

[3] Magnons and Spin Waves

Spin-spin interaction

- The Hamiltonian describing a solid typically consists of the kinetic term and the Coulomb interaction. Both of them do not explicitly depend on or affect the spin degrees of freedom of electrons. Nevertheless, we observe non-trivial correlations among spins; the spontaneous magnetization in ferro- and antiferromagnets is the most striking example.
- It is the Pauli principle that generate the spin-spin correlation even from the spin-non-mixing Hamiltonian. To see this, let us consider a simple toy model: we have only two "orbitals" $\phi_1(x)$ and $\phi_2(x)$, and only two electrons. The Hamiltonian is

$$\begin{aligned} \mathcal{H} &= K + V \\ K &= -t \sum_{\sigma} (c_{1\sigma}^{\dagger} c_{2\sigma} + c_{2\sigma}^{\dagger} c_{1\sigma}) \quad (\sigma = \uparrow, \downarrow) \\ V &= (u/2) \sum_i n_i (n_i - 1) \quad (n_i \equiv n_{i\uparrow} + n_{i\downarrow}) \\ &(t, u > 0) \end{aligned}$$

K is the simplified kinetic energy, and V mimics the Coulomb interaction. (When two electrons occupies the same orbit, it raises the energy by u .)

Obviously, neither term changes \uparrow into \downarrow or vice versa. To be more precise, the Hamiltonian is invariant under the $SU(2)$ global spin rotation:

$$\begin{pmatrix} c_{i\uparrow} \\ c_{i\downarrow} \end{pmatrix} \rightarrow \begin{pmatrix} c'_{i\uparrow} \\ c'_{i\downarrow} \end{pmatrix} \equiv \begin{pmatrix} u & -v^* \\ v & u^* \end{pmatrix} \begin{pmatrix} c_{i\uparrow} \\ c_{i\downarrow} \end{pmatrix} \quad \left(\begin{array}{l} \text{for both} \\ i=1, 2 \end{array} \right)$$

$$|u|^2 + |v|^2 = 1$$

In the case of $N=2$, we have $4C_2 = 6$ states.
In this 6-dimensional Hilbert space, the Hamiltonian is

$$\begin{array}{l} | (1\uparrow)(2\uparrow) \rangle \\ | (1\downarrow)(2\downarrow) \rangle \\ | (1\uparrow)(2\downarrow) \rangle \\ | (2\uparrow)(1\downarrow) \rangle \\ | (1\uparrow)(1\downarrow) \rangle \\ | (2\uparrow)(2\downarrow) \rangle \end{array} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -t & -t \\ 0 & 0 & 0 & 0 & -t & -t \\ 0 & 0 & -t & -t & u & 0 \\ 0 & 0 & -t & -t & 0 & u \end{bmatrix} = \mathcal{H}$$

$$(| (1\uparrow)(2\uparrow) \rangle \equiv c_{1\uparrow}^\dagger c_{2\uparrow}^\dagger | 0 \rangle, \text{ etc. })$$

Note here that we can't have states like $| (1\uparrow)(1\uparrow) \rangle$ because of the Pauli principle.

Obviously, the two "fully-aligned" states $\bar{\Psi}_1 \equiv | (1\uparrow)(2\uparrow) \rangle$ and $\bar{\Psi}_2 \equiv | (1\downarrow)(2\downarrow) \rangle$ are eigenstates with energy 0. The other 4 eigenstates can be obtained by diagonalizing the reduced Hamiltonian

$$\begin{bmatrix} 0 & 0 & -t & -t \\ 0 & 0 & -t & -t \\ -t & -t & u & 0 \\ -t & -t & 0 & u \end{bmatrix} = \overset{1}{\mathcal{H}}$$

With some arithmetics, we obtain 4 eigenstates:

$$\begin{array}{l} \bar{\Psi}_3 \equiv (| (1\uparrow)(2\downarrow) \rangle - | (2\uparrow)(1\downarrow) \rangle) / \sqrt{2} \quad \dots \quad E = 0 \\ \bar{\Psi}_5 \equiv (| (1\uparrow)(1\downarrow) \rangle - | (2\uparrow)(2\downarrow) \rangle) / \sqrt{2} \quad \dots \quad E = u \\ \bar{\Psi}_4 \equiv (| (1\uparrow)(2\downarrow) \rangle + | (2\uparrow)(1\downarrow) \rangle) / \sqrt{2} \quad \dots \quad E \doteq -\frac{4t^2}{u} \\ \bar{\Psi}_6 \equiv (| (1\uparrow)(1\downarrow) \rangle + | (2\uparrow)(2\downarrow) \rangle) / \sqrt{2} \quad \dots \quad E \doteq u + \frac{4t^2}{u} \end{array}$$

(we've assumed $u \gg t$ for the sake of simplicity.)

