Introduction to magnetic interactions

Dominik Munzar

Department of Condensed Matter Physics, Faculty of Science, Masaryk University

(Preliminary version of study materials for the course

“Physics of strongly correlated electron systems” - 2019)

1. Basic idea and outline

Many magnetic phenomena in solids described using effective Heisenberg-type Hamiltonians.
$H = \sum_{ij} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j$,
where J_{ij} are coupling parameters and
$\mathbf{S}_i$ ($\mathbf{S}_j$) is the spin operator connected with the unit i (j).
Unit: orbital, atom, aggregate of atoms etc.
$J_{ij} < 0/J_{ij} > 0$ ... tendency towards parallel/antiparallel (FM/AF) ordering.
The standard solid state Hamiltonian does not contain any Heisenberg-type terms and the common interaction between magnetic dipoles is too small (see chapter 32 in AM).
What is the origin of these terms?
In this text, the basic strategy for deriving Heisenberg-type Hamiltonians starting from the solid state Hamiltonian will be presented.
Suitable representation of the solid state hamiltonian → Neglect of terms that can be assumed not to be important in the given context and/or Construction of an effective Hamiltonian acting on the relevant subspace of the Hilbert space only → Transformation of the approximate Hamiltonian into an expression involving Heisenberg-type terms.

2. Description of a many-electron system at the Hartree-Fock level
3. Solid state Hamiltonian in the basis of the Hartree-Fock orbitals and in the basis of the Wannier orbitals
4. Approximations: direct Coulomb repulsions and (simple) exchange terms, the Hunds' rule coupling term
5. Hubbard Hamiltonian
6. Heisenberg Hamiltonian for a simple Mott insulator

2. Description of a many-electron system at the Hartree-Fock level

Solid State Hamiltonian (nuclei fixed):

$$H = \sum_i \left\{ -\frac{\hbar^2}{2m}\nabla_i^2 + v_J(\mathbf{r}_i) \right\} + \frac{1}{2}\sum_{i,j,i\neq j} \frac{e'^2}{|\mathbf{r}_i - \mathbf{r}_j|}\,, v_J(\mathbf{r}) = \sum_J -\frac{Z_J e'^2}{|\mathbf{r} - \mathbf{R}_j|}\,.$$

Hartree-Fock variational ansatz: Slater determinant consisting of Hartree-Fock-Bloch spinorbitals $\varphi_{nk}(\mathbf{r},\sigma)$.

In this section we consider a non-spin polarized case, i.e., $\varphi_{nk}(\mathbf{r},\sigma)$ can be written as $\psi_{nk}(\mathbf{r})\chi(\sigma)$, $\chi = \chi_\uparrow$ or $\chi_\downarrow$, ψ does not depend on the spin and the occupation of spin up states is the same as the occupation of spin down states.

Hartree-Fock equation for the Bloch waves:

$$\left\{ -\frac{\hbar^2}{2m}\nabla^2 + v_J(\mathbf{r}) + v_H(\mathbf{r}) + v_x \right\} \psi_{nk}(\mathbf{r}) = \epsilon_{nk}\psi_{nk}(\mathbf{r})\,,$$

where

$$v_H(\mathbf{r}) = 2\sum_{pk',\text{occ.}} \int d\mathbf{r}' \frac{e'^2 |\psi_{pk'}(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|}\,,$$

$$v_x[\psi] = -\sum_{pk',\text{occ.}} \int d\mathbf{r}' \frac{e'^2 \psi^*_{pk'}(\mathbf{r}')\psi(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \times \psi_{pk'}(\mathbf{r})\,.$$

2. Description of a many-electron system at the Hartree-Fock level

By multiplying both sides of the HF equations by $\psi^*_{nk}(\mathbf{r})$ a integrating $d\mathbf{r}$ we obtain

$$\epsilon_{nk} = \langle n\mathbf{k}|h_1|n\mathbf{k}\rangle + \sum_{pk',\text{occ.}} 2\langle n\mathbf{k}, p\mathbf{k}'\left|\frac{e'^2}{|\Delta\mathbf{r}|}\right| n\mathbf{k}, p\mathbf{k}'\rangle - \sum_{pk',\text{occ.}} \langle n\mathbf{k}, p\mathbf{k}'\left|\frac{e'^2}{|\Delta\mathbf{r}|}\right| p\mathbf{k}', n\mathbf{k}\rangle \,,$$

where

$$h_1 = -\frac{\hbar^2}{2m}\nabla^2 + v_J(\mathbf{r})\,,$$

$$\langle ab\left|\frac{e'^2}{|\Delta\mathbf{r}|}\right| cd\rangle = \int d\mathbf{r}d\mathbf{r}'\frac{e'^2\psi^*_a(\mathbf{r})\psi^*_b(\mathbf{r}')\psi_c(\mathbf{r})\psi_d(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,.$$

The sum of the second and third terms on the right hand side of the equation for ϵ_{nk} will be denoted as $\Delta\varepsilon_{nk}$.

