lett.*, vol. 62, pp. 324–327, Jan 1989.
- [11] P. W. Anderson, “Localized magnetic states in metals,” *Phys. Rev.*, vol. 124, pp. 41–53, Oct 1961.

- [12] L. Boehnke, H. Hafermann, M. Ferrero, F. Lechermann, and O. Parcollet, “Orthogonal polynomial representation of imaginary-time green’s functions,” *Phys. Rev. B*, vol. 84, p. 075145, Aug 2011.
- [13] M. Ferrero and O. Parcollet, “Triqs: “a toolkit for research on interacting quantum systems”,”
- [14] P. Werner, A. Comanac, L. de’ Medici, M. Troyer, and A. J. Millis, “Continuous-time solver for quantum impurity models,” *Phys. Rev. Lett.*, vol. 97, p. 076405, Aug 2006.
- [15] P. Werner and A. J. Millis, “Hybridization expansion impurity solver: General formulation and application to kondo lattice and two-orbital models,” *Phys. Rev. B*, vol. 74, p. 155107, Oct 2006.
- [16] A. Dirks, P. Werner, M. Jarrell, and T. Pruschke, “Continuous-time quantum monte carlo and maximum entropy approach to an imaginary-time formulation of strongly correlated steady-state transport,” *Phys. Rev. E*, vol. 82, p. 026701, Aug 2010.
- [17] J. E. Gubernatis, M. Jarrell, R. N. Silver, and D. S. Sivia, “Quantum monte carlo simulations and maximum entropy: Dynamics from imaginary-time data,” *Phys. Rev. B*, vol. 44, pp. 6011–6029, Sep 1991.
- [18] A. Papoulis, “Probability and statistics,” p. 422, 1990.
- [19] M. Jarrell and J. Gubernatis, “Bayesian inference and the analytic continuation of imaginary-time quantum monte carlo data,” *Physics Reports*, vol. 269, no. 3, pp. 133 – 195, 1996.
- [20] E. Jaynes, “The maximum entropy formalism,” pp. 15–118, 1978.
- [21] J. G. C.R. Smith, “Maximum entropy and bayesian methods in inverse problems,” vol. 14, 1985.
- [22] O. Gunnarsson, M. W. Haverkort, and G. Sangiovanni, “Analytical continuation of imaginary axis data for optical conductivity,” *Phys. Rev. B*, vol. 82, p. 165125, 2010.
- [23] C. Raas and G. S. Uhrig, “Spectral densities from dynamic density-matrix renormalization,” *The European Physical Journal B - Condensed Matter and Complex Systems*, vol. 45, no. 3, pp. 293–303, 2005.
- [24] C. J. R. Horn, “Matrix analysis,” p. 397, 1985.

- [25] R. Bryan, “Maximum entropy analysis of oversampled data problems,” *European Biophysics Journal*, vol. 18, no. 3, pp. 165–174, 1990.
- [26] O. Gunnarsson, M. W. Haverkort, and G. Sangiovanni, “Analytical continuation of imaginary axis data using maximum entropy,” *Phys. Rev. B*, vol. 81, p. 155107, 2010.
- [27] A. W. Sandvik, “Stochastic method for analytic continuation of quantum monte carlo data,” *Phys. Rev. B*, vol. 57, pp. 10287–10290, 1998.
- [28] E. Gull, D. R. Reichman, and A. J. Millis, “Bold-line diagrammatic monte carlo method: General formulation and application to expansion around the noncrossing approximation,” *Phys. Rev. B*, vol. 82, p. 075109, Aug 2010.

Appendices

Appendix A

Second Quantization

Assume that $|f_{i\sigma}\rangle$, where $i \in \{1, 2, \dots, \frac{d}{2}\}$ is orbital index and $\sigma \in \{\uparrow, \downarrow\}$ is spin, represents a basis in a Hilbert space \mathcal{H} of dimension d . If we have N distinguishable particles it's Hilbert space can be spanned by states given by the product of single particle states

$$|\Psi_\alpha^{(N)}\rangle = |f_{i_1\sigma_1}\rangle_1 |f_{i_2\sigma_2}\rangle_2 \dots |f_{i_N\sigma_N}\rangle_N = \prod_{n=1}^N |f_{i_n\sigma_n}\rangle_n \quad (\text{A.1})$$