(To be precise, $\bar{\Psi}_4, \bar{\Psi}_6$ contains a small amount of $(| (1\uparrow)(1\downarrow) \rangle + | (2\uparrow)(2\downarrow) \rangle) / \sqrt{2}$.)

matrix elements of \mathcal{H}

$$\langle (1\uparrow)(2\downarrow) | \mathcal{H} | (2\uparrow)(2\downarrow) \rangle$$

$$= \langle 0 | c_{2\downarrow} c_{1\uparrow} | K | c_{2\uparrow}^\dagger c_{2\downarrow}^\dagger | 0 \rangle$$

$$= -t \langle 0 | c_{2\downarrow} c_{1\uparrow} \underbrace{(c_{1\uparrow}^\dagger c_{2\uparrow}^\dagger)}_{\text{the only term in } K \text{ that contributes}} c_{2\uparrow}^\dagger c_{2\downarrow}^\dagger | 0 \rangle$$

the only term in K that contributes

$$= -t \operatorname{sgn} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 3 & 6 & 4 & 5 \end{pmatrix} \langle 0 | c_{1\uparrow} c_{1\uparrow}^\dagger c_{2\downarrow} c_{2\downarrow}^\dagger c_{2\uparrow} c_{2\uparrow}^\dagger | 0 \rangle$$

$$= -t (+1) \times 1 = -t$$

$$\langle (1\uparrow)(2\downarrow) | \mathcal{H} | (1\uparrow)(1\downarrow) \rangle$$

$$= -t \langle 0 | c_{2\downarrow} c_{1\uparrow} c_{2\downarrow}^\dagger c_{1\downarrow}^\dagger c_{1\uparrow}^\dagger c_{1\downarrow}^\dagger | 0 \rangle$$

$$= -t \operatorname{sgn} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 1 & 3 & 4 & 6 \end{pmatrix} = -t (+1) = -t$$

$$\langle (2\uparrow)(1\downarrow) | \mathcal{H} | (1\uparrow)(1\downarrow) \rangle$$

$$= -t \langle 0 | c_{1\downarrow} c_{2\uparrow} c_{2\uparrow}^\dagger c_{1\downarrow}^\dagger c_{1\uparrow}^\dagger c_{1\downarrow}^\dagger | 0 \rangle$$

$$= -t \operatorname{sgn} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 6 & 4 & 5 \end{pmatrix} = -t (+1) = -t$$

$$\langle (2\uparrow)(1\downarrow) | \mathcal{H} | (2\uparrow)(2\downarrow) \rangle$$

$$= -t \langle 0 | c_{1\downarrow} c_{2\uparrow} c_{1\downarrow}^\dagger c_{2\downarrow}^\dagger c_{2\uparrow}^\dagger c_{2\downarrow}^\dagger | 0 \rangle$$

$$= -t \operatorname{sgn} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 1 & 3 & 4 & 6 \end{pmatrix} = -t (+1) = -t$$

The eigenvectors of $\hat{\mathcal{H}}$

Taking the new basis as

$$\phi_1 \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \phi_2 \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \quad \phi_3 \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \phi_4 \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}$$

$$\begin{aligned} \hat{\mathcal{H}} \rightarrow \hat{\mathcal{H}}' &\equiv \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} t & t \\ t & t \\ t & u \\ t & u \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 & 1 & & \\ & 1 & 1 & \\ 1 & -1 & & \\ & & 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 2t & 0 & 0 \\ 0 & 2t & 0 & 0 \\ 2t & u & 0 & u \\ 2t & u & 0 & -u \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 0 & 4t & 0 & 0 \\ 4t & 2u & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2u \end{pmatrix} \\ &= \begin{pmatrix} 0 & 2t & & \\ 2t & u & & \\ & & 0 & 0 \\ & & 0 & u \end{pmatrix} \quad \dots \quad \phi_3 \text{ and } \phi_4 \text{ are eigenvectors} \\ & \quad \quad \quad \text{with } \epsilon = 0, u \end{aligned}$$

when $t \ll u$ ϕ_1 and ϕ_2 are also eigenvectors with $\epsilon = 0, u$.

