Contents

Chapter IV

One-dimensional Hubbard model

In this chapter we discuss the Hubbard model in one dimension[1]. In contrast to the models of localized spins that we discussed in Chapter 2, the Hubbard model has two different types of excitations, namely spin and charge excitations. This two-component structure is also reflected in the Bethe-Ansatz treatment of the model.
In the following we will first show that the Hubbard model is solvable by Bethe Ansatz for any parameter values. Then we discuss the groundstate and the structure of low-lying excitations. But first we investigate some limiting cases to get a rough understanding of the underlying physics. Then the symmetries of the Hamiltonian are found which simplifies the analysis.

IV.1 Limiting cases

The one-dimensional Hubbard model on a chain of L lattice sites is given by the Hamiltonian

$$\boxed{\mathcal{H}(t,U) = -t\sum_{j=1}^{L}\sum_{\sigma=\uparrow,\downarrow}\left(c_{j\sigma}^{\dagger}c_{j+1,\sigma} + c_{j+1,\sigma}^{\dagger}c_{j\sigma}\right) + U\sum_{j=1}^{L} n_{j\uparrow}n_{j\downarrow} \qquad \text{(IV.1.1)}}$$

where $c_{j\sigma}^{\dagger}$, $c_{j\sigma}$ are fermionic creation/annihilation operators that satisfy the standard anticommutation rules

$$\{c_{j\sigma}, c_{l\sigma'}^{\dagger}\} = \delta_{jl}\delta_{\sigma\sigma'}, \qquad \{c_{j\sigma}, c_{l\sigma'}\} = \{c_{j\sigma}^{\dagger}, c_{l\sigma'}^{\dagger}\} = 0. \tag{IV.1.2}$$

$n_{j\sigma} = c_{j\sigma}^{\dagger}c_{j\sigma}$ is the (local) particle number operator for spin σ. In the following we will always assume periodic boundary conditions.
The Hamiltonian (IV.1.1) consists of two simple parts. The first term $\mathcal{H}_t$ is the kinetic energy. For $U = 0$ the Hamiltonian reduces to that of fermions with spin and can be easily diagonalized by Fourier transformation

$$c_{j\sigma} = \frac{1}{\sqrt{L}}\sum_{k} e^{ikj}c_{k\sigma}. \tag{IV.1.3}$$

[1] For a collection of reprints of the most important papers on the exact solution of the Hubbard model, see [1].

This yields

$$\mathcal{H}_t = -2t\sum_{k,\sigma}\cos(k)n_{k\sigma} \tag{IV.1.4}$$

which is diagonal in Fourier space. The discussion is completely analogous to the investigation of spinless free fermions since the two parts for spin projections $\sigma=\uparrow$ and $\sigma=\downarrow$ are independent. For the groundstate we have to fill the N_σ states with the lowest energies $-2t\cos k$ for each spin direction σ up to the Fermi momenta $k_F^{(\sigma)} = \frac{\pi}{2}\frac{N_\sigma}{L}$. Here $N_\sigma = \sum_{j=1}^{L} n_{j\sigma}$ is the total number of particles with spin σ. We immediately see that the groundstate is metallic since creating a current-carrying state requires only an infinitesimal energy.

The case $t=0$ is sometimes called *atomic limit*. Here electrons can not move. The Hamiltonian becomes

$$\mathcal{H}_U = U\sum_{j=1}^{L} n_{j\uparrow}n_{j\downarrow} = UN_D \tag{IV.1.5}$$

where N_D is the number of doubly-occupied sites. In contrast to the free fermion limit $U=0$ this part is diagonal in real space. It has a equal-spaced discrete energy spectrum $E=0,U,2U,\ldots$. The groundstate is highly degenerate. For $U>0$ and $N\leq L$ it is given explicitly by

$$|\psi\rangle_U = c^\dagger_{j_1\sigma_1}\cdots c^\dagger_{j_N\sigma_N}|0\rangle. \tag{IV.1.6}$$

where $1\leq j_1 < j_2 < \ldots < j_N \leq L$ so that there are no doubly-occupied sites, $N_D=0$. The spins σ_l are arbitrary so that for each set of j_l this state is 2^N-fold degenerate. Obviously it describes an insulator since no motion of electrons is possible.

Summarizing we have seen that the limit $U=0$ is simple in momentum space and corresponds to a metallic case. In contrast, the limit $t=0$ is simply in real space and describes an insulator. Therefore, for finite values of both t and U we can expect a *metal-insulator transition* at some critical value U_c of the Coulomb repulsion.

A typical application of the Hubbard model is *band magnetism*. Here the same electrons are responsible for electric conductivity and magnetism. It is also called *itinerant magnetism* in contrast to spins due to localized moments that we have discussed in Chapter II.

Another application are so-called *Mott insulators*. Some transition metal oxides, in contrast to predictions of band theory. E.g. MnO has a partially filled 3d-band of Mn, but nevertheless is insulating. The reason for the failure of band theory is that the interactions between the electrons are not taken into account properly.

IV.2 Symmetries

We now investigate the symmetries of the Hubbard Hamiltonian. This will help us to identify the fundamental region of the model. Also it is essential for a discussion of the completeness of Bethe Ansatz states.

First, we have some obvius conservation laws, namely the total numbers

$$N_\sigma = \sum_{j=1}^{L} n_{j\sigma} \tag{IV.2.1}$$

of electrons with spin $\sigma =\uparrow, \downarrow$. Then, of course, also the total number $N = N_\uparrow + N_\downarrow$ of electrons is conserved. These conservation laws hold in any dimension, not only for the one-dimensional case. Another obvious symmetry is the spin-flip symmetry $\sigma \to -\sigma$.

On a bipartite lattice with sublattices A and B, e.g. a one-dimensional chain, we can perform the rotation

$$U = \exp\left(i\pi \sum_{j\in A} \sum_{\sigma} c^\dagger_{j\sigma} c_{j\sigma} \right) \tag{IV.2.2}$$

on one of the sublattices. This transformation implies

$$\begin{aligned} j \in A: \qquad & \tilde{c}_{j\sigma} = -c_{j\sigma}, \quad \tilde{c}^\dagger_{j\sigma} = -c^\dagger_{j\sigma}, \\ j \in B: \qquad & \tilde{c}_{j\sigma} = c_{j\sigma}, \quad \tilde{c}^\dagger_{j\sigma} = c^\dagger_{j\sigma}. \end{aligned} \tag{IV.2.3}$$

On both sublattices we have $\tilde{n}_{j\sigma} = n_{j\sigma}$. Therefore the Hamiltonian $\mathcal{H}(t, U)$ is transformed

$$\mathcal{H}(t, U) \quad \longrightarrow \quad \mathcal{H}(-t, U). \tag{IV.2.4}$$

This once again shows what we already know from the free fermion case, namely that the sign of the kinetic energy is not important for the spectrum. Usually one chooses $t = 1$. Another consequence of this symmetry is that the spectrum of $\mathcal{H}(t, -U)$ is just the spectrum of $\mathcal{H}(t, U)$ inverted, i.e.

$$\mathcal{H}(t, -U) \quad \longrightarrow \quad -\mathcal{H}(t, U). \tag{IV.2.5}$$

Next we discuss the *particle-hole symmetry*. Consider the transformation

$$\tilde{c}_{j\sigma} = (-1)^j c^\dagger_{j\sigma}, \qquad \tilde{c}^\dagger_{j\sigma} = (-1)^j c_{j\sigma} \tag{IV.2.6}$$

where $(-1)^j = 1$ for $j \in A$ and $(-1)^j = -1$ for $j \in B$. This transformation implies $\tilde{n}_{j\sigma} = 1 - n_{j\sigma}$ for the particle-number operators and implies for the Hamiltonian

$$\begin{aligned} \mathcal{H}(t, U) \quad & \longrightarrow \quad \mathcal{H}(t, U) + U(L - N) \\ N \text{ electrons} \quad & \longrightarrow \quad 2L - N \text{ electrons.} \end{aligned} \tag{IV.2.7}$$

One can also perform a *partial particle-hole transformation* for only one spin direction σ, such that the $-\sigma$-operators are unchanged. In this case one finds

$$\mathcal{H}(t, U) \quad \longrightarrow \quad \mathcal{H}(t, -U) - UN_{-\sigma} \tag{IV.2.8}$$

$$(N_\sigma, N_{-\sigma}) \text{ electrons} \quad \longrightarrow \quad (L - N_\sigma, N_{-\sigma}) \text{ electrons.} \tag{IV.2.9}$$

Summarizing these results, the groundstate energy $E_0(N_\uparrow, N_\downarrow; U)$ for a system with $N_\uparrow$ $\uparrow$-electrons and $N_\downarrow$ $\downarrow$-electrons satisfies

$$E_0(N_\uparrow, N_\downarrow; U) = E_0(L - N_\uparrow, L - N_\downarrow; U) - (L - N_\uparrow - N_\downarrow)U \tag{IV.2.10}$$

$$= E_0(N_\uparrow, L - N_\downarrow; -U) + N_\uparrow U \tag{IV.2.11}$$

$$= E_0(L - N_\uparrow, N_\downarrow; -U) + N_\downarrow U. \tag{IV.2.12}$$

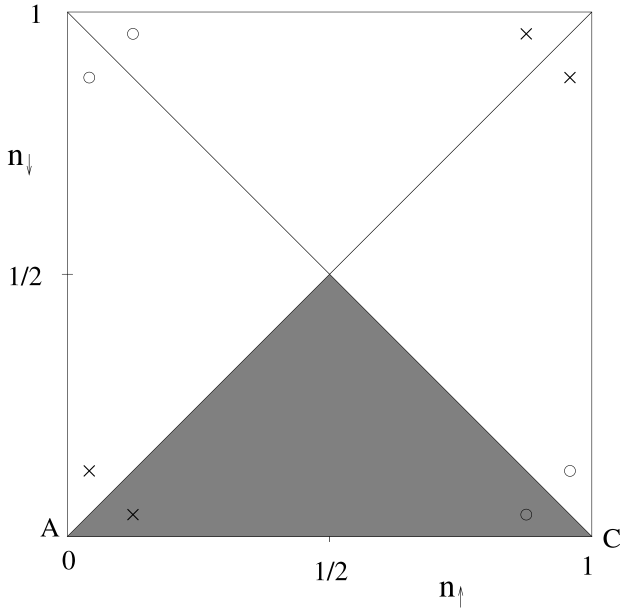

Figure IV.2.1: The density parameter space of the Hubbard model. $n_\sigma = N_\sigma/L$ are the densities of spin-σ electrons. The grey region corresponds to the fundamental region. Also special points have been indicated (see text). The circles and crosses denote parameter values that are equivalent due to the particle-hole and spinflip symmetries of the Hamiltonian, respectively.

Therefore it is sufficient to investigate the Hubbard model only in the fundamental region

$$N_\uparrow + N_\downarrow \le 1, \qquad N_\uparrow \ge N_\downarrow. \tag{IV.2.13}$$

This is shown in Fig. IV.2.1. In this figure also special points and lines are shown. Point A corresponds to the empty chain. The line BC is the so-called *half-filled band* where the particle number is $N = L$, i.e. the total particle density $n = \frac{N}{L} = 1$. The line AC is ferromagnetic since here only $\uparrow$-spins exist. AB corresponds to the subspace with magnetization 0. Furthermore sets of equivalent points are shown.

There are further symmetries which are important for the Bethe Ansatz. We have already introduced the spin operators (see Exercise 1)

$$S_j^+ = c_{j\uparrow}^\dagger c_{j\downarrow}, \qquad S_j^- = c_{j\downarrow}^\dagger c_{j\uparrow}, \qquad S_j^z = \frac{1}{2}(n_{j\uparrow} - n_{j\downarrow}) \tag{IV.2.14}$$

that satisfy the standard SU(2)-spin algebra. The global spin operators commute with the Hubbard Hamiltonian, i.e.

$$[\mathcal{H}(t,U), S^\alpha] = 0 \qquad \text{for} \quad S^\alpha = \sum_{j=1}^{L} S_j^\alpha \qquad (\alpha = +, -, z). \tag{IV.2.15}$$

This shows that the Hubbard model has a SU(2)-spin symmetry.

However, there is another SU(2)-symmetry of the model [4] that is generated by the pseudospin or $\eta-$pairing operators

$$\eta^\dagger = \sum_{j=1}^{L} (-1)^j c_{j\downarrow}^\dagger c_{j\uparrow}^\dagger, \qquad \eta = \sum_{j=1}^{L} (-1)^j c_{j\uparrow} c_{j\downarrow}, \qquad \eta^z = \frac{1}{2}\left(\hat{N} - L\right). \tag{IV.2.16}$$

These operators form a SU(2)-algebra analogous to that of the spin operators (IV.2.14) (see Exercise 39). Note that $\eta^\dagger$ creates a doubly-occupied site with momentum π.
If we rewrite the Hamiltonian as

$$\tilde{\mathcal{H}}(t,U) = -t\sum_{j=1}^{L}\sum_{\sigma=\uparrow,\downarrow}\left(c_{j\sigma}^\dagger c_{j+1,\sigma} + c_{j+1,\sigma}^\dagger c_{j\sigma}\right) + U\sum_{j=1}^{L}(n_{j\uparrow}-\frac{1}{2})(n_{j\downarrow}-\frac{1}{2}) = \mathcal{H}(t,U) - \frac{U}{2}N + \frac{U}{4}L \tag{IV.2.17}$$

one can show that

$$\left[\tilde{\mathcal{H}}(t,U), \eta^\alpha\right] = 0. \tag{IV.2.18}$$

The commutation relation with the original Hamiltonian are discussed in Exercise 39. This result is not completely surprising since the η-operators can be obtained from the spin operators by a partial particle-hole transformation.
Similar to the spin we can also define the total η-pseudo-spin

$$\eta_{\text{tot}}^2 = \frac{1}{2}\left(\eta^\dagger\eta + \eta\eta^\dagger\right) + (\eta^z)^2. \tag{IV.2.19}$$

It is easy to check that this operator commutes with $\mathcal{H}(t,U)$.
Since all η-operators commute with all spin-operators, it seems that the Hubbard model has a SU(2)$\otimes$ SU(2) symmetry. However, the two SU(2) symmetries are not completely independent. For fixed chain length L we have

$$S^z + \eta^z = \frac{1}{2}(N_\uparrow - N_\downarrow) + \frac{1}{2}(N - L) = N_\downarrow - \frac{L}{2} \tag{IV.2.20}$$

which is always integer (for L even). Therefore the symmetry group is in fact smaller. It is

$$\mathrm{SO}(4) = \mathrm{SU}(2) \otimes \mathrm{SU}(2)/\mathbb{Z}_2. \tag{IV.2.21}$$

This SO(4) symmetry plays an important role for the completeness of the Bethe Ansatz states since, as we have already seen for the XXZ model, the Bethe Ansatz typically only generates states that are highest weight state of the underlying symmetry algebra.

