

12 The Two-Dimensional Hubbard Model

Robert Eder

Karlsruhe Institute of Technology

Institute for Solid State Physics

Contents

1	The Hubbard model	2
2	The Hubbard-I approximation	3
3	The Gutzwiller wave function	9
4	Strong coupling theory	15
5	Spin waves	18
6	Single hole problem	21
7	Summary and discussion	29
A	The bosonic Bogoliubov transformation	31

1 The Hubbard model

The Hubbard model was proposed in the 1960's to describe electrons in $3d$ transition metals. In these elements, the radial wave function of the $3d$ -electrons has a very small spatial extent. Therefore, if the $3d$ shell is occupied by several electrons, these are forced to be close to one another on the average so that the electrostatic energy will be large. The energy of a given transition-metal ion therefore varies strongly with the number of electrons it contains. To study the motion of conduction electrons under the influence of this strong Coulomb repulsion Hubbard [1], Kanamori [2] and Gutzwiller [3] proposed a simplified model. Thereby both the five-fold degeneracy of the $3d$ -orbital and the presence of other bands in the solid are neglected. Rather, one considers a lattice of sites – whereby the geometry of the lattice is not really specified – with one s -like orbital at each site. Orbitals on different sites are assumed to be orthogonal, but for not too distant sites i and j there are nonvanishing matrix elements $t_{i,j}$ of the Hamiltonian between the orbitals centered on these sites. The Coulomb interaction between electrons in orbitals on different sites is neglected, but if two electrons – which then necessarily have opposite spin – occupy the same orbital the energy is assumed to increase by the large amount U to simulate the strong dependence of the energy on the occupation number. If we denote the creation operator for an electron of spin σ in the orbital at the lattice site i by $c_{i,\sigma}^\dagger$ the model thus can be written as

$$H = \sum_{i,j} \sum_{\sigma} t_{i,j} c_{i,\sigma}^\dagger c_{j,\sigma} + U \sum_i n_{i,\uparrow} n_{i,\downarrow} = H_t + H_U. \quad (1)$$

Here $n_{i,\sigma} = c_{i,\sigma}^\dagger c_{i,\sigma}$ counts the number of electrons with spin σ in the orbital at site i .

After the discovery of the cuprate superconductors in 1987 and after Zhang and Rice demonstrated [4] that the CuO_2 planes in these compounds can be described by the so-called t - J model – which is equivalent to the Hubbard model in the limit $U/t \gg 1$ – there was renewed interest in the 2-dimensional Hubbard model. However, the lightly doped Mott-insulator – which most probably is the system to be understood in order to solve the many puzzles posed by the cuprate superconductors – is still far from being solved. Accordingly, the purpose of this lecture is to present basic approximations and to discuss some of the problems which so far precluded a full solution.

We consider (1) for a two-dimensional square lattice with N sites and periodic boundary conditions. The number of electrons with spin σ in the system is denoted by N_σ – whereby we are mostly interested in the nonmagnetic case $N_\uparrow = N_\downarrow$ – so that the number of electrons is $N_e = N_\uparrow + N_\downarrow$. In the following, densities per site will be denoted n , e.g., $n_\uparrow = N_\uparrow/N$. For $n_e = 1$ we have $N_\uparrow = N_\downarrow = N/2$ so that precisely half of the \mathbf{k} -points for each spin direction are occupied and we have a half-filled band, i.e., a metal in conventional band theory. Instead it will be shown below that for sufficiently large U/t the Hubbard model describes an insulator, the so-called Mott-insulator. The region of primary concern for cuprate superconductors is $n_e \geq 0.8$, i.e., the lightly doped Mott-insulator, and $U/t \approx 10$.

2 The Hubbard-I approximation

This is the ‘defining approximation’ of the Mott-insulator by which Hubbard for the first time introduced central concepts of strongly correlated electron systems such as the two Hubbard bands [1]. In the following we first give a sloppy re-derivation which is meant to clarify the physical content of the Hubbard-I approximation and then present Hubbard’s rigorous derivation in terms of Green’s functions.

We consider the case of finite U and $t_{i,j} = 0$, $N_{\uparrow} = N_{\downarrow} = N/2$ so that $N_e = N$. The ground state has one electron per lattice site and the energy is $E = 0$. Since the spin of the electron at any given site is arbitrary this ground state is highly degenerate. We ignore this degeneracy and assume that there is a unique state $|\Psi_0\rangle$ which may be thought of as a suitable superposition of all these degenerate states and which we assume to be ‘disordered’ – it will become clear in a moment what this means.

Next we assume that a small but finite $t_{i,j}$ is switched on. Then, an electron of spin σ can be transferred from a site j to another site i resulting in an empty site at j and a double occupancy at site i . The energy thereby increases by U . The hopping process is possible only if the electron which was originally at the site i has the spin $-\sigma$ and since our initial state $|\Psi_0\rangle$ is ‘disordered’ the probability for this to be the case is $1/2$, which is the definition of ‘disordered.’ We now interpret the original state $|\Psi_0\rangle$ as the vacuum, denoted by $|0\rangle$, of our theory and the state created by the hopping process as containing a fermionic hole-like particle at j and a fermionic double-occupancy-like particle at site i : $d_{i,\sigma}^{\dagger} h_{j,-\sigma}^{\dagger} |0\rangle$. The order of the fermionic operators in this state is due to the fact that in the original hopping term the annihilation operator $c_{j,\sigma}$ which creates the hole stands to the right of the creation operator $c_{i,\sigma}^{\dagger}$ which creates the double occupancy. Moreover we assign the negative spin to the operator which creates the hole because replacement of, e.g., an \uparrow -electron by a hole decreases the z -spin by $1/2$. We obtain the following Hamiltonian to describe the holes and double occupancies:

$$H_{\text{eff},1} = \frac{1}{2} \sum_{i,j} \sum_{\sigma} \left(t_{i,j} d_{i,\sigma}^{\dagger} h_{j,-\sigma}^{\dagger} + H.c. \right) + U \sum_{i,\sigma} d_{i,\sigma}^{\dagger} d_{i,\sigma}. \quad (2)$$

Once a hole and a double occupancy have been created, each of these particles may be transported further by the hopping term. If we assume that the surplus or missing electron retains its spin, which means that the double occupancies and holes propagate without ‘leaving a trace’ of inverted spins, for example a surplus \uparrow -electron can hop from site i to site j only if the spin at site j is \downarrow . Again, we assume that the probability for this is $1/2$. We therefore can write down the second term for the effective Hamiltonian

$$H_{\text{eff},2} = \frac{1}{2} \sum_{i,j} \sum_{\sigma} t_{i,j} \left(d_{i,\sigma}^{\dagger} d_{j,\sigma} - h_{i,-\sigma}^{\dagger} h_{j,-\sigma} \right). \quad (3)$$

The negative sign of the hopping term for holes is due to the fact that the original hopping term has to be rewritten as $-t_{i,j} c_{j,\sigma} c_{i,\sigma}^{\dagger}$ to describe the propagation of a hole. Addition of (2) and (3) and Fourier transformation gives

$$H_{\text{eff}} = \sum_{\mathbf{k},\sigma} \left(\left(\frac{\varepsilon_{\mathbf{k}}}{2} + U \right) d_{\mathbf{k},\sigma}^{\dagger} d_{\mathbf{k},\sigma} - \frac{\varepsilon_{\mathbf{k}}}{2} h_{\mathbf{k},\sigma}^{\dagger} h_{\mathbf{k},\sigma} \right) + \sum_{\mathbf{k},\sigma} \frac{\varepsilon_{\mathbf{k}}}{2} \left(d_{\mathbf{k},\sigma}^{\dagger} h_{-\mathbf{k},-\sigma}^{\dagger} + H.c. \right), \quad (4)$$

where $\varepsilon_{\mathbf{k}}$ is the Fourier transform of $t_{i,j}$. Note that this now is a quadratic form where the Coulomb interaction is described by the extra energy of U for the double-occupancy-like ‘particle.’ Via the Bogoliubov transformation

$$\begin{aligned}\gamma_{-, \mathbf{k}, \sigma} &= u_{\mathbf{k}} d_{\mathbf{k}, \sigma} + v_{\mathbf{k}} h_{-\mathbf{k}, -\sigma}^{\dagger} \\ \gamma_{+, \mathbf{k}, \sigma} &= -v_{\mathbf{k}} d_{\mathbf{k}, \sigma} + u_{\mathbf{k}} h_{-\mathbf{k}, -\sigma}^{\dagger}\end{aligned}\quad (5)$$

this can be solved resulting in the dispersion relations for the lower and upper Hubbard band

$$E_{\mathbf{k}, \pm} = \frac{1}{2} \left(\varepsilon_{\mathbf{k}} + U \pm \sqrt{\varepsilon_{\mathbf{k}}^2 + U^2} \right). \quad (6)$$

In the limit $U/t \gg 1$ this simplifies to $E_{\mathbf{k}, -} = \varepsilon_{\mathbf{k}}/2$, $E_{\mathbf{k}, +} = \varepsilon_{\mathbf{k}}/2 + U$ so that the original band with dispersion $\varepsilon_{\mathbf{k}}$ is split into two bands, separated by a gap of U , each having half of the original width. For the case of particle-hole symmetry, the chemical potential is $U/2$ [5], so the lower band is completely filled and the upper one completely empty. Rather than being a metal, as expected for the situation of a half-filled band, the presence of the Coulomb interaction turns the system into an insulator. From the above we can see that this is the consequence of ‘expanding around’ the hypothetical ‘vacuum state’ $|\Psi_0\rangle$ with one electron per site so that we obtain a dilute gas of hole-like and double-occupancy-like particles that are created in pairs and propagate, whereby the double-occupancies have a large ‘energy of formation’ of U .

To compute the spectral weight of the bands we translate the electron annihilation operator as follows:

$$c_{\mathbf{k}, \sigma} = \frac{1}{\sqrt{2}} \left(d_{\mathbf{k}, \sigma} + h_{-\mathbf{k}, -\sigma}^{\dagger} \right) = \frac{1}{\sqrt{2}} \left((u_{\mathbf{k}} + v_{\mathbf{k}}) \gamma_{-, \mathbf{k}, \sigma} + (u_{\mathbf{k}} - v_{\mathbf{k}}) \gamma_{+, \mathbf{k}, \sigma} \right).$$

Namely, annihilation of an electron on a singly occupied site creates a hole, whereas annihilation on a doubly occupied site annihilates a double occupancy. The factor of $1/\sqrt{2}$ takes into account that both processes are possible with a probability of $1/2$ in the disordered state. We obtain

$$Z_{\pm}(\mathbf{k}) = \frac{1}{2} (u_{\mathbf{k}} \mp v_{\mathbf{k}})^2 = \frac{1}{2} \left(1 \pm \frac{\varepsilon_{\mathbf{k}}}{\sqrt{\varepsilon_{\mathbf{k}}^2 + U^2}} \right). \quad (7)$$

Taking again the limit $U/t \gg 1$, the spectral weight of each of the bands is only $\approx 1/2$ per \mathbf{k} -point.

Next, we derive these results in a more rigorous fashion following Hubbard’s original paper [1]. We split the electron operator into the two eigenoperators of the interaction part H_U in (1):

$$c_{i, \sigma} = c_{i, \sigma} n_{i, -\sigma} + c_{i, \sigma} (1 - n_{i, -\sigma}) = \hat{d}_{i, \sigma} + \hat{c}_{i, \sigma}, \quad (8)$$

which obey $[\hat{d}_{i, \sigma}, H_U] = U \hat{d}_{i, \sigma}$ and $[\hat{c}_{i, \sigma}, H_U] = 0$. Then we define the four time-ordered zero-temperature Green’s functions [6]

$$G_{\alpha, \beta}(\mathbf{k}, t) = -i \langle T \alpha_{\mathbf{k}, \sigma}(t) \beta_{\mathbf{k}, \sigma}^{\dagger} \rangle, \quad (9)$$

where $\alpha, \beta \in \{\hat{c}, \hat{d}\}$. These Green's functions obey the equations of motion (with $\hbar = 1$)

$$i\partial_t G_{\alpha,\beta}(\vec{k}, t) = \delta(t) \langle \{\beta_{\mathbf{k},\sigma}^\dagger, \alpha_{\mathbf{k},\sigma}\} \rangle - i \langle T[\alpha_{\mathbf{k},\sigma}, H](t) \beta_{\mathbf{k},\sigma}^\dagger \rangle.$$

The commutators $[\alpha_{\mathbf{k},\sigma}, H_U]$ are trivial but the commutators with the kinetic term H_t are involved. After some algebra, using the identity $n_{i,\sigma} = n_i/2 + \sigma S_i^z$, we find:

$$\begin{aligned} [\hat{c}_{i,\uparrow}, H_t] &= \sum_j t_{ij} \left[\left(1 - \frac{n_e}{2}\right) c_{j,\uparrow} + (c_{j,\uparrow} S_i^z + c_{j,\downarrow} S_i^-) - \frac{1}{2} c_{j,\uparrow} (n_i - n_e) + c_{j,\downarrow}^\dagger c_{i,\downarrow} c_{i,\uparrow} \right], \\ [\hat{d}_{i,\uparrow}, H_t] &= \sum_j t_{ij} \left[\frac{n_e}{2} c_{j,\uparrow} - (c_{j,\uparrow} S_i^z + c_{j,\downarrow} S_i^-) + \frac{1}{2} c_{j,\uparrow} (n_i - n_e) - c_{j,\downarrow}^\dagger c_{i,\downarrow} c_{i,\uparrow} \right]. \end{aligned} \quad (10)$$

The first term on the right-hand side describes the ‘simple’ propagation of the hole. The second term is the Clebsch-Gordan contraction of the spin-1 operator \mathbf{S}_i and the spinor $c_{j,\sigma}$ into a spin-1/2 object. It describes how a hole moves to site j but leaves behind a spin-excitation at site i . Similarly, the third term describes hopping combined with creation of a density excitation at site j whereas the last term describes the coupling to a pair-excitation (this would be important for negative U). The Hubbard-I approximation is obtained by keeping only the first term in each of the square brackets on the respective right-hand sides – obviously a rather crude approximation. After Fourier transformation we obtain

$$\begin{aligned} [\hat{c}_{\mathbf{k},\uparrow}, H] &\approx \left(1 - \frac{n_e}{2}\right) \varepsilon_{\mathbf{k}} (\hat{c}_{\mathbf{k},\uparrow} + \hat{d}_{\mathbf{k},\uparrow}) \\ [\hat{d}_{\mathbf{k},\uparrow}, H] &\approx \frac{n_e}{2} \varepsilon_{\mathbf{k}} (\hat{c}_{\mathbf{k},\uparrow} + \hat{d}_{\mathbf{k},\uparrow}) + U \hat{d}_{\mathbf{k},\uparrow} \end{aligned}$$

If we set $n_e = 1/2$ and identify

$$\begin{aligned} \sqrt{2} \hat{d}_{\mathbf{k},\sigma} &\rightarrow d_{\mathbf{k},\sigma} \\ \sqrt{2} \hat{c}_{\mathbf{k},\sigma} &\rightarrow h_{\mathbf{k},-\sigma} \end{aligned} \quad (11)$$

exactly the same equations of motion are obtained from the heuristic Hamiltonian H_{eff} (4) (the significance of the factor $\sqrt{2}$ will become clear in a moment).

