which is the lower edge of a continuum of excitations whose upper edge is bounded by

$$\omega(\mathbf{q}) = \pi \mathcal{J} \cos(q/2). \tag{8.544}$$

The continuum of excitations develops because spinons are always created in pairs, and therefore the momentum of the two spinons can be distributed in a continuum of different ways. Neutron scattering experiments on quasi-one-dimensional materials like $KCuF_3$ have corroborated the picture outlined here (see, e.g. Tennant *et al.*, 1995).

In dimensions higher than one, separating a flipped spin into a pair of kinks, or a magnon into a pair of spinons, costs energy, which thus confines spinons in dimensions $d \geq 2$.

The next section we constructs the Hubbard model from first principles and then show how the Heisenberg model can be obtained from the Hubbard model for half-filling and in the limit of strong on-site repulsion.

8.7 Hubbard model

The Hubbard model presents one of the simplest ways to obtain an understanding of the mechanisms through which interactions between electrons in a solid can give rise to insulating versus conducting, magnetic, and even novel superconducting behaviour. The preceding sections of this chapter more or less neglected these interaction or correlation effects between the electrons in a solid, or treated them summarily in a mean-field or quasiparticle approach (cf. sections 8.2 to 8.5). While the Hubbard model was first discussed in quantum chemistry in the early 1950s (Pariser and Parr, 1953; Pople, 1953), it was introduced in its modern form and used to investigate condensed matter problems in the 1960s independently by Gutzwiller (1963), Hubbard (1963), and Kanamori (1963). Their proposals of the model were motivated by different physical problem situations. Gutzwiller used the model to study the transition between metallic and insulating phases of solids, Hubbard's research focused on electron correlations in narrow energy bands of transition metals, and Kanamori's objective was the study of itinerant ferromagnetism.

Despite its simplicity, the Hubbard model has proven to be a versatile model with many applications to condensed matter systems. For a careful exposition of the materials and phenomena discussed on the basis of the Hubbard model, especially the high-temperature superconductors, see Fazekas (1999). However, in a recent editorial 2013 on the occasion of the fiftieth anniversary of the work of Gutzwiller, Hubbard, and Kanamori, a second upsurge of interest in the Hubbard model emerged occurred as a result of the experimental possibilities made available by the novel experimental techniques to trap ultracold atoms in optical lattices (see, e.g. Bloch (2005) for a review). These techniques allow the experimental investigation of the original Fermionic Hubbard (Joerdens *et al.*, 2008) model discussed here, but were also used to explore its Bosonic version (Greiner *et al.*, 2002) a number of years earlier.

This section an understanding of how the Hubbard model can be justified from a microscopic 'first principles' perspective. As in section 8.2.1, we begin with the electronic part of the 'Theory of Everything' Hamiltonian Laughlin and Pines (2000); Laughlin (1998) of condensed matter theory

$$\mathcal{H} = \mathcal{H}_{\text{kin}} + \mathcal{H}_{\text{int}} \tag{8.545}$$

$$= \sum_{\sigma} \int d^3 r \psi_{\sigma}^{\dagger}(\mathbf{r}) \left[-\frac{}{2m} \nabla^2 + U_{\text{ion}}(\mathbf{r}) \right] \psi_{\sigma}(\mathbf{r}) + \sum_{\sigma,\sigma'} \int d^3 r \int d^3 r' \psi_{\sigma}^{\dagger}(\mathbf{r}) \psi_{\sigma'}^{\dagger}(\mathbf{r}') V_{\text{ee}}(\mathbf{r} - \mathbf{r}') \psi_{\sigma'}(\mathbf{r}') \psi_{\sigma}(\mathbf{r}), \tag{8.546}$$

which, unlike Laughlin in his Nobel presentation 1998, we again have written in second quantized form with field operators ${\psi_{\sigma}}^{\dagger}(\mathbf{r})$.

This very general Hamiltonian describes electrons interacting with the potential $U_{\text{ion}}(\mathbf{r})$ of a static lattice of ions, i.e. employs the Born–Oppenheimer approximation. Thus, we neglect the motion of the ion lattice since we are only interested in the interactions of electrons and not in dynamical lattice effects, such as phonons. Moreover, the electrons interact via Coulomb repulsion

$$V_{\text{ee}}(\mathbf{r} - \mathbf{r}') \propto \frac{1}{|\mathbf{r} - \mathbf{r}'|}. \tag{8.547}$$

In second quantization language, the Hamiltonian (8.546) describes the situation adequately.

The Hamiltonian (8.546) is quite complicated still, and in the following we make a number of approximations to reach a more tractable Hamiltonian. We attempt to make the approximations in such a way that still captures the essential physics of strongly interacting electrons. The method of second quantization introduced and discussed in chapter 2 is very well adapted to this task.

The first ingredient in our analysis is *Bloch's theorem*. The lattice potential $U_{\text{ion}}(\mathbf{r})$ is periodic

$$U_{\text{ion}}(\mathbf{r} + \mathbf{R_i}) = U_{\text{ion}}(\mathbf{r}) \tag{8.548}$$

with periodicity given by a lattice vector $\mathbf{R_i}$ (assuming a primitive lattice with one atom per unit cell)

$$\mathbf{R_i} = i_1 \mathbf{a}_1 + i_2 \mathbf{a}_2 + i_3 \mathbf{a}_{\mathcal{J}} = \sum_{\mathcal{J}=1}^{3} i_{\mathcal{J}} \, \mathbf{a}_{\mathcal{J}} \tag{8.549}$$

where $\{\mathbf{a}_\mathcal{J}\}$ are a set of basis vectors of the lattice and the numbers $i_\mathcal{J} \in \mathbb{Z}$ ($\mathcal{J} = 1, 2, 3$) describe an infinite lattice or a lattice with some appropriately chosen boundary conditions. For the triple of numbers $i_\mathcal{J}$ we have already used a vector notation: $\mathbf{i}$.

Bloch's theorem states that the solution of the Schrödinger equation in a periodic lattice potential $U_{\text{ion}}(\mathbf{r})$ is given by wave functions that have the form of so-called Bloch functions $u_{\mathbf{k},\alpha}(\mathbf{r})$ and energy eigenvalues in the form of electronic bands $\epsilon_{\mathbf{k},\alpha}$.