To be more precise, $\begin{vmatrix} \epsilon & -2t \\ -2t & \epsilon - u \end{vmatrix} = \epsilon^2 - u\epsilon - 4t^2 = 0$

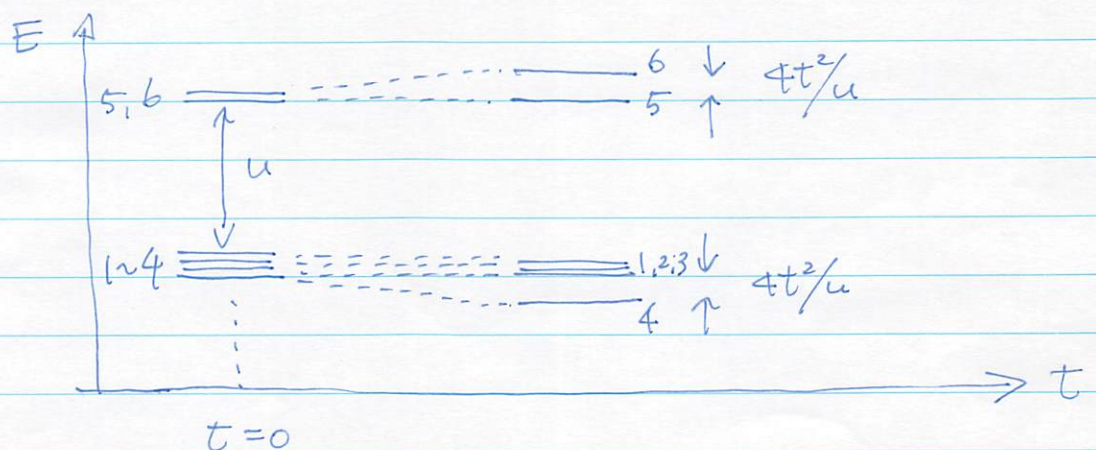
$$\rightarrow \epsilon_{\pm} = \frac{1}{2} \left(u \pm \sqrt{u^2 + 16t^2} \right) = \begin{cases} -\frac{4t^2}{u} & (\phi_1) \\ u + \frac{4t^2}{u} & (\phi_2) \end{cases}$$

$$\bar{\Psi}_3 \equiv \phi_3 \quad \bar{\Psi}_5 \equiv \phi_4$$

$$\bar{\Psi}_4 \equiv \phi_1 + \alpha \phi_2 \quad (|\alpha| \ll 1) \quad \dots \quad \epsilon = -\frac{4t^2}{u}$$

$$\bar{\Psi}_6 \equiv \phi_2 + \beta \phi_2 \quad (|\beta| \ll 1) \quad \dots \quad \epsilon = u + \frac{4t^2}{u}$$

- If we focus on the low-energy behavior of the system, we can neglect $\bar{\Psi}_5$ and $\bar{\Psi}_6$ since they are separated from the other 4 by an energy gap of u .



- In other words, the perturbative term (i.e. K) mixes $\bar{\Psi}_4$ and $\bar{\Psi}_6$ and shifts their energies by $-4t^2/u$ and $+4t^2/u$, respectively.
- It is instructive to express $\bar{\Psi}_1 \sim \bar{\Psi}_4$ in the original (coordinate-specifying) representation.

For example,

$$\begin{aligned} \bar{\Psi}_1 &= |(1\uparrow)(2\uparrow)\rangle = \frac{1}{\sqrt{2}} \left(\phi_1(x_1)\chi_{\uparrow}(x_1)\phi_2(x_2)\chi_{\uparrow}(x_2) \right. \\ &\quad \left. - \phi_2(x_1)\chi_{\uparrow}(x_1)\phi_1(x_2)\chi_{\uparrow}(x_2) \right) \\ &= \frac{1}{\sqrt{2}} (\phi_1(x_1)\phi_2(x_2) - \phi_2(x_1)\phi_1(x_2)) \chi_{\uparrow}(x_1)\chi_{\uparrow}(x_2) \\ &= \bar{\Phi}_{1,2}^{(a)}(\mathbf{x}) \chi_{\uparrow\uparrow}(\mathbf{x}) \end{aligned}$$

(x_i represents the i -th electrons degrees of freedom incl.