Using the anticommutator relations $\{\hat{d}_{i,\sigma}^\dagger, \hat{d}_{i,\sigma}\} = n_{i,-\sigma}$, $\{\hat{c}_{i,\sigma}^\dagger, \hat{c}_{i,\sigma}\} = (1 - n_{i,-\sigma})$, $\{\hat{d}_{i,\sigma}^\dagger, \hat{c}_{i,\sigma}\} = \{\hat{c}_{i,\sigma}^\dagger, \hat{d}_{i,\sigma}\} = 0$ and putting $\langle n_{i,\sigma} \rangle = n_e/2$ we obtain the Fourier transformed equations of motion:

$$\begin{pmatrix} \omega - (1 - n_e/2) \varepsilon_{\mathbf{k}} & -(1 - n_e/2) \varepsilon_{\mathbf{k}} \\ -n_e/2 \varepsilon_{\mathbf{k}} & \omega - n_e/2 \varepsilon_{\mathbf{k}} - U \end{pmatrix} \begin{pmatrix} G_{\hat{c},\hat{c}} & G_{\hat{c},\hat{d}} \\ G_{\hat{d},\hat{c}} & G_{\hat{d},\hat{d}} \end{pmatrix} = \begin{pmatrix} 1 - n_e/2 & 0 \\ 0 & n_e/2 \end{pmatrix}. \quad (12)$$

If we consider again $n_e = 1/2$ and multiply both sides of (12) by 2 we have the unit matrix on the right while all Green's functions are multiplied by a factor of 2 – which would originate from the additional factor of $\sqrt{2}$ in (11). The Hamiltonian (4) would therefore produce exactly the same equations of motion as the Hubbard-I approximation, which demonstrates the equivalence. We continue with arbitrary n_e and use the identity (which holds for any 2×2 matrix)

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

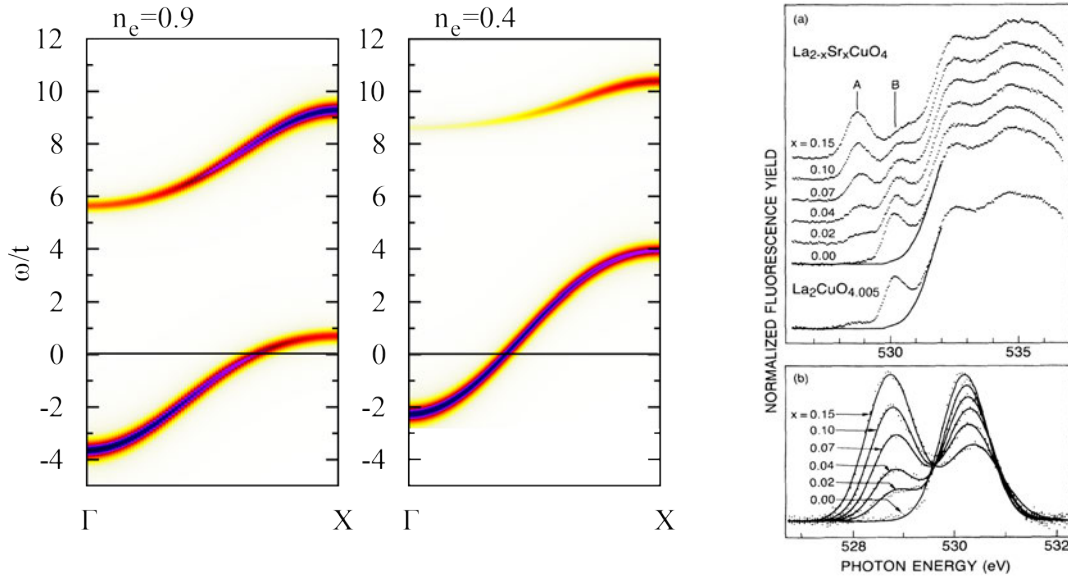


Fig. 1: Left: Single particle spectral function obtained from the Green's function (13) for two different electron densities. The Fermi energy is at zero. Right: X-ray luminescence spectra for $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ show the unoccupied part of the lower Hubbard-band (A) and the upper Hubbard-band (B). With decreasing $n_e \approx 1 - x$ the upper Hubbard-band rapidly loses intensity. Reprinted with permission from [7], Copyright 1991 by the American Physical Society.

to solve for the Green's function matrix $\mathbf{G}(\mathbf{k}, \omega)$. Recalling that $c_{\mathbf{k},\sigma} = \hat{c}_{\mathbf{k},\sigma} + \hat{d}_{\mathbf{k},\sigma}$, the standard electron Green's function $G(\mathbf{k}, \omega)$ is given by $G = G_{\hat{c},\hat{c}} + G_{\hat{c},\hat{d}} + G_{\hat{d},\hat{c}} + G_{\hat{d},\hat{d}}$ and after some algebra this can be brought to the form

$$G(\mathbf{k}, \omega) = \frac{1}{\omega - \varepsilon_{\mathbf{k}} - \Sigma(\omega)}$$

$$\Sigma(\omega) = \frac{n_e}{2} U + \frac{n_e}{2} \left(1 - \frac{n_e}{2}\right) \frac{U^2}{\omega - (1 - \frac{n_e}{2}) U} \quad (13)$$

from which also the self-energy $\Sigma(\omega)$ corresponding to the Hubbard-I approximation can be read off. In order to fix the Fermi energy E_F we write the operator of electron number as

$$\hat{N}_e = 2 \sum_i n_{i,\uparrow} n_{i,\downarrow} + \sum_i \left(n_{i,\uparrow} (1 - n_{i,\downarrow}) + n_{i,\downarrow} (1 - n_{i,\uparrow}) \right) = \sum_{i,\sigma} \left(\hat{d}_{i,\sigma}^\dagger \hat{d}_{i,\sigma} + \hat{c}_{i,\sigma}^\dagger \hat{c}_{i,\sigma} \right).$$

The expectation value of \hat{N}_e then can be expressed in terms of the Green's functions (9)

$$\begin{aligned} \langle N_e \rangle &= -2i \sum_{\mathbf{k}} \left(G_{\hat{d},\hat{d}}(\mathbf{k}, t = 0^-) + G_{\hat{c},\hat{c}}(\mathbf{k}, t = 0^-) \right) \\ &= -2i \sum_{\mathbf{k}} \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega e^{i\omega 0^+} \left(G_{\hat{d},\hat{d}}(\mathbf{k}, \omega) + G_{\hat{c},\hat{c}}(\mathbf{k}, \omega) \right) \\ &= 2 \sum_{\mathbf{k}} \int_{-\infty}^{\mu} d\omega \left(A_{\hat{d},\hat{d}}(\mathbf{k}, \omega) + A_{\hat{c},\hat{c}}(\mathbf{k}, \omega) \right), \end{aligned} \quad (14)$$

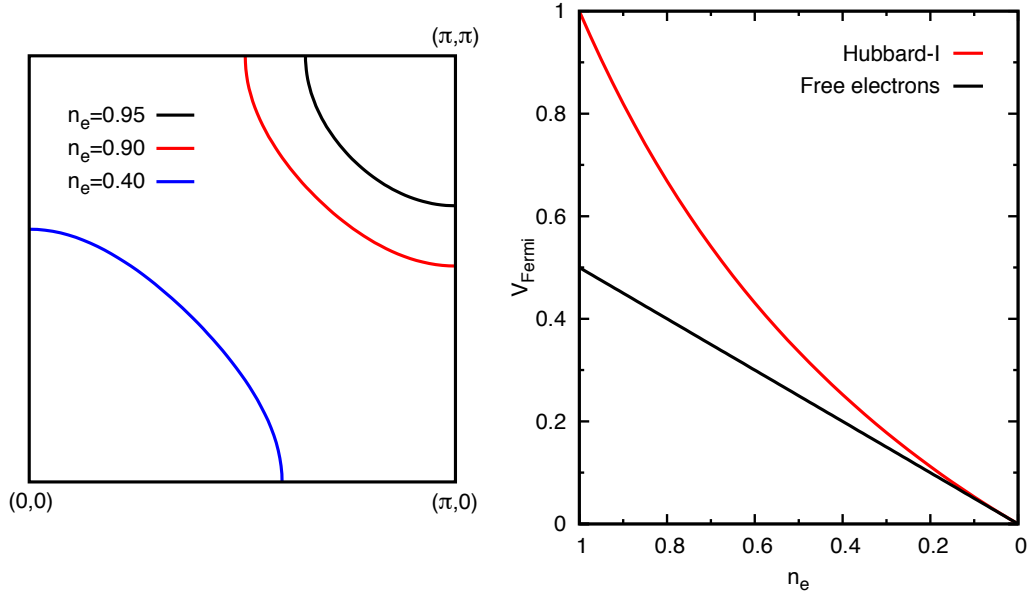


Fig. 2: Left: Fermi surface for different electron densities. Right: Fermi surface volume as a function of electron density n_e .

where $A_{\hat{d},\hat{d}}$ and $A_{\hat{c},\hat{c}}$ are the spectral densities of the respective Green's function. The resulting expression for N_e is somewhat lengthy so we do not write it explicitly. Close to half-filling it can be written as

$$N_e = \Theta(E_F - E_{\mathbf{k},-}) + \mathcal{O}(1 - n_e).$$

For $n_e = 1$ this means that the completely filled lower Hubbard band with N occupied momenta per spin-direction corresponds to N electrons (as it has to be), whereas for $n_e < 1$ it means that the lower Hubbard band is doped with holes. Figure 1 shows the spectral density obtained from the Green's function (13) for $U/t = 8$ and two different band fillings, whereas Figure 2 shows the resulting Fermi surfaces and the dependence of the Fermi surface volume, obtained from (14), on the electron density. In Figure 1, one can recognize the two Hubbard bands separated by an appreciable energy gap. For $n_e = 0.9$, i.e. close to half-filling, the Fermi energy intersects the lower Hubbard band close to (π, π) so that the Fermi surface takes the form of a small pocket around $X = (\pi, \pi)$, see Figure 2, whereby the area of the pocket is roughly proportional to the hole density $n_h = 1 - n_e$.

An interesting feature seen in Figure 1 is the transfer of spectral weight from the upper to the lower Hubbard band upon hole doping: as the electron density n_e decreases, the upper Hubbard band persists but loses weight, whereas the lower Hubbard band becomes more intense. To understand this we note first that for $n_e \leq 1$ the upper Hubbard band always belongs to the inverse photoemission or electron addition spectrum. Also, we have seen in the simplified derivation that the upper band mainly has double-occupancy character. As electrons are removed from the system, however, the probability that an added electron is placed at an occupied site to create a double occupancy becomes smaller and consequently the weight of the upper band diminishes. This doping-dependent intensity of what would be the conduction band in an ordinary

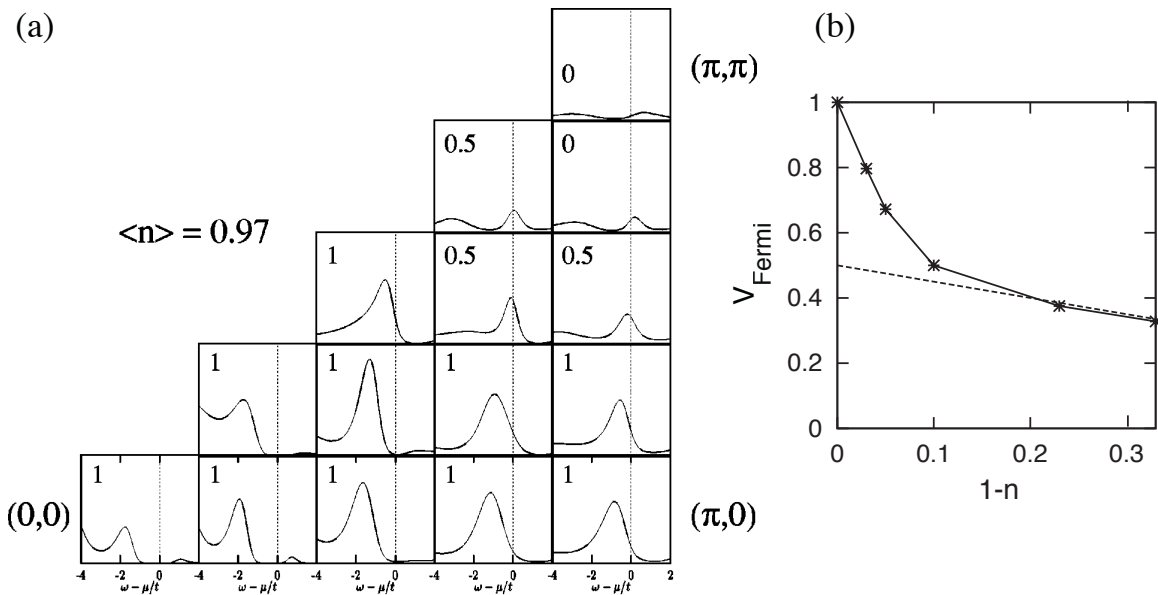


Fig. 3: *Left: Single particle spectral function $A(\mathbf{k}, \omega)$ obtained by Quantum Monte-Carlo simulations on an 8×8 cluster at $k_B T = t$. Right: Fermi surface volume (15) deduced from $A(\mathbf{k}, \omega)$ versus electron density. The dashed line corresponds to free electrons. Reprinted with permission from [8], Copyright 2000 by the American Physical Society.*

semiconductor or insulator is one of the fingerprints of strong correlations and can be observed experimentally in cuprate superconductors. An example is shown in Figure 1 [7]. It should be noted, however, that the Hubbard-I approximation considerably underestimates the decrease of the intensity of the upper Hubbard band with doping.

Figure 2 also shows the dependence of the Fermi surface volume V_{Fermi} on electron density n_e . More precisely, this is the fraction of the Brillouin zone where the lower Hubbard band is below E_F , i.e., ‘occupied.’ Also shown is the Fermi surface volume for free electrons, where $V_{\text{Fermi}} = n_e/2$. The Hubbard-I approximation gives $V_{\text{Fermi}} \rightarrow 1$, a completely filled band, as $n_e \rightarrow 1$, and approaches the free electron behavior for small n_e . This leads to a peculiar nonlinear dependence in $V_{\text{Fermi}}(n_e)$, which most probably is unphysical.