The Bloch functions are functions that have been Fourier transformed with respect to the periodic lattice $\mathbf{R_i}$. Their inverse Fourier transformed counterparts are the Wannier functions

$$\phi_{\mathbf{i}\alpha}(\mathbf{r}) \equiv \phi_\alpha(\mathbf{r} - \mathbf{R_i}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R_i}} u_{\mathbf{k},\alpha}(\mathbf{r}), \tag{8.550}$$

which are localized at site $\mathbf{R_i}$ of the ionic lattice. Bloch and Wannier functions are connected by a unitary transformation and provide equivalent descriptions of the physical situation.

We now define creation operators for electrons in a Wannier state

$$c^\dagger_{\mathbf{i}\alpha\sigma} = \int d^3r\, \phi_{\mathbf{i}\alpha}(\mathbf{r}) \psi^\dagger_\sigma(\mathbf{r}), \tag{8.551}$$

which have the inverse relation

$$\psi^\dagger_\sigma(\mathbf{r}) = \sum_{\mathbf{i},\alpha} \phi^*_{\mathbf{i}\alpha}(\mathbf{r}) c^\dagger_{\mathbf{i}\alpha\sigma}. \tag{8.552}$$

These operators, being Fermionic, obey anti-commutation relations

$$\{c_{\mathbf{j}\alpha\sigma}, c^\dagger_{\mathbf{l}\beta\sigma'}\} = \delta_{\mathbf{jl}}\delta_{\alpha\beta}\delta_{\sigma\sigma'}, \quad \{c_{\mathbf{j}\alpha\sigma}, c_{\mathbf{l}\beta\sigma'}\} = \{c^\dagger_{\mathbf{j}\alpha\sigma}, c^\dagger_{\mathbf{l}\beta\sigma'}\} = 0. \tag{8.553}$$

The Hamiltonian 8.546 becomes in the Wannier basis

$$\mathcal{H} = \sum_{\mathbf{ij}\alpha\sigma} t^\alpha_{\mathbf{ij}} c^\dagger_{\mathbf{i}\alpha\sigma} c_{\mathbf{j}\alpha\sigma} + \sum_{\mathbf{ijmn}} \sum_{\alpha\beta\mu\nu} \sum_{\sigma\sigma'} v^{\alpha\beta\mu\nu}_{\mathbf{ijmn}} c^\dagger_{\mathbf{i}\alpha\sigma} c^\dagger_{\mathbf{j}\beta\sigma'} c_{\mathbf{n}\nu\sigma'} c_{\mathbf{m}\mu\sigma} \tag{8.554}$$

with the hopping

$$t^\alpha_{\mathbf{ij}} = \left\langle \mathbf{i}\alpha \left| \left[-\frac{1}{2m}\nabla^2 + U_{\text{ion}}(\mathbf{r}) \right] \right| \mathbf{j}\alpha \right\rangle \tag{8.555}$$

$$= \int d^3r\, \phi^*_{\mathbf{i}\alpha}(\mathbf{r}) \left[-\frac{1}{2m}\nabla^2 + U_{\text{ion}}(\mathbf{r}) \right] \phi_{\mathbf{j}\alpha}(\mathbf{r}) \tag{8.556}$$

and the interaction matrix elements

$$v_{\mathbf{ijmn}}^{\alpha\beta\mu\nu} = \langle \mathbf{i}\alpha, \mathbf{j}\beta | V_{\mathrm{ee}}(\mathbf{r} - \mathbf{r}') | \mathbf{m}\mu, \mathbf{n}\nu \rangle \tag{8.557}$$

$$= \int d^3r \int d^3r' \; \phi_{\mathbf{i}\alpha}^*(\mathbf{r}) \phi_{\mathbf{j}\beta}^*(\mathbf{r}') V_{\mathrm{ee}}(\mathbf{r} - \mathbf{r}') \phi_{\mathbf{m}\mu}(\mathbf{r}) \phi_{\mathbf{n}\nu}(\mathbf{r}'). \tag{8.558}$$

Apart from the Born–Oppenheimer approximation of a static ionic lattice, we have not yet made any further approximation. We just have rewritten the Hamiltonian (8.546) in a form that is more suitable for our purpose of describing interacting localized magnetic moments.

Following Hubbard 1963, Gutzwiller (1963), and Kanamori (1963), we now make a number of simplifying assumptions. Firstly, we assume that all except the lowest band have very high energies and are, thus, energetically unavailable. Therefore, we can drop all band indices α, β, μ, ν, and the corresponding summations. Secondly, we assume that the remaining band has rotational symmetry, i.e. is an *s*-band. This implies that the hopping matrix elements depend only on the distance between the sites **i** and **j**, i.e.

$$t_{\mathbf{ij}} = t(|\mathbf{R_i} - \mathbf{R_j}|). \tag{8.559}$$

We arrive, thus, at the simplified Hamiltonian

$$\mathcal{H} = \sum_{\mathbf{ij}\sigma} t_{\mathbf{ij}} c_{\mathbf{i}\sigma}^\dagger c_{\mathbf{j}\sigma} + \sum_{\mathbf{ijmn}} \sum_{\sigma\sigma'} v_{\mathbf{ijmn}} c_{\mathbf{i}\sigma}^\dagger c_{\mathbf{j}\sigma'}^\dagger c_{\mathbf{n}\sigma'} c_{\mathbf{m}\sigma}. \tag{8.560}$$

The matrix elements decrease fast with increasing distance $|\mathbf{R_i} - \mathbf{R_j}|$, so that we can restrict the summation over sites to nearest-neighbour sites $\langle \mathbf{ij} \rangle$ and arrive at the *generalized Hubbard model*

$$\begin{aligned} \mathcal{H} = & -t \sum_{\langle \mathbf{ij} \rangle} \sum_{\sigma} \left(c_{\mathbf{i}\sigma}^\dagger c_{\mathbf{j}\sigma} + c_{\mathbf{j}\sigma}^\dagger c_{\mathbf{i}\sigma} \right) + U \sum_{\mathbf{i}} n_{\mathbf{i}\uparrow} n_{\mathbf{i}\downarrow} + V \sum_{\langle \mathbf{ij} \rangle} n_{\mathbf{j}} n_{\mathbf{j}} \\ & + X \sum_{\langle \mathbf{ij} \rangle} \sum_{\sigma} \left(c_{\mathbf{i}\sigma}^\dagger c_{\mathbf{j}\sigma} + c_{\mathbf{j}\sigma}^\dagger c_{\mathbf{i}\sigma} \right) \left(n_{\mathbf{i},-\sigma} + n_{\mathbf{j},-\sigma} \right) + \mathcal{J} \sum_{\langle \mathbf{ij} \rangle} \mathbf{S_i} \cdot \mathbf{S_j} \\ & + Y \sum_{\langle \mathbf{ij} \rangle} \left(c_{\mathbf{i}\uparrow}^\dagger c_{\mathbf{i}\downarrow}^\dagger c_{\mathbf{j}\downarrow} c_{\mathbf{j}\uparrow} + c_{\mathbf{j}\downarrow}^\dagger c_{\mathbf{j}\uparrow}^\dagger c_{\mathbf{i}\uparrow} c_{\mathbf{i}\downarrow} \right), \end{aligned} \tag{8.561}$$