Similarly,

$$\bar{\Psi}_2 = \bar{\Phi}_{12}^{(a)}(\mathbf{x}) \chi_{\downarrow\downarrow}(\mathbf{x})$$

$$\bar{\Psi}_3 = \bar{\Phi}_{12}^{(a)}(\mathbf{x}) \chi_{\uparrow\downarrow}^{(s)}(\mathbf{x}) \quad \left(\chi_{\uparrow\downarrow}^{(s)}(\mathbf{x}) \equiv (\chi_{\uparrow}(x_1)\chi_{\downarrow}(x_2) + \chi_{\downarrow}(x_1)\chi_{\uparrow}(x_2))/\sqrt{2} \right)$$

$$\bar{\Psi}_4 = \bar{\Phi}_{12}^{(s)}(\mathbf{x}) \chi_{\uparrow\uparrow}^{(a)}(\mathbf{x})$$

where $\chi_{\uparrow\downarrow}^{(s)} \equiv \frac{1}{\sqrt{2}} (\chi_{\uparrow}(x_1)\chi_{\downarrow}(x_2) + \chi_{\downarrow}(x_1)\chi_{\uparrow}(x_2))$

$$\chi_{\uparrow\downarrow}^{(a)} \equiv \frac{1}{\sqrt{2}} (\quad \quad \quad - \quad \quad \quad)$$

in the 4-dim space

Therefore, when we focus on the spin degrees of freedom, the effective Hamiltonian is simply

$$\mathcal{H}_{\text{eff}} = \begin{cases} 0 & (\text{if spin wave function is symmetric}) \\ -J & (\text{" " " asymmetric}) \end{cases}$$

with $J = +4t^2/u$.

Here, we notice the symmetric/asymmetric 2-spin wave function is nothing but the triplet/singlet wave functions. In addition, $\mathcal{S}_1 \cdot \mathcal{S}_2 = 1/4$ and $-3/4$ for the triplet and the singlet spin states respectively. Therefore, we can simply express the effective Hamiltonian as

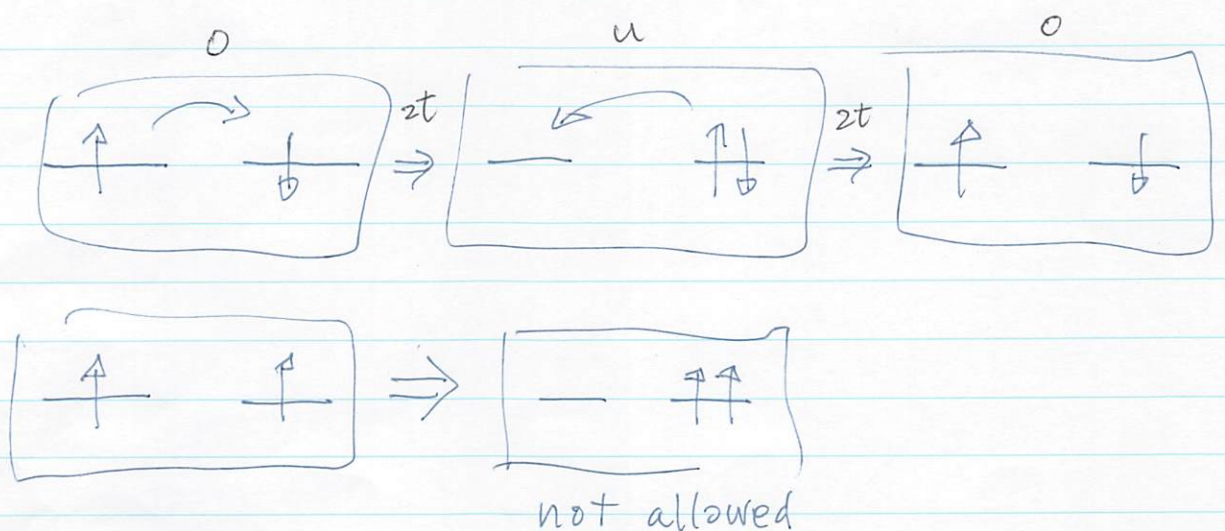
$$\mathcal{H}_{\text{eff}} = +J \mathcal{S}_1 \cdot \mathcal{S}_2 - \frac{3}{4} J$$

