Quantum Theory of Spins

Jung Hoon Han

April 3, 2012

Contents

Chapter 1

Miscellaneous

1.1 Two-state Problem

All two-state quantum-mechanical problems can be mapped onto the Hamiltonian

$$H = \hat{h} \cdot \sigma = \begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix}, \tag{1.1}$$

with $\hat{h} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ denoting the spin orientation. This is diagonalized by the unitary rotation U^+HU with

$$U = \begin{pmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2)e^{i\phi} & \cos(\theta/2)e^{i\phi} \end{pmatrix} \tag{1.2}$$

implying that the two eigenstates are

$$\begin{aligned} |1\rangle &= \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2)e^{i\phi} \end{pmatrix}, \quad E_1 = +1 \\ |2\rangle &= \begin{pmatrix} -\sin(\theta/2) \\ \cos(\theta/2)e^{i\phi} \end{pmatrix}, \quad E_1 = -1. \end{aligned} \tag{1.3}$$

1.2 d-orbitalogy

The five d-orbitals are classified using the spherical harmonics

$$\begin{aligned} Y_{22}(\theta,\phi) &= 3\sqrt{\frac{5}{96\pi}}\sin^2\theta e^{2i\phi} = |2\rangle \\ Y_{21}(\theta,\phi) &= -3\sqrt{\frac{5}{24\pi}}\sin\theta\cos\theta e^{i\phi} = |1\rangle \\ Y_{20}(\theta,\phi) &= \sqrt{\frac{5}{4\pi}}\left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right) = |0\rangle \\ Y_{2\bar{1}}(\theta,\phi) &= 3\sqrt{\frac{5}{24\pi}}\sin\theta\cos\theta e^{-i\phi} = |\bar{1}\rangle \\ Y_{2\bar{2}}(\theta,\phi) &= 3\sqrt{\frac{5}{96\pi}}\sin^2\theta e^{-2i\phi} = |\bar{2}\rangle. \end{aligned} \tag{1.4}$$

Three t_{2g} orbitals are

$$\begin{aligned} |xy\rangle &= \frac{i}{\sqrt{2}}(|\overline{2}\rangle - |2\rangle) \sim \sin^2\theta \sin\phi\cos\phi \sim xy, \\ |yz\rangle &= \frac{i}{\sqrt{2}}(|1\rangle + |\overline{1}\rangle) \sim \sin\theta\cos\theta\sin\phi \sim yz, \\ |zx\rangle &= \frac{1}{\sqrt{2}}(|\overline{1}\rangle - |1\rangle) \sim \sin\theta\cos\theta\cos\phi \sim zx. \end{aligned} \tag{1.5}$$

The two e_g orbitals are

$$\begin{aligned} |x^2-y^2\rangle &= \frac{1}{\sqrt{2}}(|2\rangle + |\overline{2}\rangle) \sim \sin^2\theta(\cos^2\phi - \sin^2\phi) \sim x^2 - y^2, \\ |3z^2-r^2\rangle &= |0\rangle \sim 3z^2 - r^2. \end{aligned} \tag{1.6}$$

1.3 p-orbitalogy

The three p-orbitals are classified using the spherical harmonics

$$\begin{aligned} Y_{11}(\theta,\phi) &= -\sqrt{\frac{3}{8\pi}}\sin\theta e^{i\phi} = |1\rangle \\ Y_{10}(\theta,\phi) &= \sqrt{\frac{3}{4\pi}}\cos\theta = |0\rangle \\ Y_{1\overline{1}}(\theta,\phi) &= \sqrt{\frac{3}{8\pi}}\sin\theta e^{-i\phi} = |\overline{1}\rangle. \end{aligned} \tag{1.7}$$

x, y, z-orbitals are given by

$$\begin{aligned} |x\rangle &= \frac{1}{\sqrt{2}}(|\overline{1}\rangle - |1\rangle) \\ |y\rangle &= \frac{i}{\sqrt{2}}(|1\rangle + |\overline{1}\rangle) \\ |z\rangle &= |0\rangle. \end{aligned} \tag{1.8}$$

Chapter 2

Path Integral for Spins

2.1 Single spin

In the coherent state representation of spins, the spin state is expressed by a spinor

$$|\mathbf{n}\rangle = \mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \quad z_1 = e^{-i\phi/2}\cos\frac{\theta}{2}, \; z_2 = e^{i\phi/2}\sin\frac{\theta}{2}. \tag{2.1}$$

The angles are those of the classical spin vector $\mathbf{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. The path integral for spin contains the term

$$\langle\dot{\mathbf{n}}|\mathbf{n}\rangle = \dot{\mathbf{z}}^\dagger\mathbf{z} = \frac{i}{2}\dot{\phi}\cos\theta. \tag{2.2}$$

For an arbitrary spin length I there is a Berry phase contribution to the action e^{-iS_B} with

$$S_B = I\int d^2\mathbf{r}dt\,(1-\cos\theta)\,\dot{\phi}. \tag{2.3}$$

The term $\int dt(1-\cos\theta)\dot{\phi}$ gives the solid angle subtended by the unit vector over time t. Equation of motion for spin follows from varying the Berry action

$$\frac{\delta S_B}{\delta\mathbf{n}} = I\dot{\mathbf{n}}\times\mathbf{n}. \tag{2.4}$$

2.2 Antiferromagnetic spin chain

For a spin chain the Berry phase term is a collection of spin-spin Berry phase action

$$S_{1\mathrm{D}}^{\mathrm{AF}} = \sum_i S_B[\mathbf{n}_i] \tag{2.5}$$

Owing to the staggered nature of spins we can write $\mathbf{n}_i = (-1)^i \mathbf{L}_i + \mathbf{M}_i$, assuming $|\mathbf{L}_i| \gg |\mathbf{M}_i|$. We can expand the Berry action as

$$\sum_i S_B[(-1)^i \mathbf{L}_i + \mathbf{M}_i] \approx \sum_i S_B[(-1)^i \mathbf{L}_i] + \sum_i \mathbf{M}_i \cdot \frac{\delta \mathbf{S}_B}{\delta \mathbf{n}_i}\Big|_{\mathbf{n}_i=(-1)^i \mathbf{L}_i}. \quad (2.6)$$

When the spin direction is reversed, the sign of the Berry phase is also reversed, $S_B[(-1)^i \mathbf{L}_i] = (-1)^i S_B[\mathbf{L}_i]$. We get

$$S_{1\mathrm{D}}^{\mathrm{AF}} \approx \sum_i (-1)^i S_B[\mathbf{L}_i] + I \sum_i \mathbf{M}_i \cdot \dot{\mathbf{L}}_i \times \mathbf{L}_i. \quad (2.7)$$

In the case of the one-dimensional chain one can manipulate the first term in the Berry phase action a little more. For a chain of length N (N=even) the first term can be re-arranged as

$$\sum_{i=1}^{N/2} (S_B[\mathbf{L}_{2i}] - S_B[\mathbf{L}_{2i-1}]) \approx \sum_{i=1}^{N/2} \frac{\delta S_B}{\delta \mathbf{L}} \cdot (\mathbf{L}_{2i} - \mathbf{L}_{2i-1}) \approx \frac{I}{2} \sum_i \dot{\mathbf{L}}_i \times \mathbf{L}_i \cdot \partial_x \mathbf{L}_i. \quad (2.8)$$

In the final expression all sites are included in the sum. The continuum limit of the Berry phase action for 1D AF chain reads

$$S_{1\mathrm{D}}^{\mathrm{AF}} \approx \frac{I}{2} \int \mathbf{L} \cdot \partial_x \mathbf{L} \times \partial_t \mathbf{L} + I \int \mathbf{M} \cdot \dot{\mathbf{L}} \times \mathbf{L}. \quad (2.9)$$

What about the Heisenberg part? By using the same strategy one can re-write the Heisenberg spin interaction as

$$\begin{aligned} H &= J \sum_i \mathbf{n}_i \cdot \mathbf{n}_{i+1} = J \sum_i \left((-1)^i \mathbf{L}_i + \mathbf{M}_i\right) \left(-(-1)^i \mathbf{L}_{i+1} + \mathbf{M}_{i+1}\right) \\ &\approx \frac{J}{2} \sum_i (\mathbf{L}_{i+1} - \mathbf{L}_i)^2 + J \sum_i \mathbf{M}_i^2 \\ &\to \frac{J}{2} \int \sum_\mu (\partial_\mu \mathbf{L})^2 + J \int \mathbf{M}^2. \end{aligned} \quad (2.10)$$

Chapter 3

1D Quantum Spin Chain

3.1 The model

The one-dimensional quantum $S = 1/2$ spin model on a periodic lattice of length N is given by

$$H = J_\perp \sum_n S_n \cdot S_{n+1} + J_z \sum_n S_n^z S_{n+1}^z. \tag{3.1}$$

3.2 Ferromagnetic case

Ground state: For $J_z < 0$ and $|J_z|/J_\perp$ sufficiently large, the ground state is the ferromagnetic one $|F\rangle$ with all the spins up. Its energy $H|F\rangle = (J_z/4)|F\rangle$ is considered as zero.

3.3 Antiferromagnetic case

The general solution of Eq. (3.1) with $J_\perp = J_z = J$ which is an eigenstate of $S^z = \sum_j S_j^z$ is

$$\begin{aligned} |\psi\rangle &= \sum_{1 \le n_1 < n_2 < \cdots < n_M \le N} \phi(n_1, n_2, \cdots, n_M) |n_1, n_2, \cdots n_M\rangle, \\ |n_1, n_2, \cdots n_M\rangle &= S_{n_1}^- \cdot S_{n_M}^- |F\rangle. \end{aligned} \tag{3.2}$$

The ansatz solution that diagonalizes the isotropic spin Hamiltonian is

$$\phi(n_1, \cdots, \cdot n_M) = \sum_P A_P \exp\left(i \sum_{j=1}^{M} k_{P(j)} n_j \right). \tag{3.3}$$

Here $k_{P(j)}$ spans all permutations of the set $\{k_1, k_2, \cdots k_M\}$, $P(\{k_1, k_2, \cdots k_M\}) = \{k_{P(1)}, k_{P(2)}, \cdots k_{P(M)}\}$, and A_P is a coefficient that depends on each permutation. The k_j's are the pseudo-momenta for each of the M down spins, and the total energy and momenta in terms of the pseudo-momenta is

$$E = J\sum_{j=1}^{M}(\cos k_j - 1), \qquad P = \sum_{j=1}^{M} k_j. \tag{3.4}$$

The coefficients A_P and the set of allowed pseudo-momenta $\{k_j\}$ can be derived by substitution of Eq. (3.3) into Eq. (3.1). One can separate the Hamiltonian as two pieces $H = H_\perp + H_z$ where

$$\begin{aligned} H_\perp &= \frac{1}{2}J\sum_{n=1}^{N}(S_n^- S_{n+1}^+ + S_n^+ S_{n+1}^-) \\ H_z &= J\sum_n S_n^z S_{n+1}^z. \end{aligned} \tag{3.5}$$

The action of $H_\perp$ on $|\psi\rangle$ is

$$\frac{1}{2}J\sum_{\{n_j\}}\phi(n_1,\cdots,n_M)\sum_{\{n_j'\}}|\{n_j'\}\rangle. \tag{3.6}$$

Here $|\{n_j'\}\rangle$ refers to all set of configurations related to the original configuration $|\{n_j\}\rangle$ by the reversal of the nearest pair of up and down spins. For example, if a given basis configuration is $|3,6\rangle$ (spin down position at 3 and 6), the set of $|\{n_j'\}\rangle$ is $|2,6\rangle, |4,6\rangle, |3,5\rangle, |3,1\rangle$. The last configuration arises in the periodic boundary condition geometry. The total number of configurations $\{n_j'\}$ arising from $\{n_j\}$ equals the the total number of spin up-down pairs, or the number of kinks. The action of H_z gives a numerical factor proportional to the number of kinks,

$$H_z|\{n_j\}\rangle = -\frac{1}{2}J \times (\text{number of kinks})|\{n_j\}\rangle. \tag{3.7}$$

The Schrodinger equation becomes ($J = 2$)

$$E\phi(\{n\}) = \sum_{\{n'\}}\left(\phi(\{n'\}) - \phi(\{n\})\right). \tag{3.8}$$

One-spin down state: The one-spin down state, $M = 1$, has Eq. (3.8) reduced to

$$\phi(n+1) + \phi(n-1) - 2\phi(n) = E\phi(n). \tag{3.9}$$

Ansatz $c(n) = Ae^{ikn}$ solves the equation with the energy given by

$$E_k = 2(\cos k - 1). \tag{3.10}$$

Imposing the periodic boundary condition $\phi(n+N) = \phi(n)$ quantizes the k values

$$k_j = \frac{2\pi}{N}I_j \tag{3.11}$$

with the allowed quantum number $I_j = 1, \cdots N$.