where the only surviving matrix elements are the hopping matrix element

$$t = -t_{\mathbf{ij}}, \tag{8.562}$$

describing single-electron hopping between nearest-neighbour sites **i** and **j**, the on-site or Hubbard repulsion[6]

$$U = v_{\mathbf{iiii}}, \tag{8.563}$$

the Coulomb interaction between electrons on neighbouring sites

$$V = v_{\mathbf{ijij}}, \tag{8.564}$$

the so-called bond-charge interaction

$$X = v_{\mathbf{iiij}}, \tag{8.565}$$

the spin-spin, or Heisenberg exchange interaction between neighbouring sites

$$\mathcal{J} = -2v_{\mathbf{ijji}}, \tag{8.566}$$

and, finally, the term describing the hopping of pairs of electrons

$$Y = v_{\mathbf{iijj}}. \tag{8.567}$$

The bond-charge interaction X describes the hopping of single electrons where the hopping depends on the occupation of the sites involved. It is proportional to the charge, i.e. the number of electrons, located on the sites of the bond $\langle \mathbf{ij} \rangle$ between sites **i** and **j**.

In the generalized Hubbard Hamiltonian (8.561), we introduced the following operators: the electron number operators at site **i**

$$n_{\mathbf{i}\sigma} = c^{\dagger}_{\mathbf{i}\sigma} c_{\mathbf{i}\sigma}, \tag{8.568}$$

and

$$n_{\mathbf{i}} = n_{\mathbf{i}\uparrow} + n_{\mathbf{i}\downarrow}, \tag{8.569}$$

and the spin-$\frac{1}{2}$ operators at site **i**

$$\mathbf{S}_{\mathbf{i}} = \frac{1}{2} \sum_{\alpha,\beta} c^{\dagger}_{\mathbf{i}\sigma} \boldsymbol{\sigma}_{\alpha\beta} c_{\mathbf{i}\sigma}, \tag{8.570}$$

where $\sigma_{\alpha\beta}$ are the components of the vector σ of Pauli matrices (cf. chapter 3).

Some of the symmetries of the generalized Hubbard Hamiltonian (8.561), i.e. some of the operators with which (8.561) commutes, are:

[6] The attractive case is also sometimes considered.

- the number of electrons

$$\mathcal{N} = \sum_{\mathbf{i}} n_{\mathbf{i}}. \tag{8.571}$$

  On a lattice of N sites we can at most place $2N$ electrons. Note the important special case of a so-called half-filled band, where in the ground state one has $N_e = \frac{N}{2}$ electrons in the lattice, one electron per site.

- the magnetization

$$\mathcal{M} = \mathcal{N}_\uparrow - \mathcal{N}_\downarrow = \sum_{\mathbf{i}} \left(n_{\mathbf{i}\uparrow} - n_{\mathbf{i}\downarrow} \right), \tag{8.572}$$

  and the

- total spin of the chain

$$\mathbf{S} = \sum_{\mathbf{i}} \mathbf{S}_{\mathbf{i}}. \tag{8.573}$$

In other words, the total number of electrons $\mathcal{N}$, the magnetization $\mathcal{M}$, and the total spin $\mathbf{S}$ are conserved, the latter because the Heisenberg interaction $\mathcal{J}$ is isotropic. This latter property of the generalized Hubbard model can, in principle, be relaxed to allow anisotropic exchange interactions with $\mathcal{J}^x \neq \mathcal{J}^y \neq \mathcal{J}^z$ in the most general case. In fact, the anisotropic Heisenberg model retains all but the spin-spin interaction of the generalized Hubbard model (8.561) . Obviously, in such a model no electron motion is possible that would require at least a hopping term t. Therefore, it cannot describe itinerant magnetism. However, it is a perfectly reasonable model to describe the interaction of localized magnetic moments in magnetic insulators where the magnetism is indeed due to the interaction between local magnetic moments. The one-dimensional version of this model, the Heisenberg quantum spin chain, is one of our prime examples of a model solvable by Bethe ansatz.

.....

EXERCISE 8.26 Generalized Hubbard model Convince yourself explicitly that the generalized Hubbard Hamiltonian (8.561) is obtained from the Hamiltonian (8.560) if we only retain nearest-neighbour interactions.

.....

Hubbard (1963) provided a simple estimate for the order of magnitude of the various matrix elements in the generalized Hubbard model (8.561) in the case of transition metals like iron, cobalt, or nickel: $t \approx 1eV, U \approx 10eV, V \approx 2 - 3eV, X \approx 0.5eV$, and $\mathcal{J}$, $Y \ll 1eV$. Motivated by these orders of magnitude, most attention was henceforth focused on the case where all terms in the generalized Hubbard model (8.561) are

neglected except the on-site Coulomb repulsion U, and the single-electron hopping t, the latter being necessary to still describe itinerant magnetism using this model.

The remaining model Hamiltonian

$$\mathcal{H} = -t \sum_{\langle \mathbf{ij} \rangle \sigma} \left(c^\dagger_{\mathbf{i}\sigma} c_{\mathbf{j}\sigma} + c^\dagger_{\mathbf{j}\sigma} c_{\mathbf{i}\sigma} \right) + U \sum_{\mathbf{i}} n_{\mathbf{i}\uparrow} n_{\mathbf{i}\downarrow} \equiv \mathcal{H}_t + \mathcal{H}_U \tag{8.574}$$

today bears the name Hubbard model, and constitutes arguably one of the models, alongside the Ising model and the Heisenberg model, to which a great deal of research effort in condensed matter physics is devoted. For calculations involving the grand canonical ensemble, a chemical potential μ and a corresponding term in the Hamiltonian

$$\mathcal{H}_\mu = -\mu \sum_{\mathbf{i}\sigma} n_{\mathbf{i}\sigma} \tag{8.575}$$

are introduced. Furthermore, an external magnetic field $\mathbf{h} = h\mathbf{e}_z$ will couple to the z-component of the total spin adding a further term

$$\mathcal{H}_h = h \sum_{\mathbf{i}} S^z_{\mathbf{i}} = \frac{h}{2} \sum_{\mathbf{i}} \left(n_{\mathbf{i}\uparrow} - n_{\mathbf{i}\downarrow} \right) \tag{8.576}$$

to the Hubbard Hamiltonian.