IV.3 Bethe Ansatz

The application of the coordinate Bethe Ansatz to the Hubbard model [2, 3] is at first rather similar to the XXZ model. However, in the XXZ model we found that the effects of the interactions are encoded in a scattering phase shift. This will be different in the Hubbard model. Here we have to take into account the internal degree of freedom of the particles, i.e. the spin σ. Therefore their interactions are described by a scattering matrix, not just a phase factor. We will see that this scattering matrix must also be integrable in some sense for the whole model to be integrable. Then one has to perform a second Bethe Ansatz in order to obtain the eigenfunctions of the Hubbard Hamiltonian.

We start by writing down the Schrödinger equation for a system of N electrons in first quantized form:

$$-t\sum_{j=1}^{L}\left[\psi_{\sigma_1\cdots\sigma_N}(x_1,\ldots,x_j-1,\ldots,x_N)+\psi_{\sigma_1\cdots\sigma_N}(x_1,\ldots,x_j+1,\ldots,x_N)\right]$$
$$+U\sum_{j<l}\delta_{x_j,x_l}\delta_{\sigma_j,-\sigma_l}\psi_{\sigma_1\cdots\sigma_N}(x_1,\ldots,x_N)=E\psi_{\sigma_1\cdots\sigma_N}(x_1,\ldots,x_N). \qquad \text{(IV.3.1)}$$

Here x_j is the position of particle j and σ_j its spin. The first sum is the kinetic energy that describes the hopping of electrons to neighbouring sites. The second term is the Coulomb interaction. It basically counts the number of doubly occupied sites $x_j = x_l$ in which case the spins σ_j and σ_l of the involved electrons have to be different due to the Pauli principle. Of course the wavefunction should also satisfy the antisymmetry condition for fermions. Note that, similar to the case of the Heisenberg model, later also unphysical amplitudes will appear in the calculations where formally two electrons with the same spin sit on the same site. However, these unphysical amplitudes do not have to bother us since the corresponding term, e.g. $\psi_{\sigma\sigma}(x,x)c^\dagger_{x\sigma}c^\dagger_{x\sigma}|0\rangle$ of the wavefunction vanishes identically due to the Fermi statistics of the c-operators! In the following we will set $t=1$.

First we have a brief look at the case $N=1$. Here wthere is only one electron and no interaction, i.e. a single free fermion. The Schrödinger equation (IV.3.1) reduces to

$$-\psi_\sigma(x-1)-\psi_\sigma(x+1)=E\psi_\sigma(x) \qquad \text{(IV.3.2)}$$

which is solved by

$$\psi_\sigma(x)=A_\sigma e^{ikx}, \qquad E=-2\cos k. \qquad \text{(IV.3.3)}$$

For the case $N=2$ interactions are possible. As in the treatment of the XXZ model we will distinguish two cases: (i) $x_1\neq x_2$, i.e. no interaction between the two electrons, and (ii) $x_1=x_2$ where the Coulomb interaction has to be taken into account. Case (i) will give the general structure of the wavefunction. Then it has to be checked whether case (ii) can be treated consistently with this structure by imposing conditions on the free parameters (amplitudes).

For $x_1\neq x_2$ the Schrödinger equation reads

$$-\psi_{\sigma_1\sigma_2}(x_1-1,x_2)-\psi_{\sigma_1\sigma_2}(x_1+1,x_2)-\psi_{\sigma_1\sigma_2}(x_1,x_2-1)-\psi_{\sigma_1\sigma_2}(x_1,x_2+1)=E\psi_{\sigma_1\sigma_2}(x_1,x_2) \qquad \text{(IV.3.4)}$$

which is solved by

$$\psi_{\sigma_1\sigma_2}(x_1,x_2)=A^{(1)}_{\sigma_1\sigma_2}e^{ik_1x_1+ik_2x_2}+A^{(2)}_{\sigma_1\sigma_2}e^{ik_2x_1+ik_1x_2} \qquad \text{(IV.3.5)}$$

with arbitrary coefficients $A^{(1)}_{\sigma_1\sigma_2}$, $A^{(2)}_{\sigma_1\sigma_2}$ and energy

$$E = -2(\cos k_1 + \cos k_2). \tag{IV.3.6}$$

For $x_1 = x_2 = x$ we have

$$\begin{aligned} -\psi_{\sigma_1\sigma_2}(x-1,x) - \psi_{\sigma_1\sigma_2}(x+1,x) - \psi_{\sigma_1\sigma_2}(x,x-1) - \psi_{\sigma_1\sigma_2}(x,x+1) \\ +U\delta_{\sigma_1,-\sigma_2}\psi_{\sigma_1\sigma_2}(x,x) = E\psi_{\sigma_1\sigma_2}(x,x). \end{aligned} \tag{IV.3.7}$$

We try to solve these equations using a wavefunction of Bethe form but with different coefficients for $x_1 < x_2$ and $x_1 > x_2$:

$$x_1 < x_2: \quad \psi_{\sigma_1\sigma_2}(x_1,x_2) = A_{\sigma_1\sigma_2}(k_1,k_2)e^{ik_1x_1+ik_2x_2} - A_{\sigma_1\sigma_2}(k_2,k_1)e^{ik_2x_1+ik_1x_2}, \tag{IV.3.8}$$

$$x_2 < x_1: \quad \psi_{\sigma_1\sigma_2}(x_1,x_2) = A_{\sigma_2\sigma_1}(k_2,k_1)e^{ik_1x_1+ik_2x_2} - A_{\sigma_2\sigma_1}(k_1,k_2)e^{ik_2x_1+ik_1x_2}. \tag{IV.3.9}$$

This two cases can be written in a unified form that also anticipates the form of the expression for $N > 2$. For given coordinates x_1, x_2 there exists a permutation[2] of $(1,2)$ for which $x_{Q_1} < x_{Q_2}$. Then we can write

$$\psi_{\sigma_1\sigma_2}(x_1,x_2) = \sum_{P\in S_2}(-1)^P A_{\sigma_{Q_1}\sigma_{Q_2}}(P,Q)e^{ik_{P_1}x_1+ik_{P_2}x_2} \qquad (\text{for } \ x_{Q_1} < x_{Q_2}) \tag{IV.3.10}$$

where S_2 is the symmetric group of all permutations of two elements and $(-1)^P$ denotes the sign of the permutation.
The continuity of the wavefunction at $x_1 = x_2$ requires

$$A_{\sigma_1\sigma_2}(k_1,k_2) - A_{\sigma_1\sigma_2}(k_1,k_2) = A_{\sigma_2\sigma_1}(k_2,k_1) - A_{\sigma_2\sigma_1}(k_1,k_2). \tag{IV.3.11}$$

Inserting the Ansatz into (IV.3.7) yields the condition

$$\begin{aligned} -2(\cos k_1 + \cos k_2)\left[A_{\sigma_1\sigma_2}(k_1,k_2) - A_{\sigma_1\sigma_2}(k_2,k_1)\right] = E\left[A_{\sigma_1\sigma_2}(k_1,k_2) - A_{\sigma_1\sigma_2}(k_2,k_1)\right] \\ +\big\{\left[A_{\sigma_2\sigma_1}(k_2,k_1) - A_{\sigma_1\sigma_2}(k_1,k_2)\right]\left(e^{ik_1} - e^{-ik_1} - e^{ik_2} + e^{-ik_2}\right) \\ -U\left[A_{\sigma_1\sigma_2}(k_1,k_2) - A_{\sigma_1\sigma_2}(k_2,k_1)\right]\big\}. \end{aligned} \tag{IV.3.12}$$

If the term in the brackets $\{\ldots\}$ vanishes, this becomes an eigenvalue equation for the eigenfunction $A_{\sigma_1\sigma_2}(k_1,k_2) - A_{\sigma_1\sigma_2}(k_2,k_1)$ with eigenvalue $-2(\cos k_1 + \cos k_2)$. The vanishing condition can be rewritten as

$$A_{\sigma_1\sigma_2}(k_2,k_1) = \frac{-U/2i}{\sin k_1 - \sin k_2 - U/2i}A_{\sigma_1\sigma_2}(k_1,k_2) + \frac{\sin k_1 - \sin k_2}{\sin k_1 - \sin k_2 - U/2i}A_{\sigma_2\sigma_1}(k_1,k_2) \tag{IV.3.13}$$

or

$$A_{\sigma_1\sigma_2}(k_2,k_1) = \sum_{\sigma_1',\sigma_2'} S^{\sigma_1\sigma_1'}_{\sigma_2\sigma_2'}(k_1,k_2)A_{\sigma_1'\sigma_2'}(k_1,k_2) \tag{IV.3.14}$$

[2]Clearly $Q = I$ for $x_1 < x_2$ and $Q = \tau$ for $x_2 < x_1$ where τ is the transposition.

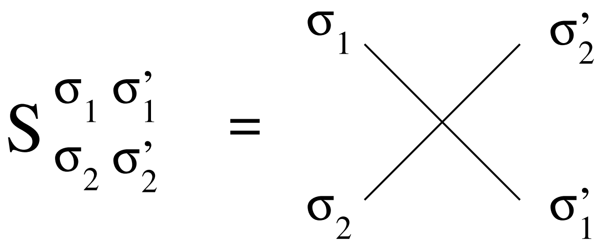

Figure IV.3.1: Graphical representation of the 2-particle scattering matrix. One can consider the horizontal axis as time direction and the vertical axis as space direction. The S-matrix then describes a scattering process of two incoming particles with spins σ_1, σ_2 into two outgoing particles with spins σ_1', σ_2'.

with the *2-particle scattering matrix*

$$S^{\sigma_1\sigma_1'}_{\sigma_2\sigma_2'}(k_1,k_2) = \frac{\sin k_1 - \sin k_2}{\sin k_1 - \sin k_2 - U/2i}\delta_{\sigma_1\sigma_1'}\delta_{\sigma_2\sigma_2'} + \frac{-U/2i}{\sin k_1 - \sin k_2 - U/2i}\delta_{\sigma_1\sigma_2'}\delta_{\sigma_2\sigma_1'}. \tag{IV.3.15}$$

Fig. IV.3.1 shows a graphical representation of the S-matrix. It can be written in a much more compact operator form

$$S(\vartheta) = \frac{\vartheta + iU/2 \cdot P_{12}}{\vartheta + iU/2} \tag{IV.3.16}$$

with $\vartheta = \sin k_1 - \sin k_2$ and the permutation operator P_{12},

$$P^{\sigma_1\sigma_1'}_{\sigma_2\sigma_2'} = \delta_{\sigma_1\sigma_2'}\delta_{\sigma_2\sigma_1'} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{IV.3.17}$$

that exchanges two particles (spins). In Fig. IV.3.2 we give an interpretation of the scattering matrix elements. These elements give the amplitudes for the two possible outcomes of the scattering process. Either the particles can pass each other, i.e. they keep their momenta and spins. Alternatively the particles could exchange their spins so that after scattering particle 1 has momentum k_1 and spin σ_2 and particle 1 momentum k_2 and spin σ_1.[3] The first process corresponds to the first part of the S-matrix (IV.3.16) which is proportional to the identy matrix (see Fig. IV.3.1), whereas the second one corresponds to the second term propertional to the permuation P_{12}. The coefficients of the terms are then the amplitudes for the two processes.
As final step, as in the case of the Heisenberg model, we have to consider the effect of the periodic boundary conditions which lead to a quantization of the momenta. In fact, we have $\psi_{\sigma_1\sigma_2}(x_1 + L, x_2) = \psi_{\sigma_1\sigma_2}(x_1, x_2)$ where it should be kept in mind that $x_1 + L > x_2$ even if $x_1 < x_2$ so that different forms of the wavefunctions have to be used. This leads to a condition of the form

$$e^{iLk_j} A_{\sigma_1\sigma_2}(k_1,k_2) = \sum_{\sigma_1',\sigma_2'} (T_j)_{(\sigma_1,\sigma_2),(\sigma_1',\sigma_2')} A_{\sigma_1'\sigma_2'} \tag{IV.3.18}$$

[3]This is just a convention, since an equivalent interpretation would be an exchange of momenta instead of spins.