Hartree-Fock based Hamiltonian used in studies of low-energy excited states:

$$H_{H.F.} = \sum_{nks}\varepsilon_{nk}c^\dagger_{nks}c_{nks}\,.$$

3. Solid state H. in the HF basis and in the W. basis, (a) HF

$$H = H_{H.F.} + V_{\text{int}} - \Delta H_{HF}\ ,\ \text{where}$$

$$V_{\text{int}} = \frac{1}{2} \sum_{k_1,k_2,k_1',k_2',n,p,q,r,s,s'} \langle n\mathbf{k}_1, p\mathbf{k}_2 \left| \frac{e'^2}{|\Delta\mathbf{r}|} \right| q\mathbf{k}_1', r\mathbf{k}_2' \rangle c^\dagger_{nk_1s} c^\dagger_{pk_2s'} c_{rk_2's'} c_{qk_1's}\ ,$$

and

$$\Delta H_{HF} = \sum_{nks} \Delta\varepsilon_{nk} c^\dagger_{nks} c_{nks} =$$

$$\sum_{nk,pk',s} \left\{ 2\langle n\mathbf{k}, p\mathbf{k}' \left| \frac{e'^2}{|\Delta\mathbf{r}|} \right| n\mathbf{k}, p\mathbf{k}' \rangle - \langle n\mathbf{k}, p\mathbf{k}' \left| \frac{e'^2}{|\Delta\mathbf{r}|} \right| p\mathbf{k}', n\mathbf{k} \rangle \right\} \nu_{pk'} c^\dagger_{nks} c_{nks}\ .$$

Here ν_{pk} is the occupation factor, in the simplest case $\nu_{pk} = 0$ or $\nu_{pk} = 1$. The term ΔH_{HF} has to be subtracted in order to avoid double counting. If the essential effects of the Coulomb interaction are included already at the H.F. level, the term $V_{\text{int}} - \Delta H_{HF}$ can be treated as a perturbation. If the interaction term is dominant this is not possible. In such cases it is often more suitable to employ a localized basis set, e.g. the Wannier basis $\{\phi_n(\mathbf{r} - \mathbf{R}_i)\}$. Transformations between the two basis sets:

$$\phi_n(\mathbf{r} - \mathbf{R}_i) = \frac{1}{\sqrt{N}} \sum_k \mathrm{e}^{-i\mathbf{k}\mathbf{R}_i} \psi_{nk}(\mathbf{r})\ , \psi_{nk}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_i \mathrm{e}^{i\mathbf{k}\mathbf{R}_i} \phi_n(\mathbf{r} - \mathbf{R}_i)\ ,$$

$$c_{nks} = \frac{1}{\sqrt{N}} \sum_i \mathrm{e}^{-i\mathbf{k}\mathbf{R}_i} c_{nis}\ ,\ c_{nis} = \frac{1}{\sqrt{N}} \sum_k \mathrm{e}^{i\mathbf{k}\mathbf{R}_i} c_{nks}\ .$$

3. Solid state H. in the HF basis and in the W. basis, (b) W.

Operators $H_{H.F.}$, $V_{\rm int}$, ΔH_{HF}:

$$H_{H.F.} = \sum_{ijns} t_{ijn} c^{\dagger}_{nis} c_{njs} \, ,$$

where

$$t_{ijn} = \frac{1}{N} \sum_{k} \varepsilon_{nk} {\rm e}^{i\mathbf{k}(\mathbf{R}_i - \mathbf{R}_j)} \, .$$

$$V_{\rm int} = \frac{1}{2} \sum_{ijkl;npqr;ss'} V(in, jp, kq, lr) c^{\dagger}_{nis} c^{\dagger}_{pjs'} c_{rls'} c_{qks} \, ,$$