Let us now compare the Hubbard-I approximation to numerical simulations. As we saw in our simplified derivation, an important assumption of the Hubbard-I approximation is the ‘disordered’ ground state. This is best realized at high temperatures, more precisely at a temperature much higher than the characteristic energy of spin excitations, which will be seen to be $J = 4t^2/U$. Figure 3 shows the result of a quantum Monte-Carlo calculation of the spectral density for an 8×8 cluster at the rather high temperature $k_B T = t$. The 8×8 cluster has the allowed momenta $(n\pi/4, m\pi/4)$ with integer m and n and Figure 3 shows the part of the spectral density near the chemical potential μ for all allowed momenta in the irreducible wedge of the Brillouin zone for an electron densities close to $n_e = 1$. Close to (π, π) a relatively well-defined peak passes through μ as (π, π) is approached and forms a relatively small hole pocket around (π, π) – similar to the prediction of the Hubbard-I approximation in Figure 1

for $n_e = 0.9$. To study V_{Fermi} , an ‘occupation number’ $n_{\mathbf{k}}$ of 1, 0.5 or 0 was assigned to each momentum \mathbf{k} , depending on whether the dispersive peak is below, more or less on, or above the chemical potential at \mathbf{k} . The fractional Fermi surface volume then is

$$V_{\text{Fermi}} = \frac{1}{64} \sum_{\mathbf{k}} n_{\mathbf{k}}, \quad (15)$$

where 64 is the number of momenta in the 8×8 cluster. The obtained estimate for V_{Fermi} is also shown in Figure 3 as a function of electron density and indeed has a rough similarity to the result for the Hubbard-I approximation.

3 The Gutzwiller wave function

The Gutzwiller wave function is the second ‘classic’ approximation for the Hubbard model. It starts from the Fermi sea $|FS\rangle$, i.e. the ground state for $U = 0$, and reduces the number of double occupancies by acting with a suitable projection operator. More precisely, the Gutzwiller wave function reads [3]

$$|\Phi_G\rangle = \prod_i (1 - \lambda n_{i,\uparrow} n_{i,\downarrow}) |FS\rangle,$$

where λ is a variational parameter to be determined by minimizing the energy E_G . First we rewrite the Fermi sea as a superposition of real space configurations. Suppressing the spin index we have

$$\begin{aligned} \prod_{j=1}^M c_{\mathbf{k}_j}^\dagger |0\rangle &= \frac{1}{\sqrt{N^M}} \sum_{i_1, i_2, i_3, \dots, i_M} \exp\left(i \sum_{j=1}^M \mathbf{k}_j \cdot \mathbf{R}_{i_j}\right) \prod_{j=1}^M c_{i_j}^\dagger |0\rangle \\ &= \frac{1}{\sqrt{N^M}} \sum_{i_1 > i_2 > i_3 \dots > i_M} \sum_{\sigma} \exp\left(i \sum_{j=1}^M \mathbf{k}_j \cdot \mathbf{R}_{i_{\sigma(j)}}\right) \prod_{j=1}^M c_{i_{\sigma(j)}}^\dagger |0\rangle \end{aligned}$$

In the second line we used the fact that instead of summing over all M -tuples of indices we may as well sum only over ordered M -tuples of indices and then sum over all $M!$ permutations σ of the M indices.

Next, in each of the products $\prod_{j=1}^M c_{i_{\sigma(j)}}^\dagger$ we permute the c_i^\dagger operators back to the ordered sequence $c_{i_1}^\dagger c_{i_2}^\dagger \dots c_{i_M}^\dagger$. The permutation that brings $\sigma(i) \rightarrow i$ obviously is σ^{-1} and since the Fermi sign of σ^{-1} is equal to that of σ we obtain

$$\begin{aligned} &\frac{1}{\sqrt{N^M}} \sum_{i_1 > i_2 > i_3 \dots > i_M} \sum_{\sigma} (-1)^\sigma \exp\left(i \sum_{j=1}^M \mathbf{k}_j \cdot \mathbf{R}_{i_{\sigma(j)}}\right) c_{i_1}^\dagger c_{i_2}^\dagger \dots c_{i_M}^\dagger |0\rangle \\ &= \frac{1}{\sqrt{N^M}} \sum_{i_1 > i_2 > i_3 \dots > i_M} D(\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_M | i_1, i_2, \dots, i_M) c_{i_1}^\dagger c_{i_2}^\dagger \dots c_{i_M}^\dagger |0\rangle, \end{aligned}$$

where the second line is the definition of the symbol $D(\mathbf{k}_j | i_j)$. From the above we see that the Fermi sea may be thought of as a superposition of real space configurations

$$c_{i_{1,\uparrow}}^\dagger c_{i_{2,\uparrow}}^\dagger c_{i_{3,\uparrow}}^\dagger \dots c_{i_{N,\uparrow}}^\dagger c_{j_{1,\downarrow}}^\dagger c_{j_{2,\downarrow}}^\dagger c_{j_{3,\downarrow}}^\dagger \dots c_{j_{N,\downarrow}}^\dagger |0\rangle$$

that are multiplied by two determinants D , one for each spin direction. Each of these real space configurations has a certain number N_d of doubly occupied sites and therefore gets an additional factor of $(1 - \lambda)^{N_d} < 1$ in the Gutzwiller wave function so that states with a larger number of double occupancies have a smaller weight as compared to the original Fermi sea. The Gutzwiller function can be decomposed into components with fixed N_d

$$|\Phi_G\rangle = \sum_{N_d} |\Phi(N_d)\rangle,$$

where $|\Phi(N_d)\rangle$ is the sum over all real-space configurations with N_d double occupancies, each multiplied by its proper prefactor. The total norm $\langle\Phi_G|\Phi_G\rangle$ can be rewritten as the sum over N_d of $W(N_d) = \langle\Phi(N_d)|\Phi(N_d)\rangle$ and we consider which N_d gives the largest contribution in this sum. To compute norms, we need to evaluate expressions such as

$$\begin{aligned} D^*(\mathbf{k}_j|i_j) D(\mathbf{k}_j|i_j) &= \sum_{\sigma,\sigma'} (-1)^\sigma (-1)^{\sigma'} \exp\left(i \sum_{j=1}^M \mathbf{k}_j \cdot (\mathbf{R}_{i_{\sigma(j)}} - \mathbf{R}_{i_{\sigma'(j)}})\right) \\ &= M! + \sum_{\sigma \neq \sigma'} (-1)^\sigma (-1)^{\sigma'} \exp\left(i \sum_{j=1}^M \mathbf{k}_j \cdot (\mathbf{R}_{i_{\sigma(j)}} - \mathbf{R}_{i_{\sigma'(j)}})\right). \end{aligned} \quad (16)$$

where in the first term we have collected the $M!$ terms with $\sigma = \sigma'$. At this point, we make an important approximation: (16) still has to be summed over $i_1, i_2, i_3 \dots i_M$. The terms for $\sigma \neq \sigma'$ thereby have a rapidly oscillating phase and a large degree of cancellation will occur in the summation. Accordingly we retain only the first term, i.e., we replace

$$D^*(\mathbf{k}_j|i_j) D(\mathbf{k}_j|i_j) \approx M!.$$

With this approximation the contribution of all states with N_d double occupancies becomes

$$W(N_d) = \frac{N_\uparrow! N_\downarrow!}{N^{N_\uparrow+N_\downarrow}} (1 - \lambda)^{2N_d} C(N_\uparrow, N_\downarrow, N_d)$$

where $C(N_\uparrow, N_\downarrow, N_d)$ is the number of ways in which N_\uparrow electrons with spin \uparrow and N_\downarrow electrons with spin \downarrow can be distributed over the N lattice sites such as to generate N_d double occupancies. This is a straightforward combinatorial problem with the result

$$C(N_\uparrow, N_\downarrow, N_d) = \frac{N!}{N_d!(N_\uparrow - N_d)! (N_\downarrow - N_d)! (N - N_\uparrow - N_\downarrow + N_d)!}.$$

Next, we take the logarithm of $W(N_d)$, use the Stirling formula $\log(N!) \approx N \log(N) - N$ and differentiate with respect to N_d . Introducing the densities $n_d = N_d/N$ etc. we obtain

$$\begin{aligned} \frac{d}{dN_d} \log(W(N_d)) &= \log\left((1 - \lambda)^2 \frac{(n_\uparrow - n_d)(n_\downarrow - n_d)}{n_d(1 - n_\uparrow - n_\downarrow + n_d)}\right), \\ \frac{d^2}{dN_d^2} \log(W(N_d)) &= -\frac{1}{N} \left(\frac{1}{n_d} + \frac{1}{n_\uparrow - n_d} + \frac{1}{n_\downarrow - n_d} + \frac{1}{1 - n_\uparrow - n_\downarrow + n_d}\right) = -\frac{c}{N} \end{aligned}$$

where $c > 0$ in the last line is of order unity. The first of these equations gives us the n_d where the contribution to the norm, $W(N_d)$ is a maximum. For the general case, the formula is somewhat involved, so we specialize to the case $n_\uparrow = n_\downarrow = 1/2$ where

$$n_{d,0} = \frac{1 - \lambda}{2(2 - \lambda)}. \quad (17)$$

For the noninteracting case $\lambda \rightarrow 0$ this gives $n_{d,0} = 1/4$ as it has to be. From the second equation we find

$$\begin{aligned} \log(W(N_d)) &= \log(W(N_{d,0})) - \frac{c}{2N} (N - N_{d,0})^2 + \dots \\ W(N_d) &= W(N_{d,0}) \exp\left(-\frac{c}{2N}(N_d - N_{d,0})^2\right) = W(N_{d,0}) \exp\left(-\frac{Nc}{2}(n_d - n_{d,0})^2\right), \end{aligned}$$

which shows that as a function of n_d the weight $W(N_d)$ is a Gaussian with a width $\propto N^{-1/2}$. This means, however, that in the thermodynamical limit only states with $n_d = n_{d,0}$ have an appreciable weight in the Gutzwiller wave function and variation of λ simply shifts this sharp peak of $W(N_d)$ to a different $n_{d,0}$. An immediate consequence is that the computation of the expectation value of the interaction Hamiltonian becomes trivial, namely $\langle H_U \rangle = NU n_{d,0}$. The expectation value of the kinetic energy is more involved. The above discussion showed that the Gutzwiller wave function is composed of real-space configurations for which the number of double occupancies is close to a certain value $N_{d,0}$, which is smaller than for the noninteracting Fermi sea. This means, however, that the expectation value of the kinetic energy is smaller as well. Namely, using again the operators \hat{d} and \hat{c} we have

$$c_{i,\sigma}^\dagger c_{j,\sigma} = \hat{d}_{i,\sigma}^\dagger \hat{d}_{j,\sigma} + \hat{c}_{i,\sigma}^\dagger \hat{d}_{j,\sigma} + \hat{d}_{i,\sigma}^\dagger \hat{c}_{j,\sigma} + \hat{c}_{i,\sigma}^\dagger \hat{c}_{j,\sigma}.$$

If the number of double occupancies is decreased, the expectation value of the first term on the right-hand side clearly must decrease. Second, since the number of electrons is constant, reducing the number of double occupancies necessarily results in a reduction of the number of empty sites by the same number so that the expectation value of the last term on the right-hand side also must decrease.

The Gutzwiller approximation assumes that these effects can be taken into account by reducing the expectation value of the kinetic energy of the uncorrelated Fermi sea by suitable renormalization factors η :

$$\frac{\langle \Phi_G | H_t | \Phi_G \rangle}{\langle \Phi_G | \Phi_G \rangle} = \sum_{\sigma} \eta_{\sigma} \langle FS, \sigma | H_t | FS, \sigma \rangle$$

where $|FS, \sigma\rangle$ is the Fermi sea for σ -electrons (if $N_\uparrow = N_\downarrow$ the two terms are of course identical). These renormalization factors η_{σ} thereby are evaluated for an ‘auxiliary wave function’ in which the determinants $D(\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_M | i_1, i_2, \dots, i_M)$ are replaced by a constant (which would have to be $\sqrt{M!}$ if the auxiliary wave function is supposed to have the same norm as the Gutzwiller wave function) and where the Fermi sign is ignored in the calculation of all matrix elements of the hopping term (this is because the Fermi sign is supposed to be taken care

of already by the filling of the uncorrelated Fermi sea according to the Pauli principle!). The evaluation of the η by combinatorial considerations is discussed in a very transparent way by Ogawa, Kanda, and Matsubara [9]. Here we use an even simpler way of calculating η by introducing four ‘book-keeping kets’ for every site i : $|i, 0\rangle$, $|i, \uparrow\rangle$, $|i, \downarrow\rangle$ and $|i, \uparrow\downarrow\rangle$. They represent in an obvious way the four possible configurations of the site i . Then we define

$$B_i = \frac{|i, 0\rangle + \alpha_\uparrow|i, \uparrow\rangle + \alpha_\downarrow|i, \downarrow\rangle + \beta|i, \uparrow\downarrow\rangle}{\sqrt{1 + \alpha_\uparrow^2 + \alpha_\downarrow^2 + \beta^2}}$$

$$|\Psi\rangle = \prod_i B_i$$

with real α_σ and β . The state $|\Psi\rangle$ has norm 1, and if it were translated into a true state of electrons, the numbers of electrons and double occupancies would be

$$\langle N_\sigma \rangle = N \frac{\alpha_\sigma^2 + \beta^2}{1 + \alpha_\uparrow^2 + \alpha_\downarrow^2 + \beta^2},$$

$$\langle N_d \rangle = N \frac{\beta^2}{1 + \alpha_\uparrow^2 + \alpha_\downarrow^2 + \beta^2}. \quad (18)$$

These equations can be reverted to give

$$\alpha_\sigma = \sqrt{\frac{n_\sigma - n_d}{1 - n_\uparrow - n_\downarrow + n_d}},$$

$$\beta = \sqrt{\frac{n_d}{1 - n_\uparrow - n_\downarrow + n_d}}. \quad (19)$$

On the other hand, $|\Psi\rangle$ does not correspond to a state with a fixed number of electrons, so we introduce

$$|\Psi'\rangle = \mathcal{P}(N_\uparrow, N_\downarrow, N_d) |\Psi\rangle,$$

where the projection operator \mathcal{P} projects onto the component of $|\Psi\rangle$ that has precisely $\langle N_\uparrow \rangle$ \uparrow -electrons, etc. Next, the representation of the electron annihilation operator $c_{i,\sigma}$ is

$$\tilde{c}_{i,\sigma} = |i, 0\rangle \langle i, \sigma| + |i, -\sigma\rangle \langle i, \uparrow\downarrow|.$$

Here a subtle detail should be noted: in the expression on the right-hand side it is assumed that a double occupancy is always converted into the state $|i, -\sigma\rangle$ with a positive sign. This would not be the case for the true fermion operator, where the sign would depend on the sequence of the two electron creation operators on the doubly occupied site. This is precisely the neglect of the Fermi sign that was mentioned above. Then, to estimate the reduction of the kinetic energy due to the reduction of the number of doubly occupied and empty sites we evaluate

$$r(\sigma, n_\uparrow, n_\downarrow, n_d) = \frac{\langle \Psi' | \tilde{c}_{i,\sigma}^\dagger \tilde{c}_{j,\sigma} | \Psi' \rangle}{\langle \Psi' | \Psi' \rangle}. \quad (20)$$