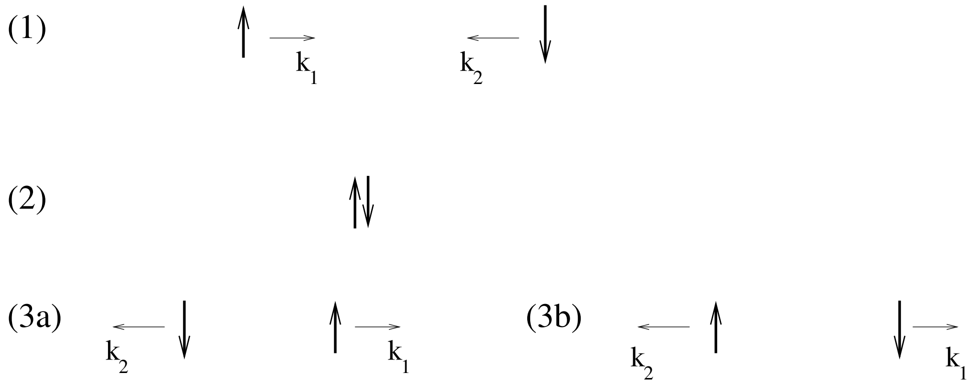

Figure IV.3.2: Interpretation of the 2-particle S-matrix of the Hubbard model. (1) Two particles with different spins $\sigma_1 =\uparrow$ and $\sigma_2 =\downarrow$ and momenta k_1 and k_2 approach each other and scatter by forming a doubly-occupied site (2). After the scattering process two situations are possible. In (3a) the $\downarrow$-spin is now to the left of the $\uparrow$-spin. This can be interpreted as an uninfluenced passing of the electrons since both particles keep their momenta and is described by the part of the scattering matrix proportional to the unit matrix I. In (3b) the ordering of spins has not changed. This can be interpreted as a scattering process where the particles (momenta) have exchanged their spins and is described by the permutation part P_{12} of the scattering matrix.

where the matrix T_j is basically a 2-scattering S-matrix. We will later give its general form explicitly.
We now generalize our treatment to the case of N particles. The Bethe Ansatz wavefunction then has the form

$$\boxed{\psi_{\sigma_1\dots\sigma_N}(x_1,\dots,x_N) = \sum_{P\in S_N} (-1)^P A_{\sigma_{Q_1},\dots,\sigma_{Q_N}}(P,Q) e^{i\sum_{j=1}^N k_{P_j}x_j} \qquad \text{(IV.3.19)}}$$

where again the permutation Q has to be determined such that $x_{Q_1} < x_{Q_2} < \ldots < x_{Q_N}$. In the following we will use $X_Q = \{x \in \mathbb{R}^N / x_{Q_1} < x_{Q_2} < \ldots < x_{Q_N}\}$ to denote the set of all points that satisfies this ordering.
It is easy to see that this wavefunction satisfies the Schrödinger equation (IV.3.1) for the noninteracting case when no two coordinates are identical. In this case the energy is given by

$$E = -2\sum_{j=1}^N \cos k_j. \qquad \text{(IV.3.20)}$$

To incorporate the interaction effects, i.e. the cases where at least two coordinates are identical, we generalize our treatment of the case $N = 2$, e.g. by requiring continuity of the wavefunction at the boundaries of the regions X_Q.
Neighbouring sectors X_Q and $X_{\tilde{Q}}$ where e.g. $x_{Q_j} = x_{\tilde{Q}_l}$ and $x_{Q_l} = x_{\tilde{Q}_j}$ are connected by the 2-particle S-matrix. For convenience we introduce the notation

$$\bar{A}_{\sigma_1,\dots,\sigma_N}(Q|P) := A_{\sigma_{Q_1},\dots,\sigma_{Q_N}}(P,Q). \qquad \text{(IV.3.21)}$$

Then the coefficients are related through the 2-particle S-matrix by

$$\bar{A}_{\sigma_1,\dots,\sigma_N}(\tilde{Q}|P) = \sum_{\sigma_j',\sigma_l'} S^{\sigma_l,\sigma_l'}_{\sigma_j,\sigma_j'}(k_j,k_l)\bar{A}_{\sigma_1,\dots,\sigma_j',\dots,\sigma_l',\dots,\sigma_N}(Q|P). \tag{IV.3.22}$$

Similar to what we observed in the treatment of the Heisenberg model such a relation allows to express all amplitudes through $\bar{A}(I|I)$, where $\bar{A}$ is vector with components $\bar{A}_{\sigma_1,\dots,\sigma_N}$, as a product of S-matrices. However, there are different ways of achieving this using (IV.3.22). The reason is that any permutation Q can be expressed by transpositions, but usually not in a unique way. One must make sure that these different ways give the same result. Then the 2-particle S-matrix has to satisfy the Yang-Baxter equation

$$\sum_{\sigma_1',\sigma_2',\sigma_3'} S^{\sigma_1\sigma_1'}_{\sigma_2\sigma_2'}(\vartheta)S^{\sigma_1'\sigma_1''}_{\sigma_3\sigma_3'}(\tilde{\vartheta})S^{\sigma_2'\sigma_2''}_{\sigma_3'\sigma_3''}(\tilde{\vartheta}-\vartheta) = \sum_{\sigma_1',\sigma_2',\sigma_3'} S^{\sigma_2\sigma_2'}_{\sigma_3\sigma_3'}(\tilde{\vartheta}-\vartheta)S^{\sigma_1\sigma_1'}_{\sigma_3'\sigma_3''}(\tilde{\vartheta})S^{\sigma_1'\sigma_1''}_{\sigma_2'\sigma_2''}(\vartheta). \tag{IV.3.23}$$

Imposing periodic boundary conditions $\psi(x_1,\dots,x_j,\dots,x_N) = \psi(x_1,\dots,x_j+L,\dots,x_N)$ for the wavefunction (IV.3.19) leads to the condition

$$e^{ik_jL}\bar{A}_{\underline{\sigma}}(I|I) = (T_j)_{\underline{\sigma},\underline{\sigma}'}\,\bar{A}_{\underline{\sigma}'}(I|I) \tag{IV.3.24}$$

where we have introduced the notation $\underline{\sigma}$ for the vector with components $\sigma_1,\dots,\sigma_N$. The 'transfer-matrices' T_j are given by

$$\begin{aligned}(T_j)_{\underline{\sigma},\underline{\sigma}'} = \sum_{(\sigma_1'',\sigma_2'',\dots,\sigma_N'')} & S^{\sigma_j\sigma_1''}_{\sigma_{j-1}\sigma_{j-1}'}(k_j,k_{j-1})S^{\sigma_1''\sigma_2''}_{\sigma_{j-2}\sigma_{j-2}'}(k_j,k_{j-2})\cdots S^{\sigma_{j-2}''\sigma_{j-1}''}_{\sigma_1\sigma_1'}(k_j,k_1)\,\cdot \\ & \cdot S^{\sigma_{j-1}''\sigma_j''}_{\sigma_N\sigma_N'}(k_j,k_N)\cdots S^{\sigma_{N-1}''\sigma_j'}_{\sigma_{j+1}\sigma_{j+1}'}(k_j,k_{j+1}).\end{aligned} \tag{IV.3.25}$$

This is the generalization of (II.3.43) to multicomponent models. The interpretation is the same as the one given for (II.3.43): Carrying particle j once around the lattice it picks up a kinematical phase k_jL. Each scattering process with the $N-1$ other particles generates a further phase shift which now depends on the spins of the particles. Therefore we have no longer just a scattering phase, but a scattering matrix. Nevertheless, all scattering processes can be reduced to 2-particle processes which is the essential property of exactly solvable models. In the case of multicomponent models it is far from trivial since now the 2-particle scattering matrix has to satisfy the Yang-Baxter equation.

In order to see whether the S-matrix (IV.3.16) of the Hubbard model satisfies the Yang-Baxter equation we rewrite it explicitly in matrix form

$$S(\vartheta) = \frac{\vartheta + iU/2\cdot P_{12}}{\vartheta+iU/2} = \frac{1}{\vartheta+iU/2}\begin{pmatrix}\vartheta+iU/2 & 0 & 0 & 0\\ 0 & \vartheta & iU/2 & 0\\ 0 & iU/2 & \vartheta & 0\\ 0 & 0 & 0 & \vartheta+iU/2\end{pmatrix}. \tag{IV.3.26}$$

We see that it has the same structure as the $\mathcal{L}$-operator (II.5.4) of the 6-vertex model. Therefore it is clear that S also satisfies the Yang-Baxter equation.

In fact, T_j can be interpreted transfer matrix of an *inhomogeneous 6-vertex model* where the vertex weight depends on the position of the vertex. In the homogeneous vertex models treated in Sec. II.5 the monodromy matrix was given by $\mathcal{T}(u) = \mathcal{L}_1(u)\mathcal{L}_2(u)\cdots\mathcal{L}_L(u)$ (see (II.5.5)). In our case it is of the form $\mathcal{T}_{\text{inhom}}(u) = \mathcal{L}_1(u-u_1^0)\mathcal{L}_2(u-u_2^0)\cdots\mathcal{L}_N(u-u_N^0)$, i.e. the argument of each $\mathcal{L}$-operator is shifted by a constant u_j^0. This is a special kind of inhomogeneity where the $\mathcal{L}$-operators are just shifted along the commutation lines (see Fig. II.4.8). Since all operators on this line commute this already indicates that this kind of inhomogeneity does not cause problems for the exact solvability of the model.

In our case the S-matrix effectively is a function of $\vartheta = \sin k_j - \sin k_l$. Therefore $\sin k_l$ takes the role of the inhomogeneity u_l^0. We can then construct the eigenvectors and eigenvalues of T_j by using the algebraic Bethe Ansatz. As reference state we have as usual $|0\rangle = |\uparrow\uparrow\ldots\uparrow\rangle$. An eigenvector with $N_\downarrow$ $\downarrow$-spins can be constructed as $\prod_{l=1}^{N_\downarrow} B(\vartheta_j)|0\rangle$ if the parameters ϑ_j satisfy the Bethe-Ansatz equations. We need the corresponding eigenvalue function $\Lambda(u)$, which is an inhomogeneous generalization of (II.5.15), only for the special value $u = u_j^0$. The reason is our S-matrix entering the definition (IV.3.25) has the argument $\sin k_j - \sin k_l$. It is given (with small modifications) by (II.5.15). We only need the eigenvalue for the argument $u = u_j^0$. Here it simplifies to

$$\Lambda(u = u_j^0; \vartheta_1, \ldots, \vartheta_{N_\downarrow}; u_1^0, \ldots, u_{N_\downarrow}^0) = \prod_{l=1}^{N_\downarrow} \frac{u_j^0 - \vartheta_l + iU/2}{u_j^0 - \vartheta_l} \tag{IV.3.27}$$

where ϑ_j denotes the Bethe-Ansatz parameters and $u_j^0 = \sin k_j$ the inhomogeneities. As already mentioned the Bethe parameters ϑ_j have to satisfy the appropriate inhomogeneous generalization of the 6-vertex model Bethe equations. We will give them explicitly below.

After introduction of the new variables λ_α with $\vartheta_j = \lambda_j - i\frac{U}{4}$, the Bethe-Ansatz equations for the Hubbard model are given by [2]

$$\boxed{\begin{aligned} e^{iLk_j} &= \prod_{\alpha=1}^{N_\downarrow} \frac{\sin k_j - \lambda_\alpha + iU/4}{\sin k_j - \lambda_\alpha - iU/4}, \qquad (j = 1, \ldots, N), && \text{(IV.3.28)} \\ \prod_{j=1}^{N} \frac{\lambda_\alpha - \sin k_j + iU/4}{\lambda_\alpha - \sin k_j - iU/4} &= -\prod_{\beta=1}^{N_\downarrow} \frac{\lambda_\alpha - \lambda_\beta + iU/2}{\lambda_\alpha - \lambda_\beta - iU/2} \qquad (\alpha = 1, \ldots, N_\downarrow). && \text{(IV.3.29)} \end{aligned}}$$

Note that these equations are only valid for $N_\downarrow \le \frac{N}{2}$ and $N \le L$. The first set of equations comes from the periodic boundary conditions (IV.3.24) where the explicit form (IV.3.27) of the eigenvalue the transfer matrix T_j has been inserted. The second set of equations has its origin in the Bethe Ansatz for T_j. Obiviously it is an inhomogeneous generalization of the equations for the isotropic Heisenberg chain. Neglecting the terms $\sin k_l$ coming from the wavenumbers they reduce to (II.6.6) (after rescaling the λ-variables by $U/2$).

A few remarks should be made regarding the Bethe Ansatz for the Hubbard model.

- In contrast to the Heisenberg model we have used the Bethe Ansatz *twice* for the Hubbard model. In a first step we have written down a Bethe Ansatz wavefunction which captures the motion of electrons. This lead to the introduction of wavenumbers k_j. However, it turn out the the motion of electrons depends on the spin configuration through the amplitudes of the wavefunction. This dependence has then been treated by a second Bethe Ansatz which allowed a diagonalization of the spin-dependent amplitudes by introducing the spin rapidities λ_α. Such a twofold application of the Bethe Ansatz is known as *nested Bethe Ansatz*.

- The Bethe Ansatz wavefunction and equations clearly show that motion and spin dynamics are not independent of each other. They are coupled in an intricate way.
- The Bethe Ansatz states are highest weight states of the SO(4)-algebra, i.e.

$$S^+|\mathrm{BA}\rangle = 0, \qquad \eta|\mathrm{BA}\rangle = 0 \tag{IV.3.30}$$

where $|\mathrm{BA}\rangle$ denotes an arbitrary Bethe Ansatz state. Note that we do not have $\eta^\dagger|\mathrm{BA}\rangle = 0$ which is related to the fact that we are a parameter regime where $\eta^z \leq 0$.

- The original solution of the Hubbard model by Lieb and Wu [2] was based on the previous Bethe Ansatz treatment of the Fermi gas with (repulsive) δ-interaction by C.N. Yang [5]. It is defined by the Hamiltonian

$$\mathcal{H} = -\sum_{j=1}^{N} \frac{\partial^2}{\partial x_j^2} + 2c\sum_{j<l} \delta(x_j - x_l) \tag{IV.3.31}$$

($c > 0$) which can be regarded as the continuum version of the Hubbard model.

In the following section we will discuss the solutions of the Bethe Ansatz equations that correspond to the groundstate and the low-lying excitations.