An alternative form for the on-site part of the Hubbard Hamiltonian is

$$\mathcal{H}'_U = U \sum_{\mathbf{i}} \left(n_{\mathbf{i}\uparrow} - \frac{1}{2} \right) \left(n_{\mathbf{i}\downarrow} - \frac{1}{2} \right) \tag{8.577}$$

such that

$$\mathcal{H}'_U = \mathcal{H}_U - \frac{U}{2} \sum_{\mathbf{i}} \left(n_{\mathbf{i}\uparrow} + n_{\mathbf{i}\downarrow} \right) + \frac{NU}{4}. \tag{8.578}$$

This form shifts the chemical potential $\mu \to \mu + U/2$ and introduces an overall additive constant $NU/4$ to the energy. However, the particle-hole symmetry discussed in section 8.7.1 is manifest in this form of the Hubbard Hamiltonian.

In summary, it is fair to say that

> the Hubbard model describes—in the simplest possible fashion—a system of interacting electrons. It can be viewed in just this way, as a toy model, or it can be viewed, a bit more realistically in the repulsive case, as a serious model of π-electrons hopping between localized Wannier orbitals in some molecule such as benzene (with $N = 6$) Heilmann and Lieb (1970); the half-filled band (one electron per site) is then especially important because it corresponds to

neutrality. The ultra-short range interaction is supposed to mimic a highly screened Coulomb potential. From the latter viewpoint it was known first in the chemistry literature as the Pariser–Parr–Pople Pariser and Parr (1953); Pople (1953) model; molecules having a bipartite structure were called 'alternant molecules'. It was a decade later that Hubbard (1963), Gutzwiller (1963), and Kanamori (1963) realized its importance for bulk matter—paraphrased from Lieb (1993)

8.7.1 Particle-hole symmetry of the Hubbard model

We now consider the bipartite lattice introduced in section 8.6.6.2.

On bipartite lattices we may consider new Fermionic creation and annihilation operators according to

$$C^{\dagger}_{\mathbf{i}\sigma} = (-1)^{\mathbf{i}} c_{\mathbf{i}\sigma} \qquad \text{and} \qquad C_{\mathbf{i}\sigma} = (-1)^{\mathbf{i}} c^{\dagger}_{\mathbf{i}\sigma} \tag{8.579}$$

where the factor $(-1)^{\mathbf{i}}$ is defined to be $(+1)$ on sub-lattice A and (-1) on the sub-lattice B. These new operators exchange the roles of creation and annihilation Fermions. The transformation to the new operators is called a particle-hole transformation, which is justified by the observation

$$C^{\dagger}_{\mathbf{i}\sigma} C_{\mathbf{i}\sigma} = 1 - c^{\dagger}_{\mathbf{i}\sigma} c_{\mathbf{i}\sigma}. \tag{8.580}$$

Among the symmetries exhibited by the Hubbard model (8.574), the particle-hole symmetry is of particular practical importance, e.g. in quantum Monte Carlo simulations.

This symmetry is manifest for the Hubbard Hamiltonian with the on-site interaction written in the form (8.577)

$$\mathcal{H}_t + \mathcal{H}'_U \to \mathcal{H}_t + \mathcal{H}'_U. \tag{8.581}$$

However, using (8.574) directly produces upon a particle-hole transformation (8.579) extra terms

$$\mathcal{H}_t + \mathcal{H}_U \to \mathcal{H}_t + \mathcal{H}_U + UN - U \sum_{\mathbf{i}} \left(n_{\mathbf{i}\uparrow} + n_{\mathbf{i}\downarrow} \right), \tag{8.582}$$

the first of which is an irrelevant constant. However, the second term acts as a shift in the chemical potential (8.575), which itself behaves as

$$\mathcal{H}_\mu \to \mathcal{H}_{-\mu} - 2\mu N \tag{8.583}$$

under a particle-hole transformation.

A simple application of the particle-hole transformation (8.579) of the Hubbard Hamiltonian (8.574) finds for the site occupation number or particle density

$$n(\mu, T) = \frac{1}{N}\langle n_\uparrow + n_\downarrow \rangle = \frac{1}{NZ}\mathrm{Tr}\left[(n_\uparrow + n_\downarrow)e^{-\beta\mathcal{H}_t - \beta\mathcal{H}_U - \beta\mathcal{H}_\mu}\right] = \frac{1}{N\beta}\frac{\partial \ln Z}{\partial \mu} \quad (8.584)$$

with

$$Z = \mathrm{Tr}\left[e^{-\beta\mathcal{H}_t - \beta\mathcal{H}_U - \beta - \beta\mathcal{H}_\mu}\right]. \quad (8.585)$$

Using the transformation behaviour of the various parts of the Hubbard Hamiltonian discussed above, under the particle-hole transformation, the site occupation number becomes

$$\frac{1}{N\beta}\frac{\partial}{\partial \mu}\ln\left[e^{-\beta\mathcal{H}_t - \beta\mathcal{H}_U - \beta\mathcal{H}_{(U-\mu)} - \beta(U-2\mu)N}\right] = -n(U-\mu, T) + 2. \quad (8.586)$$

Hence, we have the relation for the site occupation number

$$n(\mu, T) = 2 - n(U-\mu, T), \quad (8.587)$$

which implies for half-filling, i.e. $n = 1$, that we need $\mu = U/2$ for the particle-hole symmetric Hubbard model independent of temperature.

Exercise 8.27 concerns a similar application of the particle-hole symmetry. See exercise 8.28 for a discussion of further symmetries of the Hubbard Hamiltonian in the special case of one dimension.

EXERCISE 8.27 The single-site Hubbard model Consider the Hubbard Hamiltonian (8.574) together with the chemical potential term (8.575) and a magnetic field term (8.576), i.e. $\mathcal{H} = \mathcal{H}_t + \mathcal{H}_U + \mathcal{H}_\mu + \mathcal{H}_h$. Putting the hopping matrix element to zero, $t = 0$, in this Hamiltonian is tantamount to considering a collection of identical single-site Hubbard Hamiltonians, i.e. $\mathcal{H} = \sum_{\mathbf{j}} \mathcal{H}_{\mathbf{j}}$.