Two-spin down states: For $M = 2$, the case of $n_2 = n_1 + 1$ needs to be treated separately from the others. If $n_2 = n_1 + 1$, EoM gives

$$E\phi(n_1, n_1 + 1) = -2\phi(n_1, n_1 + 1) + \phi(n_1 - 1, n_1 + 1) + \phi(n_1, n_1 + 2). \quad (3.12)$$

Otherwise EoM gives

$$\begin{aligned} &E\phi(n_1, n_2) = -4\phi(n_1, n_2) \\ &+\phi(n_1 - 1, n_2) + \phi(n_1 + 1, n_2) + \phi(n_1, n_2 - 1) + \phi(n_1, n_2 + 1). \end{aligned} \quad (3.13)$$

Suppose Eq. (3.13) can be analytically continued to $n_2 = n_1 + 1$. Then subtracting out Eq. (3.12) from Eq. (3.13) for $n_1 = n, n_2 = n + 1$ would give

$$2\phi(n, n + 1) = \phi(n, n) + \phi(n + 1, n + 1). \quad (3.14)$$

One can try the solution

$$\phi(n_1, n_2) = A_1 e^{ik_1 n_1 + ik_2 n_2} + A_2 e^{ik_2 n_1 + ik_1 n_2}. \quad (3.15)$$

If Eq. (3.13) were valid everywhere, we would immediately conclude that Eq. (3.15) is the solution with the energy

$$E = 2\sum_{j=1}^{2}(\cos k_j - 1). \quad (3.16)$$

Substituting the ansatz, Eq. (3.15), into the constraint, Eq. (3.14), gives

$$-\frac{A_1}{A_2} = \frac{e^{ik_1+ik_2} + 1 - 2e^{ik_1}}{e^{ik_1+ik_2} + 1 - 2e^{ik_2}} = \frac{e^{i\frac{k_1-k_2}{2}} - \cos\frac{k_1+k_2}{2}}{e^{i\frac{k_2-k_1}{2}} - \cos\frac{k_1+k_2}{2}}. \quad (3.17)$$

The ratio $|A_1/A_2|$ is one, hence one can write $A_1 = e^{i\theta_{12}/2}$ and $A_2 = e^{i\theta_{21}/2}$, $\theta_{21} = -\theta_{12}$, and show that θ_{12} in terms of k_1, k_2 is obtained from

$$\cot\frac{\theta_{12}}{2} = \frac{\sin\frac{k_1-k_2}{2}}{\cos\frac{k_1+k_2}{2} - \cos\frac{k_1-k_2}{2}}. \quad (3.18)$$

The two particle solution becomes

$$\phi(n_1, n_2) = e^{i\theta_{12}/2 + i(k_1 n_1 + k_2 n_2)} + e^{-i\theta_{12}/2 + i(k_2 n_1 + k_1 n_2)}. \quad (3.19)$$

Imposing the periodic boundary condition becomes tricky for two particles and beyond. We choose the condition that $\phi(n_1, n_2) = \phi(n_2, n_1 + N)$, which renders

$$e^{ik_1 N} = e^{i\theta_{12}}, \quad e^{ik_2 N} = e^{i\theta_{21}}. \quad (3.20)$$

The quantum numbers are quantized accordingly,

$$k_1 N = 2\pi I_1 + \theta_{12}, \quad k_2 = 2\pi I_2 + \theta_{21}, \quad I_i = 1, \cdots, N. \quad (3.21)$$

M-spin down state: We return to the general, M-spin down state with the ansatz solution given by

$$\phi(n_1, \cdots, \cdot n_M) = \sum_P A_P \exp\left(i \sum_{j=1}^{M} k_{P(j)} n_j \right). \tag{3.22}$$

The general form of the Schrodinger equation becomes more explicitly

$$\sum_{j=1}^{M} \Big(\phi(n_1, \cdots, n_j - 1, \cdots, n_M) + \phi(n_1, \cdots, n_j + 1, \cdots, n_M) - 2\phi(n_1, \cdots, n_M) \Big) = E\phi(n_1, \cdots, n_M). \tag{3.23}$$

The terms which contains two adjacent spins of the same orientation are handled by the constraints

$$\phi(\cdots, n_j, n_j, \cdots) + \phi(\cdots, n_j + 1, n_j + 1, \cdots) = 2\phi(\cdots, n_j, n_j + 1, \cdots). \tag{3.24}$$

Inserting the M-particle wave function into this constraint yields the relation

$$\sum_P \mathcal{A}_P + \sum_P \mathcal{A}_P e^{ik_{P(j)} + ik_{P(j+1)}} = 2 \sum_P \mathcal{A}_P e^{ik_{P(j+1)}},$$
$$\mathcal{A}_P = A_P \exp\left(i \sum_{l \neq j, j \neq j+1} n_l k_{P(l)} + n_j k_{P(j)} + n_j k_{P(j+1)} \right). \tag{3.25}$$

It can be seen that only a pair of coefficients $\mathcal{A}_P$ and $\mathcal{A}_{P(j,j+1)}$ need to be involved in the equality. I define $P(j, j+1)$ as the original permutation P plus the interchange of $k_{P(j)}$ and $k_{P(j+1)}$. Hence, Eq. (**??**) reduces to the pair-wise equaiton

$$A_P + A_{P(j,j+1)} + (A_P + A_{P(j,j+1)}) e^{ik_{P(j)} + ik_{P(j+1)}} = 2(A_P e^{ik_{P(j+1)}} + A_{P(j,j+1)} e^{ik_{P(j)}})$$
$$-\frac{A_{P(j,j+1)}}{A_P} \equiv -e^{-i\theta_{P(j),P(j+1)}} = \frac{e^{ik_{P(j)} + ik_{P(j+1)}} + 1 - 2e^{ik_{P(j+1)}}}{e^{ik_{P(j)} + ik_{P(j+1)}} + 1 - 2e^{ik_{P(j)}}}. \tag{3.26}$$

The relation is satisfied by the ansatz

$$A_P = e^{i \sum_{i<j} \theta_{P(i),P(j)}/2}, \quad -e^{-i\theta_{P(i),P(j)}} = \frac{e^{ik_{P(i)} + ik_{P(j)}} + 1 - 2e^{ik_{P(j)}}}{e^{ik_{P(i)} + ik_{P(j)}} + 1 - 2e^{ik_{P(i)}}}. \tag{3.27}$$

The fact that A_P must be of the form given above is easily illustrated with an example $M = 4$. In that case, the phase factor for A_P reads $(1/2)$ times

$$\theta_{P(1),P(2)} + \theta_{P(1),P(3)} + \theta_{P(1),P(4)} + \theta_{P(2),P(3)} + \theta_{P(2),P(4)} + \theta_{P(3),P(4)}. \tag{3.28}$$

The phase factor for $A_{P(1,2)}$, for instance, is given by 1 and 2 in the above equation replaced by 2 and 1. An inspection reveals that the net difference only occurs in replacing $\theta_{P(1),P(2)}$ by $\theta_{P(2),P(1)}$, hence the ratio $A_{P(1),P(2)}/A_P$ is indeed equal to $e^{-i\theta_{P(1),P(2)}}$. The same result obtains for interchange of any other consecutive indices j and $j+1$.

Following the earlier practice we impose the periodic boundary condition

$$\phi(n_1,\cdots,n_M) = \phi(n_2,\cdots,n_M,n_1+N) \rightarrow$$

$$\sum_P A_P e^{i\sum_j k_{P(j)}n_j} = \sum_P A_P e^{ik_{P(1)}n_2+\cdots+ik_{P(M-1)}n_M+ik_{P(M)}(n_1+N)}. \tag{3.29}$$

One can define a permutation P_0 which takes

$$P_0(1) = M, \quad P_0(j+1) = j \ (j>1). \tag{3.30}$$

Then the r.h.s. of Eq. (3.29) becomes

$$\sum_P A_P e^{ik_{PP_0(2)}n_2+\cdots+ik_{PP_0(M)}n_M+ik_{PP_0(1)}(n_1+N)} =$$
$$\sum_P A_{PP_0^{-1}} e^{i\sum_j k_{P(j)}n_j+ik_{P(1)}N} = \sum_P A_P e^{i\sum_j k_{P(j)}n_j}. \tag{3.31}$$

From the last equality it follows that

$$A_P = A_{PP_0^{-1}} e^{ik_{P(1)}N}. \tag{3.32}$$

The amplitudes are written out explicitly

$$\begin{aligned} A_P &= \exp\left(\frac{i}{2}\sum_{i<j}\theta_{P(i),P(j)}\right) \\ A_{PP_0^{-1}} &= \exp\left(\frac{i}{2}\sum_{i<j}\theta_{PP_0^{-1}(i),PP_0^{-1}(j)}\right). \end{aligned} \tag{3.33}$$

As long as $j < M$, $\theta_{PP_0^{-1}(i),PP_0^{-1}(j)}$ gives another element belonging to the family $\theta_{P(i),P(j)}$ with $i<j$. When $j=M$, one gets

$$\theta_{PP_0^{-1}(i),PP_0^{-1}(M)} = \theta_{P(i+1),P(1)} = -\theta_{P(1),P(i+1)}. \tag{3.34}$$

It follows that Eq. (3.32) becomes

$$e^{ik_{P(1)}N} = e^{i\sum_{i=1}^{M-1}\theta_{P(1),P(i+1)}}, \tag{3.35}$$

or

$$e^{ik_jN} = e^{i\sum_{l\neq j}\theta_{jl}}, \tag{3.36}$$

Chapter 4

Hubbard Model

4.1 The Model

In the previous chapter we learned how to analyze electron motion in solids by employing a simple model - the tight-binding model. An easy generalization of the model also allowed us to study the disordered system. While disorder is an essential aspect of all real materials, there is an equally or in some instances even more important aspect of the many-particle system. And that is the interaction between the particles. For electrons, the interaction is due to the Coulomb force. Hubbard model, invented by John Hubbard in the 1950's, is the most representative of the models dealing with electron-electron interaction. It is written as the Hamiltonian

$$\begin{aligned} H &= T + V \\ T &= \sum_{\langle ij\rangle\sigma} t_{ij}\left(c^+_{j\sigma}c_{i\sigma} + c^+_{i\sigma}c_{j\sigma}\right) - \mu\sum_{i\sigma} c^+_{i\sigma}c_{i\sigma} \\ V &= U\sum_i n_{i\uparrow}n_{i\downarrow} \end{aligned} \tag{4.1}$$

T represents hopping of electron between a pair of sites $\langle ij\rangle$ with hopping amplitude $t_{ij} = t_{ji}$. For simplicity we take $t_{ij} = -t$ for all the bonds. Unlike in the previous chapter we include electron spin σ explicitly because, after all, real electrons do carry spin. Diagonalizing T in the absence of the interaction V leads to $\sum_{k\sigma}\epsilon_k c^+_{k\sigma}c_{k\sigma}$, where ϵ_k is the energy of the eigenstate $|k\rangle$. Thus T alone describes delocalized electrons that form a band inside the solid. This has been the essence of the discussion of previous chapter.