IV.4 Groundstate and metal-insulator transition

In logarithmic form the Bethe Ansatz equations read

$$Lk_j = 2\pi I_j + \sum_{\beta=1}^{N_\downarrow} \theta(2\sin k_j - 2\lambda_\beta), \tag{IV.4.1}$$

$$\sum_{j=1}^{N} \theta(2\sin k_j - 2\lambda_\alpha) = 2\pi J_\alpha - \sum_{\beta=1}^{N_\downarrow} \theta(\lambda_\alpha - \lambda_\beta) \tag{IV.4.2}$$

where we have introduced

$$\theta(x) = -2\arctan(2x/U) \tag{IV.4.3}$$

As in the case of the Hubbard model the allowed values of the Bethe Ansatz quantum numbers I_j and J_α depend on the parities of L, N and $N_\downarrow$:

$$I_j \equiv \frac{N_\downarrow}{2} \mod 1, \qquad J_\alpha \equiv \frac{N - N_\downarrow + 1}{2} \mod 1, \tag{IV.4.4}$$

which e.g. means that the I_j are integers, if $N_\downarrow$ is even, and half-odd integers if $N_\downarrow$ is odd.
As in the Heisenberg model we can use the logarithmic form of the Bethe Ansatz equations to express the momentum P of the eigenstate through the Bethe quantum numbers I_j and J_α. Summing over the first equation and using the second we then obtain (see Exercise 40)

$$P = \sum_{j=1}^{N} k_j = \frac{2\pi}{L}\left(\sum_{j=1}^{N} I_j + \sum_{\alpha=1}^{N_\downarrow} J_\alpha\right). \tag{IV.4.5}$$

IV.4.1 Groundstate

For the groundstate we have to choose the Bethe quantum numbers

$$I_j = j - \frac{N+1}{2}, \qquad J_\alpha = \alpha - \frac{N_\downarrow + 1}{2} \tag{IV.4.6}$$

with $j = 1, \ldots, N$ und $\alpha = 1, \ldots, N_\downarrow$ (see Exercise 40).
We introduce the density functions

$$\rho(k_j) = \lim_{L\to\infty} \frac{1}{L(k_{j+1} - k_j)}, \qquad \sigma(\lambda_\alpha) = \lim_{L\to\infty} \frac{1}{L(\lambda_{\alpha+1} - \lambda_\alpha)}. \tag{IV.4.7}$$

The density is then given by (with the abbreviation $\Delta k_j = k_{j+1} - k_j$)

$$n = \frac{N}{L} = \frac{1}{L}\sum_{j=1}^{N} 1 = \sum_{j=1}^{N} \rho(k_j)\Delta k_j \longrightarrow \int_{-Q}^{Q} \rho(k)dk. \tag{IV.4.8}$$

Sometimes instead of density one uses the "filling ν" which is related to the density via $n = 2\nu$. Therefore half-filling means $n = 1$.
Thus we see that the density is directly related to the interval length $2Q$ over which the k_j are distributed. One can show that the density $n(Q)$ is a monotonically increasing function of Q. Obviously we have $n(Q = 0) = 0$. Later we will see that $n(Q = \pi) = 1$.
In a similar way (see Exercise 40) we can derive the following integral equations for the groundstate densities from the logarithmic Bethe Ansatz equations:

$$2\pi\rho(k) = 1 - 2\cos k \int_{-B}^{B} \theta'(2\sin k - 2\lambda)\sigma(\lambda)d\lambda \tag{IV.4.9}$$

$$2\pi\sigma(\lambda) = \int_{-B}^{B} \theta'(\lambda - \tilde{\lambda})\sigma(\tilde{\lambda})d\tilde{\lambda} - 2\int_{-Q}^{Q} \theta'(2\sin k - 2\lambda)\rho(k)dk \tag{IV.4.10}$$

with

$$\theta'(x) = -\frac{U/2}{x^2 + (U/2)^2}. \tag{IV.4.11}$$

For $N_\downarrow = N_\uparrow$ we have $B = \infty$. Solving the second equation by Fourier transform, the first equation can be reduced to

$$\boxed{2\pi\rho(k) = 1 + \cos k \int_{-Q}^{Q} h(\sin k - \sin \tilde{k})\rho(\tilde{k})d\tilde{k} \qquad \text{(IV.4.12)}}$$

where the integral kernel $h(x)$ is given by

$$\begin{aligned} h(x) &= \frac{2}{U}\,\mathrm{Re}\left[\psi\left(1 + i\frac{x}{U}\right) - \psi\left(\frac{1}{2} + i\frac{x}{U}\right)\right] \\ &= U\sum_{m=1}^{\infty}(-1)^{m-1}\frac{m}{x^2 + (\frac{mU}{2})^2} = \int_{-\infty}^{\infty}\frac{e^{i\omega x}}{1 + e^{|\omega|U/2}}d\omega \end{aligned} \qquad \text{(IV.4.13)}$$

with the Digamma-function $\psi(x) = \frac{d}{dx}\Gamma(x)$.
For $Q < \pi$ equation (IV.4.12) can only be solved numerically. For $Q = \pi$, however, we can find a closed analytic expression. Here we follow the treatment in [6] that only uses symmetry properties of the functions involved[4]. The integral equation then reads

$$2\pi\rho_0(k) = 1 + \cos k \int_{-\pi}^{\pi} h(\sin k - \sin \tilde{k})\rho_0(\tilde{k})d\tilde{k}. \qquad \text{(IV.4.14)}$$

First, we see that obviously

$$\rho_0(k) = \rho_0(k + 2\pi) = \rho_0(-k). \qquad \text{(IV.4.15)}$$

Furthermore, by replacing k with $\pi - k$ in (IV.4.14) we find

$$2\pi\rho_0(\pi - k) = 1 - \cos k \int_{-\pi}^{\pi} h(\sin k - \sin \tilde{k})\rho_0(\tilde{k})d\tilde{k} \qquad \text{(IV.4.16)}$$

where we have made use of the well-known properties of the trigonometric functions. Adding the equations (IV.4.14) and (IV.4.16) leads to

$$\rho_0(k) + \rho_0(\pi - k) = \frac{1}{\pi}. \qquad \text{(IV.4.17)}$$

The symmetries (IV.4.15) and (IV.4.17) are sufficent to determine the explicit form of the density function $\rho_0(k)$:

$$\begin{aligned} 2\pi\rho_0(k) &= 1 + \cos k \int_{-\pi/2}^{3\pi/2} h(\sin k - \sin \tilde{k})\rho_0(\tilde{k})d\tilde{k} \\ &= 1 + \cos k \int_{-\pi/2}^{\pi/2} h(\sin k - \sin \tilde{k})\rho_0(\tilde{k})d\tilde{k} + \cos k \int_{\pi/2}^{3\pi/2} h(\sin k - \sin \tilde{k})\rho_0(\tilde{k})d\tilde{k} \\ &= 1 + \cos k \int_{-\pi/2}^{\pi/2} h(\sin k - \sin \tilde{k})(\rho_0(\tilde{k}) + \rho(\pi - k))d\tilde{k} \\ &= 1 + \frac{\cos k}{\pi}\int_{-\pi/2}^{\pi/2} h(\sin k - \sin \tilde{k})d\tilde{k} \\ &= 1 + \frac{\cos k}{2\pi}\int_{-\pi}^{\pi} h(\sin k - \sin \tilde{k})d\tilde{k}. \end{aligned} \qquad \text{(IV.4.18)}$$

[4]The standard treatment uses a Fourier analysis of the integral equation.

In the first step we have used the 2π-periodicity which implies that any integral over an interval of length 2π gives the same result. Then the integral is splitted into two parts of length π where in the second integral we substitute $k \to \pi - k$. Then we can use (IV.4.17) to eliminate the density function from the integral.
Using the symmetries it is also straightforward to show the $Q = \pi$ corresponds to half-filling:

$$\begin{aligned} n &= \int_{-\pi}^{\pi} \rho_0(k)dk = \int_{-\pi/2}^{\pi/2} \rho_0(k)dk + \int_{\pi/2}^{3\pi/2} \rho_0(k)dk \\ &= \int_{-\pi/2}^{\pi/2} (\rho_0(k) + \rho_0(\pi - k))dk = \int_{-\pi/2}^{\pi/2} \frac{1}{\pi} dk = 1. \end{aligned} \tag{IV.4.19}$$

Therefore $Q = \pi$ corresponds to half-filling.
Finally we express the energy through the density function:

$$\boxed{E_0 = -2\sum_{j=1}^{N} \cos k_j = -2L\int_{-Q}^{Q} \rho(k)\cos k dk. \qquad \text{(IV.4.20)}}$$

At half-filling we get explicitly

$$\boxed{\begin{aligned} E_0(n=1) &= -\frac{L}{2\pi^2}\int_{-\pi}^{\pi} dk \int_{-\pi}^{\pi} d\tilde{k}\; h(\sin k - \sin\tilde{k})\cos^2 k \\ &= -4L\int_{-\infty}^{\infty} \frac{J_0(\omega)J_1(\omega)}{\omega\left(1+e^{\omega U/2}\right)} d\omega. \qquad \text{(IV.4.21)} \end{aligned}}$$

The last form is the classical one derived in [2] which expresses the groundstate energy in terms of Bessel functions $J_\alpha(\omega)$. Due to the repulsive nature of the interactions it is larger than the energy

$$E_0^{(\mathrm{ff})}(n) = -\frac{4L}{\pi}\sin\left(\frac{\pi}{2}n\right) \tag{IV.4.22}$$

corresponding to the limit $U \to 0$.

IV.4.2 Metal-insulator transition

We have already argued that the Hubbard model at $t = 0$ is an insulator, but has metallic properties for $U = 0$. Therefore an important question addressed by Lieb and Wu in their solution of the Hubbard model [2] was the possible existence of a metal-insulator transition. Since this transition in our case is controlled by the interaction strength one usually speaks of a *Mott transition* to distinguish it from other types of metal-insulator transitions.
It is clear (as we will see later explicitly) that the Hubbard model away from half-filling will be a metal. Therefore only the case of half-filling is relevant for a possible Mott transition. In the following we will show that already the results derived so far for the groundstate are sufficient to discuss this question.

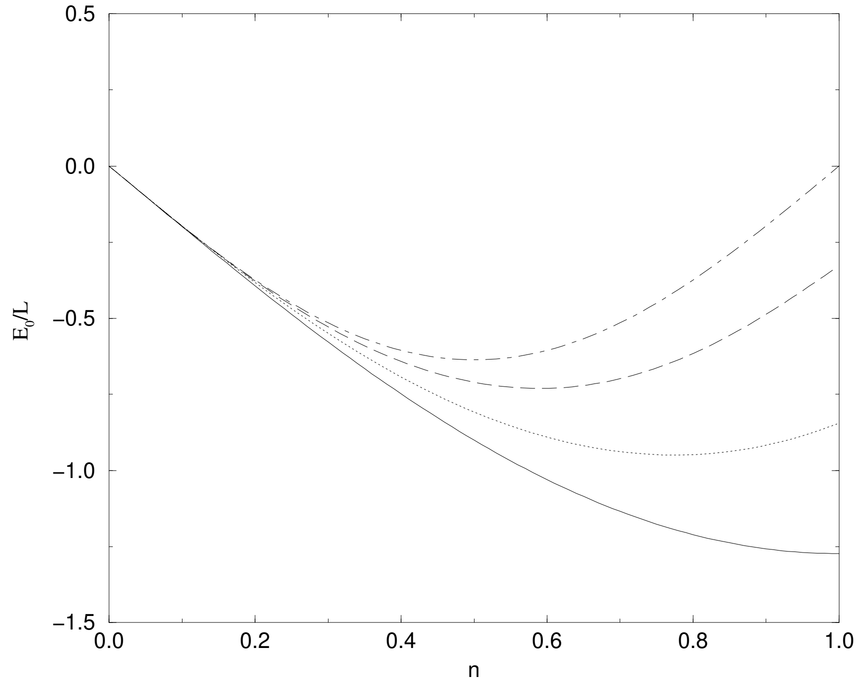

Figure IV.4.1: Groundstate energy per site E_0/L of the Hubbard model for $U = 8$ (dashed line) and $U = 2$ (dotted line) as a function of the particle density n. For comparison also the free fermion result for $U = 0$ (full line) and the strong coupling limit $U = \infty$ (dash-dotted line) are given.

A criterion for a metal-insulator transition is the behaviour of the chemical potential μ in the vicinity of $n = 1$. To be more specific, one needs to determine the chemical potentials $\mu_\pm$ for adding or removing one particle from the half-filled band. Then we can distinguish insulating from conducting behaviour:

$$\begin{aligned} \text{metal} \quad &: \qquad \mu_+ > \mu_- \,, \\ \text{insulator} \quad &: \qquad \mu_+ < \mu_- . \end{aligned} \tag{IV.4.23}$$

The origin of this criterion can be understood within the band picture. An insulator has a completely filled conduction band and an empty valence band that are separated by an energy gap $\Delta > 0$. So adding a particle at least requires an energy Δ whereas removing an electron costs almost no energy. This is different in a metal which has either a partially filled band or overlapping conduction and valence bands (see Fig. IV.4.2).

There are two different ways of calculating $\mu_\pm$. The first one is thermodynamical and uses the

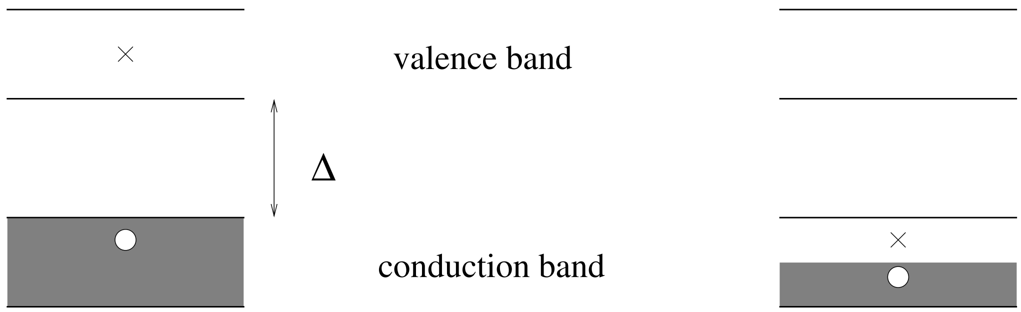

Figure IV.4.2: Explanation of the criterion (IV.4.23) for the distinction between metals and insulators within the band picture. The left part shows the situation in typical insulator where adding an electron ($\times$) costs at least the energy Δ whereas removing an electron ($\circ$) is much cheaper. This is different in a metal (right part).

relation with the groundstate energy density $e_0(n) = \frac{1}{L}E_0(N)$:

$$\mu = \frac{\partial e_0(n)}{\partial n}. \tag{IV.4.24}$$

The other approach is microscopic and based on the investigation of particle-hole excitations. Here we will use the first approach, although the second method will be discussed later in Sec. IV.5.2.
Since we have already derived the groundstate energy we can in principle calculate

$$\mu_\pm = \lim_{n\to 1\pm} \frac{\partial e_0(n)}{\partial n}. \tag{IV.4.25}$$

However, our calculation in Sec. IV.4.1 was restricted to the case $N \leq L$ or $n \leq 1$. In order to determine μ_+ we therefore have to make use of the symmetries derived in Sec. IV.2. From (IV.2.12) we find

$$e_0(n) = e_0(2-n) - (1-n)U \tag{IV.4.26}$$

which, after taking the derivative with respect to n and taking the limit $n \to 1-$, yields

$$\mu_+ + \mu_- = U. \tag{IV.4.27}$$

Therefore our criterion (IV.4.23) now reads that the system is a conductor if

$$\mu_- < \frac{U}{2}. \tag{IV.4.28}$$

Therefore we only have to calculate μ_-. Explicitly one finds [2] that, as a function of the Coulomb interaction U,

$$\mu_-(U) = 2 - 4\int_0^\infty \frac{J_1(\omega)}{\omega\left(1+e^{\omega U/2}\right)} d\omega. \tag{IV.4.29}$$

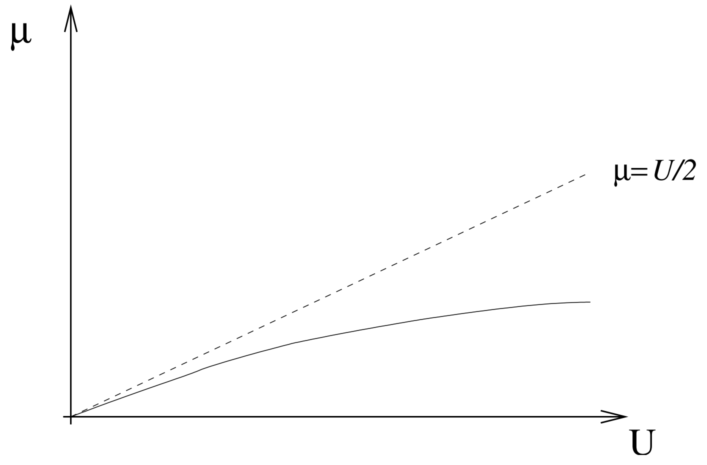

Figure IV.4.3: Illustration of the proof that $\mu_- < U/2$. The broken line is $\mu = U/2$. The chemical potential $\mu_-(U)$ satisfies $\mu_-(U=0) = 0$ and $\mu'_-(U=0) = 1/2$. Therefore it is clear that it is below the line $\mu = U/2$ at least for small values of U. Since $\mu''_-(U) < 0$ it must lie below this line everywhere.