Calculate the partition function of the single-site Hubbard Hamiltonian

$$Z = \mathrm{Tr}\left[\exp(-\beta\mathcal{H}_{\mathbf{j}})\right] \quad (8.588)$$

and the site occupation number

$$\rho(\mu, h, T) \equiv \langle n_\uparrow + n_\downarrow \rangle \equiv Z^{-1}\mathrm{Tr}\left[\left(n_\uparrow + n_\downarrow\right)\exp(-\beta\mathcal{H}_{\mathbf{j}})\right], \quad (8.589)$$

which are, of course, independent of the site index $\mathbf{j}$, as a function of chemical potential μ, temperature $T = 1/\beta$, and magnetic field h.

Convince yourself that, in the limit $T \to 0$ and for $h = 0$, the site occupation number $\rho(\mu, T)$ changes in steps: $\rho(\mu) = 0$ for $\mu < -U/2$, $\rho(\mu) = 1$ for $-U/2 < \mu < U/2$,

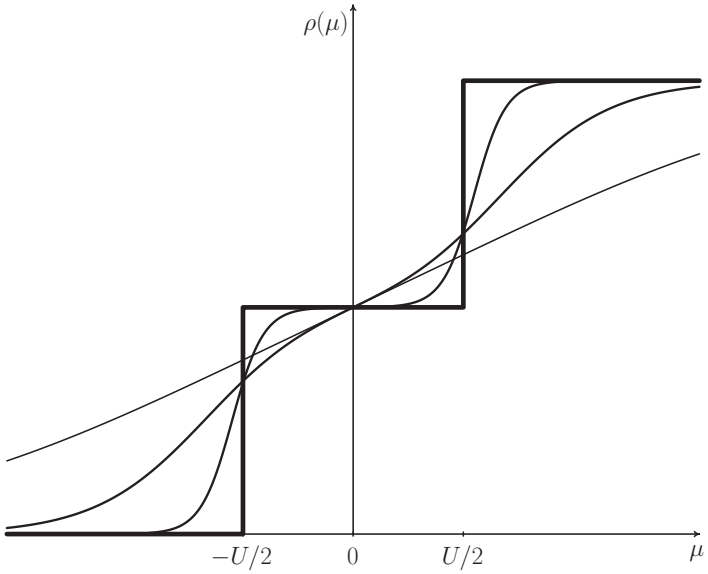

Figure 8.8 *Site occupation number Hubbard model. The so-called Mott plateau between $\mu = -U/2$ and $\mu = U/2$ with occupation number $\rho = 1$, i.e. half-filling, develops for decreasing temperature (thicker lines correspond to lower temperatures). In the limit $\mathsf{T} \to \mathsf{0}$ the occupation is $\rho = 0$ for $\mu < -U/2$. It jumps to $\rho = 1$ at $\mu = -U/2$. A cost in energy of U has to be expended to add a second electron.*

and $\rho(\mu) = 2$ for $\mu > U/2$. For finite temperature T, these steps are washed out but still clearly discernible for low temperatures (see figure 8.8).

.....

In one dimension, the Hubbard model can be solved by Bethe ansatz. Investigating the properties of the one-dimensional Hubbard model is therefore a very useful exercise.

.....

EXERCISE 8.28 The one-dimensional Hubbard model The Hubbard model in one dimension describes N_e electrons on a lattice of N lattice sites that are allowed to hop to nearest-neighbour sites. Double occupancy of a site with two electrons (which, of course, must have opposite spin orientation) costs an energy U. The second quantized Hamiltonian for this model is

$$\mathcal{H} = -t \sum_{\substack{j=1,\\ \sigma=\pm 1}}^{N} \left\{ \psi_\sigma^\dagger(x_j)\psi_\sigma(x_j + a) + \psi_\sigma^\dagger(x_j)\psi_\sigma(x_j - a) \right\} + \frac{U}{2} \sum_{\substack{j=1,\\ \sigma=\pm 1}}^{N} \left\{ \psi_\sigma^\dagger(x_j)\psi_\sigma(x_j)\psi_{-\sigma}^\dagger(x_j)\psi_{-\sigma}(x_j) \right\}, \tag{8.590}$$

where $x_j = ja, j = 1, 2, \ldots N$, and a is the lattice spacing.

The Fermi operators obey anti-commutation relations

$$\left\{\psi_\sigma(x_j), \psi^\dagger_{\sigma'}(x_l)\right\} = \delta_{\sigma,\sigma'}\delta_{jl}. \tag{8.591}$$

a) Diagonalize the hopping part of this Hamiltonian by Fourier transformation

$$\psi_\sigma(x) = \frac{1}{\sqrt{N}} \sum_k \exp(ikx) c_{k,\sigma} \tag{8.592}$$

and determine the allowed k-values assuming periodic boundary conditions $x_{j+N} = x_j$.

b) Show that

$$\mathcal{M}_\uparrow = \sum_j \psi^\dagger_\uparrow(x_j)\psi_\uparrow(x_j) \tag{8.593}$$

and $\mathcal{M}_\downarrow$ defined analogously are conserved quantities, i.e. commute with $\mathcal{H}$. Therefore $\mathcal{M}_\uparrow$ and $\mathcal{M}_\downarrow$ provide quantum numbers $M_\uparrow$ and $M_\downarrow$ of the spectrum of $\mathcal{H}$

$$E = E(M_\uparrow, M_\downarrow; t, U). \tag{8.594}$$

c) Determine the symmetries of $E = E(M_\uparrow, M_\downarrow; t, U)$ under the transformations

$$\psi_\sigma(x_j) = (-1)^j c_{j,\sigma} \tag{8.595}$$

and

$$\psi_\uparrow(x_j) = (-1)^j c^\dagger_{j,\uparrow}, \tag{8.596}$$

$$\psi_\downarrow(x_j) = c_{j,\downarrow}. \tag{8.597}$$

Are the new operators in both cases again Fermi operators?

8.8 Heisenberg model

Section 8.6 on magnetism, especially section 8.6.4 on the electrostatic origin of the exchange interaction in the hydrogen molecule, discussed the Heisenberg model as a model for localized magnetism. There is also another route to this model as an approximative limit starting from the Hubbard model, arguably one of the central models in the physics of strongly interacting quantum matter, and that is connected in the limit of large on-site interaction U to two other important models, the t–$\mathcal{J}$ model and, eventually, the Heisenberg model.