So far our auxiliary wave function has not brought about much simplification because the presence of the projection operator \mathcal{P} makes the computation of r very tedious. It is straightforward to see, however, that if $|\Psi\rangle$ is decomposed into components of fixed N_\uparrow , N_\downarrow , and N_d , only those components with values of N_\uparrow , N_\downarrow , and N_d that deviate by at most $N^{-1/2}$ from the average values (18) have an appreciable weight. This means, however, that \mathcal{P} simply can be dropped so that we replace $|\Psi'\rangle \rightarrow |\Psi\rangle$ in (20). Then, since $|\Psi\rangle$ is normalized, the denominator can be dropped. Since $|\Psi\rangle$ is a product state, the expectation value of the two operators factorizes, and since all sites are equivalent and the coefficients α_σ and β are real, the expectation values of $\tilde{c}_{i,\sigma}^\dagger$ and $\tilde{c}_{j,\sigma}$ are identical. So

$$\begin{aligned} r(\sigma, n_\uparrow, n_\downarrow, n_d) &= \langle \Psi | \tilde{c}_{i,\sigma}^\dagger | \Psi \rangle^2 = \left(\frac{\alpha_\sigma + \alpha_{-\sigma}\beta}{1 + \alpha_\uparrow^2 + \alpha_\downarrow^2 + \beta^2} \right)^2 \\ &= \left(\frac{\sqrt{n_\sigma - n_d} \sqrt{1 - n_\uparrow - n_\downarrow + n_d} + \sqrt{n_d} \sqrt{n_{-\sigma} - n_d}}{\sqrt{n_\sigma(1 - n_\sigma)}} \right)^2, \end{aligned}$$

where the second line has been obtained by inserting (19). In this way we have expressed $r(n_\sigma, n_d)$ in terms of n_d , which in turn is given as a function of λ by (17). Lastly, we divide r by the value for $U \rightarrow 0$ where $n_d = n_\uparrow \cdot n_\downarrow$ to obtain the proper limiting value for $U = 0$ and finally obtain

$$\eta(\sigma, n_\uparrow, n_\downarrow, n_d) = \left(\frac{\sqrt{n_\sigma - n_d} \sqrt{1 - n_\uparrow - n_\downarrow + n_d} + \sqrt{n_d} \sqrt{n_{-\sigma} - n_d}}{\sqrt{n_\sigma(1 - n_\sigma)}} \right)^2. \quad (21)$$

In varying the energy it is actually easier to switch from λ to n_d as variational parameter. Specializing to the paramagnetic case $n_\uparrow = n_\downarrow$, the energy per site thus becomes

$$e_G = \eta(n_\sigma, n_d) t_0 + n_d U. \quad (22)$$

where $e_G = E_G/N$ and t_0 is the (kinetic) energy of the Fermi sea per site. Using (21) this is now readily minimized with respect to n_d .

The Gutzwiller wave function gives us strictly speaking only the ground state energies and some ground state expectation values, but not a band structure. However, we may consider states like

$$|\Phi_G(\mathbf{k})\rangle = \prod_i (1 - \lambda n_{i,\uparrow} n_{i,\downarrow}) c_{\mathbf{k},\uparrow} |FS\rangle,$$

i.e. a state with one hole in the Fermi sea (it is understood that \mathbf{k} is an occupied momentum). The Fermi sea with a hole has energy $E_{FS} - \varepsilon_{\mathbf{k}}$. It thus seems plausible that the energy of $|\Phi_G(\mathbf{k})\rangle$ is $E_G - \tilde{\varepsilon}_{\mathbf{k}}$, i.e., the energy of the Gutzwiller wave function minus the ‘quasiparticle energy.’ Performing the variational procedure for the new state $|\Phi_G(\mathbf{k})\rangle$ amounts to replacing $e_G \rightarrow e_G - \tilde{\varepsilon}_{\mathbf{k}}/N$, $t_0 \rightarrow t_0 - \varepsilon_{\mathbf{k}}/N$, $n_\uparrow \rightarrow n_\uparrow - 1/N$ and $n_d \rightarrow n_d + \delta n_d/N$ where δn_d is the as-yet-unknown shift of n_d . Inserting into (22) and expanding we find

$$e_G - \frac{1}{N} \tilde{\varepsilon}_{\mathbf{k}} = \left(\eta - \frac{1}{N} \frac{\partial \eta}{\partial n_\uparrow} + \frac{1}{N} \frac{\partial \eta}{\partial n_d} \delta n_d \right) \left(t_0 - \frac{1}{N} \varepsilon_{\mathbf{k}} \right) + n_d U + \frac{1}{N} \delta n_d U.$$

The terms of zeroth order in $1/N$ cancel due to (22), and collecting the first order terms gives

$$\tilde{\varepsilon}_{\mathbf{k}} = \eta \varepsilon_{\mathbf{k}} + t_0 \frac{\partial \eta}{\partial n_{\uparrow}} - \left(\frac{\partial \eta}{\partial n_d} t_0 + U \right) \delta n_d.$$

The last term on the right-hand side vanishes because the expression in the bracket is $\frac{de_G}{dn_d}$. The second term on the right-hand side is a \mathbf{k} -independent shift that can be absorbed into a shift of E_F . The quasiparticle dispersion $\tilde{\varepsilon}_{\mathbf{k}}$ therefore follows the original dispersion but is renormalized by the factor $\eta < 1$. This is an effect known as ‘correlation narrowing.’

Next we consider the ground state momentum distribution function, i.e., the ground state expectation value $n_{\mathbf{k}} = 2 \langle c_{\mathbf{k},\uparrow}^\dagger c_{\mathbf{k},\uparrow} \rangle$. This may be obtained as the functional derivative of the ground state energy with respect to $\varepsilon_{\mathbf{k}}$, which means under a change $t_{ij} \rightarrow t_{ij} + \delta t_{ij}$ so that $\varepsilon_{\mathbf{k}} \rightarrow \varepsilon_{\mathbf{k}} + \delta \varepsilon_{\mathbf{k}}$, the change of the ground state energy is

$$e_G \rightarrow e_G + 2 \sum_{\mathbf{k}} n_{\mathbf{k}} \delta \varepsilon_{\mathbf{k}}.$$

From (22), we obtain the variation of e_G as

$$\delta e_G = 2\eta \sum_{\mathbf{k}} n_{\mathbf{k}}^{(0)} \delta \varepsilon_{\mathbf{k}} + \delta n_d \left(\frac{\partial \eta}{\partial n_d} t_0 + U \right),$$

where $n_{\mathbf{k}}^{(0)} = \Theta(E_F - \varepsilon_{\mathbf{k}})$ is the momentum distribution of the Fermi sea. Again, the second term on the right-hand side vanishes due to the extremum condition for n_d so that $n_{\mathbf{k}} = \eta n_{\mathbf{k}}^{(0)}$. This cannot be entirely correct, however, because we have the sum-rule $2 \sum_{\mathbf{k}} n_{\mathbf{k}} = N_e$, and since this is fulfilled by $n_{\mathbf{k}}^{(0)}$ and $\eta < 1$, it cannot be fulfilled for $n_{\mathbf{k}}$. The solution is that the ‘missing $n_{\mathbf{k}}$ ’ takes the form of a \mathbf{k} -independent additive constant, which then has to be $(1 - \eta)n_e/2$. In fact, for any $\varepsilon_{\mathbf{k}}$ that can be represented by hopping integrals $t_{i,j}$, one has $\sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} = 0$, so such a \mathbf{k} -independent additive constant would not contribute to the variation of e_G . The momentum distribution obtained by the Gutzwiller approximation thus has step of magnitude η at the position of the original Fermi surface. Let us now consider in more detail the case $n_\sigma = 1/2$ in which the Mott-insulator should be realized for large U/t . We find from Eq. (21)

$$\eta(n_d) = 16 n_d \left(\frac{1}{2} - n_d \right).$$

Minimizing (22) this gives

$$n_d = \frac{1}{4} - \frac{U}{32|t_0|}$$

whereby we have taken into account that $t_0 < 0$ for a half-filled band. Starting from the noninteracting value $1/4$, n_d decreases linearly with U and reaches zero at the critical value $U_c = 8|t_0|$. For $n_d = 0$ we have $\eta = 0$, so that the bandwidth of the quasiparticles becomes zero, i.e., the band mass diverges, and the step in the momentum distribution vanishes as well.

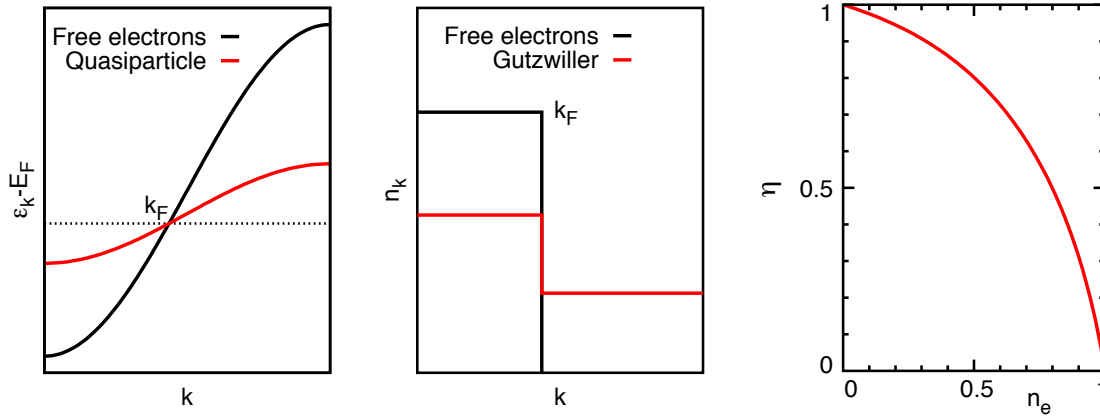


Fig. 4: Dispersion relation (left) and momentum distribution function (middle) obtained from the Gutzwiller wave function compared to the free electron case. The right part shows the dependence $\eta(n_e)$ for the two-dimensional Hubbard model with $U/t = 16$.

This is commonly interpreted as a metal-to-insulator transition as a function of increasing U , the so-called Brinkman-Rice transition [10]. Brinkman and Rice also could show that the magnetic susceptibility diverges at the transition as one would expect for a diverging effective mass.

Let us now consider the two-dimensional model with nearest-neighbor hopping $-t$. Then, $t_0 = -1.621 t$ so that the critical $U_c = 12.969 t$. Figure 4 then shows the dependence of η on n_e for $U/t = 16$, i.e., for $U > U_c$. As $n_e \rightarrow 1$ the renormalization factor $\eta \rightarrow 0$ so that both the bandwidth and the step in $n_{\mathbf{k}}$ vanish for the half-filled band. The Hubbard-I approximation and the Gutzwiller wave function thus give completely different predictions about what happens when the half-filled band case is approached by increasing the electron density for constant U/t : whereas the Hubbard-I approximation predicts a lower Hubbard band with (almost) constant bandwidth and a hole-pocket-like Fermi surface with a volume $\propto (1 - n_e)$ so that the Fermi surface vanishes at $n_e \rightarrow 1$, the Gutzwiller wave function predicts a Fermi surface with a volume equal to that obtained for free electrons, but with a vanishing bandwidth and spectral weight as $n_e \rightarrow 1$.

4 Strong coupling theory

The approximations we have considered so far – Hubbard-I and Gutzwiller wave functions – neglect the coupling of the electrons to the collective excitations of the strongly correlated electron system. This interaction with collective excitations on the one hand has a massive impact on the dispersion and lifetimes of the electrons, but on the other hand is very hard to treat. In the following, we illustrate this effect by studying the problem of a single hole in a quantum antiferromagnet, a problem for which a reasonably accurate solution is possible. As a first step, we derive an effective Hamiltonian that describes only the lower Hubbard band.

As usual, we consider the limit $U/t \gg 1$. In this limit most of the sites are occupied by at most one electron of either spin direction and double occupancies exist only as short-lived interme-

diate states. In this section we derive an effective Hamiltonian that operates in the sector of the Hilbert space with at most one electron per site but takes into account the effect of the ‘virtual’ double occupancies by suitable correction terms. This can be achieved by a technique called canonical perturbation theory. The basic assumption of this approximation is that the Hilbert space can be decomposed into ‘sectors’ that are energetically well separated. For the Hubbard model an obvious decomposition is according to the number N_d of double occupancies. Accordingly, we decompose the Hamiltonian into a part H_0 having matrix elements only between states in the same sector and a part H_1 that connects different sectors [11]:

$$\begin{aligned} H_0 &= \sum_{i,j} \sum_{\sigma} t_{i,j} \left(\hat{c}_{i,\sigma}^{\dagger} \hat{c}_{j,\sigma} + \hat{d}_{i,\sigma}^{\dagger} \hat{d}_{j,\sigma} \right) + U \sum_i n_{i,\uparrow} n_{i,\downarrow}, \\ H_1 &= \sum_{i,j} \sum_{\sigma} t_{i,j} \left(\hat{d}_{i,\sigma}^{\dagger} \hat{c}_{j,\sigma} + \hat{c}_{i,\sigma}^{\dagger} \hat{d}_{j,\sigma} \right). \end{aligned} \quad (23)$$

Obviously, H_0 does not change the number of double occupancies whereas H_1 decreases or increases this number by one. We now consider unitary transformations within the Hilbert space that act on the states $|\Psi\rangle$ and operators \hat{O} as follows

$$\begin{aligned} |\Psi'\rangle &= e^S |\Psi\rangle \\ \hat{O}' &= e^S \hat{O} e^{-S} = \hat{O} + [S, \hat{O}] + \frac{1}{2!} [S, [S, \hat{O}]] + \frac{1}{3!} [S, [S, [S, \hat{O}]]] + \dots, \end{aligned}$$

where the second line is the Baker-Campbell-Hausdorff theorem. Unitarity of e^S requires that the so-called generator S is anti-Hermitian, $S^{\dagger} = -S$ (for this reason one can often see this written as $S \rightarrow iS'$ with a Hermitian S' in the literature). We now seek a generator S such that the ‘inter sector part’ H_1 in (23) is eliminated from the transformed Hamiltonian H' . The approach obviously makes sense only if S is small so that the expansion of the transformed operators can be terminated after some low order, usually second order in S . Since

$$H' = H_0 + H_1 + [S, H_0] + [S, H_1] + \frac{1}{2!} [S, [S, H_0]] + \frac{1}{2!} [S, [S, H_1]] + \dots$$

the generator S obviously has to fulfill $H_1 + [S, H_0] = 0$ in order to eliminate H_1 . To second order in S the transformed Hamiltonian then becomes

$$\begin{aligned} H' &= H_0 + [S, H_1] + \frac{1}{2!} [S, [S, H_0]] + \frac{1}{2!} [S, [S, H_1]] + \frac{1}{3!} [S, [S, [S, H_0]]] + \dots \\ &= H_0 - \frac{1}{2} [S, H_1] + \frac{1}{2!} [S, [S, H_1]] + \dots \end{aligned}$$