This can be derived [3] by solving the integral equation (IV.4.12) near half-filling, i.e. for $Q = \pi - \epsilon$ by an expansion in ϵ. We now have

$$\mu_-(U=0) = 2 - 2\int_0^\infty \frac{J_1(\omega)}{\omega} d\omega = 0 \tag{IV.4.30}$$

and

$$\mu'_-(U=0) = \frac{1}{2}\int_0^\infty J_1(\omega) d\omega = \frac{1}{2} \tag{IV.4.31}$$

where the prime denotes the derivative with respect to the argument. Furthermore it can be shown that

$$\mu''_-(U=0) < 0. \tag{IV.4.32}$$

This is sufficient to see that the curve $\mu_-(U)$ lies always below $\mu = U/2$. From $\mu_-(U) = 0$ and $\mu'_-(U) = \frac{1}{2}$ we see that initially both curves have the same slope. However, the slope of $\mu_-(U)$ decreases since $\mu''_-(U=0) < 0$. Therefore we have the situation illustrated in Fig. IV.4.3 and thus

$$\mu_-(U) < \frac{U}{2} \qquad \text{for all } U > 0. \tag{IV.4.33}$$

Then the Hubbard model at half-filling is an insulator for any $U > 0$ and $U_c = 0$. No Mott transition occurs which is reflected in the title[5] of the Lieb-Wu paper [2].

Another important quantity that also appears in the calculation of asymptotics of correlation functions based on bosonization or conformal field theory is the so-called *charge susceptibility* χ_c defined by

$$\frac{1}{\chi_c} = \frac{\partial \mu}{\partial n} = \frac{\partial^2 e_0(n)}{\partial n^2}. \tag{IV.4.34}$$

[5] *Absence of Mott transition in an exact solution of the short-range, one-band model in one dimension*

It is also an indicator for metal-insulator transitions and for the Hubbard model it diverges for $n \to 1$.

IV.5 Excited states

In the following subsections we discuss the various excitations that are possible in the Hubbard model. Basically two types can be distinguished: those where the distribution of the λ_α-parameters differs qualitatively from the groundstate and those where the k_j-distribution is qualitatively different. The first type is called *spin excitations*, the latter *charge excitations*.

IV.5.1 Spin excitations

Since we have seen that the second set of Bethe Ansatz equations basically comes from an inhomogeneous Heisenberg model it is natural to assume that the structure of spin excitations is similar to those found there. Indeed we can consider again two different types of excitations related to qualitative changes of the λ_α-distribution. In the first type we flip a spin, i.e. we go to the case $N_\downarrow = \frac{N}{2} - 1$ where all λ_α are real. The second type is creating a complex pair $\lambda^\pm$. It turns out that the allowed two strings have the form $\lambda^\pm = \lambda_0 \pm i\frac{U}{4}$. Note that through the "backflow" also the distribution of the wavenumbers k_j is changed slightly in comparison to the groundstate. This is an effect of the coupling of the variables through the Bethe Ansatz equations.
The analysis of the Bethe Ansatz equations follows closely the treatment of the corresponding case for the antiferromagnetic Heisenberg model. Again it turns out that we have two-parametric spin excitation comprising degenerate singlet- and triplet states. The elementary excitation can be interpreted as a *spinon*.
The excitation energy of low-lying spin excitations is

$$\Delta E_s = E_s - E_0 = \sum_{\alpha=1}^{\nu} \epsilon_s(\theta_\alpha) \tag{IV.5.1}$$

where ν has to be even. For half-filling $n = 1$ the dispersion $\epsilon_s(p_s)$ of the spinons is given explicitly by

$$\boxed{\begin{aligned} \epsilon_s(\theta) &= \frac{2}{U}\int_{-\pi}^{\pi} \frac{\cos^2 k}{\cosh\left(\frac{2\pi}{U}(\sin k - \theta)\right)} dk, && \text{(IV.5.2)} \\ p_s(\theta) &= \frac{1}{\pi}\int_{-\pi}^{\pi} \arctan\left(e^{\frac{2\pi}{U}(\sin k - \theta)}\right) dk. && \text{(IV.5.3)} \end{aligned}}$$

For general filling $n \neq 1$ a closed form does not exist. However, one can derive expansion, e.g. in the Coulomb interaction U. Up to first order in $1/U$ one finds

$$\epsilon_s(p_s) = \frac{2\pi n}{U}\left(1 - \frac{\sin(2\pi n)}{2\pi n}\right)\sin\left(\frac{p_s}{n}\right). \tag{IV.5.4}$$

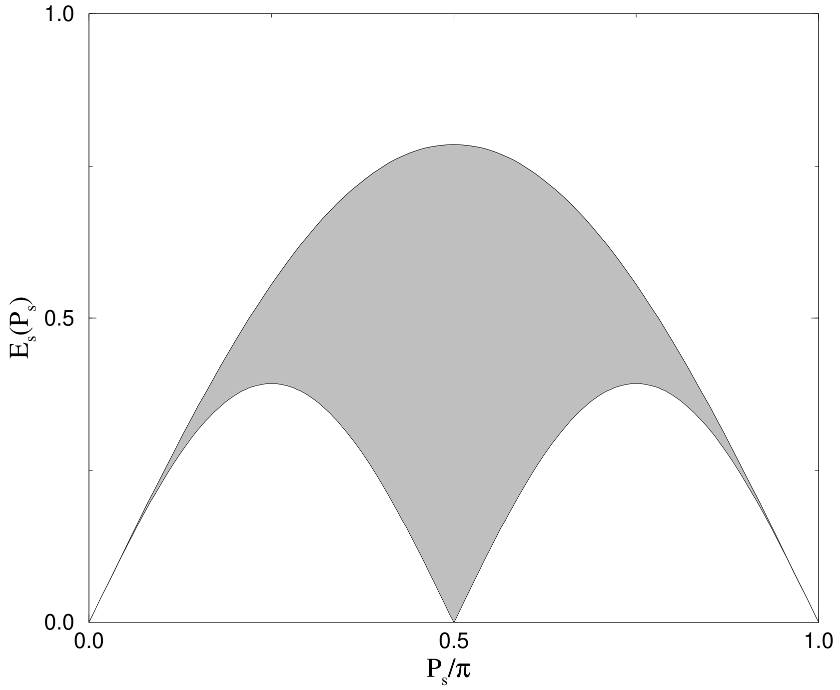

Figure IV.5.1: Typical 2-spinon continuum for the case $U = 8$ and quarter-filling $(n = 1/2)$. The excitation energy vanishes at $p_s = 0, 2k_F, 4k_F = 0, \pi n, 2\pi n$.

Note that for half-filling this strong coupling expansion reduces to the spinon dispersion (II.6.65) of the isotropic Heisenberg antiferromagnet with coupling $J = \frac{1}{U}$.[6] In fact this can be proven for Hubbard models in any dimension by a perturbative expansion.
Fig. IV.5.1 shows a typical 2-spinon continuum. Note that the spin excitations are always gapless for any density n and any coupling strength U. The excitation energy vanishes at k= $0, 2k_F, 4k_F$ with $k_F = \frac{\pi}{2}n$.

IV.5.2 Particle-hole excitations

Particle-hole excitations involve a qualitatitive change of the distribution of the wavenumbers k_j. They are very similar to particle-hole excitations in a system of free fermions. A wavenumber k^h that is occupied in the groundstate is removed thus creating a *hole*. To conserve the particle number a particle with momentum k^p has to be added. This can be done only in a region that is not occupied in the groundstate. Therefore we have $k^h \in [-Q, Q]$ and $k^p \in [-\pi, -Q[\cup]Q, \pi]$

[6]For a Heisenberg model written in terms of spin operators rather than Pauli matrices and with t not set to 1 the identification is $J = \frac{4t^2}{U}$.

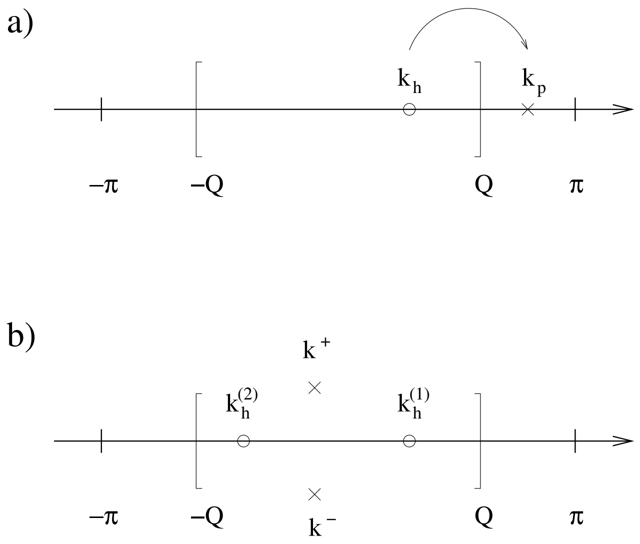

Figure IV.5.2: Distribution of k-wavenumbers for a) particle-hole excitations and b) excitations with double-occupations.

(see Fig. IV.5.2). Since $Q = \pi$ at half-filling, such excitations are not possible[7].
The excitation energy and momentum of a particle-hole excitation where the wavenumbers $k_1^h, \ldots, k_{\nu_{ph}}^h$ are removed from the groundstate distribution (holes) and the wavenumbers $k_1^p, \ldots, k_{\nu_{ph}}^p$ are added (particles) are given by

$$E_{ph} - E_0 = \sum_{\alpha=1}^{\nu_{ph}} \epsilon(k_\alpha^p) - \sum_{\alpha=1}^{\nu_{ph}} \epsilon(k_\alpha^h), \tag{IV.5.5}$$

$$P_{ph} - P_0 = \sum_{\alpha=1}^{\nu_{ph}} p(k_\alpha^p) - \sum_{\alpha=1}^{\nu_{ph}} p(k_\alpha^h). \tag{IV.5.6}$$

Here the function $\epsilon(k)$ is defined by the integral equation

$$\boxed{2\pi\epsilon(k) - \int_{-Q}^{Q} h(\sin k - \sin \tilde{k}) \cos \tilde{k} \epsilon(\tilde{k}) d\tilde{k} = -2\pi(2\cos k + \mu) \qquad \text{(IV.5.7)}}$$

where the chemical potential μ is fixed by the additional condition

$$\epsilon(Q) = \epsilon(-Q) = 0. \tag{IV.5.8}$$

The momentum can be expressed by the groundstate density (IV.4.12) of the wavenumbers:

$$p(k) = 2\pi \int_0^k \rho(\tilde{k}) d\tilde{k}. \tag{IV.5.9}$$

[7] Although hole excitations that change the particle number are possible.

The function $\epsilon(k)$ is called *dressed energy*. The reason is that the groundstate energy can be obtained by occupying all states with negative dressed energy:

$$E_0 = \frac{L}{2\pi}\int_{-Q}^{Q}\epsilon(k)dk. \tag{IV.5.10}$$

This description is complementary to that given in (IV.4.20). There the groundstate energy was by occupying the *bare energy* band $-2\cos k$ and the information of the interaction was encoded in the density $\rho(k)$ of the wavenumbers which was constant ($\rho_{\rm ff}(k) = \frac{1}{2\pi}$) for noninteracting fermions. In contrast in (IV.5.10) the density of the k_j is $\frac{1}{2\pi}$, but now the interaction effects have been included through a renormalization of the energy band from $-2\cos k$ to $\epsilon(k)$.
Fig. IV.5.3 shows the typical form of the energy continuum of particle-hole excitations with $\nu_{ph} = 1$. Note that these excitations do not possess a gap. The excitation energy vanishes at $4k_F = 2\pi n$. For large U we an can again give an expansion of the energy and momentum:

$$\epsilon(k) = -2\cos k - \frac{4n\ln 2}{U}\left(1+\frac{\sin(2\pi n)}{2\pi n}\right) - \mu, \tag{IV.5.11}$$

$$p(k) = k + \frac{4n\ln 2}{U}\sin k. \tag{IV.5.12}$$

The expression for the energy clearly shows that the dressed energy is closely related to the bare energy $-2\cos k$.