Let's examine the properties of the V-term. First it is completely local, *i.e.* it only acts at a given site i: $n_{i\sigma} = c^+_{i\sigma}c_{i\sigma}$ (no sum on σ) counts the number of electrons of spin σ at site i. This being an electron model, $n_{i\sigma}$ can only be 0 or 1. When a given site i is occupied by zero electron ($n_{i\uparrow} = n_{i\downarrow} = 0$) or by only one electron of spin σ ($n_{i\sigma} = 1$, $n_{i\bar{\sigma}} = 0$, $n_{i\uparrow}n_{i\downarrow} = 0$), V has an expectation value of zero. When the site is doubly occupied, *i.e.* $n_{i\uparrow} = n_{i\downarrow} = 1$, the site has an energy increase of U. We may therefore understand U as the Coulomb repulsion energy of the two electrons. The challenge we must face now is to understand the properties of, and to solve, the model Hamiltonian (4.1) in the simultaneous presence of T and V.

state	n_i	E
$\|0\rangle$	0	0
$c_{i\uparrow}^\dagger\|0\rangle$, $c_{i\downarrow}^\dagger\|0\rangle$	1	$-\mu$
$c_{i\downarrow}^\dagger c_{i\uparrow}^\dagger\|0\rangle$	2	$U-2\mu$

Table 4.1: Possible atomic states and their energies

In the atomic limit $t_{ij}\to 0$, the Hamiltonian reduces to the atomic form $H=\sum_i H_i$, $H_i=-\mu n_i+Un_{i\uparrow}n_{i\downarrow}$, that lacks communication between neighboring sites. Diagonalizing H_i is trivial since each fixed occupation of electron gives an eigenstate. Eigenstates and their corresponding energies are

Suppose the chemical potential is chosen to satisfy $0<\mu<\frac{U}{2}$. Then among the three possible eigenenergies, only the $-\mu$, corresponding to single occupation, lies below zero energy. According to the Fermi factor $F(E)=\left(e^{\beta E}+1\right)^{-1}$ that determines the occupation probability of a given eigenstate, only the $n_i=1$ state will occur at zero temperature. That means each site of the lattice will be occupied by one, and only one electron at zero temperature. Since both spin states give rise to the same energy, each electron occupying the site can be either spin up or spin down. For N electrons occupying N different lattice sites, this leads to enormous ground state degeneracy $2\times 2\times\cdots\times 2=2^N$. In other words, any state of the spins given by $|\sigma_1\sigma_2\cdots\sigma_N\rangle$, $\sigma_i=\pm 1$, is an eigenstate having the same energy $-\mu N$.

Next we raise the question: when we add the hopping part T to the Hamiltonian will the eigenstates of V alone remain eigenstates, and will their energies differ due to the hopping effect?

Before we answer these equations we re-visit the foundations of quantum mechanics. The basic postulates of quantum mechanics is that there are a set of states, labelled $|m\rangle$, that together form a complete, orthonormal set, $\sum_m |m\rangle\langle m|=1$, $\langle m|n\rangle=\delta_{mn}$. Each $|m\rangle$ is not necessarily an eigenstate of H, but due to the completeness any eigenstate of H can be written as a linear combination $|\psi\rangle=\sum_m c_m|m\rangle$,

$$H\left(\sum_m c_m|m\rangle\right)=E\left(\sum_m c_m|m\rangle\right). \tag{4.2}$$

By taking projection on both sides with $\langle n|$, we get

$$\sum_m c_m\langle n|H|m\rangle=Ec_n. \tag{4.3}$$

We will call matrix element $\langle n|H|m\rangle=V_{nm}$, then we have a linear matrix equation

$$\sum_m V_{nm}c_m=Ec_n. \tag{4.4}$$

By diagonalizing the matrix V_{nm}, the eigenvectors $\{c_m\}$ are obtained. Then the eigenstate of the Hamiltonian is expressed as $|\psi\rangle=\sum_m c_m|m\rangle$, having the energy E.

We apply precisely this strategy to solve the Hubbard model. And before doing so, we must first identify the basis states to be used in the diagonalization

process. For a given site i there are four possible states, listed in Table 4.1, corresponding to zero, one (of spin up and down), and two electron occupation, which we will label $|0_i\rangle$, $|\sigma_i\rangle$, and $|d_i\rangle$ in obvious notation. When there are N sites, the basis states are given by

$$|\alpha_1\rangle|\alpha_2\rangle \cdots |\alpha_N\rangle. \tag{4.5}$$

where each α_i is one of the four possible states. Working out V_{nm} corresponds to calculating the overlap of H between two arbitrary basis states (4.5). For N sites, there are 4^N such basis states, and the size of the matrix V_{nm} is $4^N \times 4^N$. Even for $N=10$, the dimension of the matrix is $4^{10} \times 4^{10} \simeq 10^6 \times 10^6$!! Diagonalizing such a huge matrix is impossible, even with the world's best computer available right now. In this chapter we will try to diagonalize H up to $N=4$.

To get a flavor of the steps needed to diagonalize H, we start with the two-site Hubbard model

$$H = Un_{1\uparrow}n_{1\downarrow} + Un_{2\uparrow}n_{2\downarrow} - \mu \sum_\sigma (n_{1\sigma} + n_{2\sigma}) - t \sum_\sigma \left(c^+_{2\sigma}c_{1\sigma} + c^+_{1\sigma}c_{2\sigma}\right). \tag{4.6}$$

First there are sixteen basis states given as the direct product

$$\{|0_1\rangle, |\uparrow_1\rangle, |\downarrow_1\rangle, |d_1\rangle\} \otimes \{|0_2\rangle, |\uparrow_2\rangle, |\downarrow_2\rangle, |d_2\rangle\}. \tag{4.7}$$

We know how to work out the matrix element of T already. Since T conserves the total spin, nonzero element only occurs between states such as

$$|\sigma_1, 0_2\rangle \leftrightarrow |0_1, \sigma_2\rangle, \quad |d_1, 0_2\rangle \leftrightarrow |\sigma_1, \overline{\sigma}_2\rangle, \quad \text{or} \quad |d_1, \sigma_2\rangle \leftrightarrow |\overline{\sigma}_1, d_2\rangle. \tag{4.8}$$

You should carefully work out the matrix element for each of these.

Next, the interaction term V has a nonzero element if and only if a site is doubly occupied. Thus

$$\begin{aligned}
&\langle d_1, d_2|V|d_1, d_2\rangle = U + U = 2U \\
&\langle d_1, \alpha_2|V|d_1, \alpha_2\rangle = U \quad (\alpha_2 \neq d_2) \\
&\langle \alpha_1, d_2|V|\alpha_1, d_2\rangle = U \quad (\alpha_1 \neq d_1) \\
&\langle \alpha_1, \alpha_2|V|\alpha_1, \alpha_2\rangle = 0 \quad (\alpha_1 \neq d_1, \alpha_2 \neq d_2)
\end{aligned} \tag{4.9}$$

In this way we can work out all the matrix elements among 16 basis states. The 16×16 matrix can be diagonalized easily, and will yield 16 energy levels. The lowest energy and the corresponding eigenvector defines the ground state of the Hubbard Hamiltonian.

4.2 Mean-field Theory and Antiferromagnetism

4.2.1 Square lattice

The previous section described a strategy as to how one would exactly diagonalize the Hubbard Hamiltonian for a fixed size and particle number. Undoubtedly the process will have to stop and, with the present capacity of computing power, stop at a rather miserably level of the system size. The thermodynamic behavior

which often emerges only as the limit of the infinite system size is taken is hard to obtain by direct numerical diagonalization. If one has a good guess what the ground state of the model will be, on the other hand, one can often define a mean-field theory that easily produces such a ground state in a self-consistent manner. Below we show how one such phase, namely the antiferromagnetic ground state, is "predicted" from the Hubbard model.

First one rewrites the interaction term of the Hubbard model

$$H = -t\sum_{i,j,\sigma} c^+_{j\sigma}c_{j\sigma} + U\sum_i n_{i\uparrow}n_{i\downarrow} - \mu\sum_i n_i \tag{4.10}$$

using the simple identity

$$n_\uparrow n_\downarrow = \frac{1}{4}(n_\uparrow + n_\downarrow)^2 - \frac{1}{4}(n_\uparrow - n_\downarrow)^2 \equiv \frac{1}{4}n^2 - \frac{1}{4}m^2, \tag{4.11}$$

$$Un_{i\uparrow}n_{i\downarrow} = \frac{U}{4}(n_i^2 - m_i^2). \tag{4.12}$$

We will carry out the meanfield decoupling at half-filling, $\langle n_i\rangle = 1$.

$$\frac{U}{4}\left(n_i^2 - m_i^2\right) \to \frac{U}{2}\langle n_i\rangle n_i - \frac{U}{2}\langle m_i\rangle m_i = \frac{U}{2}n_i - \frac{U}{2}s_i m_i. \tag{4.13}$$

The average of the magnetic moment is denoted s_i. The meanfield Hubbard Hamiltonian is

$$H_{MF} = -t\sum_{ij\sigma} c^+_{j\sigma}c_{i\sigma} + \left(\frac{U}{2} - \mu\right)\sum_i n_i - \frac{U}{2}\sum_i s_i m_i. \tag{4.14}$$

To describe antiferromagnetism we assume the staggered field $s_i = (-1)^i s$ where s is the size of the magnetic moment at each site. At half-filling the effective chemical potential $\mu_{eff} = \mu - \frac{U}{2}$ must equal zero. Hence

$$H_{MF} = -t\sum_{ij\sigma} c^+_{j\sigma}c_{i\sigma} - \eta\sum_i (-1)^i m_i = \sum_{k\sigma}\epsilon_k c^+_{k\sigma}c_{k\sigma} - \eta\sum_{k\sigma}\sigma c^+_{k+Q\sigma}c_{k\sigma}, \tag{4.15}$$

where $\eta = Us/2$. The last line follows from Fourier transform. Q implies the AF wavevector (π,π)(2D) or (π,π,π)(3D). We divide the original Brillouin zone $[-\pi,\pi]\otimes[-\pi,\pi]$ into two: one covers the region bounded by the four lines $k_x + k_y = \pi$, $k_x + k_y = -\pi$, $k_x - k_y = \pi$, $k_x - k_y = -\pi$, and the other, the remaining part of the BZ. The first region, shaded in Fig. **??**, is the reduced Brillouin zone (RBZ).

Now one can rewrite the Hamiltonian

$$\begin{aligned} H_{MF} &= {\sum_{k\sigma}}' \left(\epsilon_k c^+_{k\sigma}c_{k\sigma} + \epsilon_{k+Q}c^+_{k+Q\sigma}c_{k+Q\sigma}\right) - \eta{\sum_{k\sigma}}'\sigma\left(c^+_{k+Q\sigma}c_{k\sigma} + c^+_{k\sigma}c_{k+Q\sigma}\right) \\ &= {\sum_{k\sigma}}' \begin{pmatrix} c^+_{k\sigma} & c^+_{k+Q\sigma}\end{pmatrix}\begin{pmatrix}\epsilon_k & -\eta\sigma \\ -\eta\sigma & \epsilon_{k+Q}\end{pmatrix}\begin{pmatrix} c_{k\sigma} \\ c_{k+Q\sigma}\end{pmatrix}. \end{aligned} \tag{4.16}$$

$\sigma = \pm 1$ corresponds to $\uparrow, \downarrow$ spins. The 2×2 matrix given above can be diagonalized exactly with the canonical transformation

$$\begin{pmatrix} c_{k\sigma} \\ c_{k+Q\sigma} \end{pmatrix} = \begin{pmatrix} \cos\theta_k & \sin\theta_k \\ -\sin\theta_k & \cos\theta_k \end{pmatrix} \begin{pmatrix} \gamma_{1k\sigma} \\ \gamma_{2k\sigma} \end{pmatrix}, \tag{4.17}$$

provided we take

$$\sin 2\theta_k = \frac{\sigma\eta}{E_k}, \quad \cos 2\theta_k = \frac{\epsilon_k - \epsilon_{k+Q}}{2E_k}, \quad E_k = \sqrt{\eta^2 + \frac{1}{4}(\epsilon_k - \epsilon_{k+Q})^2}. \tag{4.18}$$

In the standard cubic(square) lattice $\epsilon_{k+Q} = -\epsilon_k$, and thus $\sin 2\theta_k = \sigma\eta/E_k$, $\cos 2\theta_k = \epsilon_k/E_k$, $E_k = \sqrt{\epsilon_k^2 + \eta^2}$. Using this transformation the Hamiltonian is brought to a diagonalized form:

$$H = \sum_{k\sigma}{}' E_k[\gamma_{1\sigma}^+ \gamma_{1k\sigma} - \gamma_{2k\sigma}^+ \gamma_{2k\sigma}], \tag{4.19}$$

assume $\epsilon_k = -\epsilon_{k+Q}$. We have managed to reduce the Hamiltonian to the diagonalized form, and obtain the meanfield energies $\pm E_k$. Furthermore, we have explicitly derived a relation between the quasiparticle operators (γ) and the original, electron operators. The eigenvectors of the matrix act as a "connection" between the two operators. Since $E_k > 0$, all the k-states in the RBZ for the lower energy branch are occupied, and completely empty for the upper branch at zero temperature, in agreement with the original prescription of half-filling. The ground state is expressed as

$$|\mathrm{GS}\rangle = \prod_k{}' \gamma_{2k\uparrow}^+ \gamma_{2k\downarrow}^+ |0\rangle. \tag{4.20}$$