While the Hubbard model describes itinerant magnetic behaviour of the conduction band electrons for arbitrary filling of the band, the Heisenberg model is a model of fully localized strongly interacting spins that arises for an exactly half-filled conduction band. The t–$\mathcal{J}$ model occupies an intermediate position where the conduction band is away from half-filling and electrons can still hop between lattice sites because the penalty for double occupancy of a site is not too large.

In the following two sections we derive these two models. Our starting point is the Hamiltonian $\mathcal{H} = \mathcal{H}_t + \mathcal{H}_U$, (8.574), of the Hubbard model introduced in section 8.7. We focus on the Hubbard model on a bipartite lattice, which simplifies the discussion considerably. There are various techniques available for these derivations, see, e.g. Fazekas (1999), chapters 4 and 5, or Auerbach (1994), chapter 3. The method we choose for the derivation, the Schrieffer–Wolff transformation, was introduced by Schrieffer and Wolff (1966) in connection with quantum impurity models to derive the Kondo model from the Anderson model. However, in order to derive the Kondo Hamiltonian from the Anderson Hamiltonian in section 8.9, we employ a different method.

The Schrieffer–Wolff transformation is a unitary transformation used in many instances to extract the effective low-energy physics of strongly interacting quantum systems. It can be used if the energy scales in a Hamiltonian can be chosen such that only a subset of the states available in principle can be reached for low energies.

Under the condition of strong on-site repulsion $U \gg t$, double occupancy in the Hubbard model is suppressed. Depending then on the filling N_e/N of the N lattice sites with a total number of $N_e = N_\uparrow + N_\downarrow$ electrons we obtain either the t–$\mathcal{J}$ model for filling $N_e/N \neq 1$, i.e. away from half-filling, or the Heisenberg model for filling $N_e/N = 1$, i.e. at half-filling. We concentrate on the case of less than half-filling, $N_e/N < 1$. Close to half-filling most of the N sites of the lattice sites are filled with only one electron in the ground state configuration. At exactly half-filling, single occupancy of all N lattice sites defines the ground state. However, these ground states are highly degenerate.

8.8.1 From the Hubbard to the *t–J* model: non-half filled band case

We begin by looking at two sites only. By doing this, we obtain an overview of the most important processes possible in the Hubbard model and how they contribute as the parameters U, t, and the filling factor $\nu = N_e/N$ vary.

There are $4^2 = 16$ possible states to consider for two sites with $N_e = 0, 1, 2, 3, 4$ electrons. As discussed in the previous section, the number of electrons can be adjusted by a variation of the chemical potential μ. Here, we assume appropriate values of μ such that the lattice exhibits the desired regimes of electron filling.

For the strong-coupling limit $U \gg t$ we are interested in here, it is natural to start our discussion in the extreme limit, where $t = 0$, i.e. the atomic limit. In this limit, there are no hopping processes. The ground state consists of a lattice of singly occupied lattice sites and is highly degenerate. If we now switch on hopping, i.e. $t > 0$, the ground state degeneracy gets lifted and the four hopping processes

(a) $\underline{\uparrow\ \ }\ \underline{\ \ } \Longrightarrow \underline{\ \ }\ \underline{\uparrow\ \ }$ (8.598)

(b) $\underline{\uparrow\downarrow\ }\ \underline{\downarrow\ } \Longrightarrow \underline{\downarrow\ }\ \underline{\uparrow\downarrow\ }$ (8.599)

(c) $\underline{\uparrow\ \ }\ \underline{\downarrow\ } \Longrightarrow \underline{\ \ }\ \underline{\uparrow\downarrow\ }$ (8.600)

(d) $\underline{\uparrow\downarrow\ }\ \underline{\ \ } \Longrightarrow \underline{\downarrow\ }\ \underline{\uparrow\ \ }$ (8.601)

and their spin reversed versions must be considered. While in the hopping processes (a) and (b) the number of doubly occupied states remain unchanged, process (c) creates a new doubly occupied site, and process (d) lowers this number by one. Therefore, only the hopping processes (c) and (d) contribute an on-site energy U.

These considerations can now be used as input into the Schrieffer–Wolff transformation where we need a separation of energy scales in the Hamiltonian. Our goal with this unitary transformation is to eliminate the high-energy contributions from the Hubbard Hamiltonian and thus obtain an effective low-energy description.

We thus split up the hopping part $\mathcal{H}_t$ of the Hubbard Hamiltonian (8.574) into pieces that describe the different hopping processes

$$\mathcal{H}_t = \mathcal{H}_{t,0} + \mathcal{H}_{t,2} + \mathcal{H}_{t,d+} + \mathcal{H}_{t,d-} \equiv \mathcal{H}_t^{(0)} + \mathcal{H}_t^{(\pm)} \tag{8.602}$$

where

$$\mathcal{H}_{t,0} = -t \sum_{\langle \mathbf{ij} \rangle \sigma} \left\{ (1 - n_{\mathbf{i}-\sigma})\, c_{\mathbf{i}\sigma}^\dagger c_{\mathbf{j}\sigma} \left(1 - n_{\mathbf{j}-\sigma}\right) + \left(1 - n_{\mathbf{j}-\sigma}\right) c_{\mathbf{j}\sigma}^\dagger c_{\mathbf{i}\sigma}\, (1 - n_{\mathbf{i}-\sigma}) \right\} \tag{8.603}$$

represents processes where unoccupied sites hop by one lattice site, while

$$\mathcal{H}_{t,2} = -t \sum_{\langle \mathbf{ij} \rangle \sigma} \left\{ n_{\mathbf{i}-\sigma}\, c_{\mathbf{i}\sigma}^\dagger c_{\mathbf{j}\sigma}\, n_{\mathbf{j}-\sigma} + n_{\mathbf{j}-\sigma}\, c_{\mathbf{j}\sigma}^\dagger c_{\mathbf{i}\sigma}\, n_{\mathbf{i}-\sigma} \right\} \tag{8.604}$$

characterizes processes where doubly occupied sites hop by one lattice site. The Hamilton operator

$$\mathcal{H}_{t,d+} = -t \sum_{\langle \mathbf{ij} \rangle \sigma} \left\{ n_{\mathbf{i}-\sigma}\, c_{\mathbf{i}\sigma}^\dagger c_{\mathbf{j}\sigma} \left(1 - n_{\mathbf{j}-\sigma}\right) + n_{\mathbf{j}-\sigma}\, c_{\mathbf{j}\sigma}^\dagger c_{\mathbf{i}\sigma}\, (1 - n_{\mathbf{i}-\sigma}) \right\} \tag{8.605}$$

indicates hopping processes that *increase* the number of doubly occupied sites by one, while

$$\mathcal{H}_{t,d-} = -t \sum_{\langle \mathbf{ij} \rangle \sigma} \left\{ (1 - n_{\mathbf{i}-\sigma})\, c^{\dagger}_{\mathbf{i}\sigma} c_{\mathbf{j}\sigma}\, n_{\mathbf{j}-\sigma} + \left(1 - n_{\mathbf{j}-\sigma}\right) c^{\dagger}_{\mathbf{j}\sigma} c_{\mathbf{i}\sigma}\, n_{\mathbf{i}-\sigma} \right\} \tag{8.606}$$

corresponds to hopping processes that *decrease* the number of doubly occupied sites by one.