IV.5.3 Excitations with double occupations

There is a second way of qualitatively changing the distribution of the k_j, namely by allowing complex wavenumbers. It turns out that in the simplest case 2-strings $k^\pm = \kappa \pm i\chi$ can appear such that

$$\sin k^\pm = \lambda \pm i\frac{U}{4}. \tag{IV.5.13}$$

Here the λ is actually one occuring in the distribution of the λ_α.
The excitations described by complex wavenumbers become relevant at half-filling. As already discussed, for $n = 1$ particle-hole excitations are no longer possible. Therefore excitations with complex wavenumbers are the only (particle-number conserving) electronic excitations that are possible. In fact, it has been argued in [7–9] that these states have real double occupations instead of just virtual ones as the other states discussed so far.
For each complex pair $k^\pm$ we have to create to holes k_1^h, k_2^h in the groundstate distribution of the wavenumbers (see Fig. IV.5.2b). As for the Heisenberg model the energies can be solely expressed through the holes k_1^h, k_2^h. Energy and momentum of an excitation with D_c 2-strings $k_j^\pm$ are given by

$$E_D - E_0 = \sum_{\alpha=1}^{2D_c}\epsilon_c(k_\alpha^h), \tag{IV.5.14}$$

$$P_D - P_0 = \sum_{\alpha=1}^{2D_c}p_c(k_\alpha^h), \tag{IV.5.15}$$

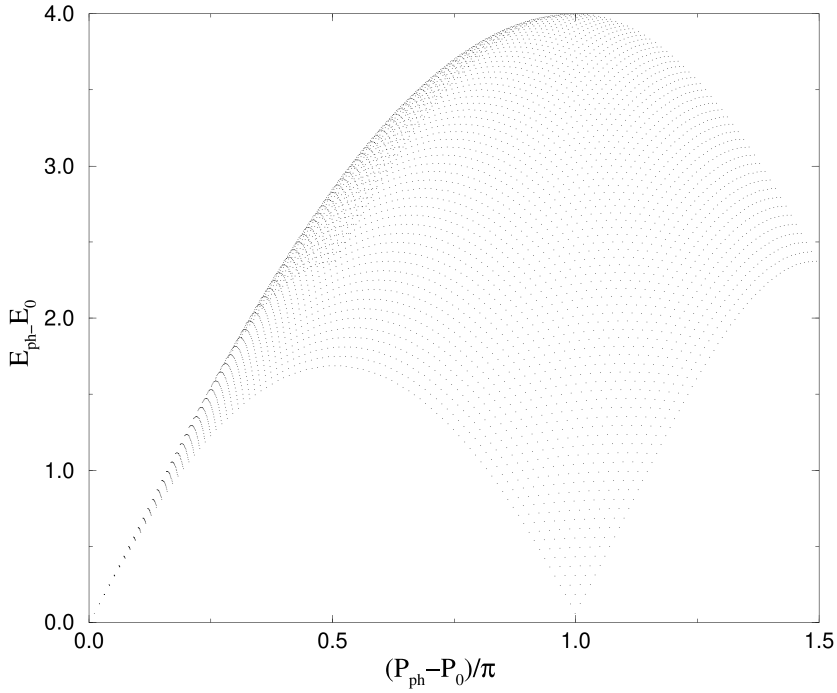

Figure IV.5.3: Energy continuum of particle-hole excitations at quarter-filling ($n = 1/2$) and $U = 8$. The four branches (starting from the lowest branch at small momenta and proceeding counterclockwise) that form the boundary of the continuum are obtained by fixing 1) $k^p = Q$, 2) $k^h = -Q$, and 3) $k^p = \pi$. Only the upper limit boundary for momenta $0 \leq P_{ph} - P_0 < \pi$ can not be obtained in such a simple way due to the curvature of the function $\epsilon(k)$ and has to be determined numerically. The excitation energy vanishes at $p = 4k_F = 2\pi n$. The maximal momentum is given by $\pi + 2k_F$.

where the functions $\epsilon_c(k)$ and $p_c(k)$ are related to those for particle-hole excitations (see (IV.5.7) and (IV.5.9)):

$$\epsilon_c(k) = \frac{U}{2} + \epsilon(k), \tag{IV.5.16}$$

$$p_c(k) = -p(k). \tag{IV.5.17}$$

Since the ϵ_c-functions occur in pairs, each pair has a leading contribution of U to the energy. This is a strong indication that these excited states involve doubly-occupied sites.

At half-filling, the energies can be calculated explicitly. Fig. IV.5.4 shows a typically continuum of these excitations. It is important to notice that for any value $U > 0$ these excitations have a finite energy gap Δ. It is explicitly given by

$$\Delta = 2\epsilon_c(\pi) = U - 4 + \frac{2}{\pi}\int_{-\pi}^{\pi} h(\sin k)\cos^2 k dk. \tag{IV.5.18}$$

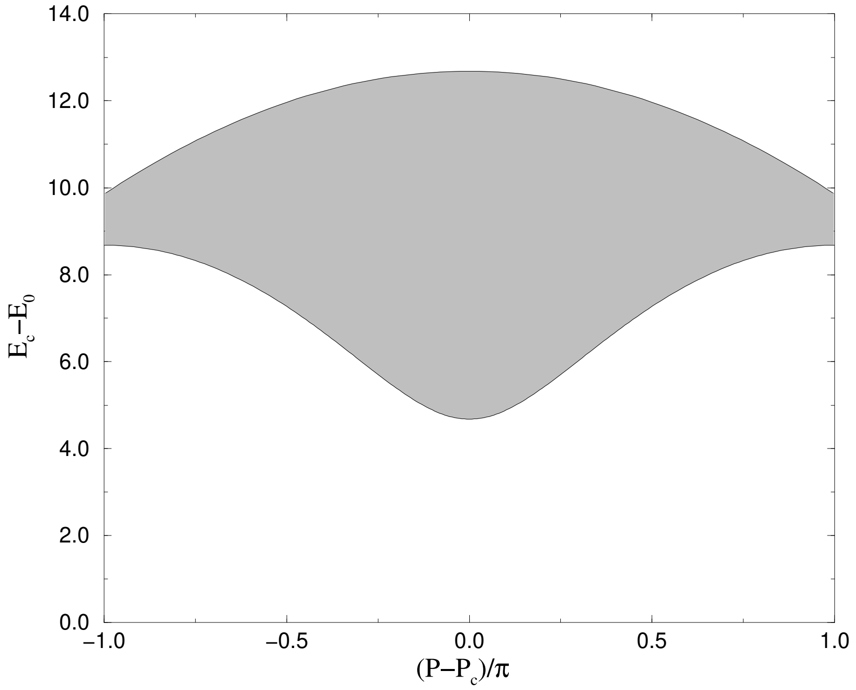

Figure IV.5.4: Energy continuum of excitations with double occupations at half-filling ($n = 1$) and $U = 8$.

For large and small U the gap behaves as

$$U \to \infty: \qquad \Delta \sim U - 4, \tag{IV.5.19}$$

$$U \to 0: \qquad \Delta \sim \frac{8}{\pi}\sqrt{U} e^{-2\pi/U}. \tag{IV.5.20}$$

For large U the energy consists of a contribution U from the double occupancy and an energy gain 4 of kinetic energy. For small U the gap shows a strongly nonanalytic behaviour.
Finally it should be noted that complex k do not necessarily correspond to bound states of elementary excitations in the sense we have encountered this for magnons in the ferromagnetic Heisenberg model in Sec. II.6.1. Here we see that the spectrum of the excitation can be interpreted as scattering states leading to a continuum of states (see Fig. IV.5.4). However, the states are bound states of electrons in the sense that the wavefunction decays exponentially.

IV.5.4 Classification of low-lying excitations

In the previous subsections we have derived the different elementary excitations of the (repulsive) Hubbard model. The low-lying spectrum is then given by an indenpendent superposition of theses excitations:

$$\boxed{\begin{aligned} E &= E_0 + \sum_{\alpha=1}^{\nu} \epsilon_s(\theta_\alpha) + \sum_{\alpha=1}^{\nu_p} \epsilon(k_\alpha^p) - \sum_{\alpha=1}^{\nu_h} \epsilon(k_\alpha^h) + \sum_{\alpha=1}^{2D_c} \epsilon_c(\bar{k}_\alpha^h), && \text{(IV.5.21)} \\ P_{ph} &= P_0 + \sum_{\alpha=1}^{\nu} p_s(\theta_\alpha) + \sum_{\alpha=1}^{\nu_p} p(k_\alpha^h) - \sum_{\alpha=1}^{\nu_h} p(k_\alpha^p) + \sum_{\alpha=1}^{2D_c} p_c(\bar{k}_\alpha^h). && \text{(IV.5.22)} \end{aligned}}$$

Here the number of ν of spinons has to be even. In addition, we have allowed for different number ν_p and ν_h of particle and hole excitations, i.e. for a change of the total particle number by $\nu_p - \nu_h$. Finally we have introduced D_c exciations with double occupations that are parametrized by the hole positions $\bar{k}_\alpha^h$. The above expression is valid as long as the number of each excitation type is of the order 1.

The fact that the different excitations can form an independent superposition points to a certain independence of spin and charge excitations. This can be made more precise with the help of the SO(4) symmetry that we will use in the following to classify the excitations.

We have already seen that the spin excitations are rather similar to that of the isotropic Heisenberg antiferromagnet. Therefore we might interpret the elementary spin excitation as a *spinon* with disperson $\epsilon_s(p_s)$. It can be shown [10] that it corresponds to a spin-$\frac{1}{2}$ representation of the spin-SU(2) algebra SU(2)$_s$, i.e. we have a spin-$\uparrow$ and a spin-$\downarrow$ spinon. In addition, it forms a singlet with respect to the pseudospin-SU(2) algebra SU(2)$_\eta$. Therefore one can say that spinons have spin $\frac{1}{2}$, but do not carry a charge. The latter is not so surprising if one considers the close relation to the spinons in the Heisenberg antiferromagnet where we do not have any charge particles but only localized spins.

The elementary charge excitations are called *holons* and *antiholon*. They form singlets with respect to SU(2)$_s$ and a spin-$\frac{1}{2}$ representation of SU(2)$_\eta$. Thus we are a led to the interpretation of holons/antiholons as particles with charge[8] $\pm e$, but no spin. The dispersion of the holon is given by $\epsilon_h(k) = \epsilon_c(k)$ and $p_h(k) = p_c(k)$ whereas the antiholon has dispersion $\epsilon_{ah}(k) = \epsilon_c(k)$ and $p_{ah}(k) = p_c(k) - \pi$. Here $\epsilon_c(k)$ and $p_c(k)$ are given by (IV.5.16) and (IV.5.17).

The low-lying spectrum of the Hubbard model can now be viewed as scattering states of these elementary excitations. Due to the $\mathbb{Z}_2$-symmetry they can only occur pairwise. Note the unusual character of the excitations. In contrast to the real particles (electrons) they either have spin or charge, but not both. This phenomenon is quite common in (one-dimensional) strongly correlated electron system and called *spin-charge separation*. In the context of the exact solution of the Hubbard model it can be given a precise meaning, as we have shown above. In the framework of bosonisation general conclusions from spin-charge separation can be derived.

IV.6 Attractive Hubbard model

So far we have treated only repulsive interactions explicitly. The properties of the attractive model with $U < 0$ can in principle be derived using the symmetries discussed in Sec. IV.2. In

[8]It is clear that the holon, corresponding to removing an electron, has charge $-e$, where e is the charge of an electron.

this section we want to give a short overview of the properties of Hubbard models with attractive interactions. In Exercise 42 it is shown how the Bethe Ansatz solutions are related. If we have a solution $\{k_j, \lambda_\alpha\}$ for interaction U, then $\{\pi + k_j, -\lambda_\alpha\}$ is a solution for $-U$. These are called *complementary solutions*.

First of all it might seem unusual to consider attractive Coulomb interactions. However, these can occur as *effective* interactions, when other interactions are present, e.g. electron-phonon coupling or bond-charge interactions.

If we put $U < 0$ the excitations with double occupations have a negative "excitation energy". This points at an instability towards the formation of pairs which we also expect from a general discussion of the limit $-U \gg 1$. Indeed it turns out the groundstate of an attractive Hubbard model has all wavenumbers paired to 2-strings $k_j^{\pm}$ of the form (IV.5.13). For the groundstate energy we find, as expected from the symmetries,

$$E_0(U < 0) = E_0(|U|) - \frac{|U|}{2}L. \tag{IV.6.1}$$

The excitation spectrum essential can be obtained from that of the repulsive model by exchanging spin and charge degrees of freedom. The reason is that U and $-U$ models are related by a partial particle-hole excitations that basically exchanges the roles of the spin and pseudospin. Spin excitations in the attractive model always have a finite excitation gap. The reason behind this is the structure of the groundstate. In order to make a spin excitation one first has to break up a 2-string $k^{\pm}$ which costs a finite energy of the order $|U|$. In contrast, the charge excitations are always gapless.

Chapter V

Conformal invariance and correlation functions

Although the Bethe Ansatz provides an explicit form of the exact eigenfunctions this form is so complicated that it is virtually impossible to calculate correlation function[1]. Even the calculation of the norm, which is the simplest possible correlation function, is not possible in general. In order to get information about the behaviour of correlations one has to combine the Bethe Ansatz with other approaches, e.g. bosonisation. A very successful approach is the so-called *conformal invariance* that allows to calculate the general form of correlation functions from simple symmetries. The Bethe Ansatz is then needed to calculate e.g. the exact values of exponents. This is the main topic of this Chapter.
Historically, conformal invariance has been applied in two different fields [12]. One is statistical physics or, more precisely, the theory of critical phenomena. This is the approach that we will follow here. The other main line of application is *conformal field theory* which has been used a lot in string theory. The reason is that the trajectories of spins in space-time form surfaces rather than lines as in the case of point-particles. Therefore one is interested in the behaviour of fields on these surfaces. It turns out that here one can make use of the powerful machinery of complex variable theory.

V.1 Conformal invariance in statistical physics

From the theory of phase transition we know that at a *critical point* a system exhibits *scale invariance*. The reason is an diverging correlation length which means that there is no longer an intrinsic length scale in the system, as e.g. a typical cluster size away from the critical point. It looks the same on each length scale and thus can be called self-similar.
On a mathematical level, scale invariance means that correlation functions under rescaling $\mathbf{r} \to \frac{\mathbf{r}}{b}$ satisfy

$$\langle \phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2)\cdots\rangle = b^{-\alpha_1}b^{-\alpha_2}\cdots\left\langle \phi_1\left(\frac{\mathbf{r}_1}{b}\right)\phi_2\left(\frac{\mathbf{r}_1}{b}\right)\cdots\right\rangle . \qquad \text{(V.1.1)}$$

[1] Although some progess has been made in special cases [11].

Here the α_j are called the *scaling dimension* of the operator ϕ_j. In the renormalization group scaling operators that are conjugate to scaling fields play an important role. For those one even has

$$\phi(\mathbf{r}) = b^{-\alpha}\phi\left(\frac{\mathbf{r}}{b}\right) \tag{V.1.2}$$

on the operator level. These scaling relations usually can not be proven in a rigorous sense but are expected to hold quite generally for systems at a critical point which have only short-ranged interactions[2].