where $\alpha = \{i_1\sigma_1 i_2\sigma_2 \dots i_N\sigma_N\}$ is a super index containing all information about the N particles. There are d^N possible configurations so there are d^N orthonormal $|\Psi_\alpha^{(N)}\rangle$. The dimension of the Hilbert space for N particles are thus d^N . This N -body Hilbert space $\mathcal{H}^{(N)}$ is described by a tensor product of N Hilbert spaces

$$\mathcal{H}^{(N)} = \bigotimes_{n=1}^N \mathcal{H}_n \quad (\text{A.2})$$

The exponential scaling of the Hilbert space with the number of particles makes many-body quantum problems hard to solve exactly.

Indistinguishable particles have the property that exchanging two particles should leave the system unchanged up to a complex phase factor. This is not necessary true for a state in eq A.1. For fermions this particle exchange factor is -1 . In order to make sure that this is fulfilled one can apply the antisymmetrization operator S_- on the many-body state $|\Psi_\alpha^{(N)}\rangle$

$$|\Psi_n^{A,(N)}\rangle = S_- |\Psi_\alpha^{(N)}\rangle \equiv \mathcal{N}_A \sum_p \text{sgn}(p) |f_{i_{p(1)}\sigma_{p(1)}}\rangle_1 |f_{i_{p(2)}\sigma_{p(2)}}\rangle_2 \dots |f_{i_{p(N)}\sigma_{p(N)}}\rangle_N \quad (\text{A.3})$$

where n is a superindex containing all information of how occupied each single particle states is, but not by which particles. An easier way of representing

the antisymmetric state in equation A.3 is in the occupation number basis

$$|n\rangle = |n_{1\uparrow}n_{1\downarrow}n_{2\uparrow}n_{2\downarrow}\dots n_{d/2\uparrow}n_{d/2\downarrow}\rangle \quad (\text{A.4})$$

where $n_{i\sigma} \in \{0, 1\}$ for fermions and is the occupation number for state $|f_{i\sigma}\rangle$.

So far the particle number N has been fix. If the system is allowed to have any particle number, the occupation number basis representation becomes even more advantageous. This Hilbert space, called the Fock space, is just the sum of antisymmetrized N -particle Hilbert spaces

$$\mathcal{H}_{\text{Fock}} = \bigoplus_{N=0}^{\infty} S_- \mathcal{H}^{(N)} \quad (\text{A.5})$$

To create, modify and destroy many-body states, the creation $c_{i\sigma}^\dagger$ and annihilation $c_{i\sigma}$ operators are used. They obey the canonical anticommutation relations

$$\{c_{i\sigma}^\dagger, c_{j\sigma'}\} = \delta_{ij}\delta_{\sigma\sigma'} \quad (\text{A.6})$$

$$\{c_{i\sigma}^\dagger, c_{j\sigma'}^\dagger\} = \{c_{i\sigma}, c_{j\sigma'}\} = 0 \quad (\text{A.7})$$

The basis state in equation A.4 can be expressed by creation operators

$$\begin{aligned} |n\rangle &= |n_{1\uparrow}n_{1\downarrow}n_{2\uparrow}n_{2\downarrow}\dots n_{d/2\uparrow}n_{d/2\downarrow}\rangle = \prod_{i=1}^{d/2} \prod_{\sigma \in \{\uparrow, \downarrow\}} (c_{i\sigma}^\dagger)^{n_{i\sigma}} |0\rangle \\ &= (c_{1\uparrow}^\dagger)^{n_{1\uparrow}} (c_{1\downarrow}^\dagger)^{n_{1\downarrow}} (c_{2\uparrow}^\dagger)^{n_{2\uparrow}} (c_{2\downarrow}^\dagger)^{n_{2\downarrow}} (c_{d/2\uparrow}^\dagger)^{n_{d/2\uparrow}} (c_{d/2\downarrow}^\dagger)^{n_{d/2\downarrow}} |0\rangle \end{aligned} \quad (\text{A.8})$$

Since the creation operators anticommute, see equation A.6, the ordering in equation A.8 is important. Any ordering is possible but one has to decide for one convention and then be consistent with that to avoid sign errors.