where

$$V(in, jp, kq, lr) = \int d\mathbf{r} d\mathbf{r}' \frac{e'^2 \phi^*_n(\mathbf{r} - \mathbf{R}_i) \phi^*_p(\mathbf{r}' - \mathbf{R}_j) \phi_q(\mathbf{r} - \mathbf{R}_k) \phi_r(\mathbf{r}' - \mathbf{R}_l)}{|\mathbf{r} - \mathbf{r}'|} \, .$$

$$\Delta H_{HF} = \sum_{ijkl,np,s} \left[2V(in, jp, kn, lp) - V(in, jp, lp, kn) \right] \nu_{jlp} c^{\dagger}_{nis} c_{nks} \, ,$$

where

$$\nu_{jlp} = \frac{1}{N} \sum_{k} \nu_{pk} {\rm e}^{i\mathbf{k}(\mathbf{R}_j - \mathbf{R}_l)} \, .$$

4. Approximations: direct Coulomb repulsions and exchange terms ...

The expression for $V_{\rm int}$ on the previous slide is fairly complicated. For this reason, usually only two classes of coupling terms are taken into account.
(a) Direct Coulomb repulsions, i.e., terms with $(in) = (kq)$ a $(jp) = (lr)$;
(b) Exchange terms, i.e., terms with $(in) = (lr)$ a $(jp) = (kq)$.
Ad (a).

$$V_{\rm c} = \frac{1}{2} \sum_{ij,np,ss',(ins)\neq(jps')} V_{ijnp} n_{nis} n_{pjs'} \, ,$$

where

$$V_{ijnp} = \int d\mathbf{r} d\mathbf{r}' \frac{e'^2 |\phi_n(\mathbf{r} - \mathbf{R}_i)|^2 \, |\phi_p(\mathbf{r}' - \mathbf{R}_j)|^2}{|\mathbf{r} - \mathbf{r}'|}$$

is the so called Coulomb integral and $n_{nis} = c^\dagger_{nis} c_{nis}$. For a single band (n), i.e., just one W. orbital per atom, we obtain the Hubbard interaction term:

$$V_{\rm c} = \sum_i U n_{ni\uparrow} n_{ni\downarrow}$$

with $V_{ijnn} = \delta_{ij} U$.

4. Approximations: direct Coulomb repulsions and exchange terms ...

Ad (b).

$$V_{\rm ex} = \frac{1}{2} \sum_{ijnp, in \neq jp, ss'} J_{ijnp} c^{\dagger}_{nis} c^{\dagger}_{pjs'} c_{nis'} c_{pjs} \, ,$$

where

$$J_{ijnp} = \int d\mathbf{r} d\mathbf{r}' \frac{e'^2 \phi^*_n(\mathbf{r}-\mathbf{R}_i) \phi^*_p(\mathbf{r}'-\mathbf{R}_j) \phi_p(\mathbf{r}-\mathbf{R}_j) \phi_n(\mathbf{r}'-\mathbf{R}_i)}{|\mathbf{r}-\mathbf{r}'|}$$

is the so called exchange integral. The term with $in = jp$ is excluded because it is contained already in the direct Coulomb term. The expressiom for $V_{\rm ex}$ can be simplified by introducing spin operators $\mathbf{S}_{in} = (S^x_{in}, S^y_{in}, S^z_{in})$:

$$S^{x/y/z}_{in} = \frac{1}{2} (c^+_{in\uparrow}, c^+_{in\downarrow}) \Big(\sigma_{x/y/z} \Big) \begin{pmatrix} c_{in\uparrow} \\ c_{in\downarrow} \end{pmatrix} \, , \sigma_x = \begin{pmatrix} 0 1 \\ 1 0 \end{pmatrix} \, , \sigma_y = \begin{pmatrix} 0 -i \\ i \ 0 \end{pmatrix} \, , \sigma_z = \begin{pmatrix} 1 \ 0 \\ 0 -1 \end{pmatrix} \, .$$

Physical meaning of $\mathbf{S}_{in}$: it is a one particle operator, that can be expressed as $\sum_k \mathbf{s}_k P_{k,in}$, with the sum running over all electrons, $\mathbf{s}_k$ is the spin operator of the k-th electron (in units of $\hbar$) and $P_{k,in}$ is the projector on the subspace corresponding to the W. orbital $\phi_n(\mathbf{r}-\mathbf{R}_i)$. After some manipulations we obtain

$$V_{\rm ex} = - \sum_{ijnp, in \neq jp} J_{ijnp} [\mathbf{S}_{in} \cdot \mathbf{S}_{jp} + \frac{1}{4} n_{in} n_{jp}] \, .$$