The inverse transformation

$$\begin{pmatrix} \gamma_{1k\sigma} \\ \gamma_{2k\sigma} \end{pmatrix} = \begin{pmatrix} \cos\theta_k & -\sin\theta_k \\ \sin\theta_k & \cos\theta_k \end{pmatrix} \begin{pmatrix} c_{k\sigma} \\ c_{k+Q\sigma} \end{pmatrix}, \tag{4.21}$$

with

$$\cos\theta_k = \sqrt{\frac{1}{2}\left(1 + \frac{\varepsilon_k}{E_k}\right)}, \quad \sin\theta_k = \sigma\sqrt{\frac{1}{2}\left(1 - \frac{\varepsilon_k}{E_k}\right)}, \tag{4.22}$$

gives

$$\begin{aligned} \gamma_{2k\sigma}^+ &= u_k c_{k\sigma}^+ + \sigma v_k c_{k+Q\sigma}^+, \\ u_k &= \sqrt{\frac{1}{2}\left(1 - \frac{\varepsilon_k}{E_k}\right)}, \quad v_k = \sqrt{\frac{1}{2}\left(1 + \frac{\varepsilon_k}{E_k}\right)}. \end{aligned} \tag{4.23}$$

The ground state is expressed as

$$|\mathrm{GS}\rangle = \prod_k \left(u_k c_{k\uparrow}^+ + v_k c_{k+Q\uparrow}^+\right)\left(u_k c_{k\downarrow}^+ - v_k c_{k+Q\downarrow}^+\right)|0\rangle. \tag{4.24}$$

The real-space wave function is a product of Slater determinants, $D[\phi_\uparrow(k,r)]D[\phi_\downarrow(k,r)]$:

$$\begin{aligned}\phi_\uparrow(k_i,r_j) &= e^{ik_i\cdot r_j}u_k + e^{i(k_i+Q)\cdot r_j}v_k \\ \phi_\downarrow(k_i,r_j) &= e^{ik_i\cdot r_j}u_k - e^{i(k_i+Q)\cdot r_j}v_k.\end{aligned} \tag{4.25}$$

The mean-field condition on η reads

$$\eta = U\eta \sum_k{}' \frac{1}{E_k}\tanh\left(\frac{E_k}{2T}\right). \tag{4.26}$$

This equation has an exact, $T=0$ solution in the case of the constant density of states $D(\epsilon)=1/(2D)$,

$$s = \frac{2D/U}{\sinh[2D/U]}, \tag{4.27}$$

The transition temperature T_N is given when $s=0$, $E_k=|\epsilon_k|$.

In this section I have given a rather detailed derivation of the gap equation for the antiferromagnetic ordering. The BCS problem is entirely similar to this in its mathematical structure. Technically, if you know how to solve the AF problem, then you are (almost) equally well prepared for the BCS or any other kind of mean-field problems of condensed matter physics.

The antiferromagnetic ordering at (π,π) introduced a spatial modulation in an initially translationally invariant Hamiltonian that led to (i) doubling of the unit cell and halving of the Brillouin zone and (ii) folding of the band into two subbands and a gap between the bands. Full occupation of the lower band and a complete absence in the upper band leads to the opening of the gap at half-filling.

4.2.2 Triangular lattice

For the triangular lattice the ordered spins no longer lie along a common axis. The general belief is that the ordered moments form a collinear state, with a 120° angle between adjacent pair of spins. To capture this in the mean-field theory one must first generalize the mean-field decoupling of the Hubbard interaction to the spin-rotation-invariant form

$$n_{i\uparrow}n_{i\downarrow} \to \frac{U}{2}\langle n_i\rangle n_i - \frac{U}{2}\langle \mathbf{s}_i\rangle\cdot\mathbf{s}_i = \frac{U}{2}n_i - \frac{U}{2}\mathbf{s}_i\cdot\mathbf{m}_i. \tag{4.28}$$

The average moment $\mathbf{m}_i$ are taken to lie within the plane, and it is useful to use the characterization in the complex notation

$$m_{ix} + im_{iy} = m_i = m_0 e^{i\mathbf{Q}\cdot\mathbf{r}_i} \tag{4.29}$$

Taking $\mathbf{Q}=(4\pi/3)\hat{x}$ produces the right spin orientation of the triangular lattice where the lattice sites are given by $\mathbf{r}_i = p\hat{x} + q(\hat{x}/2 + \sqrt{3}\hat{y}/2)$ for a pair of integer (p,q). The initial configuration m_0 is an arbitrary complex number of unit magnitude or less. The Zeeman term becomes

$$-\frac{U}{2}\sum_i (m_i^* s_i^- + m_i s_i^+) = -\frac{Um}{2}(s_{\mathbf{Q}}^- + s_{\mathbf{Q}}^+) = -\frac{Um}{2}\sum_k (c_{k+Q\uparrow}^+ c_{k\downarrow} + c_{k\downarrow}^+ c_{k+Q\uparrow}). \tag{4.30}$$

The Brillouin zone of the triangular lattice can be broken up into three pieces of equal area which are connected to each other by the translation of $+\mathbf{Q}$ or $-\mathbf{Q}$ in momentum space. The Hamiltonian within the reduced Brillouin zone becomes

$$
\begin{aligned}
&\sum_{k\sigma}{}' \left(\epsilon_k c^+_{k\sigma} c_{k\sigma} + \epsilon_{k+Q} c^+_{k+Q,\sigma} c_{k+Q,\sigma} + \epsilon_{k-Q} c^+_{k-Q,\sigma} c_{k-Q,\sigma} \right) \\
&-\eta \sum_{k}{}' \left(c^+_{k+Q\uparrow} c_{k\downarrow} + c^+_{k-Q\uparrow} c_{k+Q\downarrow} + c^+_{k\uparrow} c_{k-Q\downarrow} + h.c. \right) \\
&= \sum_{k}{}' \begin{pmatrix} c^+_{k\uparrow} \\ c^+_{k\downarrow} \\ c^+_{k+Q\uparrow} \\ c^+_{k+Q\downarrow} \\ c^+_{k-Q\uparrow} \\ c^+_{k-Q\downarrow} \end{pmatrix}^T \begin{pmatrix} \epsilon_k & 0 & 0 & 0 & 0 & -\eta \\ 0 & \epsilon_k & -\eta & 0 & 0 & 0 \\ 0 & -\eta & \epsilon_{k+Q} & 0 & 0 & 0 \\ 0 & 0 & 0 & \epsilon_{k+Q} & -\eta & 0 \\ 0 & 0 & -\eta & 0 & \epsilon_{k-Q} & 0 \\ -\eta & 0 & 0 & 0 & 0 & \epsilon_{k-Q} \end{pmatrix} \begin{pmatrix} c_{k\uparrow} \\ c_{k\downarrow} \\ c_{k+Q\uparrow} \\ c_{k+Q\downarrow} \\ c_{k-Q\uparrow} \\ c_{k-Q\downarrow} \end{pmatrix} \\
&= \sum_{k}{}' \begin{pmatrix} c^+_{k\uparrow} \\ c^+_{k-Q\downarrow} \\ c^+_{k+Q\uparrow} \\ c^+_{k\downarrow} \\ c^+_{k-Q\uparrow} \\ c^+_{k+Q\downarrow} \end{pmatrix}^T \begin{pmatrix} \epsilon_k & -\eta & 0 & 0 & 0 & 0 \\ -\eta & \epsilon_{k-Q} & 0 & 0 & 0 & 0 \\ 0 & 0 & \epsilon_{k+Q} & -\eta & 0 & 0 \\ 0 & 0 & -\eta & \epsilon_k & 0 & 0 \\ 0 & 0 & 0 & 0 & \epsilon_{k-Q} & -\eta \\ 0 & 0 & 0 & 0 & -\eta & \epsilon_{k+Q} \end{pmatrix} \begin{pmatrix} c_{k\uparrow} \\ c_{k-Q\downarrow} \\ c_{k+Q\uparrow} \\ c_{k\downarrow} \\ c_{k-Q\uparrow} \\ c_{k+Q\downarrow} \end{pmatrix}
\end{aligned}
\tag{4.31}
$$

where $\eta = mU/2$. Although the Hamiltonian looks six-dimensional at first, there is a pairwise scattering of one type of fermion into another, and no scattering among the fermions belonging to different pairs. It can be diagonalized by using

$$
\begin{pmatrix} c^+_{k\uparrow} \\ c^+_{k-Q\downarrow} \\ c^+_{k+Q\uparrow} \\ c^+_{k\downarrow} \\ c^+_{k-Q\uparrow} \\ c^+_{k+Q\downarrow} \end{pmatrix} = \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 & 0 & 0 & 0 & 0 \\ \sin\theta_1 & \cos\theta_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \cos\theta_2 & -\sin\theta_2 & 0 & 0 \\ 0 & 0 & \sin\theta_2 & \cos\theta_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cos\theta_3 & -\sin\theta_3 \\ 0 & 0 & 0 & 0 & \sin\theta_3 & \cos\theta_3 \end{pmatrix} \begin{pmatrix} a_{1\uparrow} \\ a_{1\downarrow} \\ a_{2\uparrow} \\ a_{2\downarrow} \\ a_{3\uparrow} \\ a_{3\downarrow} \end{pmatrix}
\tag{4.32}
$$

with

$$
\tan 2\theta_1 = \frac{2\eta}{\epsilon_{k-Q} - \epsilon_k}, \quad \tan 2\theta_2 = \frac{2\eta}{\epsilon_k - \epsilon_{k+Q}}, \quad \tan 2\theta_3 = \frac{2\eta}{\epsilon_{k+Q} - \epsilon_{k-Q}}. \tag{4.33}
$$

The six eigenenergies are given by

$$
\begin{aligned}
E^{\pm}_{1k} &= \frac{1}{2}[(\epsilon_k + \epsilon_{k-Q}) \pm \sqrt{(\epsilon_k - \epsilon_{k-Q})^2 + 4\eta^2}] \\
E^{\pm}_{2k} &= \frac{1}{2}[(\epsilon_{k+Q} + \epsilon_k) \pm \sqrt{(\epsilon_{k+Q} - \epsilon_k)^2 + 4\eta^2}]
\end{aligned}
$$

$$E^{\pm}_{3k} = \frac{1}{2}[(\epsilon_{k-Q} + \epsilon_{k+Q}) \pm \sqrt{(\epsilon_{k-Q} - \epsilon_{k+Q})^2 + 4\eta^2}].$$

The bare dispersion is given by $\epsilon_k = -2t[\cos k_x + \cos(k_x/2 + \sqrt{3}k_y/2) + \cos(k_x/2 - \sqrt{3}k_y/2)] - \mu$.

Once we fix the chemical potential at zero, $\mu = 0$, one can see that $E^{+}_{\alpha k} > 0$ but $E^{-}_{\alpha k} < 0$ for all $\alpha = 1, 2, 3$. (See Fig. **??**.) That means as long as we take $\mu = 0$ the three lowest bands are completely filled, and the three highest completely empty, with the net electron density of one per site, or half-filling. The gap magnitude is given by $\min(E^{+}_{\alpha k} - E^{-}_{\alpha k}) = 2\eta$.

In the triangular lattice, the antiferromagnetic ordering at $\mathbf{Q} = (4\pi/3, 0)$ introduces a spatial modulation that led to (i) tripling of the unit cell and 1/3-ing of the Brillouin zone and (ii) folding of the band into six subbands and a gap separating the lower three bands from the upper three. Full occupation of the lower bands and a complete absence in the upper bands leads to the opening of the gap at half-filling. The reduced Brillouin zone has the shape of a hexagon with the edge-to-edge separation exactly equalling $4\pi/3$.

4.3 Projection to truncated Hilbert space

When the on-site interaction energy U is very large, double occupation of any given quantum state is realistically forbidden. It is as if the Hilbert space for each atomic site is restricted to three, consisting of empty, or singly-occupied with either spin-up or spin-down. Within such a restricted Hilbert space the Hubbard term no longer exists because $n_{i\uparrow}n_{i\downarrow}$ acting on such a space will identically give zero.