Thus, the first two parts, $\mathcal{H}_{t,0}$ and $\mathcal{H}_{t,2}$, describe the hopping processes that leave the double occupancy unchanged. The last two parts, $\mathcal{H}_{t,d+}$ and $\mathcal{H}_{t,d-}$, change the double occupancy by $+1$ and -1, respectively, and thus change the energy by U. Moreover, we can also write $\mathcal{H}_t$ as two parts: a part $\mathcal{H}_t^{(0)}$ that leaves double occupancy unchanged, and a part $\mathcal{H}_t^{(\pm)}$ that changes double occupancy by one.

We now apply the Schrieffer–Wolff transformation to the present situation where we want to eliminate the high-energy contributions of the term $\mathcal{H}_t^{(\pm)}$ in the Hubbard Hamiltonian in order to generate an effective Hamiltonian.

The Schrieffer–Wolff transformation is a unitary canonical transformation of the form

$$\tilde{\mathcal{A}} = e^{\mathcal{S}} \mathcal{A} e^{-\mathcal{S}} \tag{8.607}$$

where $\mathcal{S}$, called the action operator, is required to be an anti-Hermitian operator, i.e.

$$\mathcal{S}^{\dagger} = -\mathcal{S} \tag{8.608}$$

such that, if $\mathcal{A}$ is Hermitian, then so is $\tilde{\mathcal{A}}$. Expectation values remain unchanged if the state is also transformed as

$$|\tilde{\psi}\rangle = e^{\mathcal{S}} |\psi\rangle. \tag{8.609}$$

In particular, eigenvalues of $\mathcal{A}$ and $\tilde{\mathcal{A}}$ are identical.

.....

EXERCISE 8.29 Proof of the Schrieffer–Wolff transformation formula The Schrieffer–Wolff transformation gains its usefulness from the expansion

$$\begin{aligned} \tilde{\mathcal{A}} &= e^{\mathcal{S}} \mathcal{A} e^{-\mathcal{S}} \\ &= \mathcal{A} + [\mathcal{S}, \mathcal{A}] + [\mathcal{S}, [\mathcal{S}, \mathcal{A}]] + \ldots = \mathcal{A} + \sum_{n=1}^{\infty} [\mathcal{S}, [\mathcal{S}, \ldots [\mathcal{S}, \mathcal{A}] \ldots]]_n \end{aligned} \tag{8.610}$$

where the index $n = 1, 2, \ldots$ signifies an n-fold commutator.

Prove the validity of this expansion.

.....

Here, we thus are looking for an operator $\mathcal{S}$ such that the part of the Hubbard Hamiltonian $\mathcal{H}_t^{(\pm)}$ that changes the number of doubly occupied sites is eliminated from the Hubbard Hamiltonian in lowest order

$$\tilde{\mathcal{H}} = e^{\mathcal{S}} \left(\mathcal{H}_t^{(0)} + \mathcal{H}_t^{(\pm)} + \mathcal{H}_U \right) e^{-\mathcal{S}} = \mathcal{H}_t^{(0)} + \mathcal{H}_U + \mathcal{O}\left(\frac{t^2}{U} \right). \tag{8.611}$$

Using the expansion derived in exercise 8.29, we can write a transformed Hamiltonian

$$\tilde{\mathcal{H}} = \mathcal{H} + [\mathcal{S}, \mathcal{H}] + \frac{1}{2} [\mathcal{S}, [\mathcal{S}, \mathcal{H}]] + \ldots \tag{8.612}$$

$$= \mathcal{H}_U + \mathcal{H}_t^{(0)} + \mathcal{H}_t^{(\pm)} + [\mathcal{S}, \mathcal{H}_U] + \left[\mathcal{S}, \mathcal{H}_t^{(0)} \right] + \left[\mathcal{S}, \mathcal{H}_t^{(\pm)} \right] + \frac{1}{2} [\mathcal{S}, [\mathcal{S}, \mathcal{H}]] + \ldots . \tag{8.613}$$

The first task now consists of finding an anti-Hermitian operator $\mathcal{S}$ such that

$$\mathcal{H}_t^{(\pm)} + [\mathcal{S}, \mathcal{H}_U] = \mathcal{H}_{t,d+} + \mathcal{H}_{t,d-} + [\mathcal{S}, \mathcal{H}_U] = 0 \tag{8.614}$$

and, hence, to lowest order, the Schrieffer–Wolff transformed operator does not change the number of doubly occupied sites.

In order to find this operator, we observe that

$$\left[\mathcal{H}_{t,d\pm}, \mathcal{H}_U \right] = \mp U \mathcal{H}_{t,d\pm}, \tag{8.615}$$

which suggests that the explicit form

$$\mathcal{S} = \frac{1}{U} \left(\mathcal{H}_{t,d+} - \mathcal{H}_{t,d-} \right) \tag{8.616}$$

satisfies the requirement (8.614).

EXERCISE 8.30 Commutators in the Schrieffer–Wolff expansion Convince yourself of the commutator relations (8.615).

However, with this choice of $\mathcal{S}$, the commutator

$$\left[\mathcal{S}, \mathcal{H}_t^{(0)} \right] \tag{8.617}$$

again introduces hopping processes that change the double occupancy of sites. There are two ways out of this predicament. The first is to consider the orders of magnitude of the terms. The eliminated terms are $\mathcal{O}(t)$, while the terms of (8.617) are of order $\mathcal{O}(t^2/U)$, i.e. of one order of magnitude smaller than the eliminated ones.