4. Approximations: Discussion of the exchange term

$$V_{\text{ex}} = -\sum_{ijnp, in \neq jp} J_{ijnp}[\mathbf{S}_{in} \cdot \mathbf{S}_{jp} + \frac{1}{4} n_{in} n_{jp}]\,.$$

Remarks:

(i) In some cases, the occupations of the Wannier orbitals can be assumed to be fixed, the operators n_{in} a n_{ip} can then be replaced with their eigenvalues and only spin degrees of freedom remain active. In the following considerations we assume that this is the case.

(ii) J_{ijnp} is a positive real number. V_{ex} thus prefers configurations with parallel spins (FM).

(iii) Interpretation: electrons with parallel spins avoid each other better than electrons with antiparallel spins, a consequence of the Pauli principle.

(iv) J_{ijnp} decreases approximately exponentially with increasing distance $|\mathbf{R}_i - \mathbf{R}_j|$. The elements J_{iinp} are therefore much larger than elements with $j \neq i$, and the latter can be usually neglected. We obtain

$$V_{\text{ex}} \approx -\sum_{inp, n \neq p} J_{iinp} \mathbf{S}_{in} \cdot \mathbf{S}_{ip}\,.$$

The term containing n_{in} a n_{jp} has been omitted. The formula allows us to interpret the first Hund's rule, the expression on the r. h. s. is therefore labelled as the H. rule coupling term.

(v) The latter plays an essential role in the theory of ferromagnetism, but it is not the only player. Realistic models include, in addition, the one-particle term (the effective kinetic energy) and the on-site Coulomb repulsions. A description of a common ferromagnet in terms of a Heisenberg-type Hamiltonian is a phenomenological one.

5. Hubbard Hamiltonian

Here we address the one-band Hamiltonian with the Hubbard interaction term, the so called Hubbard Hamiltonian:

$$H = \sum_{ijs} t_{ij} c^{\dagger}_{is} c_{js} + \sum_{i} U n_{i\uparrow} n_{i\downarrow} \,.$$

The band index has been omitted for simplicity.

(a) Physical interpretation of the first term.
Consider the case of $|\Psi(0)\rangle = c^{\dagger}_{j}|0\rangle$. Using the Schrödinger equation we obtain

$$|\Psi(t)\rangle = c^{\dagger}_{j}|0\rangle - \frac{it}{\hbar}\sum_{i} t_{ij} c^{\dagger}_{i}|0\rangle$$

for sufficiently small values of t. The amplitude of the transition ("hopping") from j do i per unit time is thus $-(i/\hbar)t_{ij}$.
t_{ij} ... "hopping matrix element"
$\sum_{ijs} t_{ij} c^{\dagger}_{is} c_{js}$ ... "hopping term", "effective kinetic energy"

(b) Physical interpretation of the second term.
On-site Coulomb repulsions between electrons with antiparallel spins. The term is used to describe correlations between spin up and spin down electrons, that are not included in the H. F. theory. ,,Effective potential energy".

5. Hubbard Hamiltonian, estimates of U ...

(c) Estimates of U and of some other matrix elements.
Consider $3d$ orbitals of an atom of a 4th-period transition metal.
Rough estimate of the expectation value of the interaction between the nucleus and a $3d$ electron:

$$\langle V \rangle \approx -\frac{2{Z^*}^2 \mathrm{Ry}}{n^2}\,,$$

where Z^* is the effective charge of the nucleus, $Z^* \approx 5 - 10$, and $n = 3$. We obtain $\langle V \rangle \approx 50 - 200\,\mathrm{eV}$.

Rough estimate of U:

$$U \approx \frac{2Z^* \mathrm{Ry}}{n^2} \approx 15 - 30\,\mathrm{eV}\,.$$

It can be seen that the values of U are comparable with the conduction band bandwith of the order of tens of eV.