It is often the case that the effective Hamiltonian acting within a restricted Hilbert space looks quite different from the original Hamiltonian defined in the larger space. Being able to systematically carry out the transformation to the restricted space is of practical value and the techniques are presented in this section using the Hubbard model as the starting point.

$$H = \sum_{ij\sigma} T_{ji} c^{+}_{j\sigma} c_{i\sigma} + U \sum_{i} n_{i\uparrow} n_{i\downarrow}. \tag{4.34}$$

We have used T_{ji} above which reduces to $T_{ji} = -t$ in the case of the nearest-neighbor hopping only. The following derivation is largely taken from Yoshioka, Girvin, and MacDonald.

Using the fact that the electron occupation plus the hole occupation at any site for spin σ adds to one, $n_{i\sigma} + h_{i\sigma} = 1$, one can split the kinetic term into three pieces, $T = T_0 + T_1 + T_{-1}$ where

$$\begin{aligned} T_1 &= \sum_{ij\sigma} T_{ji} n_{j\bar{\sigma}} c^{+}_{j\sigma} c_{i\sigma} h_{i\bar{\sigma}} \\ T_{-1} &= \sum_{ij} T_{ji} h_{j\bar{\sigma}} c^{+}_{j\sigma} c_{i\sigma} n_{i\bar{\sigma}} \\ T_0 &= \sum_{ij\sigma} T_{ji} h_{j\bar{\sigma}} c^{+}_{j\sigma} c_{i\sigma} h_{i\bar{\sigma}} + \sum_{ij\sigma} T_{ji} n_{j\bar{\sigma}} c^{+}_{j\sigma} c_{i\sigma} n_{i\bar{\sigma}}. \end{aligned} \tag{4.35}$$

T_1 increases the number of doubly occupied site by one because it has a non-zero matrix element if and only if the initial state has the configuration $|\uparrow_i\downarrow_j\rangle$, or

$|\downarrow_i\uparrow_j\rangle$. Then T_1 acting on the initial state produces $|0_i d_j\rangle$ or $|d_i 0_j\rangle$, d implying the doubly occupied site. Similarly, T_{-1} decrease the number of doubly occupied sites by one by acting on the initial state $|0_i d_j\rangle$ or $|d_i 0_j\rangle$. T_0 does not change the number of double occupancy because it acts only on the initial state $|0_i\sigma_j\rangle$, $|\sigma_i 0_j\rangle$, or $|d_i\sigma_j\rangle$, $|\sigma_i d_j\rangle$, $\sigma_i =\uparrow,\downarrow$.

Now consider a canonical transformation

$$H^{'} = e^{iS} H e^{-iS} \tag{4.36}$$

using $iS = (T_1 - T_{-1})/U$. $T_{\pm 1}$ satisfies the commutation $[V, T_1] = UT_1$, $[V, T_{-1}] = -UT_{-1}$ where $V = U\sum_i n_{i\uparrow} n_{i\downarrow}$ is the Hubbard term. By a straightforward expansion of $H^{'}$ up to second order in S and keeping the leading terms in $O(1/U)$ we get

$$H^{'} = T_0 + V + \frac{1}{U}\left([T_1, T_{-1}] + [T_1, T_0] + [T_0, T_{-1}]\right). \tag{4.37}$$

When we let the Hamiltonian $H^{'}$ act between states with no doubly occupied sites it is further reduced to

$$H^{'} = T_0 - \frac{1}{U} T_{-1} T_1 \tag{4.38}$$

because the other two terms would change the number of doubly occupied sites. Write out $T_{-1}T_1$ as

$$\sum_{ijkl\sigma\lambda} T_{lk} T_{ji} h_{l\overline{\lambda}} c^{+}_{l\lambda} c_{k\lambda} n_{k\overline{\lambda}} n_{j\overline{\sigma}} c^{+}_{j\sigma} c_{i\sigma} h_{i\overline{\sigma}}. \tag{4.39}$$

Excluding the three-site interaction we get $k = j$, $l = i$,

$$T_{-1}T_1 = \sum_{ij\sigma\lambda} T_{ij} T_{ji} h_{i\overline{\lambda}} c^{+}_{i\lambda} c_{j\lambda} n_{j\overline{\lambda}} n_{j\overline{\sigma}} c^{+}_{j\sigma} c_{i\sigma} h_{i\overline{\sigma}}. \tag{4.40}$$

A short consideration gives that the operators have the following action on the possible spin states on sites i, j:

$$T_{-1}T_1 = \sum_{ij} t^2_{ij} \left(|\uparrow_i\downarrow_j\rangle - |\downarrow_i\uparrow_j\rangle\right)\left(\langle\uparrow_i\downarrow_j| - \langle\downarrow_i\uparrow_j|\right) = 2t^2_{ij}|s_{ij}\rangle\langle s_{ij}|. \tag{4.41}$$

The symbol s_{ij} refers to the singlet state in the $\langle ij\rangle$-bond. For less than half-filled case it is possible to have an empty site either at i or j, for which $T_{-1}T_1$ would give zero. Taking this into account, $|s_{ij}\rangle\langle s_{ij}|$ may be re-written

$$|s_{ij}\rangle\langle s_{ij}| = \frac{1}{4} n_i n_j - s_i \cdot s_j \tag{4.42}$$

using the spin-$\frac{1}{2}$ notation. In all, we get

$$\begin{aligned} H^{'} &= T_0 + \frac{2}{U}\sum_{ij} t^2_{ij}(s_i \cdot s_j - \frac{1}{4} n_i n_j), \\ T_0 &= \sum_{ij} T_{ji} h_{j\overline{\sigma}} c^{+}_{j\sigma} c_{i\sigma} h_{i\overline{\sigma}}. \end{aligned} \tag{4.43}$$

When we do not confine ourselves to a subspace of no double occupancy, one other term generated by the canonical transformation in Eq. (4.37) survive.

$$H^{'} = T_0 + V + \frac{1}{U}\left(T_1 T_{-1} - T_{-1} T_1\right). \tag{4.44}$$

The new term induces pair hopping.

4.4 Spin dependent hopping and DM interaction

Under time reversal operation the spin-up and spin-down operators transform as

$$c_{i\uparrow} \to c_{i\downarrow}, \quad c_{i\downarrow} \to -c_{i\uparrow}. \tag{4.45}$$

The most general fermion hopping between a pair of sites $\langle ij \rangle$ obeying the time-reversal symmetry can be written down using the spinor notation $\psi_i = \begin{pmatrix} c_{i\uparrow} \\ c_{i\downarrow} \end{pmatrix}$,

$$H_{ij} = t\psi_i^\dagger[\cos\theta + i(\hat{d}\cdot\sigma)\sin\theta]\psi_j = t\psi_i^\dagger[e^{i\theta\hat{d}\cdot\sigma}]\psi_j. \tag{4.46}$$

Here $\hat{d}$ is an arbitrary 3-dimensional real unit vector. The time-reversal operation renders $\psi_i \to \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\psi_i = i\sigma_y\psi_i$ and one can easily see that the above expression for H_{ij} is invariant under such a change of ψ_i plus the conjugation of $i\hat{d}\cdot\sigma$ to $-i\hat{d}\cdot\sigma^*$. Note that $\sigma^* = (\sigma_x, -\sigma_y, \sigma_z)$. Most generally t, θ and $\hat{d}$ will depend on the bond, so we can write

$$H_{ij} = t_{ij}\psi_i^\dagger[e^{i\theta_{ij}\hat{d}_{ij}\cdot\sigma}]\psi_j. \tag{4.47}$$

To ensure that the Hamiltonian remains Hermitian we will require $\hat{d}_{ji} = -\hat{d}_{ij}$. Let's consider a specialized case where $\hat{d}_{ij}$ is uniform, so that we can take $\hat{d}_{ij} = +\hat{d}$ for $j = i + \hat{x}$ and $j = i + \hat{y}$, and $\theta_{ij} = \theta$, $t_{ij} = t$ everywhere. (I'm assuming a square lattice.) Then the Hamiltonian can be written in the form

$$H = \sum_i t\Big(\psi_{i+x}^\dagger e^{i\theta\hat{d}\cdot\sigma} + \psi_{i-x}^\dagger e^{-i\theta\hat{d}\cdot\sigma} + \psi_{i+y}^\dagger e^{i\theta\hat{d}\cdot\sigma} + \psi_{i-y}^\dagger e^{-i\theta\hat{d}\cdot\sigma}\Big)\psi_i. \tag{4.48}$$

The superexchange calculation in the case of general hopping matrix can be worked out from the two-site model

$$t_{ij}\psi_i^\dagger[e^{i\theta_{ij}\hat{d}_{ij}\cdot\sigma}]\psi_j + Un_i(n_i-1) + Un_j(n_j-1). \tag{4.49}$$

Following Aharonny et al.[1] we can eliminate the unitary matrix by the transformation

$$\psi_i \to [e^{-i(\theta_{ij}/2)\hat{d}_{ij}\cdot\sigma}]\psi_i, \quad \psi_j \to [e^{i(\theta_{ij}/2)\hat{d}_{ij}\cdot\sigma}]\psi_j. \tag{4.50}$$

[1] L. Shekhtman, O. Entin-Wohlman, and Ammon Aharony, Phys. Rev. Lett. **69**, 836 (1992).

The Hubbard interaction $U\sum_i n_i(n_i-1)$ is invariant under the proposed transformation. Then in the rotated basis we obtain the superexchange Hamiltonian

$$J_{ij}S_i\cdot S_j,\quad J_{ij}=\frac{t_{ij}^2}{U}. \tag{4.51}$$

Now we must rotate the spins back to the original basis.

$$\begin{aligned} S_i &\rightarrow \frac{1}{2}\psi_i^\dagger e^{-i(\theta_{ij}/2)\hat{d}_{ij}\cdot\sigma}\sigma e^{i(\theta_{ij}/2)\hat{d}_{ij}\cdot\sigma}\psi_i = \cos\theta_{ij}S_i+(1-\cos\theta_{ij})\hat{d}_{ij}(\hat{d}_{ij}\cdot S_i)-\sin\theta_{ij}(\hat{d}_{ij}\times S_i)\\ S_j &\rightarrow \frac{1}{2}\psi_j^\dagger e^{-i(\theta_{ij}/2)\hat{d}_{ij}\cdot\sigma}\sigma e^{i(\theta_{ij}/2)\hat{d}_{ij}\cdot\sigma}\psi_j = \cos\theta_{ij}S_j+(1-\cos\theta_{ij})\hat{d}_{ij}(\hat{d}_{ij}\cdot S_j)+\sin\theta_{ij}(\hat{d}_{ij}\times S_j). \end{aligned} \tag{4.52}$$

We have invoked the identities

$$(\hat{d}\cdot\sigma)\sigma(\hat{d}\cdot\sigma)=2\hat{d}(\sigma\cdot\hat{d})-\sigma,\quad \sigma(\hat{d}\cdot\sigma)-(\hat{d}\cdot\sigma)\sigma=2i(\hat{d}\times\sigma), \tag{4.53}$$

to show that

$$e^{-i(\theta/2)\hat{d}\cdot\sigma}\sigma e^{i(\theta/2)\hat{d}\cdot\sigma}=(\cos\theta)\sigma+(1-\cos\theta)\hat{d}(\hat{d}\cdot\sigma)-\sin\theta(\hat{d}\times\sigma). \tag{4.54}$$

Now the inner product $S_i\cdot S_j$ becomes

$$S_i\cdot S_j\rightarrow(\cos 2\theta_{ij})S_i\cdot S_j-(\sin 2\theta_{ij})\hat{d}_{ij}\cdot(S_i\times S_j)+(1-\cos 2\theta_{ij})(S_i\cdot\hat{d}_{ij})(S_j\cdot\hat{d}_{ij}). \tag{4.55}$$

The terms generated are superexchange, DM, and Kitaev interactions, respectively.

4.5 Orbitally degenerate Hubbard model

Hubbard model treats a localized site subject to a strong Coulomb repulsion. In real materials, the atomic site treated within the Hubbard framework has not one, but several orbitals associated with the site. For example, transition metals have partially occupied d-shells that play a dominant role in the magnetic properties. The d-orbitals are fairly localized, so that the local Coulomb repulsion within an orbital and between different orbitals centered at the same site are considered important. For such a system, the concept of Hubbard interaction needs to be generalized. In addition, multiorbital systems have the Hund interaction which tends to align the spins for different orbitals. One may write down a general interaction term between local charge densities and the spins,

$$V=\frac{U_1}{2}\sum_\alpha n_\alpha^2+\frac{U_2}{2}\sum_{\alpha\neq\beta}n_\alpha n_\beta-J_1\sum_{\alpha\neq\beta}s_\alpha\cdot s_\beta+J_2\sum_{\alpha\neq\beta}d_\alpha^+d_\beta. \tag{4.56}$$

We have the intra(inter)-orbital interaction energy U_1 (U_2), the Hund coupling energy J_1, and J_2 which represents the process of moving two electrons simultaneously from α-orbital to β-orbital. All four terms shown above are invariant under the global spin rotation.