For the Hubbard model an additional complication occurs: namely H_0 in (23) contains two terms of different orders of magnitude, i.e., a part of the hopping term $\propto t_{ij}$ and the Coulomb term $H_U \propto U \gg t_{ij}$. Accordingly, we demand $H_1 + [S, H_U] = 0$. Using $[H_U, \hat{d}_{i,\sigma}^{\dagger}] = U \hat{d}_{i,\sigma}^{\dagger}$ it is easy to see that S is given by

$$S = \sum_{i,j} \sum_{\sigma} \frac{t_{i,j}}{U} \left(\hat{d}_{i,\sigma}^{\dagger} \hat{c}_{j,\sigma} - \hat{c}_{i,\sigma}^{\dagger} \hat{d}_{j,\sigma} \right),$$

so that indeed $S \sim \frac{t}{U} \ll 1$ and the truncation of the expansion in powers of S is meaningful. The correction term to the Hamiltonian then becomes

$$H'_c = \frac{1}{2} [S, H_1] = \frac{1}{2} \sum_{i,j,l} \sum_{\sigma,\sigma'} \frac{t_{i,j} t_{l,m}}{U} \left[\hat{d}_{i,\sigma}^\dagger \hat{c}_{j,\sigma} - \hat{c}_{i,\sigma}^\dagger \hat{d}_{j,\sigma}, \hat{d}_{l,\sigma'}^\dagger \hat{c}_{m,\sigma'} + \hat{c}_{l,\sigma'}^\dagger \hat{d}_{m,\sigma'} \right] \quad (24)$$

Expanding the commutator on the right-hand side will produce a considerable number of terms. However, keeping in mind the goal of the present calculation, namely the derivation of an effective Hamiltonian that describes the lower Hubbard band, i.e., the sector of the Hilbert space with $N_d = 0$, most of these terms can be discarded. Namely each of the products of four operators resulting from writing out (24) contains two \hat{d} -operators, and these must appear exactly in the sequence $\hat{d}_{i,\sigma'} \hat{d}_{i,\sigma}^\dagger$. This is because neither the initial nor the final state must contain a double occupancy and the above combination is the only one that obeys this constraint (for the same reason, we neglect the commutator of the kinetic part $\propto t_{ij}$ of H_0 with S ; this would produce terms with one \hat{d} -operator). Dropping the undesired terms we obtain

$$\begin{aligned} H'_c &= -\frac{1}{2} \sum_{i,j,l} \sum_{\sigma,\sigma'} \frac{t_{i,j} t_{l,i}}{U} \hat{c}_{l,\sigma'}^\dagger \hat{d}_{i,\sigma'} \hat{d}_{i,\sigma}^\dagger \hat{c}_{j,\sigma} - \frac{1}{2} \sum_{i,j,m} \sum_{\sigma,\sigma'} \frac{t_{i,j} t_{j,m}}{U} \hat{c}_{i,\sigma}^\dagger \hat{d}_{j,\sigma} \hat{d}_{j,\sigma'}^\dagger \hat{c}_{m,\sigma'} \\ &= -\sum_{i,j,l} \sum_{\sigma,\sigma'} \frac{t_{i,l} t_{l,j}}{U} \hat{c}_{i,\sigma}^\dagger \hat{d}_{l,\sigma'} \hat{d}_{l,\sigma}^\dagger \hat{c}_{j,\sigma} \end{aligned}$$

where the second line was obtained by exchanging $(i, l) \rightarrow (l, i)$ in the first term and $(j, m) \rightarrow (l, j)$ in the second. Next we use $\hat{d}_{l,\uparrow} \hat{d}_{l,\uparrow}^\dagger = n_{l,\downarrow}$ and $\hat{d}_{l,\downarrow} \hat{d}_{l,\downarrow}^\dagger = -S_l^+$ and find

$$H' = -\sum_{i,j,l} \frac{t_{i,l} t_{l,j}}{U} \left((\hat{c}_{i,\uparrow}^\dagger n_{l,\downarrow} \hat{c}_{j,\uparrow} - \hat{c}_{i,\downarrow}^\dagger S_l^+ \hat{c}_{j,\uparrow}) + (\hat{c}_{i,\downarrow}^\dagger n_{l,\uparrow} \hat{c}_{j,\downarrow} - \hat{c}_{i,\uparrow}^\dagger S_l^- \hat{c}_{j,\downarrow}) \right)$$

In the special case where $i = j$ there is an additional factor of 2 because the ‘intermediate’ double occupancy may be formed either at i or at j . The respective terms become

$$H' = -2 \sum_{i,j} \frac{t_{i,j}^2}{U} (n_{i,\uparrow} n_{j,\downarrow} + n_{i,\downarrow} n_{j,\uparrow} - S_i^+ S_j^- - S_i^- S_j^+) = 4 \sum_{i,j} \frac{t_{i,j}^2}{U} \left(\mathbf{S}_i \cdot \mathbf{S}_j - \frac{n_i n_j}{4} \right)$$

where we used $n_\uparrow = n/2 + S^z$ and $n_\downarrow = n/2 - S^z$. With the abbreviation $J_{ij} = 4t_{i,j}^2/U$, the complete effective Hamiltonian for the lower Hubbard band thus becomes

$$\begin{aligned} H_{\text{sc}} &= \sum_{i,j} \sum_{\sigma} t_{i,j} \hat{c}_{i,\sigma}^\dagger \hat{c}_{j,\sigma} + \sum_{i,j} J_{i,j} \left(\mathbf{S}_i \cdot \mathbf{S}_j - \frac{n_i n_j}{4} \right) \\ &\quad - \sum_{i,j,l} \frac{t_{i,l} t_{l,j}}{U} \left((\hat{c}_{i,\uparrow}^\dagger n_{l,\downarrow} \hat{c}_{j,\uparrow} - \hat{c}_{i,\downarrow}^\dagger S_l^+ \hat{c}_{j,\uparrow}) + (\uparrow \leftrightarrow \downarrow) \right), \end{aligned} \quad (25)$$

The first two terms together are called the t - J model. It has been proposed by Zhang and Rice as an effective model for the CuO_2 planes of cuprate superconductors [4]. The second line is frequently referred to as the ‘conditional hopping terms’ or ‘three-site hopping terms.’

5 Spin waves

We first recall the goal of the present discussion, which is the study of the interaction between electrons and collective modes in a doped Mott insulator. In the following, we discuss the most important collective excitations of the undoped Mott insulator, namely spin waves. Spin waves are frequently discussed using the Holstein-Primakoff transformation [12] but for the case of spin 1/2 a simpler and more transparent derivation is possible, which is outlined below.

We consider the strong-coupling Hamiltonian (25) for the case of exactly one electron per site. Since no empty site is present, the terms in (25) that transport an electron from one site to another can be dropped and we are left with

$$H_{sc} = \sum_{i,j} J_{i,j} \left(\mathbf{S}_i \cdot \mathbf{S}_j - \frac{n_i n_j}{4} \right)$$

We assume that there are nonvanishing $t_{i,j}$, and hence also $J_{i,j}$, only between nearest neighbors. The terms $\propto n_i n_j$ give only an unimportant constant shift of $-J/4$ per bond, so we omit them. Finally we arrive at the Heisenberg antiferromagnet (note that by definition $J_{i,j} > 0$)

$$H_{HAF} = J \sum_{\langle i,j \rangle} \left(S_i^z S_j^z + \frac{1}{2} (S_i^+ S_j^- + S_i^- S_j^+) \right), \quad (26)$$

where $\langle i,j \rangle$ denotes a sum over pairs of nearest neighbors. If only the term $\propto S_i^z S_j^z$ were present, the ground state of (26) would be the Néel state, shown in Figure 5 (a). In this state, the square lattice is divided into two sublattices and all sites of the A -sublattice are occupied by an \uparrow -electron, whereas all sites of the B -sublattice are occupied by a \downarrow -electron. However, the Néel state is not an eigenstate of the total Hamiltonian: acting, e.g., with the term $S_i^- S_j^+$ contained in the second term of (26), the spins at the sites i and j are inverted, see Figure 5 (b), and the resulting state is orthogonal to the Néel state. To deal with these so-called quantum fluctuations we proceed as follows: we interpret the Néel state as the vacuum $|0\rangle$ and we model an inverted spin on the site i of the A sublattice as the presence of a Boson, created by a_i^\dagger . Similarly, an inverted spin on the site j of the B sublattice is modelled by the presence of a Boson created by b_j^\dagger . The state Figure 5 (b) thus would be $a_i^\dagger b_j^\dagger |0\rangle$. We use Bosons to represent the inverted spins because the spin-flip operators acting on different sites commute. Since any given spin can be inverted only once, a state like $(a_i^\dagger)^2 |0\rangle$ is meaningless. Accordingly, we have to impose the additional constraint that at most one Boson can occupy a given site. We call this the hard-core constraint. An inverted spin on either sublattice is parallel to its $z = 4$ nearest neighbors and the energy changes from $-J/4$ to $+J/4$ for each of these z bonds. Accordingly, we ascribe an ‘energy of formation’ of $zJ/2$ to each Boson. The transverse part creates pairs of inverted spins on nearest neighbors, with the matrix element being $J/2$ and we can thus write down the following Hamiltonian to describe the quantum fluctuations:

$$H_{SW} = \frac{zJ}{2} \left(\sum_{i \in A} a_i^\dagger a_i + \sum_{i \in B} b_i^\dagger b_i \right) + \frac{J}{2} \sum_{i \in A} \sum_{\mathbf{n}} \left(a_i^\dagger b_{i+\mathbf{n}}^\dagger + b_{i+\mathbf{n}} a_i \right). \quad (27)$$

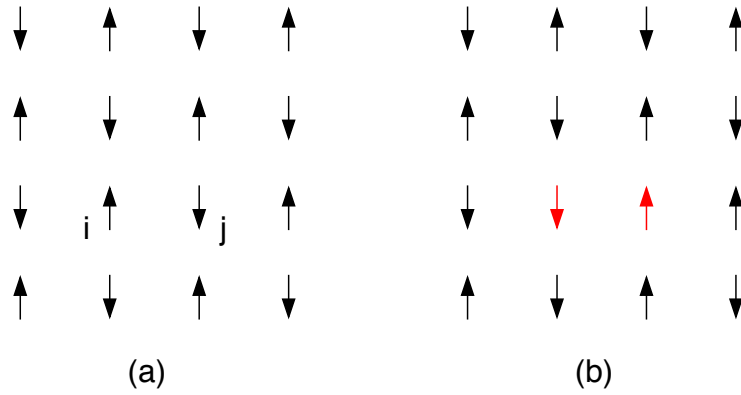


Fig. 5: The Néel state (a) is not the ground state of the Heisenberg antiferromagnet. By acting, e.g., with the term $J/2 S_i^- S_j^+$ in (26) the state (b) is generated, which is orthogonal to (a).

Here \mathbf{n} are the z vectors which connect a given site with its z nearest neighbors. The Hamiltonian (27) is in quadratic form, but we recall that the Bosons are not free particles but have to obey the hard-core constraint. However, we now simply ignore this and treat the Bosons as free particles – we will return to this issue later on. Fourier transformation of (27) gives

$$\begin{aligned}
 H_{SW} &= \frac{zJ}{2} \sum_{\mathbf{k}} \left(a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + b_{\mathbf{k}}^\dagger b_{\mathbf{k}} + \gamma_{\mathbf{k}} (a_{\mathbf{k}}^\dagger b_{-\mathbf{k}}^\dagger + b_{-\mathbf{k}} a_{\mathbf{k}}) \right), \\
 \gamma_{\mathbf{k}} &= \frac{1}{z} \sum_{\mathbf{n}} e^{i\mathbf{k}\cdot\mathbf{n}} = \frac{1}{4} \left(2 \cos(k_x) + 2 \cos(k_y) \right)
 \end{aligned} \tag{28}$$

where \mathbf{k} is a wave vector in the antiferromagnetic Brillouin zone. We can solve (28) by a Bosonic Bogoliubov transformation, i.e., we make the ansatz

$$\begin{aligned}
 \gamma_{a,\mathbf{k}}^\dagger &= u_{\mathbf{k}} a_{\mathbf{k}}^\dagger + v_{\mathbf{k}} b_{-\mathbf{k}} \\
 \gamma_{b,-\mathbf{k}}^\dagger &= u_{\mathbf{k}} b_{-\mathbf{k}}^\dagger + v_{\mathbf{k}} a_{\mathbf{k}}.
 \end{aligned} \tag{29}$$

Demanding that $[\gamma_{a,\mathbf{k}}, \gamma_{a,\mathbf{k}}^\dagger] = [\gamma_{b,\mathbf{k}}, \gamma_{b,\mathbf{k}}^\dagger] = 1$ gives the condition $u_{\mathbf{k}}^2 - v_{\mathbf{k}}^2 = 1$. An equation for $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ is obtained by demanding $[H, \gamma_{a,\mathbf{k}}^\dagger] = \omega_{\mathbf{k}} \gamma_{a,\mathbf{k}}^\dagger$. This is explained in detail in the Appendix. Using the formulae from the Appendix, we find the spin wave dispersion and the coefficients $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$:

$$\omega_{\mathbf{k}} = \frac{zJ}{2} \sqrt{1 - \gamma_{\mathbf{k}}^2}, \quad u_{\mathbf{k}} = \sqrt{\frac{1 + \nu_{\mathbf{k}}}{2\nu_{\mathbf{k}}}}, \quad v_{\mathbf{k}} = \sqrt{\frac{1 - \nu_{\mathbf{k}}}{2\nu_{\mathbf{k}}}},$$

where $\nu_{\mathbf{k}} = \sqrt{1 - \gamma_{\mathbf{k}}^2}$. For $\mathbf{k} \rightarrow 0$ we have $\gamma_{\mathbf{k}} \rightarrow 1 - (k_x^2 + k_y^2)/4$ so that $\omega_{\mathbf{k}} \rightarrow \sqrt{2}J|\mathbf{k}|$, i.e., the spin waves have a cone-shaped dispersion and reach zero frequency at $\mathbf{k} = (0, 0)$ but also at $\mathbf{k} = (\pi, \pi)$. To compute observables we revert the transformation

$$\begin{aligned}
 a_{\mathbf{k}}^\dagger &= u_{\mathbf{k}} \gamma_{a,\mathbf{k}}^\dagger - v_{\mathbf{k}} \gamma_{b,-\mathbf{k}}, \\
 b_{-\mathbf{k}} &= -v_{\mathbf{k}} \gamma_{a,\mathbf{k}}^\dagger + u_{\mathbf{k}} \gamma_{b,-\mathbf{k}}.
 \end{aligned} \tag{30}$$

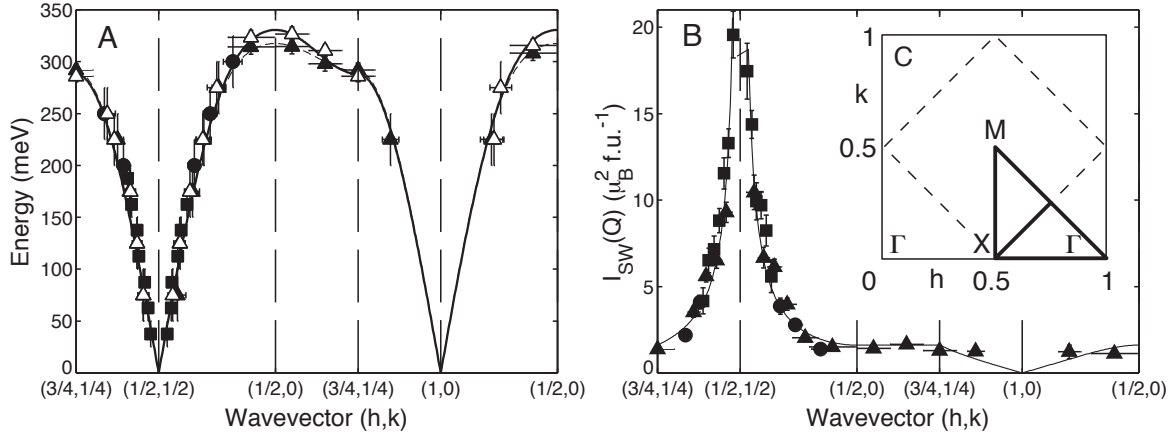


Fig. 6: Comparison of spin-wave theory to experiments on La_2CuO_4 . Panel A shows the dispersion of the frequency $\omega_{\mathbf{k}}$ of the spin waves, panel B the \mathbf{k} -dependence of the peak intensity, which is easily expressed in terms of the coefficients $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$. The dots are experimental data, the lines the prediction of spin-wave theory. Reprinted with permission from [14], Copyright 2001 by the American Physical Society.