Rough estimate of matrix elements of the type $\langle i,j \left| \frac{e'^2}{|\Delta \mathbf{r}|} \right| i,j \rangle$, where i, j are nearest neighbours. For the sake of simplicity, let us assume that the Wannier orbital are fairly localized. Then we have

$$\langle i,j \left| \frac{e'^2}{|\Delta \mathbf{r}|} \right| i,j \rangle \approx \frac{2\mathrm{Ry}}{R[a_0]}\,,$$

where R is the distance between nearest neighbours. For transition metals we have $R \approx 5a_0$ and $\langle i,j \left| \frac{e'^2}{|\Delta \mathbf{r}|} \right| i,j \rangle \approx 5\,\mathrm{eV}$. The Coulomb interaction between two electrons at i or between an electron at i and an electron at j will be further reduced by screening effects.

5. Hubbard Hamiltonian, estimates of U ...

Rough estimate of matrix elements of the type $\langle i,i \left| \frac{e'^2}{|\Delta\mathbf{r}|} \right| i,j \rangle$, with i,j nearest neighbours.

$$\langle i,i \left| \frac{e'^2}{|\Delta\mathbf{r}|} \right| i,j \rangle = \int d\mathbf{r} d\mathbf{r}' \frac{e'^2 \phi^*(\mathbf{r}-\mathbf{R}_i)\phi^*(\mathbf{r}'-\mathbf{R}_i)\phi(\mathbf{r}-\mathbf{R}_i)\phi(\mathbf{r}'-\mathbf{R}_j)}{|\mathbf{r}-\mathbf{r}'|} \,.$$

This expression can be viewed as the interaction energy of a charge cloud of the density $e\phi^*(\mathbf{r}-\mathbf{R}_i)\phi(\mathbf{r}-\mathbf{R}_i)$ and a charge cloud of the density $e\phi^*(\mathbf{r}-\mathbf{R}_i)\phi(\mathbf{r}-\mathbf{R}_j)$, the so called overlap charge density. The absolute value of the overlap charge q is approximately by an order of magnitude lower than $|e|$. Therefore

$$\langle i,i \left| \frac{e'^2}{|\Delta\mathbf{r}|} \right| i,j \rangle \approx \frac{2q\mathrm{Ry}}{0.5R[a_0]} \approx 1\,\mathrm{eV} \,.$$

The above estimates allow us to conclude that the matrix element U is indeed considerably larger than other Coulomb interaction matrix elements and that it is likely to play the main role in the physics of local electronic correlations.

5. Hubbard Hamiltonian, general remarks

(d) General remarks on the Hubbard Hamiltonian

- It is the simplest possible Hamiltonian allowing one to address correlations between spin up and spin down electrons.

- Fundamental role in theory of magnetism and superconductivity.

- Analytical solutions only in the 1D case and in the ∞D case, for a basic information, see chapter 12.5.3 in Fulde's textbook.

- Effective kinetic energy (T in the following) and the effective potential energy (V_H in the following) play here similar roles as the kinetic energy and the potential energy in one-particle quantum mechanics (delocalizing electrons and localizing electrons, respectively).

5. Hubbard Hamiltonian, limiting cases

(e) Limiting cases.
(α) First limiting case: $U = 0$.
(β) Second limiting case: $t = 0$.
Ad (α). This is the so called “band limit”. Exact solution: Slater determinant consisting of H. F. B. states.
In the following (including part (β)), for concreteness, we consider a simple one-dimensional lattice with lattice parameter a, N lattice points and N electrons. One-particle states: $\psi_k\chi_\uparrow$, $\psi_k\chi_\downarrow$, energies: $\varepsilon_k = 2t\cos(ka)$, quasiparticle operators: c_{ks}, $c^\dagger_{ks}$.
The ground state is nondegenerate,

$$|\Psi_\alpha\rangle = \Pi_{-\frac{\pi}{2a}<k<\frac{\pi}{2a}} c^\dagger_{k\uparrow} c^\dagger_{k\downarrow}|0\rangle \,.$$

Ad (β). This is the so called “atomic limit”. In the ground state, there is one electron at each lattice site - this minimizes the effective potential energy. The charge distribution is thus given, the spin distribution, however, may be arbitrary, and the ground state is 2^N degenerate,

$$|\Psi_\beta\rangle = \Pi_{i=1}^{N} c^\dagger_{i\uparrow}/c^\dagger_{i\downarrow}|0\rangle \,.$$

5. Hubbard Hamiltonian, Mott transition

(f) Mott transition.

$$|\Psi_\alpha\rangle = \Pi_{-\frac{\pi}{2a}<k<\frac{\pi}{2a}} c^\dagger_{k\uparrow} c^\dagger_{k\downarrow}|0\rangle \, , \, |\Psi_\beta\rangle = \Pi_{i=1}^{N} c^\dagger_{i\uparrow}/c^\dagger_{i\downarrow}|0\rangle$$