In the absence of crystal field splitting, all orbital states must be *degenerate*. Real-space rotation mixes different orbital states α, but the interaction must be written in a manner independent of the choice of real-space axis. The requirement of rotational invariance dictates the relations between the coefficients, U_1, U_2, J_1, and J_2.

From Dworin and Narath [Phys. Rev. Lett. **25**, 1287 (1970)] we find that the choice

$$U_1 = U + J,\ U_2 = U - J/2,\ J_1 = J,\ J_2 = 0 \tag{4.57}$$

for the $l = 2$, five-fold degenerate d-orbitals. The sum over α and β goes from -2 to $+2$. This choice of parameters was adopted in later works such as M. J. Rozenberg [Phys. Rev. B **55**, 4855 (1997)] and J. E. Han *et al.* [Phys. Rev. B **58**, 4199 (1998)].

On the other hand, Fujimori [Phys. Rev. B **51**, 12880 (1995)] and Koga *et al.* [Phys. Rev. Lett. **92**, 216402 (2004)] adopted a different convention

$$U_1 = U + 2J,\ U_2 = U - J/2,\ J_1 = J = J_2. \tag{4.58}$$

It is possible that the choice of rotation-invariant interaction is not unique.

4.6 Four-boson Theory

We are interested in implementing a numerical scheme for finding the self-consistent solution of the n-fold degenerate Hubbard model using the four-boson theory (and its n-fold generalization) introduced by Kotliar and Ruckenstein (KR). Below we work out the self-consistent equations which arise in the $n = 1$ case, for (I) the non-uniform situation (assuming no symmetries) and for (II) a bi-partite case with magnetization.

4.6.1 Non-uniform Case

For the non-uniform situation the effective Hamiltonian to be solved is

$$\begin{aligned} H &= \sum_{\langle ij\rangle\sigma} -\frac{t_{ij}}{m_i m_j}(d_j^+ p_{j-\sigma} + p_{j\sigma}^+ e_j)(p_{i-\sigma}^+ d_i + e_i^+ p_{i\sigma}) f_{j\sigma}^+ f_{i\sigma} + \sum_{i\sigma}(\rho_{i\sigma} - \mu) f_{i\sigma}^+ f_{i\sigma} \\ &+ \sum_i U d_i^+ d_i - \sum_i \lambda_i (e_i^+ e_i + \sum_\sigma p_{i\sigma}^+ p_{i\sigma} + d_i^+ d_i - 1) - \sum_{i\sigma} \rho_{i\sigma}(p_{i\sigma}^+ p_{i\sigma} + d_i^+ d_i) \end{aligned} \tag{4.59}$$

where we have introduced the abbreviation $m_i = \sqrt{n_{i\sigma}(1 - n_{i\sigma})},\ n_{i\sigma} = \langle f_{i\sigma}^+ f_{i\sigma}\rangle$. In the KR theory this factor is expressed in terms of bosons, e.g. $n_{i\sigma} \to 1 - e_i^+ e_i - p_{i-\sigma}^+ p_{i-\sigma}$, and becomes equal to the present definition only if the constraints are satisfied exactly. Our scheme has the advantage that it will simplify the minimization equation considerably. Fermionic part of the Hamiltonian is diagonalized as $f_{i\sigma} = \sum_{m=1}^{N} u_{mi\sigma}\psi_{m\sigma}$ for each species of spin σ. The number of lattice sites is N.

Minimizing the free energy with respect to d_i^+, $p_{i\sigma}^+$ lead to the condition[2]

$$d_i = \frac{1}{\sum_\sigma \rho_{i\sigma} + \lambda_i - U}\left(-\sum_{j\sigma}\frac{t_{ij}}{m_i m_j} p_{i-\sigma}(p_{j-\sigma}^+ d_j + e_j^+ p_{j\sigma})\langle f_{i\sigma}^+ f_{j\sigma}\rangle\right), \quad (4.60)$$

$$p_{i\sigma} = \frac{1}{\lambda_i + \rho_{i\sigma}}\left(-\sum_{j}\frac{t_{ij}}{m_i m_j}[e_i(p_{j-\sigma}^+ d_j + e_j^+ p_{j\sigma})\langle f_{i\sigma}^+ f_{j\sigma}\rangle + d_i(d_j^+ p_{j\sigma} + p_{j-\sigma}^+ e_j)\langle f_{j-\sigma}^+ f_{i-\sigma}\rangle]\right). \quad (4.61)$$

Minimization with respect to e_i^+ leads to

$$\lambda_i e_i = -\sum_{j\sigma}\frac{t_{ij}}{m_i m_j} p_{i\sigma}(d_j^+ p_{j-\sigma} + p_{j\sigma}^+ e_j)\langle f_{j\sigma}^+ f_{i\sigma}\rangle. \quad (4.62)$$

We can calculate d_i and $p_{i\sigma}$ from Eqs. (4.60)-(4.61) and, using $e_i^+ e_i = 1 - d_i^+ d_i - \sum_\sigma p_{i\sigma}^+ p_{i\sigma}$[3], obtain e_i. Then Eq. (4.62) is used for λ_i.

$\rho_{i\sigma}$ imposes the constraint $\langle f_{i\sigma}^+ f_{i\sigma}\rangle = p_{i\sigma}^+ p_{i\sigma} + d_i^+ d_i$. They can be independently obtained by diagonalization the fermion Hamiltonian and from Eqs. (4.60)-(4.61). If $\langle f_{i\sigma}^+ f_{i\sigma}\rangle$ turns out to be greater than $p_{i\sigma}^+ p_{i\sigma} + d_i^+ d_i$ one raises (by hand) the value of $\rho_{i\sigma}$ in the next diagonalization step. This will lower the expectation value $\langle f_{i\sigma}^+ f_{i\sigma}\rangle$ in the next iteration. Iteration will continue until the identity is achieved.

Chemical potential is adjusted (also by hand) to satisfy the global constraint

$$\frac{1}{N}\sum_{i\sigma}\langle f_{i\sigma}^+ f_{i\sigma}\rangle = \nu \quad (4.63)$$

for the filling factor ν. Fermionic averages are given by

$$\langle f_{j\sigma}^+ f_{i\sigma}\rangle = \sum_m u_{mj\sigma}^* u_{mi\sigma} F(E_{m\sigma}). \quad (4.64)$$

$F(E_{m\sigma})$ is the Fermi distribution function corresponding to energy $E_{m\sigma}$.

4.6.2 Uniform Case

We specialize to the case of the above Hamiltonian where operators can take on two different values on sublattice A and B. One can write

$$f_{i\sigma} = \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}_i} f_{A\mathbf{k}\sigma}$$
$$f_{j\sigma} = \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}_j} f_{B\mathbf{k}\sigma} \quad (4.65)$$

[2]Formally this can be achieved by differentiating H with respect to all the variables, and setting them equal to zero. Expressions involving fermions are replaced by their finite-temperature averages. Although in the actual calculation complex fields such as d_i^+ will be treated as a real number, and thus $d_i^+ = d_i$, it is convenient during the derivation to keep the complex nature of the boson fields.

[3]In invoking this relation we are already carrying out the minimization with respect to λ_i.

for i and j belonging to sublattice A and B, respectively. $\mathbf{k}$ is restricted to half the Brillouin zone. An example is the diamond-shaped space covered by connecting $(\pi,0)-(0,\pi)-(-\pi,0)-(0,-\pi)$ in the BZ.

Introducing similar notations for other variables, the effective Hamiltonian becomes

$$\begin{aligned} H &= \sum_{\mathbf{k}\sigma} \epsilon_{\mathbf{k}} (z^+_{B\sigma} z_{A\sigma} f^+_{B\mathbf{k}\sigma} f_{A\mathbf{k}\sigma} + z^+_{A\sigma} z_{B\sigma} f^+_{A\mathbf{k}\sigma} f_{B\mathbf{k}\sigma}) \\ &\quad + \sum_{\mathbf{k}\sigma} (\rho_{A\sigma} - \mu) f^+_{A\mathbf{k}\sigma} f_{A\mathbf{k}\sigma} + \sum_{\mathbf{k}\sigma} (\rho_{B\sigma} - \mu) f^+_{B\mathbf{k}\sigma} f_{B\mathbf{k}\sigma} \\ &\quad + \frac{N}{2} \times \{ U d^+_A d_A - \lambda_A (e^+_A e_A + \sum_\sigma p^+_{A\sigma} p_{A\sigma} + d^+_A d_A - 1) \\ &\quad - \sum_\sigma \rho_{A\sigma} (p^+_{A\sigma} p_{A\sigma} + d^+_A d_A) + (A \to B) \}. \end{aligned} \tag{4.66}$$

Expressions for z are

$$\begin{aligned} z_{A\sigma} &= \frac{1}{\sqrt{n_{A\sigma}(1-n_{A\sigma})}} (p^+_{A-\sigma} d_A + e^+_A p_{A\sigma}) \\ z_{B\sigma} &= \frac{1}{\sqrt{n_{B\sigma}(1-n_{B\sigma})}} (p^+_{B-\sigma} d_B + e^+_B p_{B\sigma}). \end{aligned} \tag{4.67}$$

The eigenenergies for the fermion part are given by[4]

$$E_{\mathbf{k}\sigma\pm} = \frac{1}{2}(\rho_{A\sigma} + \rho_{B\sigma}) - \mu \pm \sqrt{\frac{1}{4}(\rho_{A\sigma} - \rho_{B\sigma})^2 + \epsilon^2_{\mathbf{k}} |z_{A\sigma}|^2 |z_{B\sigma}|^2}. \tag{4.68}$$

Free energy of the system becomes

$$\begin{aligned} F = &\sum_{\mathbf{k}\sigma\pm} \left[E_{\mathbf{k}\sigma\pm} - T \ln(1 + e^{\beta E_{\mathbf{k}\sigma\pm}}) \right] + \frac{N}{2} \times \\ &\left\{ U d^+_A d_A - \lambda_A (e^+_A e_A + \sum_\sigma p^+_{A\sigma} p_{A\sigma} + d^+_A d_A - 1) - \sum_\sigma \rho_{A\sigma} (p^+_{A\sigma} p_{A\sigma} + d^+_A d_A) + (A \to B) \right\}. \end{aligned} \tag{4.69}$$

Minimization of F takes place with respect to $d^+_A, d^+_B, e^+_A, e^+_B, p^+_{A\sigma}, p^+_{B\sigma}, \rho_{A\sigma}, \rho_{B\sigma}, \lambda_A, \lambda_B$, and the chemical potential μ. $\partial F / \partial d^+_{A,B} = 0$ gives

$$\begin{aligned} d_A &= \frac{1}{\rho_{A\uparrow} + \rho_{A\downarrow} + \lambda_A - U} \times \frac{2}{N} \times \sum_{\mathbf{k}\sigma\pm} F(E_{\mathbf{k}\sigma\pm}) \frac{\partial E_{\mathbf{k}\sigma\pm}}{\partial d^+_A} \\ d_B &= \frac{1}{\rho_{B\uparrow} + \rho_{B\downarrow} + \lambda_B - U} \times \frac{2}{N} \times \sum_{\mathbf{k}\sigma\pm} F(E_{\mathbf{k}\sigma\pm}) \frac{\partial E_{\mathbf{k}\sigma\pm}}{\partial d^+_B}. \end{aligned} \tag{4.70}$$

[4]Since for each $\mathbf{k}$ and σ there is mixing between A and B, one ends up diagonalizing a 2×2 matrix, whose engenvalues are distinguished as $E_{\mathbf{k}\sigma\pm}$.