The scaling relation (V.1.1) implies that correlations at the critical point are described by homogeneous functions. This already fixes the form of any 2-point function that only depends on $|\mathbf{r}_1 - \mathbf{r}_2|$ (see Exercise 43):

$$\langle\phi(\mathbf{r}_1)\phi(\mathbf{r}_2)\rangle \propto \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|^{\alpha}}. \tag{V.1.3}$$

These correlations decay algebraically where the decay is characterized by the scaling dimension. In [13] it has been emphasized that in most cases at a critical point an even higher symmetry is realized, the so-called *conformal invariance* [14]. In fact, a critical system should not only be invariant under scaling transformations, but under all transformations that preserve the angle of intersection of arbitrary curves. As in the case of scale invariance this can not be proven rigorously, but it is expected to hold for any critical system that can be described by a Hamiltonian that is invariant under translations and rotations and has sufficiently short-ranged interactions.

Conformal transformation conserve angles of interception and have therefore no skewing component (see Fig. V.1.1). Formally an infinitesimal conformal transformation can be written in the form $\mathcal{T}(\mathbf{r}+\epsilon) = \mathcal{T}(\mathbf{r}) + M(\mathbf{r})\cdot\epsilon$ where $M(\mathbf{r})$ is a space-dependent matrix without skewing component, i.e. a combination of a rotation and dilatation. It is important to note that these are *local* transformations instead of the global rescaling with an $\mathbf{r}$-independent factor b in (V.1.1).

In d dimension a general conformal transformation is composed of the following elementary transformations:

- *translations*: $\mathbf{r} \to \mathbf{r} + \mathbf{a}$
  Translations are obviously described by d parameters, the components of $\mathbf{a}$.
- *rotations*: $\mathbf{r} \to D\mathbf{r}$
  The rotation matrix $D \in \mathrm{SO}(d)$ has $\frac{1}{2}d(d-1)$ parameters.
- *dilatations*: $\mathbf{r} \to \lambda\mathbf{r}$
  These transformations are responsible for the scale invariance and are characterized by one parameter λ.
- *inversions*: $\mathbf{r} \to \frac{\mathbf{r}}{r^2}$
  This transformation describes the inversion at the surface of a sphere and has d parameters. This can be seen more clearly if one considers the more general *special conformal transformation* $\mathbf{r} \to \frac{\mathbf{r}+r^2\mathbf{a}}{1+2\mathbf{a}\cdot\mathbf{r}+a^2r^2}$ which is a composition of an inversion and a translation in the sense that $\frac{\mathbf{r}'}{r'^2} = \frac{\mathbf{r}}{r^2} + \mathbf{a}$.

[2] This must not necessary mean a finite interaction range.

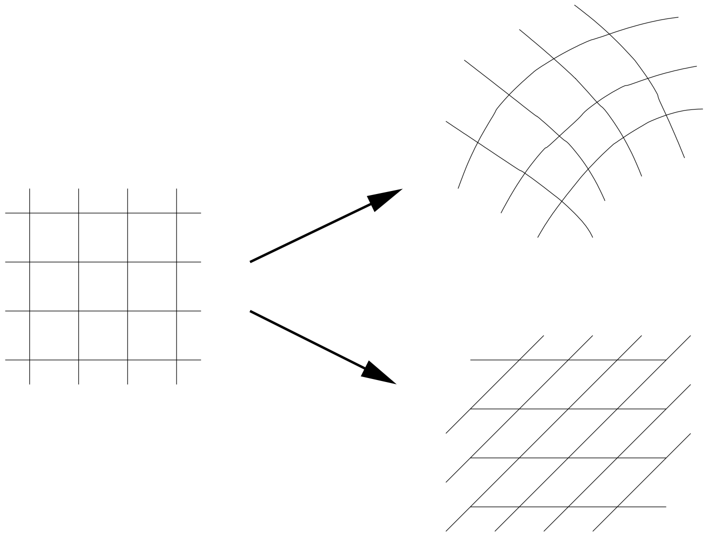

Figure V.1.1: Examples of a conformal (top) and nonconformal (bottom) transformations. The nonconformal transformation has a skewing component that local changes angles.

Thus we see that the group of conformal transformations is in general finite-dimensional. It is described by $\frac{1}{2}(d+1)(d+2)$ parameters, i.e. 10 parameters in $d = 3$. Therefore it does not give very strong further restrictions on the general form of correlation functions.
This is different in two dimensions ($d = 2$). Here we can identify $\mathbb{R}^2$ with the complex plane $\mathbb{C}$ by considering complex coordinates $z = x + iy$ instead of $\mathbf{r} = (x, y)$. Furthermore from complex variable theory we know the each holomorphic function $f(z)$ with $\frac{df(z)}{dz} \neq 0$ generates a conformal transformation of the plane. Therefore in two dimensions the conformal group is infinite-dimensional. This imposes much stronger conditions on the correlation functions as in the case of finite-dimensional groups.
In order to discuss these restrictions, we first generalize the transformation behaviour (V.1.1) to local transformations $\mathbf{r} \to \mathbf{r}'$:

$$\langle \phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2)\rangle = (b(\mathbf{r}_1))^{-\alpha_1} (b(\mathbf{r}_2))^{-\alpha_2} \langle \phi_1 (\mathbf{r}_1{}') \, \phi_2 (\mathbf{r}_1{}') \cdots \rangle \, . \tag{V.1.4}$$

where

$$(b(\mathbf{r}_1))^d = \det \left(\frac{\partial \mathbf{r}}{\partial \mathbf{r}'} \right) . \tag{V.1.5}$$

In $d = 2$ we can regard $z = x + iy$ and $\bar{z} = x - iy$ as independent variables instead of x and y. This will allows us to make full use of powerful methods from complex analysis. Similar to (V.1.2) we now have scaling operators ϕ that satisfy

$$\phi(z, \bar{z}) = \lambda^{\Delta_\phi} \phi(\lambda z, \bar{z}) \lambda^{\bar{\Delta}_\phi} \phi(z, \lambda \bar{z}) \tag{V.1.6}$$

where $\Delta_\phi, \bar{\Delta}_\phi > 0$ are called *conformal dimensions*. These are related to the so-called *scaling dimension* x_ϕ and the *(conformal) spin* s_ϕ by

$$x_\phi = \Delta_\phi + \bar{\Delta}_\phi, \qquad s_\phi = \Delta_\phi - \bar{\Delta}_\phi. \tag{V.1.7}$$

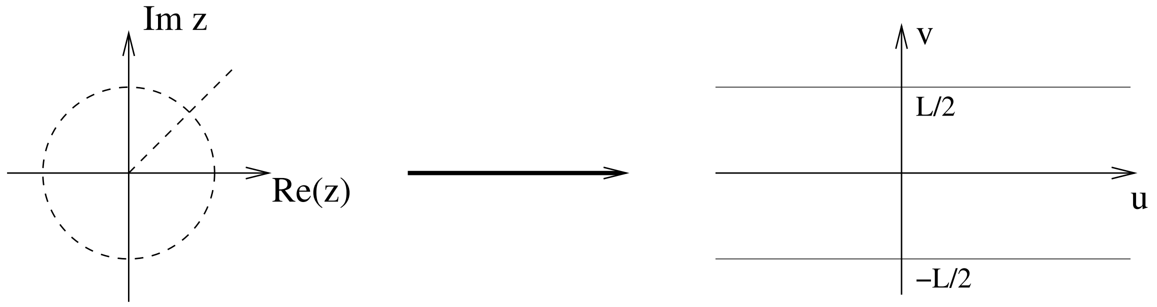

Figure V.2.1: Conformal mapping of the plane to a cylinder.

The reason for calling s_ϕ *spin* lies in the fact that it can be shown to be either integer or half-odd integer.

V.2 Application: Critical behaviour in restricted geometries

In this Section we want to give a brief example for one important type of applications of conformal invariance. It can be used to determine the "critical" behaviour of systems in special geometries once the critical behaviour e.g. in the plane is known. This is possible if the two geometries involved are related by a conformal mapping $w(z)$.
As an example we consider a classical system on a cylinder. The system is infinite in one direction and finite in the other, with periodic boundary conditions for the latter (see Fig. V.2.1). Using the holomorphic function

$$w(z) = \frac{L}{2\pi} \ln z \tag{V.2.1}$$

the complex plane is mapped onto the cylinder $\mathbb{R} \times \left[-\frac{L}{2}, \frac{L}{2}\right]$. For convenience we use the convention that $z = |z|e^{i\varphi}$ with $\varphi \in]-\pi, \pi]$ so that the cylinder lies symmetric with respect to the u-axis, where $w = u + iv$ in the cylindrical coordinates (see Fig. V.2.1).
If we have a model that is critical in the plane, we already know that its 2-point correlations satisfy

$$\langle \phi(z_1)\phi(z_2)\rangle = \frac{1}{|z_1 - z_2|^{2x_\phi}}. \tag{V.2.2}$$

Here we have assumed a convenient normalization of the operator ϕ. The question know is how the correlation functions in the cylinder geometry behave. This can be elucidated from (V.1.4). Assuming that the conformal spin s_ϕ is zero, we have

$$\langle \phi(z_1)\phi(z_2)\rangle = |z'(w_1)|^{-x_\phi} |z'(w_2)|^{-x_\phi} \langle \phi(w_1)\phi(w_2)\rangle \tag{V.2.3}$$

which relates the correlation functions in the plane (variable z) and the cylinder (variable w). A straightforward calculation using (V.2.1) then yields

$$\langle \phi(w_1)\phi(w_2)\rangle = \frac{\left(\frac{2\pi}{L}\right)^{2x_\phi}}{\left[2\cosh\left(\frac{2\pi}{L}(u_1 - u_2)\right) - 2\cos\left(\frac{2\pi}{L}(v_1 - v_2)\right)\right]^{x_\phi}}. \tag{V.2.4}$$

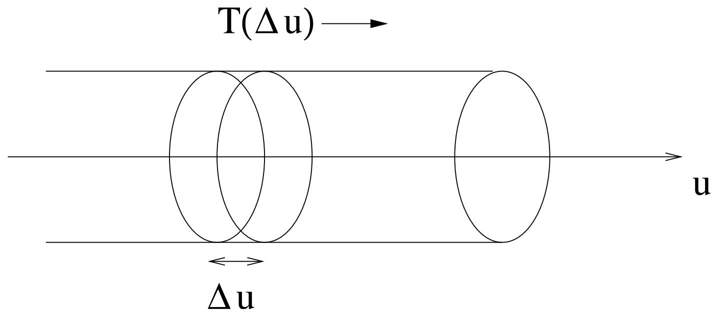

Figure V.3.1: Definition of the transfer matrix on a cylinder.

One sees that the leading term in u-direction decays as $e^{-\frac{2\pi}{L}|u_1-u_2|x_\phi}$. Comparing with the general form $e^{-|u_1-u_2|/\xi}$ we can read off the correlation length

$$\xi = \frac{L}{2\pi x_\phi}. \tag{V.2.5}$$

This shows that the system in the restricted geometry has no longer algebraically decaying correlation functions and is thus not critical. It has a finite correlation length of the order L, i.e. even in the infinitely extended direction u the system 'feels' the finiteness of the geometry orthogonal to it.
For later use we finally give the full expansion of the correlation function (V.2.4):

$$\langle\phi(w_1)\phi(w_2)\rangle = \left(\frac{2\pi}{L}\right)^{2x_\phi} \sum_{m\bar{m}=1}^{\infty} a_m a_{\bar{m}} e^{-\frac{2\pi}{L}(x_\phi+m+\bar{m})(u_1-u_2)} e^{\frac{2\pi i}{L}(m-\bar{m})(v_1-v_2)} \tag{V.2.6}$$

where the coefficients are given by $a_m = \frac{\Gamma(x+m)}{\Gamma(x)m!}$.

V.3 Finite-size corrections and correlation functions

We now want to apply the framework developed so far to one-dimensional quantum systems. As we have seen previously in Sec. II.4. these can mapped to two-dimensional classical systems so that such an application is rather natural. In our cylindrical geometry we can define a transfer matrix T as shown in Fig. V.3.1 by considering a transfer in the infinitely extended direction. This transfer matrix in turn defines a one-dimensional quantum Hamiltonian $\mathcal{H}$ through the relation

$$T(\Delta u) = e^{-\Delta u \mathcal{H}}. \tag{V.3.1}$$

We now choose an orthonormal basis of $\mathcal{H}$ with

$$\begin{aligned} \mathcal{H}|n,k\rangle &= E_n|n,k\rangle, && \text{(V.3.2)} \\ P|n,k\rangle &= k_n|n,k\rangle, && \text{(V.3.3)} \end{aligned}$$

where we have introduced the momentum operator in u-direction which is given by

$$P = -i\frac{\partial}{\partial u}. \tag{V.3.4}$$

Then we give a spectral decomposition of the 2-point function

$$\langle\phi(w_1)\phi(w_2)\rangle = \sum_{|n,k\rangle}\langle 0|\phi(w_1)|n,k\rangle\langle n,k|\phi(w_2)|0\rangle. \tag{V.3.5}$$

In the Hilbert space of our quantum Hamiltonian $\mathcal{H}$ the scaling operators ϕ become proper operators $\hat{\phi}$. We then have a Heisenberg representation (in imaginary time) of the form [15]

$$\phi(u,v) \longrightarrow \hat{\phi}(u,v) = e^{-\mathcal{H}u}\hat{\phi}(v)e^{\mathcal{H}u}. \tag{V.3.6}$$

Using this representation we can rewrite the spectral decomposition:

$$\begin{aligned}\langle\phi(w_1)\phi(w_2)\rangle &= \sum_{|n,k\rangle}\langle 0|\hat{\phi}(v_1)|n,k\rangle\langle n,k|\phi(\hat{v}_2)|0\rangle e^{-(E_n-E_0)(u_1-u_2)} \\ &= \sum_{|n,k\rangle}\left|\langle 0|\hat{\phi}|n,k\rangle\right|^2 e^{-(E_n-E_0)(u_1-u_2)-ik_n(v_1-v_2)}. \end{aligned} \tag{V.3.7}$$

In the first step we have made use of (V.3.6) and in the second step we have translated the arguments v_1, v_2 using the momentum operator (V.3.4).
If we compare spectral decomposition (V.3.7) with our previous result (V.2.6) we find

$$E_n - E_0 = \frac{2\pi}{L}(x_\phi + m + \bar{m}), \tag{V.3.8}$$

$$k_n = \frac{2\pi}{L}(m - \bar{m}). \tag{V.3.9}$$

This shows that the conformal dimension (a quantity that can be determined in the infinite plane) controls the smallest excitation energy in the cylinder geometry. Above this energy $\frac{2\pi}{L}x_\phi$ we have an equidistant spectrum. This is usually called a *conformal tower*. We have to emphasize that implicitly we have so far assumed that the velocity v_F of elementary excitations of $\mathcal{H}$ is normalized to 1.
The most interesting aspect of the result (V.3.8), (V.3.9) is that the behaviour of correlations functions is related in a very simple way to the behaviour of the excitations. The only problem left is that usually E_0 depends on the length L of the system. However, it is possible to show [16, 17] that

$$E_0 = Le_0 - \frac{\pi c}{6L} \tag{V.3.10}$$

where e_0 is the density of the groundstate energy. The derivation of (V.3.10) makes it necessary to study the behaviour of the theory under infinitesimal non-conformal transformations. This leads to the introduction of the so-called *stress tensor*[3] $\mathcal{T}$ describes the reaction of the system to

[3]Sometimes also called *energy-momentum tensor*.

non-conformal transformations. It can then be shown that the 2-point correlations of the stress tensor are rather simple:

$$\langle \mathcal{T}(z)\mathcal{T}(z')\rangle = \frac{c/2}{(z-z')^4}. \tag{V.3.11}$$

The constant c appearing in (V.3.11) and (V.3.10) is called *conformal anomaly* or *central charge*. It is a universal property of the system and in some sense specifies its universality class. However, from its definition it is a measure of the reaction to non-conformal transformations. It is usually a number of order 1.