As an application, let us consider the ground state energy. The energy of the Néel state, which was the zero of energy for the spin wave Hamiltonian (28), was $-J/4$ per bond, and there are $zN/2$ bonds in the system. To this we add the expectation value of (28) calculated in the ground state, i.e., the vacuum for the γ 's, thereby using (30). We obtain

$$\langle H_{sw} \rangle = \frac{zJ}{2} \sum_{bfk} (2v_{\mathbf{k}}^2 - 2\gamma_{\mathbf{k}}u_{\mathbf{k}}v_{\mathbf{k}}) = \frac{zJN}{4} \left[\frac{2}{N} \sum_{\mathbf{k}} \left(\sqrt{1 - \gamma_{\mathbf{k}}^2} - 1 \right) \right].$$

The expression in the square bracket can be converted to an integral over the antiferromagnetic zone and evaluated numerically. The result is that for the two-dimensional square lattice the ground state energy per bond is lowered from $-0.25 J$ to $-0.328974 J$ due to the quantum fluctuations. Monte-Carlo simulations for the two-dimensional Heisenberg antiferromagnet give a ground state energy of $-0.33 J$ per bond [13]. The spin waves can also be observed experimentally by inelastic neutron scattering. An example is shown in Figure 6. It is quite obvious that the agreement with experiment is excellent and in fact spin-wave theory is an extraordinarily successful description of many properties of magnetic Mott insulators.

To conclude this section, we return to the issue of the hard-core constraint which the a^\dagger and b^\dagger Bosons had to obey and which we simply ignored. To address this question, we calculate the density of these Bosons, i.e.,

$$n_a = \frac{2}{N} \sum_{\mathbf{k}} \langle a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \rangle = \frac{2}{N} \sum_{\mathbf{k}} v_{\mathbf{k}}^2 = \frac{2}{N} \sum_{\mathbf{k}} \frac{1 - \nu_{\mathbf{k}}}{2\nu_{\mathbf{k}}}.$$

Numerical evaluation for a 2D square lattice gives $n_a = 0.19$. The probability that two of the Bosons occupy the same site and violate the constraint therefore is $\approx n_a^2 = 0.04 \ll 1$ and our assumption of relaxing the constraint is justified *a posteriori*.

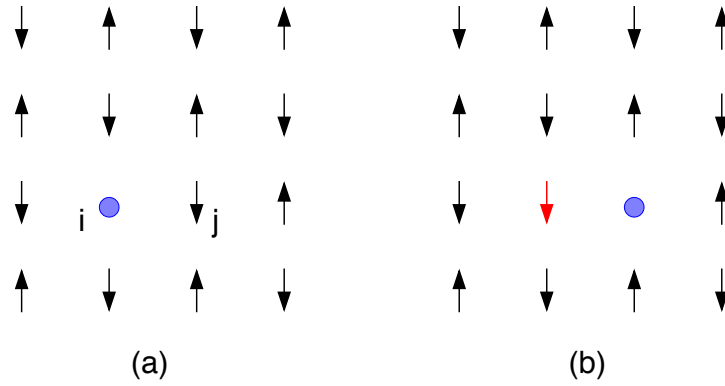


Fig. 7: A hole hopping in the Néel state displaces spins, which thus become misaligned.

6 Single hole problem

In the preceding section, we considered the strong coupling Hamiltonian (25) for $N_e = N$, where it reduces to the Heisenberg antiferromagnet, and we studied the collective excitations of the Mott insulator, i.e., the spin waves. Now we consider the case $N_e = N - 1$, i.e., a single mobile hole in an antiferromagnet, and we study the interaction between this hole and the spin waves. Again we assume that there is a nonvanishing hopping element, denoted by $-t$, only between nearest neighbors, and for simplicity we drop the three-site hopping terms in (25); that means we study the t - J model. Since we saw that the ground state still has antiferromagnetic order – although reduced – we again start from the Néel state and assume that the electron on site i , belonging to the \uparrow -sublattice, is removed – see Figure 7 (a). Then, the hopping terms in (25) become active and an electron from a neighboring site j can hop to i , resulting in the state in Figure 7 (b). Since this electron has ‘switched sublattices’, however, its spin now is opposite to that of the electron originally at site i . This inverted spin at site i then may be viewed as a spin wave as discussed in the preceding section. In other words, the propagating hole ‘radiates off’ spin waves and thus is coupled to the spin excitations. This process would be described precisely by the term $c_{j,\downarrow} S_i^-$ in the commutator relations (10) – which were neglected in the Hubbard-I approximation. In the following we will see, however, that the coupling of the single hole to the spin excitations modifies the ‘band structure’ drastically.

We continue to use the Bosons a_i^\dagger and b_j^\dagger defined in the preceding section and introduce an additional ‘particle’ namely a hole created by the fermionic creation operators $h_{a,i}^\dagger$ for $i \in A$ and $h_{b,j}^\dagger$ for $j \in B$. We have introduced two species of hole creation operators, $h_{a,i}^\dagger$ and $h_{b,j}^\dagger$, because we continue to use the antiferromagnetic Brillouin zone, which in turn necessitates the two-sublattice structure. However, both particles simply stand for a ‘hole.’ Since we are considering only a single hole the statistics of h_i^\dagger moreover is irrelevant and we might as well describe it by a Bosonic operator. The two states shown in Figure 7 would then be expressed as $h_{a,i}^\dagger |0\rangle$ and $h_{b,j}^\dagger a_i^\dagger |0\rangle$, and generalizing this we can immediately write down the Hamiltonian

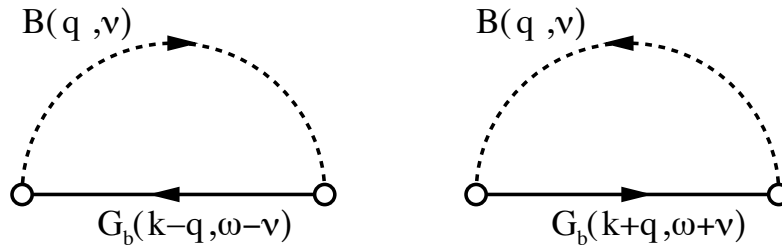


Fig. 8: Simplest Feynman diagrams for the self-energy of a hole.

for the interaction between the hole and the spin defects:

$$H_{\text{int}} = t \sum_{i \in A} \sum_{\mathbf{n}} \left(h_{b,i+\mathbf{n}}^\dagger h_{a,i} a_i^\dagger + H.c. \right) + t \sum_{j \in B} \sum_{\mathbf{n}} \left(h_{a,j+\mathbf{n}}^\dagger h_{b,j} b_j^\dagger + H.c. \right).$$

Note that the hopping integral for a hole is $+t$. Upon Fourier transformation this becomes

$$H_{\text{int}} = \sqrt{\frac{2}{N}} \sum_{\mathbf{k}, \mathbf{q}} \left(\left(\varepsilon_{\mathbf{k}-\mathbf{q}} h_{b,\mathbf{k}-\mathbf{q}}^\dagger h_{a,\mathbf{k}} a_{\mathbf{q}}^\dagger + H.c. \right) + (a \leftrightarrow b) \right).$$

where $\varepsilon_{\mathbf{k}} = 2t(\cos(k_x) + \cos(k_y))$. Next, we use the inverse Bogoliubov transformation (19) to replace a^\dagger and b^\dagger by γ_i^\dagger and add the spin-wave Hamiltonian to describe the dynamics of the γ_i^\dagger . The total Hamiltonian for the coupled hole-magnon system then reads

$$H_{\text{tot}} = \sqrt{\frac{2}{N}} \sum_{\mathbf{k}, \mathbf{q}} \left(\left(M(\mathbf{k}, \mathbf{q}) h_{b,\mathbf{k}-\mathbf{q}}^\dagger h_{a,\mathbf{k}} \gamma_{a,\mathbf{q}}^\dagger + H.c. \right) + (a \leftrightarrow b) \right) + \sum_{\mathbf{q}} \omega_{\mathbf{q}} \left(\gamma_{a,\mathbf{q}}^\dagger \gamma_{a,\mathbf{q}} + \gamma_{b,\mathbf{q}}^\dagger \gamma_{b,\mathbf{q}} \right), \quad (31)$$

where $M(\mathbf{k}, \mathbf{q}) = \varepsilon_{\mathbf{k}-\mathbf{q}} u_{\mathbf{q}} - \varepsilon_{\mathbf{k}} v_{\mathbf{q}}$. Obviously the first term describes how a hole with momentum \mathbf{k} ‘radiates off’ a spin wave with momentum \mathbf{q} thereby changing its own momentum to $\mathbf{k} - \mathbf{q}$. The Hamiltonian (31) no longer is a quadratic form and requires more sophisticated techniques for its solution. More precisely, we will use the so-called self-consistent Born approximation (SCB) to derive an equation for the self-energy of the hole. We again define the time-ordered zero-temperature Green’s functions [6]

$$\begin{aligned} G_{\alpha}(\mathbf{k}, t) &= -i \langle T h_{\alpha,\mathbf{k}}(t) h_{\mathbf{k}}(0)^\dagger \rangle \\ B_{\alpha}(\mathbf{q}, t) &= -i \langle T \gamma_{\alpha,\mathbf{q}}(t) \gamma_{\mathbf{q}}(0)^\dagger \rangle, \end{aligned}$$

where $\langle \dots \rangle$ denotes the expectation value in the empty state containing neither spin waves nor a hole. The simplest Feynman diagrams for the self-energy of the hole are shown in Fig. 8, and the corresponding expression is [6]

$$\Sigma_a(\mathbf{k}, \omega) = \frac{i}{2\pi} \frac{2}{N} \sum_{\mathbf{q}} \int d\nu M^2(\mathbf{k}, \mathbf{q}) [B_a(\mathbf{q}, \nu) G_b(\mathbf{k}-\mathbf{q}, \omega-\nu) + B_a(\mathbf{q}, \nu) G_b(\mathbf{k}+\mathbf{q}, \omega+\nu)]. \quad (32)$$

Since a single hole in a macroscopic system is not expected to have any effect on the spin wave spectrum we may use the noninteracting Green's function for the spin waves, which is given by

$$B_\alpha^{(0)}(\mathbf{q}, t) = -ie^{-i\omega_{\mathbf{q}}t} \left(\Theta(t) \langle \gamma_{\alpha, \mathbf{q}} \gamma_{\alpha, \mathbf{q}}^\dagger \rangle + \Theta(-t) \langle \gamma_{\alpha, \mathbf{q}}^\dagger \gamma_{\alpha, \mathbf{q}} \rangle \right) = -i\Theta(t) e^{-i\omega_{\mathbf{q}}t},$$

with Fourier transform $B_\alpha^{(0)}(\mathbf{q}, \omega) = (\omega - \omega_{\mathbf{q}} + i0^+)^{-1}$. Next, we note that (with $\alpha \in a, b$)

$$0 = \langle h_{\alpha, \mathbf{k}}^\dagger h_{\alpha, \mathbf{k}} \rangle = -iG_\alpha(\mathbf{k}, t = 0^-) = -\frac{i}{2\pi} \int_{-\infty}^{\infty} d\omega e^{i\omega 0^+} G_\alpha(\mathbf{k}, \omega).$$

Due to the presence of the factor $e^{i\omega 0^+}$ we can close the integration path around the upper ω half-plane. Since the integral must vanish, we can conclude that the hole Green's function is analytical in the upper half-plane (this also would have followed from the fact that for the special case of an empty ground state the time-ordered Green's function is equal to the retarded one, and the retarded Green's function is analytical in the upper half-plane). We now insert $B_\alpha^{(0)}(\mathbf{q}, \omega)$ into the expression for the self-energy, (32), and obtain

$$\Sigma_a(\mathbf{k}, \omega) = \frac{i}{2\pi} \frac{2}{N} \sum_{\mathbf{q}} \int d\nu M^2(\mathbf{k}, \mathbf{q}) \left[\frac{G_b(\mathbf{k} - \mathbf{q}, \omega - \nu)}{\nu - \omega_{\mathbf{q}} + i0^+} + \frac{G_b(\mathbf{k} + \mathbf{q}, \omega + \nu)}{\nu - \omega_{\mathbf{q}} + i0^+} \right].$$

Since $G_b \propto |\omega|^{-1}$ for large $|\omega|$ the integrand behaves like $|\nu|^{-2}$ and we can close the integration contour around either the upper or the lower ν -half-plane. For the first term we choose the lower half-plane, where the integrand has a pole at $\nu = \omega_{\mathbf{q}} - i0^+$. If ν has a negative imaginary part and ω is real, the frequency argument $\omega - \nu$ of G_b in this term has a positive imaginary part, and since G_b is regular in the upper half-plane there are no singularities from the factor of G_b . In the second term, we close the integration contour around the upper half-plane. For ν in the upper half-plane, both $B(\nu)$ and $G(\omega + \nu)$ are regular. The second term thus vanishes and the result is

$$\Sigma_a(\mathbf{k}, \omega) = \frac{2}{N} \sum_{\mathbf{q}} M^2(\mathbf{k}, \mathbf{q}) G_b(\mathbf{k} - \mathbf{q}, \omega - \omega_{\mathbf{q}})$$