$\partial F/\partial p^+_{A,B\sigma} = 0$ gives

$$\begin{aligned} p_{A\sigma} &= \frac{1}{\rho_{A\sigma}+\lambda_A} \times \frac{2}{N} \times \sum_{\mathbf{k}\sigma' \pm} F(E_{\mathbf{k}\sigma' \pm}) \frac{\partial E_{\mathbf{k}\sigma' \pm}}{\partial p^+_{A\sigma}} \\ p_{B\sigma} &= \frac{1}{\rho_{B\sigma}+\lambda_B} \times \frac{2}{N} \times \sum_{\mathbf{k}\sigma' \pm} F(E_{\mathbf{k}\sigma' \pm}) \frac{\partial E_{\mathbf{k}\sigma' \pm}}{\partial p^+_{B\sigma}}. \end{aligned} \tag{4.71}$$

$\partial F/\partial e^+_{A,B} = 0$ gives

$$\begin{aligned} \lambda_A e_A &= \frac{2}{N} \times \sum_{\mathbf{k}\sigma \pm} F(E_{\mathbf{k}\sigma \pm}) \frac{\partial E_{\mathbf{k}\sigma \pm}}{\partial e^+_{A}} \\ \lambda_B d_B &= \frac{2}{N} \times \sum_{\mathbf{k}\sigma \pm} F(E_{\mathbf{k}\sigma \pm}) \frac{\partial E_{\mathbf{k}\sigma \pm}}{\partial e^+_{B}}. \end{aligned} \tag{4.72}$$

$\partial F/\partial \rho_{A,B\sigma} = 0$ gives

$$\begin{aligned} p^+_{A\sigma} p_{A\sigma} + d^+_A d_A &= \frac{2}{N} \sum_{\mathbf{k}\pm} F(E_{\mathbf{k}\sigma \pm}) \frac{\partial E_{\mathbf{k}\sigma \pm}}{\partial \rho_{A\sigma}} \\ p^+_{B\sigma} p_{B\sigma} + d^+_B d_B &= \frac{2}{N} \sum_{\mathbf{k}\pm} F(E_{\mathbf{k}\sigma \pm}) \frac{\partial E_{\mathbf{k}\sigma \pm}}{\partial \rho_{B\sigma}}. \end{aligned} \tag{4.73}$$

Writing out the derivatives explicitly gives

$$\begin{aligned} d_A &= \frac{1}{\rho_{A\uparrow}+\rho_{A\downarrow}+\lambda_A - U} \times \frac{1}{N} \times \sum_{\mathbf{k}\sigma\pm} \pm F(E_{\mathbf{k}\sigma\pm}) \frac{\epsilon^2_{\mathbf{k}}}{\sqrt{D}} \frac{|z_{B\sigma}|^2 p_{A-\sigma}(p^+_{A-\sigma} d_A + e^+_A p_{A\sigma})}{n_{A\sigma}(1-n_{A\sigma})} \\ d_B &= \frac{1}{\rho_{B\uparrow}+\rho_{B\downarrow}+\lambda_B - U} \times \frac{1}{N} \times \sum_{\mathbf{k}\sigma\pm} \pm F(E_{\mathbf{k}\sigma\pm}) \frac{\epsilon^2_{\mathbf{k}}}{\sqrt{D}} \frac{|z_{A\sigma}|^2 p_{B-\sigma}(p^+_{B-\sigma} d_B + e^+_B p_{B\sigma})}{n_{B\sigma}(1-n_{B\sigma})}. \end{aligned} \tag{4.74}$$

$$\begin{aligned} p_{A\sigma} &= \frac{1}{\rho_{A\sigma}+\lambda_A} \times \frac{1}{N} \times \sum_{\mathbf{k}\pm} \pm F(E_{\mathbf{k}\sigma\pm}) \frac{\epsilon^2_{\mathbf{k}}}{\sqrt{D}} \times \\ &\left[\frac{|z_{B-\sigma}|^2 d_A(p_{A\sigma} d^+_A + e_A p^+_{A-\sigma})}{n_{A-\sigma}(1-n_{A-\sigma})} + \frac{|z_{B\sigma}|^2 e_A(p_{A\sigma} e^+_A + d_A p^+_{A-\sigma})}{n_{A\sigma}(1-n_{A\sigma})} \right] \\ p_{B\sigma} &= \frac{1}{\rho_{B\sigma}+\lambda_B} \times \frac{1}{N} \times \sum_{\mathbf{k}\pm} \pm F(E_{\mathbf{k}\sigma\pm}) \frac{\epsilon^2_{\mathbf{k}}}{\sqrt{D}} \times \\ &\left[\frac{|z_{A-\sigma}|^2 d_B(p_{B\sigma} d^+_B + e_B p^+_{B-\sigma})}{n_{B-\sigma}(1-n_{B-\sigma})} + \frac{|z_{A\sigma}|^2 e_B(p_{B\sigma} e^+_B + d_B p^+_{B-\sigma})}{n_{B\sigma}(1-n_{B\sigma})} \right]. \end{aligned} \tag{4.75}$$

$$\lambda_A e_A = \frac{1}{N} \times \sum_{\mathbf{k}\sigma\pm} \pm F(E_{\mathbf{k}\sigma\pm}) \frac{\epsilon_{\mathbf{k}}^2}{\sqrt{D}} \frac{|z_{B\sigma}|^2 p_{A\sigma}(p_{A-\sigma} d_A^+ + e_A p_{A\sigma}^+)}{n_{A\sigma}(1-n_{A\sigma})}$$
$$\lambda_B e_B = \frac{1}{N} \times \sum_{\mathbf{k}\sigma\pm} \pm F(E_{\mathbf{k}\sigma\pm}) \frac{\epsilon_{\mathbf{k}}^2}{\sqrt{D}} \frac{|z_{A\sigma}|^2 p_{B\sigma}(p_{B-\sigma} d_B^+ + e_B p_{B\sigma}^+)}{n_{B\sigma}(1-n_{B\sigma})}. \tag{4.76}$$

$$p_{A\sigma}^+ p_{A\sigma} + d_A^+ d_A = \frac{1}{N} \sum_{\mathbf{k}\pm} F(E_{\mathbf{k}\sigma\pm}) \left(1 \pm \frac{\rho_{A\sigma} - \rho_{B\sigma}}{2\sqrt{D}}\right)$$
$$p_{B\sigma}^+ p_{B\sigma} + d_B^+ d_B = \frac{1}{N} \sum_{\mathbf{k}\pm} F(E_{\mathbf{k}\sigma\pm}) \left(1 \mp \frac{\rho_{A\sigma} - \rho_{B\sigma}}{2\sqrt{D}}\right) \tag{4.77}$$

where $D = (1/4)(\rho_{A\sigma} - \rho_{B\sigma})^2 + \epsilon_{\mathbf{k}}^2 |z_{A\sigma}|^2 |z_{B\sigma}|^2$. Condition on the chemical potential reads

$$\frac{1}{N} \sum_{\mathbf{k}\sigma\pm} F(E_{\mathbf{k}\sigma\pm}) = \nu. \tag{4.78}$$

Strategy for updating the variables are as follows. From Eqs. (4.74) and (4.75) one obtains updated values for $d_A, d_B, p_{A\sigma}$, and $p_{B\sigma}$. Using $e_A^+ e_A = 1 - d_A^+ d_A - \sum_\sigma p_{A\sigma}^+ p_{A\sigma}$ and the corresponding relation for B sublattice one obtains e_A, e_B. From Eq. (4.76), using known values of e_A, e_B from above one obtains λ_A, λ_B. Finally, Eq. (4.77) determines $\rho_{A\sigma} - \rho_{B\sigma}$. The even combination $\rho_{A\sigma} + \rho_{B\sigma}$ leads to the uniform shift of the chemical potential (see Eq. (4.68)), or it can be absorbed into μ. Taking the difference of the two equations in Eq. (4.77) gives

$$p_{A\sigma}^+ p_{A\sigma} + d_A^+ d_A - p_{B\sigma}^+ p_{B\sigma} + d_B^+ d_B = \langle f_{A\sigma}^+ f_{A\sigma} \rangle - \langle f_{B\sigma}^+ f_{B\sigma} \rangle$$
$$= (\rho_{A\sigma} - \rho_{B\sigma}) \times \frac{1}{N} \sum_{\mathbf{k}\pm} \pm \frac{F(E_{\mathbf{k}\sigma\pm})}{\sqrt{D}}. \tag{4.79}$$

The occupation of electrons with spin σ will be different on the two sublattices provided $\rho_{A\sigma} \neq \rho_{B\sigma}$, indicating antiferromagnetic component of spin. In practice I think one should set $\rho_{A\sigma} = -\rho_{B\sigma} = \rho_\sigma$ from the outset. This simplifies the equation to

$$p_{A\sigma}^+ p_{A\sigma} + d_A^+ d_A - p_{B\sigma}^+ p_{B\sigma} + d_B^+ d_B = \rho_\sigma \times \frac{2}{N} \sum_{\mathbf{k}\pm} \pm \frac{F(E_{\mathbf{k}\sigma\pm})}{\sqrt{D}} \tag{4.80}$$

for $D = \rho_\sigma^2 + \epsilon_{\mathbf{k}} |z_{A\sigma}|^2 |z_{B\sigma}|^2$ and $E_{\mathbf{k}\sigma\pm} = -\mu \pm \sqrt{D}$.

The ground state can be (i) uniform and paramagnetic, (ii) uniform and spin-polarized (for nonzero ρ_σ) or (iii) possess staggered magnetic moment.

Chapter 5

Spin Interactions

5.1 Higher-order exchange processes

The Heisenberg spin exchange model describes spin interactions of the type $S_i \cdot S_j$ for two neighboring spins. In real materials the interactions between spin degrees of freedom are undoubtedly more complicated, but it is a highly nontrivial matter to figure out what sort of terms, or processes, contribute significantly to the Hamiltonian.

It is a curious fact of quantum mechanics that, for $S = 1/2$ spins, this exchange term has a close connection with the pair exchange operator P_{ij} which simply swaps the quantum states located at i and j; $P_{ij}|\alpha_i\beta_j\rangle = |\beta_i\alpha_j\rangle$. Here α and β may represent any of the available quantum states. For $S = 1/2$, one can readily verify

$$P_{ij} = 2\left(S_i \cdot S_j + \frac{1}{4}\right). \tag{5.1}$$

To prove that the two operator expressions are identical requires that the action of each operator on any given state yields exactly the same new state whether we use P_{ij} or $2(S_i \cdot S_j + 1/4)$. So, let's prove it. With $S = 1/2$, there are only four available states, $|\uparrow\uparrow\rangle$, $|\uparrow\downarrow\rangle$, $|\downarrow\uparrow\rangle$, $|\downarrow\downarrow\rangle$, so our task is reduced to checking the validity of Eq. (5.1) for each of the four states. For $|\uparrow\uparrow\rangle$ one gets

$$P_{ij}|\uparrow\uparrow\rangle = |\uparrow\uparrow\rangle.$$

Now, let's see what happens when we act with the r.h.s. of Eq. (5.1)

$$2\left(S_i \cdot S_j + \frac{1}{4}\right)|\uparrow\uparrow\rangle = 2\left(S_i^z S_j^z + \frac{1}{2}\right)|\uparrow\uparrow\rangle = 2\left(\frac{1}{2}\frac{1}{2} + \frac{1}{4}\right)|\uparrow\uparrow\rangle = |\uparrow\uparrow\rangle.$$

OK, so both operators produce the same state, $|\uparrow\uparrow\rangle$, when acting on $|\uparrow\uparrow\rangle$. For $|\uparrow\downarrow\rangle$, one has $P_{ij}|\uparrow\downarrow\rangle = |\uparrow\downarrow\rangle$, because the up- and down-spins simply swap their positions. Before acting with the r.h.s. of Eq. (5.1) again, first remember that

$$S_i \cdot S_j = S_i^z S_j^z + \frac{1}{2}(S_i^+ S_j^- + S_i^- S_j^+). \tag{5.2}$$

After collecting the result of action with each of the three operators of Eq. (5.2) on $|\uparrow\downarrow\rangle$ we obtain

$$\begin{aligned} & 2\left(S_i^z S_j^z + \frac{1}{2}S_i^+ S_j^- + \frac{1}{2}S_i^- S_j^+ + \frac{1}{4}\right)|\uparrow\downarrow\rangle \\ = \;& 2\left(-\frac{1}{4}|\uparrow\downarrow\rangle + 0 + \frac{1}{2}|\uparrow\downarrow\rangle + \frac{1}{4}|\uparrow\downarrow\rangle\right) \\ = \;& |\uparrow\downarrow\rangle. \end{aligned}$$

Again, we obtain agreement. The case for $|\uparrow\downarrow\rangle$ and $|\downarrow\downarrow\rangle$ states proceeds similarly and need not be reproduced here. Hence we learn that the cherished Heisenberg exchange model may equally well be written

$$H = \frac{J}{2}\sum_{\langle ij\rangle} P_{ij}.$$

In reality, there is no reason to just stop with exchanging two particles at a time. Why not exchange the spins of three, or even four sites at once. For example, the 3-site exchange process is depicted in Fig. **??**.