Summarizing our findings and generalizing slightly by allowing also for unnormalized Hamiltonians with $v_F \neq 1$ we have the following important result:

$$\boxed{\begin{aligned} E_0 &= Le_0 - \frac{\pi v_F}{6L}c, && \text{(V.3.12)}\\ E_n - E_0 &= \frac{2\pi v_F}{L}(\Delta_\phi + \bar{\Delta}_\phi + m + \bar{m}), && \text{(V.3.13)}\\ k_n &= \frac{2\pi}{L}(\Delta_\phi - \bar{\Delta}_\phi + m - \bar{m}) + 2k_F D. && \text{(V.3.14)} \end{aligned}}$$

Here we have also included excitations where D particles are transfered from one Fermi point k_F to the other at $-k_F$ and allowed for a non-zero (conformal) spin $s_\phi = \Delta_\phi - \bar{\Delta}_\phi$.

The importance of the above results comes from the fact that the asymptotic behaviour of correlation functions is related to the *finite-size corrections* of order $1/L$ to the energy spectrum. Usually the exact or numerical calculation is very difficult. We have already mentioned that within the framework of the Bethe Ansatz it is practically impossible to calculate even the simplest correlations. On the other hand, it is usually much easier to determine the finite-size corrections. Once these are known, either from analytical or numerical investigations, by comparison with the above predictions of conformal invariance one can read of the scaling dimensions x_ϕ and thus determine the correlation functions e.g. via (V.2.2). We stress that this is in no way restricted to exactly solvable models. Above program should work for any Hamiltonian that satisfies the assumptions underlying conformal invariance. This especially means that it is critical, i.e. has gapless excitations that are characterized by a Fermi velocity v_F.

Finally, a few remarks.

- The conformal anomaly c can in principle by measured since it is related to the low-temperature behaviour of the specific heat C via

$$C = \frac{\pi k_B^2 T}{3\hbar v_F} \cdot c. \tag{V.3.15}$$

- As already mentioned, c is usually a number of the order 1. For the models we have discussed here it turns out that $c = 1$ for each critical excitation. This usually implies that the critical exponents depend on the coupling constants and are not completely fixed by c. The latter happens for the so-called *unitary minimal theories* when c has the form

$$c = 1 - \frac{6}{p(p+1)} \qquad (p = 2, 3, \ldots). \tag{V.3.16}$$

In this case the critical exponents are rational numbers determined by c. Note that the Ising model corresponds to $p = 3$ and thus has $c = 1/2$.

- The calculation of the finite-size corrections is possible within the framework of the Bethe Ansatz. It is, however, usually quite involved. One has to keep in mind that finite-size corrections in principle have two different origins. One is the size-dependence of the Bethe Ansatz wavenumbers through the Bethe Ansatz equations. Secondly, an error is introduced if the Bethe equations are reformulated as integral equations. The latter effect can be taken into account by using the so-called *Euler-Maclaurin formula*

$$\sum_{j=1}^{n} f(j) \approx \int_0^n f(x)dx - \frac{1}{2}[f(0) + f(n)] + \frac{1}{12}[f'(n) - f'(0)] \tag{V.3.17}$$

  when deriving the integral equations.

- Usually it is not possible to derive an explicit relation between physical operators (e.g. $\mathbf{S}_j$ or $c^\dagger_{j\sigma}$) and the scaling operators $\hat{\phi}$. In order to know which scaling operators $\hat{\phi}_Q(z, \bar{z})$ contribute to a correlation function of a given physical operator $\psi(x, t)$ one would need a relation of the form [11]

$$\psi(x, t) = \sum_Q A(Q)\hat{\phi}_Q(z, \bar{z}) \tag{V.3.18}$$

  where Q stands for the set of quantum numbers related to the excitation under consideration. Usually the coefficients $A(Q)$ can not be determine explicitly. However, that is not necessary for many applications, e.g. the determination of the asymptotic behaviour of correlation functions. Here it is sufficient to know which operators contribute in (V.3.18) which can be decided from the relevant quantum numbers through selection rules.

- So far we have discussed only models which have only one type of excitation. For models with several types of critical excitations, e.g. the Hubbard model with $n < 1$, our results have to be generalized. This will be done in the following subsection. However, if only one excitation is critical and the others have a finite gap, e.g. the Hubbard model at half-filling, the results derived so far can still be applied.

V.3.1 Examples

In this subsection we briefly list some explicit result and also discuss the generalization to models with several types of critical excitations.

For the anisotropic Heisenberg model in the critical region $|\Delta| \leq 1$ the finite-size corrections derived from the Bethe Ansatz equations are given by

$$\begin{aligned} E_0 &= Le_0 - \frac{\pi v_F}{6L}\,, && \text{(V.3.19)} \\ E_n - E_0 &= \frac{2\pi v_F}{L}\left(\frac{(\Delta M)^2}{2\theta} + \frac{1}{2}\theta D^2 + m + \bar{m}\right), && \text{(V.3.20)} \\ k_n &= \frac{2\pi}{L}(D\Delta M + m - \bar{m}) + 2k_F D && \text{(V.3.21)} \end{aligned}$$

where ΔM is the change in the number of $\downarrow$-spins compared to the groundstate. Note that here $\Delta M = O(1)$ is assumed. The quantity θ is given by the solution of an integral equation. For vanishing magnetic field we have

$$\theta = \frac{1}{1 - \frac{\gamma}{\pi}} \qquad \text{(V.3.22)}$$

where $\cos\gamma = \Delta$, as usual. It can be related to the Fermi velocity (II.6.82) and the susceptibility (II.6.99) through

$$\theta = \frac{\pi}{2} v_F \chi. \qquad \text{(V.3.23)}$$

Similar relations also hold in other models.

By comparing with the predictions (V.3.12)–(V.3.14) of conformal field theory we can read of the central charge and the conformal dimensions as

$$c = 1, \qquad x = \Delta_\phi + \bar{\Delta}_\phi = \left(1 - \frac{\gamma}{\pi}\right)\frac{(\Delta M)^2}{2} + \frac{D^2}{2\left(1 - \frac{\gamma}{\pi}\right)}. \qquad \text{(V.3.24)}$$

Thus the set Q of quantum numbers that appears in (V.3.18) is here given by $Q = (\Delta M, D, m, \bar{m})$. Next let us illustrate how these results can be used to determine the asymptotic behaviour of (time-independent) correlation functions of the Heisenberg model in the critical region. In general we obtain by combining the generalization[4]

$$\langle \phi(z_1, \bar{z}_1)\phi(z_2, \bar{z}_2)\rangle = \frac{1}{|z_1 - z_2|^{2\Delta_\phi}|\bar{z}_1 - \bar{z}_2|^{2\bar{\Delta}_\phi}}. \qquad \text{(V.3.25)}$$

of (V.2.2) with (V.3.18) the expression

$$\boxed{\langle \psi(x,t)\psi(0,0)\rangle = \sum_Q C(Q) e^{-i2k_F Dx} \frac{1}{|x - iv_F t|^{2\Delta_\phi^+}} \cdot \frac{1}{|x + iv_F t|^{2\Delta_\phi^-}} \qquad \text{(V.3.26)}}$$

if we identify $z = ix + v_F t$ as the complex variable in the cylinder geometry. Here we have introduced

$$\Delta_\phi^+ = \Delta_\phi + m, \qquad \Delta_\phi^- = \bar{\Delta}_\phi + \bar{m}. \qquad \text{(V.3.27)}$$

First we look at the correlation $\langle \sigma_{j+r}^z \sigma_j^z \rangle$ with $r \gg 1$. Since σ^z does not change the particle number (magnetization), only operators with quantum numbers $Q = (\Delta M, D, m, \bar{m})$ that have

[4] Here we allow for a nonvanishing conformal spin.

$\Delta M = 0$ contribute. The leading terms come from the choices $Q = (0,0,1,0)$, $Q = (0,0,0,1)$, and $Q = (0,\pm 1,0,0)$. Inserting these into the spectral decomposition (V.3.7) taking into account (V.3.12)–(V.3.14) one obtains

$$\langle \sigma^z_{j+r} \sigma^z_j \rangle = \frac{A}{r^2} + \frac{B}{r^\theta} \cos 2k_F r + \dots \tag{V.3.28}$$

where the other contribution vanish faster for large r.
As a second correlation we study $\langle \sigma^-_{j+r} \sigma^+_j \rangle$. Obviously the physical operators change the particle number by $\Delta M = 1$. Therefore the leading contribution to the decay of the correlation function comes from the quantum numbers $Q = (1,0,0,0)$. Thus we have

$$\langle \sigma^-_{j+r} \sigma^+_j \rangle = \frac{C}{r^{1/\theta}} + \dots \tag{V.3.29}$$

As already mentioned before, the results (V.3.12) and (V.3.13) also apply to the Hubbard model at half-filling which has only one critical excitation, namely the spin excitations. In this case one finds

$$c = 1, \qquad x = \frac{(\Delta M)^2}{2} + \frac{D^2}{2} \tag{V.3.30}$$

which is just the isotropic antiferromagnetic limit $\gamma = 0$ of (V.3.24).
Away from half-filling the Hubbard model has critical charge and spin excitations. In this case the predictions of conformal field theory are modified. One finds

$$\boxed{\begin{aligned} E_0 &= Le_0 - \frac{\pi}{6L} \sum_\alpha v_\alpha c_\alpha, && \text{(V.3.31)} \\ E_n - E_0 &= \frac{2\pi}{L} \sum_\alpha (\Delta_\alpha + \bar{\Delta}_\alpha + m_\alpha + \bar{m}_\alpha), && \text{(V.3.32)} \\ k_n &= \frac{2\pi}{L} \sum_\alpha \left(\Delta_\alpha - \bar{\Delta}_\alpha + m_\alpha - \bar{m}_\alpha + 2k_F^{(\alpha)} D_\alpha \right). && \text{(V.3.33)} \end{aligned}}$$

Here the sum extends over all critical excitations which are characterized by their Fermi velocities v_α and their Fermi momenta $k_F^{(\alpha)}$.
For multicomponent systems the comparison of explicit results with the predictions (V.3.31) and (V.3.32) is not so easy since it is given by a sum of terms. One has to identify which terms correspond to each other in order to determine the conformal dimensions etc. Here sometimes selection rules are of help. What makes things worse is that the explicit form of the finite-size corrections is usually quite complicated and involve a coupling of the various excitations. Therefore we do not give the explicit results for the Hubbard model away from half-filling here. We only mention here that again the two central charges of charge and spin excitations are $c_c = c_s = 1$. Results for the finite-size corrections and conformal dimensions can be found in [18, 19].

Bibliography

[1] F.H.L. Eßler, V.E. Korepin (Eds.): *Exactly Solvable Models of Strongly Correlated Electrons*, (World Scientific, 1994)

[2] E.H. Lieb, F.Y. Wu: Phys. Rev. Lett. **20**, 1445 (1968); Erratum: Phys. Rev. Lett. **21**, 192 (1968)

[3] E.H. Lieb, F.Y. Wu: Physica **A321**, 1 (2003)

[4] C.N. Yang: Phys. Rev. Lett. **63**, 2144 (1989)

[5] C.N. Yang: Phys. Rev. Lett. **19**, 1312 (1967)

[6] A. Klümper, A. Schadschneider, J. Zittartz: Z. Phys. **B78**, 99 (1990)

[7] F. Woynarovich: J. Phys. **C15**, 85 (1982)

[8] F. Woynarovich: J. Phys. **C15**, 97 (1982)

[9] F. Woynarovich: J. Phys. **C16**, 5293 (1983)

[10] F.H.L. Eßler, V.E. Korepin: Phys. Rev. Lett. **72**, 908 (1994); Nucl. Phys. **426**, 505 (1994)

[11] V.E. Korepin, N.M. Bogoliubov, A.G. Izergin: *Quantum Inverse Scattering Method and Correlation Functions* (Cambridge University Press, 1993)

[12] P. di Francesco, P. Mathieu, D. Sénéchal: *Conformal Field Theory*, Springer (1996)

[13] A.A. Belavin, A.M. Polykaov, A.B. Zamolodchikov: Nucl. Phys. **B241**, 333 (1984)

[14] J. Cardy: in *Phase Transitions and Critical Phenomena, Vol. 11* (Eds. C. Domb, J.L. Lebowitz), Academic Press (1987)

[15] J. Cardy: Nucl. Phys. **B270**, 186 (1986)

[16] H.W.J. Blöte, J.L. Cardy, M.P. Nighingale: Phys. Rev. Lett. **56**, 742 (1986)

[17] I. Affleck: Phys. Rev. Lett. **56**, 746 (1986)

[18] F. Woynarovich: J. Phys. **A22**, 4243 (1989)

[19] H. Frahm, V.E. Korepin: Phys. Rev. **B42**, 10553 (1990); **B43**, 5653 (1991)