For finite U and t none of the state vectors $|\Psi_\alpha\rangle$, $|\Psi_\beta\rangle$ is an eigenvector. We can, however, pose the following question: Which of the two vectors provides a better description of the ground state (i.e., a lower value of $\langle H\rangle$)? We obtain

$$\langle\Psi_\alpha|T|\Psi_\alpha\rangle = \frac{4tN}{\pi} \, , \, \langle\Psi_\alpha|V_H|\Psi_\alpha\rangle = \frac{NU}{4} \, , \, \langle\Psi_\alpha|H|\Psi_\alpha\rangle = \frac{4tN}{\pi} + \frac{NU}{4} \, .$$

$$\langle\Psi_\beta|T|\Psi_\beta\rangle = 0 \, , \, \langle\Psi_\beta|V_H|\Psi_\beta\rangle = 0 \, , \, \langle\Psi_\beta|H|\Psi_\beta\rangle = 0 \, .$$

Clearly, for $U < 16|t|/\pi$ ($U > 16|t|/\pi$), the state vector $|\Psi_\alpha\rangle$ (any of the state vectors $|\Psi_\beta\rangle$) represents a better variational ansatz. Recall that $t < 0$ (remember the corresponding description of the hydrogen molecule) and that $16|t|/\pi$ is very approximately equal to the bandwidth $W = 4|t|$. Results of more sophisticated calculations demonstrate that
for $W > U$ the electrons are delocalized and the system behaves as a metal, in agreement with the rule that an odd number of electrons per unit cell implies a metallic behaviour,
for $W < U$ the electrons are localized and the system behaves as an insulator, the so called Mott insulator, in contrast with the above mentioned rule.
This is the so called Mott criterion and the transition between the metallic state and the insulating state is denoted as the Mott transition. The essential property of the Mott insulator is that the charge degrees of freedom are frozen and only the spins remain “alive”.

6. Heisenberg Hamiltonian for a simple Mott insulator

In order to derive an effective spin Hamiltonian, we consider first, for the sake of simplicity, the case of $N = 2$, we have two sites and two electrons. The Hubbard Hamiltonian reads

$$H = T + V_H \, , \, T = \sum_s t(c^\dagger_{2s} c_{1s} + c^\dagger_{1s} c_{2s}) \, , \, V_H = U n_{1\uparrow} n_{1\downarrow} + U n_{2\uparrow} n_{2\downarrow} \, .$$

We shall limit ourselves to the case of $U >> t$. The interaction term (i.e., the Hamiltonian of the atomic limit) can be considered as the "unperturbed" Hamiltonian and the kinetic energy term as a perturbation. Solutions of the unperturbed problem are

$$|\uparrow|\uparrow\rangle \, , \, |\uparrow|\downarrow\rangle \, , \, |\downarrow|\uparrow\rangle \, , \, |\downarrow|\downarrow\rangle \ldots E = 0 \text{ subspace}\, L$$

$$|\uparrow\downarrow|-\rangle \, , \, |-|\uparrow\downarrow\rangle \ldots E = U \text{ subspace}\, H \, .$$

Next we focus on state vectors thats evolve, "after the application of the perturbation", from the low-energy subspace L. Our aim is to find an effective Hamiltonian H_{ef} acting on L such that, *if* $|\Psi\rangle$ *satisfies* $H|\Psi\rangle = E|\Psi\rangle$, $H_{ef} P_L |\Psi\rangle = E P_L |\Psi\rangle$.
Here P_L is the projector on the subspace L.

Assume thus that $|\Psi\rangle$ satisfies $H|\Psi\rangle = E|\Psi\rangle$. The Hamiltonian can be expressed in terms of the projection operators:

$$H = P_L H P_L + P_L H P_H + P_H H P_L + P_H H P_H = P_L V_H P_L + P_H V_H P_H + P_H T P_L + P_L T P_H$$

and $|\Psi\rangle = |\Psi_L\rangle + |\Psi_H\rangle$, where $|\Psi_L\rangle = P_L|\Psi\rangle$ and $|\Psi_H\rangle = P_H|\Psi\rangle$. After inserting this into Eq. $H|\Psi\rangle = E|\Psi\rangle$ we obtain