A more formal way to deal with this problem is to consider our choice of $\mathcal{S}$ as the first term in an expansion in the small parameter t/U and introduce a second order correction

$$\mathcal{S} \to \mathcal{S} + \mathcal{S}^{(2)}, \tag{8.618}$$

which is of order $(t/U)^2$ and which thus eliminates the unwanted term

$$\left[\mathcal{S}^{(2)}, \mathcal{H}_U\right] = -\left[\mathcal{S}, \mathcal{H}_{t,0}\right] = -\frac{1}{U}\left[\mathcal{H}_{t,d+} - \mathcal{H}_{t,d-}, \mathcal{H}_{t,0}\right]. \tag{8.619}$$

The remaining first-order commutators $\left[\mathcal{S}, \mathcal{H}_t^{(\pm)}\right]$ of the expansion (8.613) are of order of magnitude

$$\left[\mathcal{S}, \mathcal{H}_{t,d\pm}\right] = \frac{1}{U}\left[\mathcal{H}_{t,d+}, \mathcal{H}_{t,d-}\right] = \mathcal{O}(t^2/U), \tag{8.620}$$

which is the same order of magnitude as we obtain for

$$\frac{1}{2}\left[\mathcal{S}, [\mathcal{S}, \mathcal{H}_U]\right] = -\frac{1}{U}\left[\mathcal{H}_{t,d+}, \mathcal{H}_{t,d-}\right] = \mathcal{O}(t^2/U), \tag{8.621}$$

while we have for the remaining parts of the second-order in the expansion (8.613) as order of magnitude

$$[\mathcal{S}, [\mathcal{S}, \mathcal{H}]] = [\mathcal{S}, [\mathcal{S}, \mathcal{H}_U]] + \mathcal{O}(t^3/U^2). \tag{8.622}$$

Thus, collecting terms with order lower than $\mathcal{O}(t^3/U^2)$, the Schrieffer–Wolff transformed Hamiltonian of the Hubbard model becomes the effective Hamiltonian

$$\tilde{\mathcal{H}} = \mathcal{H}_{\mathbf{eff}} = \mathcal{H}_t^{(0)} + \mathcal{H}_U + \frac{1}{U}\left[\mathcal{H}_{t,d+}, \mathcal{H}_{t,d-}\right] + \mathcal{O}(t^3/U^2). \tag{8.623}$$

The main remaining task now is to work out the term $\frac{1}{U}\left[\mathcal{H}_{t,d+}, \mathcal{H}_{t,d-}\right]$ in the effective Hamiltonian. We are interested in the low-energy regime near half-filling where most sites are singly occupied except for a small number of unoccupied sites or holes. Therefore, the product $\mathcal{H}_{t,d+}\mathcal{H}_{t,d-}$ does not contribute as it would require doubly occupied sites. For the same reason, $\mathcal{H}_U$ does not give a contribution. We are left with $\mathcal{H}_{t,d-}\mathcal{H}_{t,d+}$, which creates a doubly occupied site and destroys one. This can be done in two ways, where either the product of operators acts only on two neighbouring sites (**i, j**) creating and annihilating a doubly occupied site, or involving three adjacent sites (**i, j, k**) where the doubly occupied site is created and annihilated on the middle site **j**.

Exercise 8.31 focuses is on the two-site process. The three-site process is discussed in Fazekas (1999), chapter 5, where the useful tool of Hubbard (projection) operators is introduced and used to discuss the t–$\mathcal{J}$ model. In Fazekas (1999), the discussion is

framed in the notions of lower and upper Hubbard band, which we have not introduced. The t–$\mathcal{J}$ model is then the effective description of the low-energy physics or the physics of the lower Hubbard band.

EXERCISE 8.31 *t–J* model Considering only the two-site processes, show that the effective Hamiltonian (8.623) can be written as

$$\mathcal{H}_{t\mathcal{J}} = -t \sum_{\langle \mathbf{ij}\sigma \rangle} \Big((1 - n_{\mathbf{i}-\sigma})\, c^{\dagger}_{\mathbf{i}\sigma} c_{\mathbf{j}\sigma} \left(1 - n_{\mathbf{j}-\sigma}\right) + \left(1 - n_{\mathbf{j}-\sigma}\right) c^{\dagger}_{\mathbf{j}\sigma} c_{\mathbf{i}\sigma} \left(1 - n_{\mathbf{i}-\sigma}\right) \Big) + \frac{4t^2}{U} \sum_{\langle \mathbf{ij} \rangle} \left(\mathbf{S_i} \cdot \mathbf{S_j} - \frac{1}{4} n_{\mathbf{i}} n_{\mathbf{j}} \right), \qquad n_{\mathbf{i}} = n_{\mathbf{i}\uparrow} + n_{\mathbf{i}\downarrow}, \tag{8.624}$$

which is the Hamiltonian of the t–$\mathcal{J}$ model. The first term reflects the contributions of $\mathcal{H}_t^{(0)}$ in (8.623). Recall the local spin operators for electrons calculated in exercise 3.11.

The t–$\mathcal{J}$ Hamiltonian (8.624) will have to be supplemented by three-site terms. However, these are often neglected in applications of the model.

8.8.2 From the Hubbard to Heisenberg model: half-filled band case

As previously mentioned, the exactly half-filled band case is characterized by a ground state with one electron per site, i.e. there are no empty sites or holes. In this situation the first part of (8.623), stemming from $\mathcal{H}_{t,0}$, which corresponds to hopping of empty sites, ceases to contribute to the effective Hamiltonian and we are left with a pure antiferromagnetic Heisenberg Hamiltonian

$$\mathcal{H} = \frac{4t^2}{U} \sum_{\langle \mathbf{ij} \rangle} \mathbf{S_i} \cdot \mathbf{S_j} \equiv \mathcal{J} \sum_{\langle \mathbf{ij} \rangle} \mathbf{S_i} \cdot \mathbf{S_j} \tag{8.625}$$

with $\mathcal{J} > 0$. This is the case of the t–$\mathcal{J}$ model with $n_{\mathbf{i}} = n_{\mathbf{i}\uparrow} + n_{\mathbf{i}\downarrow} = 1$. We omitted the constant term arising from the density-density interaction $n_{\mathbf{i}} n_{\mathbf{j}}$ in (8.623). For further discussion of the consequences of this model's provenance from the Hubbard model, interested readers are referred to chapter 5 of Fazekas (1999).

8.9 Magnetic quantum impurity models

An important avenue to probe strongly interacting quantum matter consists in the study of the effects of quantum impurities. Quantum impurity problems present theoretical and experimental settings in which to analyse various aspects of strong electron correlations. In particular, magnetic quantum impurities induce an effective electron-electron interaction that increases as the energy scale is lowered, the so-called Kondo effect. The