Since we expect that $G_a = G_b = G$ and hence $\Sigma_a = \Sigma_b = \Sigma$ and $G^{-1}(\mathbf{k}, \omega) = \omega - \Sigma(\mathbf{k}, \omega)$ we finally obtain the self-consistency equation for $\Sigma(\mathbf{k}, \omega)$ [15]

$$\Sigma(\mathbf{k}, \omega) = \frac{2}{N} \sum_{\mathbf{q}} \frac{M^2(\mathbf{k}, \mathbf{q})}{\omega - \omega_{\mathbf{q}} - \Sigma(\mathbf{k} - \mathbf{q}, \omega - \omega_{\mathbf{q}})}.$$

This can be solved numerically on a finite \mathbf{k} -mesh and ω -grid [15], whence the hole spectral density $A(\mathbf{k}, \omega) = -\frac{1}{\pi} \text{Im} G(\mathbf{k}, \omega + i0^+)$ of the hole Green's function can be calculated. This can in principle be compared to the ARPES spectra of a Mott-insulator. Figure 9 shows the result of such a calculation, taken from Ref. [15]. The spectrum is spread out over an energy range of several t , and a considerable amount of spectral weight is distributed over this range. The spread-out weight is frequently referred to as 'incoherent continua.' At the bottom of the spectrum, around $\omega \approx -2t$, there is a relatively intense isolated peak, often referred to as the 'quasiparticle peak.' The corresponding eigenstate is the ground state of the hole for the respective momentum. Figure 9 also shows $A(\mathbf{k}, \omega)$ obtained by Lanczos diagonalization of a

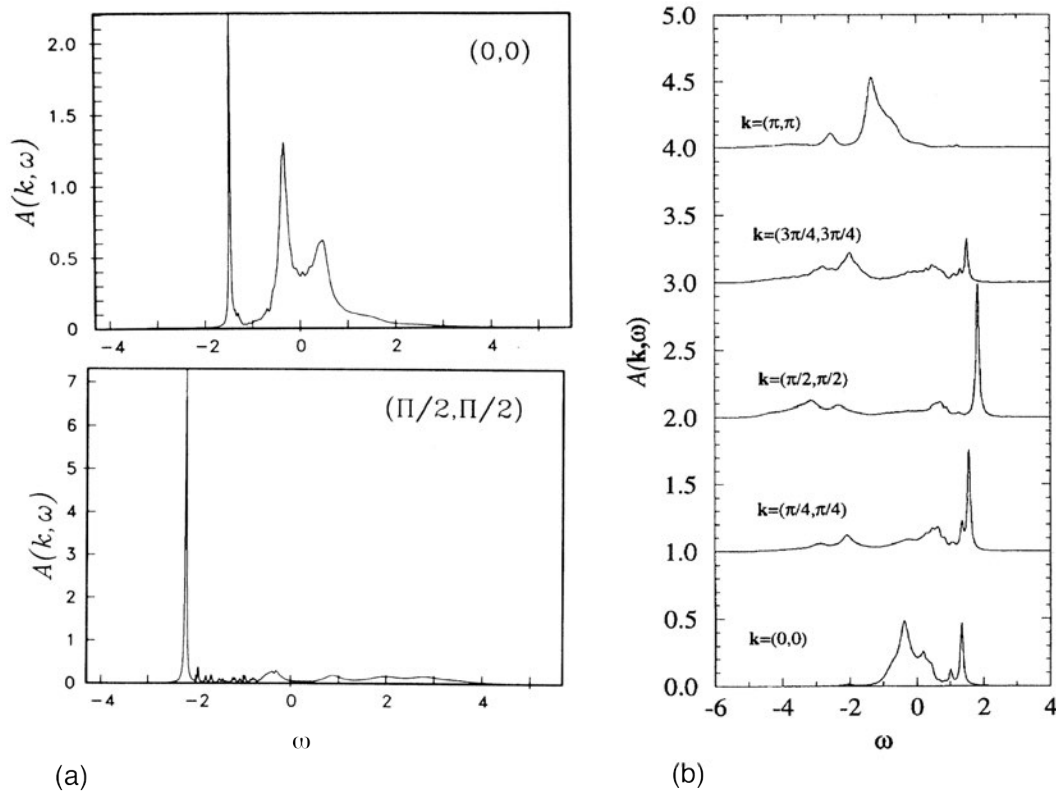


Fig. 9: (a) Single-particle spectral function $A(\mathbf{k}, \omega)$ obtained from the SCB for the t - J model [15] with $J/t = 0.3$ for $\mathbf{k} = (0, 0)$ and $\mathbf{k} = (\pi/2, \pi/2)$. (b) Single particle spectral function of the t - J model obtained by exact diagonalization of a 32-site cluster [16]. Note that this Figure has an inverted ω -axis and actually shows $A(\mathbf{k}, -\omega)$. Repinted with permission from [15], Copyright 1991 by the American Physical Society and from [16] Copyright 1995 by the American Physical Society

32-site cluster [16]. Note that the Lanczos spectra are ‘upside down’ as compared to the SCB spectra, which means the quasiparticle peak is at the top, rather than the bottom of the spectra. The Lanczos spectra also show the quasiparticle peak and the incoherent continua, and in fact even the \mathbf{k} -dependence of the incoherent continua shows some similarity with the results of the SCB-approximation. Figure 10 compares the dispersion of the quasiparticle peak as obtained by Lanczos (dots) and SCB (line). Obviously the agreement is very good. The whole spectrum – quasiparticle peak plus incoherent continua – would now replace the lower Hubbard band in the Mott-insulator, whereby the quasiparticle peak would form the top of the photoemission spectrum, as in the Lanczos spectra. Comparing with the prediction of the Hubbard-I approximation, the width of the quasiparticle band is reduced drastically. Moreover one can see from Figure 10 that the band maximum is shifted from (π, π) to $(\pi/2, \pi/2)$. The hole pocket predicted by the Hubbard-I approximation therefore should form around this momentum.

The self-consistent Born approximation in fact not only agrees very well with numerical spectra but also with experiments on insulating cuprates. After the publication of ARPES spectra for the antiferromagnetic Mott-insulator $\text{Sr}_2\text{CuO}_2\text{Cl}_2$ by Wells *et al.* [17], considerable theoretical effort was put into reproducing these spectra, and after adding additional hopping integrals

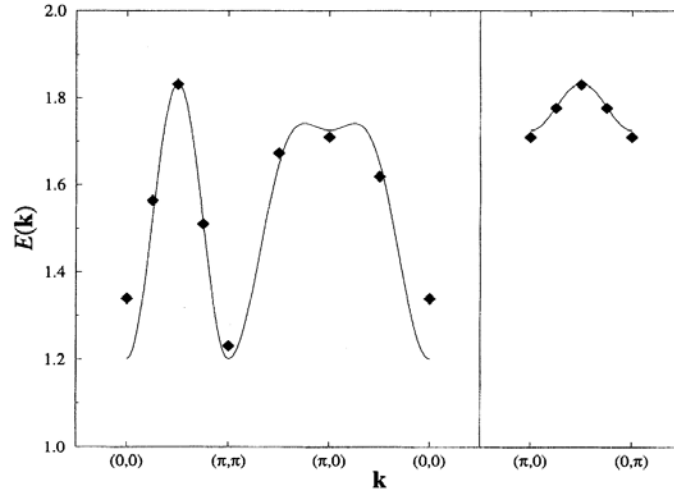


Fig. 10: Dispersion of the quasiparticle peak from the SCB (line) and by Lanczos diagonalization of a 32-site cluster. Reprinted with permission from [16], Copyright 1995 by the American Physical Society.

between 2nd and 3rd nearest neighbors, good agreement with experiment could be achieved using the self-consistent Born approximation; see for example Ref. [18].

The self-consistent Born approximation thus is quite successful but somewhat technical. For this reason we now give a simplified discussion of the states that form the quasiparticle band. We decompose the t - J Hamiltonian as $H = H_0 + H_1$ whereby

$$\begin{aligned}
 H_0 &= -t \sum_{\langle i,j \rangle} \sum_{\sigma} \left(\hat{c}_{i,\sigma}^{\dagger} \hat{c}_{j,\sigma} + H.c \right) + J \sum_{\langle i,j \rangle} \left(S_i^z S_j^z - \frac{n_i n_j}{4} \right), \\
 H_1 &= \frac{J}{2} \sum_{\langle i,j \rangle} (S_i^+ S_j^- + H.c.).
 \end{aligned}
 \tag{33}$$

In the absence of any hole, the ground state of H_0 is again the Néel state with energy NJ , and we choose this as the zero of energy. We assume that a hole is created at some site i , which raises the energy by $zJ/2$. By the action of the hopping term, the hole then starts to propagate. As was discussed above, however, in every step the hole shifts one electron to the opposite sublattice where its spin is opposite to the Néel order; see Figure 11. The hole thus

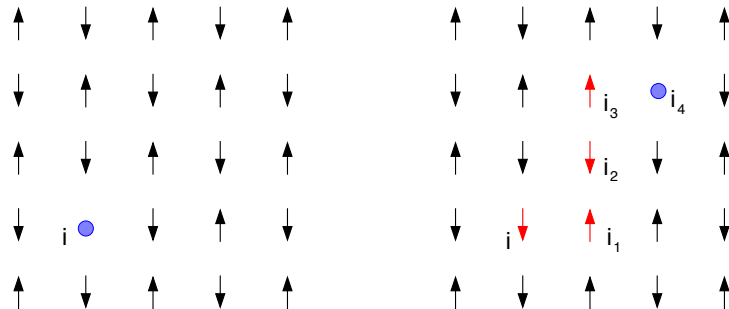


Fig. 11: A hole hopping in the Néel state creates a ‘string’ of misaligned spins.

leaves behind a trace of misaligned spins so that the magnetic energy increases roughly linearly with the distance travelled by the hole. We call such a state that is created by a hole hopping in the Néel state a ‘string’ and denote it by $|i_0, i_1, \dots, i_\nu\rangle$, where $i_1, i_2, i_{\nu-1}$ are the sites which the hole has visited, and the hole itself is at the site i_ν . In the first step the hole generates z different string states, whereas any subsequent hop from any string of length ν generates $z - 1$ strings of length $\nu + 1$. The number of strings of length ν thus is $n_\nu = z(z - 1)^{\nu-1}$ for $\nu \geq 1$, while $n_0 = 1$. Since each displaced spin is parallel to $z - 2$ neighbors – compare Figure 11 – the magnetic energy increases by $J(z - 2)/2$. The only exception is the first hop away from i where the energy increases by $J(z - 1)/2$. Accordingly, the exchange energy for a string of length $\nu > 0$ is

$$I_\nu = \frac{(z + 1)J}{2} + \nu \frac{(z - 2)J}{2}, \quad (34)$$

where $I_0 = zJ/2$. It may happen that the path that the hole has taken is folded or self-intersecting, and in this case (34) is not correct. However, it will be correct for *most* possible paths of the hole: in particular, it is correct for $\nu \leq 2$, so we will use this expression. Neglecting the possibility of self-intersection or folding of the string is an approximation known as the Bethe lattice. Since the magnetic energy increases linearly with the number of hops the hole has taken, we conclude that the hole is self-trapped. To describe the resulting localized state we make the *ansatz*

$$|\Psi_i\rangle = \sum_{\nu=0}^{\infty} \alpha_\nu \sum_{i_1, i_2, \dots, i_\nu} |i, i_1, i_2, \dots, i_\nu\rangle \quad (35)$$

where it is understood that the second sum runs only over those ν -tuples of sites that correspond to a true string starting at i . Since we assume that the magnetic energy is the same for all strings of length ν , the coefficient α_ν also depends only on the length of the string. The α_ν in (35) are to be determined by minimizing the energy. The norm and magnetic energy are

$$\langle \Psi_i | \Psi_i \rangle = \sum_{\nu=0}^{\infty} n_\nu |\alpha_\nu|^2, \quad (36)$$

$$\langle \Psi_i | H_I | \Psi_i \rangle = \sum_{\nu=0}^{\infty} n_\nu I_\nu |\alpha_\nu|^2. \quad (37)$$

To obtain the expectation value of the kinetic energy, we consider a string of length $\nu \geq 1$ with coefficient α_ν . By acting with the hopping term, we obtain $z - 1$ strings of length $\nu + 1$ with coefficient $\alpha_{\nu+1}$ and one string of length $\nu - 1$ with coefficient $\alpha_{\nu-1}$. For $\nu = 0$, we obtain z strings of length 1. In this way, we find

$$\langle \Psi_i | H_t | \Psi_i \rangle = t \left(z \alpha_0 \alpha_1 + \sum_{\nu=1}^{\infty} n_\nu \alpha_\nu (\alpha_{\nu-1} + (z - 1) \alpha_{\nu+1}) \right) = 2t \sum_{\nu=0}^{\infty} n_{\nu+1} \alpha_\nu \alpha_{\nu+1}. \quad (38)$$

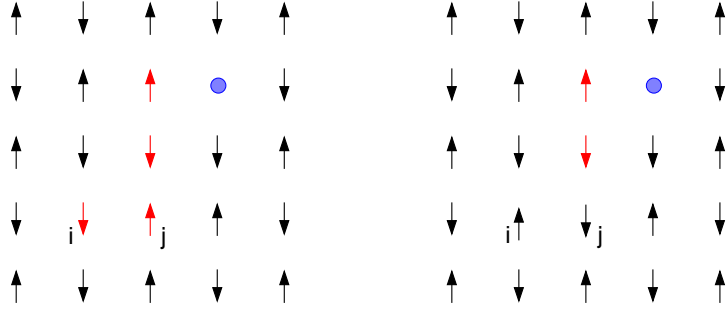


Fig. 12: By acting with the term $J/2 S_i^+ S_j^-$, the first two defects created by the hole can be ‘healed’ and the starting point of the string be shifted to a neighbor.