6. Heisenberg Hamiltonian for a simple Mott insulator

$$(P_L V_H P_L + P_H V_H P_H + P_H T P_L + P_L T P_H)(|\Psi_L\rangle + |\Psi_H\rangle) = E(|\Psi_L\rangle + |\Psi_H\rangle)\,.$$

Let us act on this equation first with P_L and second with P_H. We obtain:

$$P_L V_H P_L|\Psi_L\rangle + P_L T P_H|\Psi_H\rangle = E|\Psi_L\rangle\,,$$

$$P_H V_H P_H|\Psi_H\rangle + P_H T P_L|\Psi_L\rangle = E|\Psi_H\rangle\,.$$

We use the second equation to express $|\Psi_H\rangle$,

$$|\Psi_H\rangle = \frac{1}{E - P_H V_H P_H} P_H T P_L|\Psi_L\rangle\,,$$

and insert this into the first equation. We obtain

$$\left\{ P_L V_H P_L + P_L T P_H \frac{1}{E - P_H V_H P_H} P_H T P_L \right\} |\Psi_L\rangle = E|\Psi_L\rangle\,.$$

Using $P_L V_H P_L = 0$ a $P_H V_H P_H = P_H U P_H$ we further obtain

$$\frac{P_L T P_H T P_L}{E - U}|\Psi_L\rangle = E|\Psi_L\rangle\,.$$

The last equation is still, within the given model, exact. For $E << U$ (applies to the low energy sector), we can neglect E as compared to U in the denominator and we obtain

$$H_{ef}|\Psi_L\rangle = E|\Psi_L\rangle \text{ with } H_{ef} = -\frac{P_L T P_H T P_L}{U}\,.$$

6. Heisenberg Hamiltonian for a simple Mott insulator

Next we express H_{ef} in terms of spin operators. Using explicit representations of the projectors P_L and P_H we obtain

$$H_{ef} = -\frac{1}{U} \sum_{i,j,i\in L,j\in L} |i\rangle\langle i| T \left[\sum_{k\in H} |k\rangle\langle k| \right] T |j\rangle\langle j| \,. \tag{2}$$

Using $\sum_{k\in H} = 1 - \sum_{k\in L}$ and the fact that matrix elements of T on the subspace L are equal to zero we further obtain

$$H_{ef} = -\frac{P_L T^2 P_L}{U} = -\frac{1}{U} P_L \left[\sum_s t(c^\dagger_{2s} c_{1s} + c^\dagger_{1s} c_{2s}) \right]^2 P_L =$$

$$= P_L J \left(\mathbf{S}_1 \cdot \mathbf{S}_2 - \frac{n_1 n_2}{4} \right) P_L \,,$$

where $J = 4t^2/U$. In the final expression, the symbols P_L are usually omitted.

In a more general case of N lattice sites and N electrons we obtain

$$H_{ef} = \frac{1}{2} P_L \sum_{i,j,i\neq j} J_{ij} \left(\mathbf{S}_i \cdot \mathbf{S}_j - \frac{n_i n_j}{4} \right) P_L \,,$$

where $J_{ij} = 4t^2_{ij}/U$. Again, the symbols P_L are usually omitted.

6. Heisenberg Hamiltonian for a simple Mott insulator

Concluding remarks:

(i) As in the case of the spin representation of V_{ex}, it is not necessary to consider the $n_i n_j$ term of the effective Hamiltonian, in conjunction with the projectors, it provides a constant contribution.

(ii) The interaction constants J_{ij} are positive, the (ij) term therefore prefers configurations with antiparallel spins at sites i a j (AF).

(iii) Interpretation. For simplicity, consider $N = 2$. In a configuration with parallel spins, none of the electrons can hop to the neighbouring site, this is not allowed by the Pauli principle. In a configuration with antiparallel spins, any electron can "visit" the neighbouring site. This is not blocked by the Pauli principle. Of course, the electrons are still localized by U, but the probabilities of the visits are finite. It implies that configurations with antiparallel spins exhibit a lower kinetic energy than those with parallels spins. The mechanism is therefore labelled as "**kinetic exchange**" . In real Mott insulators the hoppings proceed via intermediate atoms and the mechanism is labelled as "**superexchange interaction**".

More rigorously: matrix elements of T between states of the subspace L and states of the subspace H are nonzero. Degenerate perturbation theory provides the corresponding effective interaction on L.

(iv) The values of the interaction constants J_{ij} can be in principle obtained by first principles calculations.