⤴ this is usually omitted since it simply shifts the energy

Hand-waving interpretation of the result

We can understand the antiferromagnetic interaction derived above as follows:

If each atom is occupied by an electron, each electron can only virtually hop to the neighbor sites. This hopping contributes to the kinetic energy of electrons, therefore, it reduces the total energy. However, if the neighbor site is already occupied by another electron with the same spin, the electron cannot hop there. Therefore, to gain the kinetic energy, electrons at nearest neighbor sites tend to have opposite spins.



$$\left(\begin{array}{l} \text{2nd order perturbation} \\ \Delta = -\frac{(2t)^2}{u} \end{array} \right)$$

Ferromagnetism

The above arguments only explains the antiferromagnetic interaction.

This is due to the simplification we made on the Coulomb interaction.

While we considered only on-site repulsion, the Coulomb interaction affects the nearest neighbor, next nearest neighbor, and so on.

So in general circumstances, the parallel alignment of spins may be more favorable than the antiparallel one because parallel alignment "forces" electrons farther apart, which may result in smaller contribution from the Coulomb term, which decays as a function of the distance.

We'll not go into the detailed discussion of the ferromagnetism. We'll simply assume that the exchange interaction J can be negative (ferromagnetic) in certain conditions.

- In spite of the rather technical appearance, the implication of the above arguments is quite general. In the above arguments, the Pauli principle breaks equivalence of spin-configurations giving rise to spin-spin correlation. However, the $SU(2)$ invariance of the original Hamiltonian demands the same invariance for the effective Hamiltonian. This story applies to more realistic and complicated cases.

- In particular, if we consider an effective spin Hamiltonian for the system in which one electron is bound to every atom and its spin interacts only with its nearest neighbors, the only possible effective spin Hamiltonian is

$$\mathcal{H} = -J \sum_{\langle ij \rangle} \mathbb{S}_i \cdot \mathbb{S}_j \quad \left(\begin{array}{l} S = 1/2 \\ \text{Heisenberg model} \end{array} \right)$$

where summation is taken over all pairs of nearest neighbors, and \mathbb{S}_i is the spin of the electron bound to the i -th atom.

- The spin anisotropy such as

$$\mathbb{S}_i \cdot \mathbb{S}_j \rightarrow \mathbb{S}_i \cdot \mathbb{S}_j + \alpha S_i^z S_j^z$$

is not allowed because it violates the symmetry, although it can arise through the spin-orbit coupling, a relativistic effect, which we'll not discuss in this course.

o The ground state of the ferromagnetic Heisenberg model

The ground state of the ferromagnetic Heisenberg model is the "all-up" state, i.e.,

$$|\bar{\Psi}\rangle \equiv |\uparrow\uparrow\uparrow \dots \uparrow\rangle = \bigotimes_i |\uparrow\rangle_i$$

or $S_i^z |\bar{\Psi}\rangle = +\frac{1}{2} |\bar{\Psi}\rangle$ for all i .

We can easily confirm that this state is a ground state by observing that this state "saturates" all the pair Hamiltonian at the same time;

$$-J \mathbf{S}_i \cdot \mathbf{S}_j |\bar{\Psi}\rangle = -\frac{J}{4} |\bar{\Psi}\rangle \quad (J > 0 \text{ for any } i \text{ and } j)$$

where $-J/4$ is the smallest eigenvalue of the operator $-J \mathbf{S}_i \cdot \mathbf{S}_j$. (In general, for a Hamiltonian of the form $\mathcal{H} = \sum_i \mathcal{H}_i$, if a state $|\bar{\Psi}\rangle$ is the ground state of all \mathcal{H}_i 's, then $|\bar{\Psi}\rangle$ must be the ground state of \mathcal{H} .)

Obviously, any global $SU(2)$ -rotation of $|\bar{\Psi}\rangle$ is also a ground state of the ferromagnetic Heisenberg model.