Then we demand

$$0 = \frac{\partial E_{\text{loc}}}{\partial \alpha_\nu} = \frac{\partial}{\partial \alpha_\nu} \frac{\langle \Psi_i | H_0 | \Psi_i \rangle}{\langle \Psi_i | \Psi_i \rangle} = \frac{1}{\langle \Psi_i | \Psi_i \rangle^2} \left[\frac{\partial \langle \Psi_i | H_0 | \Psi_i \rangle}{\partial \alpha_\nu} \langle \Psi_i | \Psi_i \rangle - \langle \Psi_i | H_0 | \Psi_i \rangle \frac{\partial \langle \Psi_i | \Psi_i \rangle}{\partial \alpha_\nu} \right]$$

$$= \frac{1}{\langle \Psi_i | \Psi_i \rangle} \left[\frac{\partial \langle \Psi_i | H_0 | \Psi_i \rangle}{\partial \alpha_\nu} - E_{\text{loc}} \frac{\partial \langle \Psi_i | \Psi_i \rangle}{\partial \alpha_\nu} \right]$$

whence

$$\frac{\partial \langle \Psi_i | H_0 | \Psi_i \rangle}{\partial \alpha_\nu} - E_{\text{loc}} \frac{\partial \langle \Psi_i | \Psi_i \rangle}{\partial \alpha_\nu} = 0.$$

Inserting equations (37), (38), and (36) we obtain [19]

$$I_\nu \alpha_\nu + t (\alpha_{\nu-1} + (z-1) \alpha_{\nu+1}) = E_{\text{loc}} \alpha_\nu \quad \text{for } \nu \geq 1$$

$$I_0 \alpha_0 + z t \alpha_1 = E_{\text{loc}} \alpha_0.$$

This is a non-Hermitian eigenvalue problem, but by introducing $\beta_\nu = \sqrt{n_\nu} \alpha_\nu$ it can be made Hermitian. So far it seems that the hole in the Néel state is localized. However, it is easy to see that the part H_1 in (33) that was neglected so far can assist the trapped hole in escaping from the string potential, see Figure 12. Namely, by acting on the first two sites of a string, the spins that were inverted by the hole are inverted a second time and thus fit with the Néel order again:

$$H_1 |i, i_1, i_2, i_3, \dots, i_\nu\rangle = \frac{J}{2} |i_2, i_3, \dots, i_\nu\rangle$$

The initial site of the string thus is shifted to a (2, 0)- or (1, 1)-like neighbor while simultaneously the length ν is increased or decreased by two. This suggests that there is a nonvanishing matrix element of H_1 between states $|\Psi_i\rangle$ and $|\Psi_j\rangle$, and it is straightforward to see that

$$\langle \Phi_i | H_1 | \Phi_{i+2\hat{x}} \rangle = J \sum_{\nu=0}^{\infty} (z-1)^\nu \alpha_\nu \alpha_{\nu+2} = J \cdot m,$$

$$\langle \Phi_i | H_1 | \Phi_{i+\hat{x}+\hat{y}} \rangle = 2J \cdot m.$$

The factor of 2 in the second matrix element is due to the fact that a string to a (1, 1)-like neighbor can pass either through (1, 0) or (0, 1) and the contributions from these two different paths are additive.

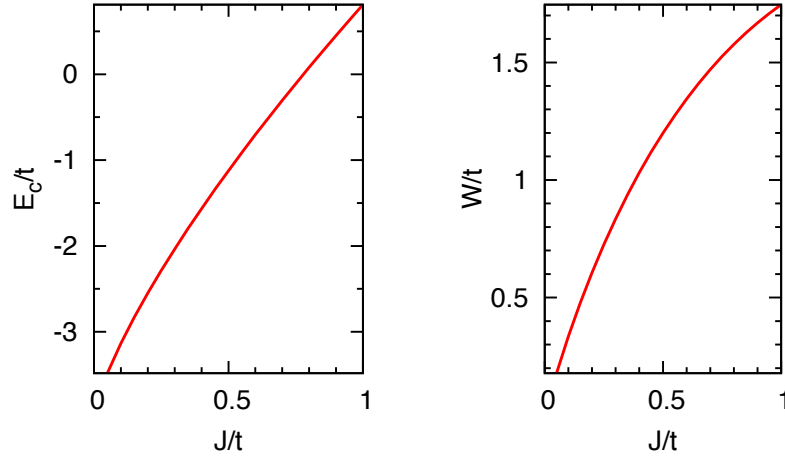


Fig. 13: Center of gravity E_c and bandwidth W of the quasiparticle band as obtained from the variational calculation in the string basis.

If the full Hamiltonian $H_0 + H_1$ is taken into account, the hole therefore can propagate through the entire lattice, and we describe this by the Bloch state

$$|\Psi_{\mathbf{k}}\rangle = \sqrt{\frac{2}{N}} \sum_{i \in A} e^{-i\mathbf{k} \cdot \mathbf{R}_j} |\Phi_i\rangle. \quad (39)$$

Since the matrix element for a $(1, 1)$ -like neighbor is twice that for a $(2, 0)$ -like neighbor we obtain the dispersion

$$\begin{aligned} E(\mathbf{k}) &= E_{\text{loc}} + 2Jm \cdot 4 \cos(k_x) \cos(k_y) + Jm \cdot 2(\cos(2k_x) + \cos(2k_y)) \\ &= E_{\text{loc}} + 4Jm (\cos(k_x) + \cos(k_y))^2 - 4Jm. \end{aligned}$$

This has a degenerate minimum along the line $k_y = \pi - k_x$ (i.e. $(\pi, 0) \rightarrow (0, \pi)$) and the symmetry equivalent lines, the maxima are at $(0, 0)$ and (π, π) . This is very similar to the dispersion in Figure 10 (note that the dispersion from SCB is upside down in this Figure!). The center of the band is $E_c = E_{\text{loc}} + 4Jm$, where E_{loc} is relatively large, of order $2-3t$. The width of the band is $W = 16mJ$. Both quantities are shown in Figure 13. At $J/t = 0.3$ we have $E_c = -2t$ and $W = 0.83t$ – in fact Figure 9 (a) shows that the center of the band is around $-2t$ and Figure 10 shows that the bandwidth is $\approx 0.7t$. The trial wave function (39) thus describes the single hole ground state that gives rise to the quasiparticle peak quite well.

The above derivation shows that due to the interaction with the spin waves the motion of the hole in an antiferromagnet is very different from free propagation: the hole executes a rapid ‘zig-zag’ motion around a given site i with hopping integral t under the influence of the string potential thereby forming the quasi-localized states $|\Psi_i\rangle$. The spin-flip part of the Hamiltonian then enables tunneling between such quasi-localized states with the much smaller energy scale $\propto J$, which accordingly determines the bandwidth of the quasiparticle.

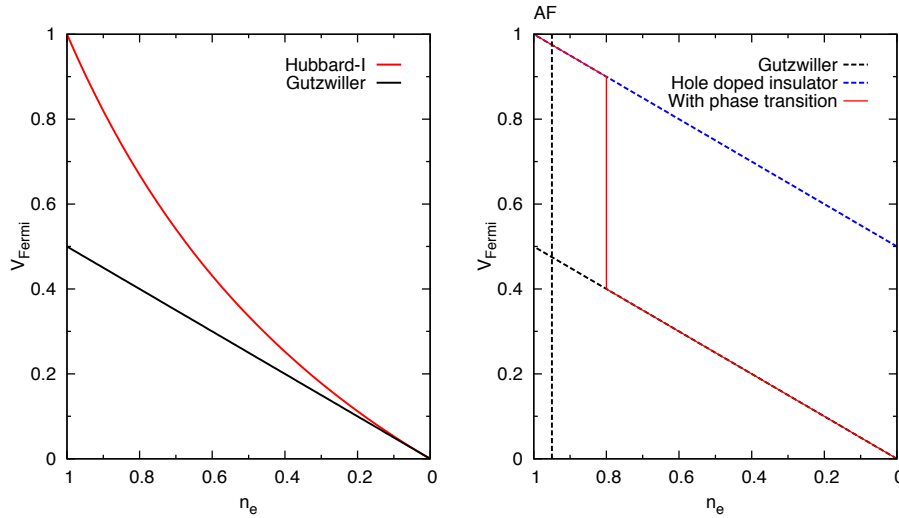


Fig. 14: Left: Fermi surface volume as obtained from the Hubbard-I approximation and the Gutzwiller wave function. Right: a possible compromise with a phase transition between two phases with different $V_{\text{Fermi}}(n_e)$.

7 Summary and discussion

Since a generally accepted theory of the lightly doped 2-dimensional Hubbard model does not exist so far, so that also the cuprate superconductors are not really understood as yet, maybe the best one can do at present is to outline the problems that would have to be solved.

The first one of these is the Fermi surface close to half-filling. As we have seen, the Hubbard-I approximation and the Gutzwiller wave function predict completely different behavior close to half-filling: a hole-like Fermi surface with a volume $\propto n_h = 1 - n_e$ in the lower Hubbard-band whose volume tends to zero as $n_e \rightarrow 1$ versus a free-electron-like Fermi surface with volume $n_e/2$ formed by a band whose mass diverges as $n_e \rightarrow 1$. It should be noted that for n_e very close to 1 antiferromagnetic order sets in, but superconductivity is observed at values of n_e where there is no more antiferromagnetism. The more relevant question therefore is, which Fermi surface is realized up to the onset of antiferromagnetism. A possible compromise between the two approximations could be as shown in Figure 14: near $n_e = 1$ but outside the antiferromagnetic doping range there are hole pockets with a volume that is strictly proportional to the hole number $n_h = 1 - n_e$, i.e., the doped Hubbard-band, and then at some critical density a phase transition occurs to a phase where the Fermi surface volume is $n_e/2$. This might be one scenario which the Hubbard-I approximation ‘tries to reproduce.’ Viewed this way, the ‘pseudogap phase’ of cuprate superconductors could be identified with the hole-pocket phase and the quantum critical point, which is surrounded by the superconducting dome, corresponds to the transition to the free-electron-like Fermi surface. A theory which is supposed to describe this transition first of all must reproduce the two Hubbard bands – otherwise the hole-doped lower Hubbard band cannot be reproduced. Next, the two different phases and the transition between them would have to be described, which is a considerable problem because there is no obvious order parameter for the transition between a paramagnetic small Fermi surface and a paramagnetic large Fermi surface.

The second major problem is that – as we have seen above – the particles in the Hubbard bands are heavily renormalized due to their interaction with, mainly, spin excitations. The quasi-particle band produced by the self-consistent Born approximation or the string wave function describes a very complicated state in which the hole is heavily dressed by spin excitations. It is this strong dressing with spin waves which leads for example to the shift of the band maximum from (π, π) , where it is predicted by the Hubbard-I approximation, to $(\pi/2, \pi/2)$, where the self-consistent Born approximation puts it. It is likely that similar effects will also occur in the lightly doped Mott insulator, and clearly it is rather hopeless to try and reproduce these heavily renormalized quasiparticles by any ‘simple’ theory, such as a mean-field theory. From the calculation with the string states it appears that long-range antiferromagnetic order is not absolutely necessary – rather a similar formation of strings may already occur in a state that has antiferromagnetic correlations with a range of a few lattice spacings. However, for the situation without antiferromagnetic order but only short ranged antiferromagnetic correlations, there is no state analogous to the Néel state that would allow for a similarly simple development as in the single-hole theory.

A The bosonic Bogoliubov transformation

We consider the following Hamiltonian, whereby a^\dagger and b^\dagger are bosonic operators that satisfy $[a, a^\dagger] = [b, b^\dagger] = 1$:

$$H = \varepsilon (a^\dagger a + b^\dagger b) + V a^\dagger b^\dagger + V^* b a, \quad (40)$$

To solve this we make the ansatz

$$\gamma_1^\dagger = u a^\dagger + v b \quad (41)$$

$$\gamma_2^\dagger = u b^\dagger + v a \quad (42)$$

In order for the γ 's to fulfil the bosonic commutator relations $[\gamma_i, \gamma_i^\dagger] = 1$, we must have $|u|^2 - |v|^2 = 1$. The relation $[\gamma_1^\dagger, \gamma_2^\dagger] = 0$ on the other hand is fulfilled automatically. Moreover, we require that when expressed in terms of the γ 's the Hamiltonian should take the simple form

$$H = \sum_{i=1}^2 \omega \gamma_i^\dagger \gamma_i + \text{const}$$

which means that the γ 's obey

$$[H, \gamma_i^\dagger] = \omega \gamma_i^\dagger \quad (43)$$

We now insert (40) and γ_1^\dagger from (42) into (43), use the commutator relations for a^\dagger and b^\dagger , and equate the coefficients of a^\dagger and b on both sides of (43). This gives the non-Hermitian eigenvalue problem

$$\begin{pmatrix} \varepsilon & -V \\ V^* & -\varepsilon \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \omega \begin{pmatrix} u \\ v \end{pmatrix}$$

which is easily solved to give

$$\begin{aligned} \omega &= \sqrt{\varepsilon^2 - |V|^2} \\ u &= \frac{V}{\sqrt{2\omega(\varepsilon - \omega)}} \\ v &= \sqrt{\frac{\varepsilon - \omega}{2\omega}} \end{aligned}$$

The eigenvalue problem has the property that if (u, v) is an eigenvector with eigenvalue ω then (v^*, u^*) is an eigenvector with eigenvalue $-\omega$. This guarantees that automatically $[H, \gamma_1] = -\omega \gamma_1$. Since H is symmetric under $a \leftrightarrow b$ the requirement $[H, \gamma_2^\dagger] = \omega \gamma_2^\dagger$ gives the same system of equations. To apply this to the spin wave Hamiltonian (28), we obviously need to set $\varepsilon = 1$ and $V = \gamma_{\mathbf{k}}$.

References

- [1] J. Hubbard, Proc. Roy. Soc. London A **276** 238 (1963)
- [2] A. Kanamori, Progr. Theor. Phys. **30**, 275 (1963)
- [3] M.C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963)
- [4] F.C. Zhang and T.M. Rice, Phys. Rev. B **37**, 3759 (1988)
- [5] See the lecture by A. Mielke in this School.
- [6] A.L. Fetter and J.D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, San Francisco, 1971)
- [7] C.T. Chen, F. Sette, Y. Ma, M.S. Hybertsen, E.B. Stechel, W.M.C. Foulkes, M. Schuler, S.-W. Cheong, A.S. Cooper, L.W. Rupp, B. Batlogg, Y.L. Soo, Z.H. Ming, A. Krol, and Y.H. Kao, Phys. Rev. Lett. **66**, **104** (1991)
- [8] C. Gröber, R. Eder, and W. Hanke, Phys. Rev. B **62** 4336 (2000)
- [9] T. Ogawa, K. Kanda, and T. Matsubara, Progr. Theor. Phys. **53**, 614 (1975)
- [10] W.F. Brinkman and T.M. Rice, Phys. Rev. B **2**, 4302 (1970)
- [11] K.A. Chao, J. Spalek, and A.M. Oles, J. Phys. C **10**, 1271 (1972)
- [12] P.W. Anderson, Phys. Rev. **86**, 694 (1952)
- [13] J.D. Reger and A.P. Young, Phys. Rev. B **37**, 5978(R) (1988)
- [14] R. Coldea, S.M. Hayden, G. Aeppli, T.G. Perring, C.D. Frost, T.E. Mason, S.-W. Cheong, and Z. Fisk, Phys. Rev. Lett. **86**, 5377 (2001)
- [15] G. Martinez and P. Horsch, Phys. Rev. B **44**, 317 (1991)
- [16] P.W. Leung and R.J. Gooding Phys. Rev. B **52**, R15711(R) (1995)
- [17] B.O. Wells, Z.-X. Shen, A. Matsuura, D.M. King, M.A. Kastner, M. Greven, and R.J. Birgeneau, Phys. Rev. Lett. **74**, 964 (1995)
- [18] O.P. Sushkov, G.A. Sawatzky, R. Eder and H. Eskes, Phys. Rev. B. **56**, 11769 (1997)
- [19] R. Eder und K.W. Becker, Z. Phys. B **78**, 219 (1990)