The exchange process of this type is denoted P_{ijk}, in obvious generalization of the two particle exchange. In turn, the three particle exchange can be decomposed as a series of two-particle exchanges:

$$P_{ijk} = P_{ik}P_{ij}.$$

Proof : Label the states of i, j, k sites by A, B, C and denote the state $|A_iB_jC_k\rangle$. The cyclic exchange P_{ijk} produces a state $|C_iA_jB_k\rangle$. Acting on the original state $|A_iB_jC_k\rangle$ with P_{ij} gives $|B_iA_jC_k\rangle$. Acting on this state with P_{ik} gives $|C_iA_jB_k\rangle$, same as $P_{ijk}|A_iB_jC_k\rangle$. Acting with $P_{ji}P_{jk}$ or $P_{kj}P_{ki}$ leads to the same state. There is also the clockwise permutation of the states P_{ijk}^{-1} which renders

$$P_{ijk}^{-1}|A_iB_jC_k\rangle = |B_iC_jA_k\rangle$$

The same result is achieved if we act with $P_{ij}P_{ik}$, hence

$$P_{ijk}^{-1} = P_{ij}P_{ik} = P_{jk}P_{ji} = P_{ki}P_{kj}.$$

In any sensible physical system the two processes - clockwise and anti-clockwise - occur with same amplitude, and thus the most general 3-site exchange process is described by a term in the Hamiltonian

$$\sum_{\langle ijk\rangle}(P_{ijk} + P_{ijk}^{-1}) = \frac{1}{3}\sum_{\langle ijk\rangle}(P_{ik}P_{ij} + P_{ji}P_{jk} + P_{kj}P_{ki} + P_{ij}P_{ik} + P_{jk}P_{ji} + P_{ki}P_{kj}).$$

Insert $P_{ij} = 2(S_i \cdot S_j + 1/4)$ into this, one gets

$$\sim \sum_{\langle ijk\rangle}(S_i \cdot S_j)(S_i \cdot S_k) + (S_j \cdot S_k)(S_j \cdot S_i) + (S_k \cdot S_i)(S_k \cdot S_j) \tag{5.3}$$

plus 2-site spin interactions which are readily absorbed into the Heisenberg term. So, allowing a 3-particle exchange generates a new spin interaction term in the

Hamiltonian of the form shown in Eq. (5.3). The general spin Hamiltonian involving 2- and 3-particle exchanges can be written down

$$H = J_1 \sum_{\langle ij \rangle} S_i \cdot S_j$$
$$+J_2 \sum_{\langle ijk \rangle} (S_i \cdot S_j)(S_i \cdot S_k) + (S_j \cdot S_k)(S_j \cdot S_i) + (S_k \cdot S_i)(S_k \cdot S_j). \quad (5.4)$$

Generalization to 4-particle exchange is also obvious, and is left to the reader.

Chapter 6

Spin Representation and Spin Excitation

There are several kinds of representations for spin operators. Some of them are exclusively used for $S = 1/2$, and others are applicable for general S. They also lead to different ways of obtaining the spin excitations of a given spin interaction model.

6.1 Semiclassical Equation of Motion

6.1.1 Ferromagnet

In the ferromagnetic spin model, how each spin will dance in coordination with the others is determined by the Hamiltonian

$$H = -J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j, \tag{6.1}$$

and the equation of motion for the individual spin $\mathbf{S}_i$ that follows from it. We first remind ourselves of the fundamental commutation relations of spin

$$[S_i^\alpha, S_j^\beta] = i\delta_{ij}\epsilon_{\alpha\beta\gamma}S_i^\gamma \quad (\hbar \equiv 1). \tag{6.2}$$

From this one can calculate, for instance,

$$\frac{dS_i^x}{dt} = -i[S_i^x, H] = -J \sum_{j \in i} \left(S_j^y S_i^z - S_j^z S_i^y \right) = -J \sum_{j \in i} \left(S_j \times S_i \right)^x. \tag{6.3}$$

Combined with the equation of motion for the y- and z-components, we get

$$\frac{d\mathbf{S}_i}{dt} = -J \sum_{j \in i} \mathbf{S}_j \times \mathbf{S}_i. \tag{6.4}$$

This equation is of a typical, nonlinear type because the quantities we must solve for, $\mathbf{S}_i$, appears as a product on the r.h.s. In other words, one can never

solve the above equation *exactly*. However we must remind ourselves that it is the small fluctuation away from the ground state that concerns us, and as such we can decompose the spin operator as $\mathbf{S}_i = \langle \mathbf{S}_i \rangle + \delta \mathbf{S}_i = \mathbf{m}_i + \delta \mathbf{S}_i$, where $\mathbf{m}_i$ represents the ground state spin average. For a ferromagnetic ground state, $\mathbf{m}_i = S\hat{z}$. A small fluctuation means $|\langle \mathbf{S}_i \rangle| \gg |\delta \mathbf{S}_i|$. In making such a comparison of magnitudes we must be careful to remember that $\mathbf{m}_i$ is a number, but $\delta \mathbf{S}_i$ is not. So, it's not altogether clear what exactly we mean by the magnitude of $\delta \mathbf{S}_i$. Nevertheless, we can at least make the substitution $\mathbf{S}_i = \mathbf{m}_i + \delta \mathbf{S}_i$ as a matter of formal definition, and proceed to plug it into Eq. (6.11).

$$\frac{d}{dt}\delta \mathbf{S}_i = -J \sum_{j \in i} \big(\mathbf{m}_j + \delta \mathbf{S}_j \big) \times \big(\mathbf{m}_i + \delta \mathbf{S}_i \big) \approx$$
$$-J \Big(\sum_{j \in i} \mathbf{m}_j \Big) \times \mathbf{m}_i - J \Big(\sum_{j \in i} \mathbf{m}_j \Big) \times \delta \mathbf{S}_i - J \Big(\sum_{j \in i} \delta \mathbf{S}_j \Big) \times \mathbf{m}_i \quad (6.5)$$

ignoring terms of order $\big(\delta \mathbf{S}\big)^2$. The sum $\sum_{j \in i}$ comprises all the nearest-neighbor sites j for i. Once again, the justification is that $\delta \mathbf{S}$ is a small quantity, and $(\delta \mathbf{S})^2$, even smaller. We take the quantization axis $\hat{z}$ and write $\mathbf{m}_i = \mathbf{m}\hat{z}$ everywhere. Then[1]

$$\frac{d}{dt}\delta \mathbf{S}_i = -Jzm\hat{z} \times \delta \mathbf{S}_i - Jm \Big(\sum_j \delta \mathbf{S}_j \Big) \times \hat{z}. \quad (6.6)$$

The $\hat{z}$-component of this equation reads $\frac{d}{dt}\delta \mathbf{S}_i^z = 0$, hence the fluctuation in the $\hat{z}$-component of spin is zero, $S_i^z\big(t\big) = m$. For the $\hat{x}$-and $\hat{y}$-components we get

$$\begin{aligned} \frac{d}{dt} S_i^x &= Jzm S_i^y - Jm \sum_{j \in i} S_j^y, \\ \frac{d}{dt} S_i^y &= -Jzm S_i^x + Jm \sum_{j \in i} S_j^x. \end{aligned} \quad (6.7)$$

We have dropped the δ symbol in front. Introducing the complex notation $Z_i \equiv S_i^x + iS_i^y$ simplifies the equation a lot:

$$\frac{dZ_i}{dt} = iJm \left(\sum_{j \in i} Z_j - zZ_i \right). \quad (6.8)$$

We will solve this equation using the ansatz

$$Z_i = Z_0 e^{i\mathbf{k}\cdot\mathbf{r}_i - i\omega_{\mathbf{k}} t}, \quad (6.9)$$

which yields the self-consistency condition

$$\omega_{\mathbf{k}} = Jm \left(z - \sum_{j \in i} e^{i\mathbf{k}\cdot(\mathbf{r}_j - \mathbf{r}_i)} \right). \quad (6.10)$$

[1] I apologize for the multiple use of z here: z is the number of neighbors, or the coordination number, whereas $\hat{z}$ is the unit vector along the z-axis.

For a 3-dimensional cubic lattice, $\sum_{j\in i} e^{i\mathbf{k}\cdot(\mathbf{r}_j-\mathbf{r}_i)} = 2(\cos k_x + \cos k_y + \cos k_z)$, and with $z = 6$, $\omega_{\mathbf{k}} = 2Jm(3 - \sum_\alpha \cos k_\alpha)$. For 2D square lattice we will get $\omega_k = 2Jm(2 - \cos k_x - \cos k_y)$. This is the desired spin wave energy for a spin disturbance associated with a wavevector $\mathbf{k}$.

6.1.2 Antiferromagnet

For an antiferromagnet, the equation of motion reads

$$\frac{d\mathbf{S}_i}{dt} = J\sum_{j\in i}\mathbf{S}_j \times \mathbf{S}_i. \tag{6.11}$$

The linearized equation of motion is

$$\frac{d}{dt}\delta\mathbf{S}_i = J\Big(\sum_{j\in i}\mathbf{m}_j\Big)\times\delta\mathbf{S}_i + J\Big(\sum_{j\in i}\delta\mathbf{S}_j\Big)\times\mathbf{m}_i. \tag{6.12}$$

We take the quantization axis $\hat{z}$ and write $\mathbf{m}_i = (-1)^i\mathbf{m}\hat{z}$. Then, the whole square lattice can be divided into two sublattices: $i \in \mathcal{A}$ if $(-1)^i = 1$, and $i \in \mathcal{B}$ if $(-1)^i = -1$. A lattice for which such a division into two sublattices can be carried out is known as a bipartite lattice.

We write $\delta\mathbf{S}_i$ as $\mathbf{S}_i$ and note that $\mathbf{m}_j$ for the neighbors j surrounding the site i is always opposite to $\mathbf{m}_i$: $\mathbf{m}_j = -\mathbf{m}_i$.

$$\frac{d}{dt}\mathbf{S}_i = zJ\mathbf{S}_i\times\mathbf{m}_i + J\Big(\sum_{j\in i}\mathbf{S}_j\Big)\times\mathbf{m}_i. \tag{6.13}$$

Finally, we note that the staggered nature of the average magnetization $\mathbf{m}_i$ implies that the spin fluctuation $\mathbf{S}_i$ must contain both uniform and staggered components. With this knowledge, one can write $\mathbf{S}_i$ as

$$\mathbf{S}_i = \mathbf{S}_i^0 + (-1)^i\mathbf{S}_i^1. \tag{6.14}$$

Inserting this expression into the equation of motion and writing down the equations separately for the uniform and the staggered components, one obtains

$$\begin{aligned}
\frac{d}{dt}\mathbf{S}_i^0 &= zJ\mathbf{S}_i^1\times\mathbf{m}_i^s - J\Big(\sum_{j\in i}\mathbf{S}_j^1\Big)\times\mathbf{m}_i^s \\
\frac{d}{dt}\mathbf{S}_i^1 &= zJ\mathbf{S}_i^0\times\mathbf{m}_i^s + J\Big(\sum_{j\in i}\mathbf{S}_j^0\Big)\times\mathbf{m}_i^s
\end{aligned} \tag{6.15}$$

The staggered magnetization $\mathbf{m}_i^s = (-1)^i\mathbf{m}_i = m\hat{z}$ is uniform. Since all the quantities appearing in the above equation is slowly varying, one can write $\mathbf{S}_i^0 = \mathbf{S}^0 e^{i\mathbf{k}\cdot\mathbf{r}_i - i\omega t}$, $\mathbf{S}_i^1 = \mathbf{S}^1 e^{i\mathbf{k}\cdot\mathbf{r}_i - i\omega t}$, and obtain

$$\begin{aligned}
-i\omega\mathbf{S}^0 &= mzJ\Big(\mathbf{S}^1\times\hat{z} - \varepsilon_{\mathbf{k}}\mathbf{S}^1\times\hat{z}\Big), \\
-i\omega\mathbf{S}^1 &= mzJ\Big(\mathbf{S}^0\times\hat{z} + \varepsilon_{\mathbf{k}}\mathbf{S}^0\times\hat{z}\Big).
\end{aligned} \tag{6.16}$$

where we write $z\varepsilon_{\mathbf{k}} = \sum_{j\in i} e^{i\mathbf{k}\cdot(\mathbf{r}_j-\mathbf{r}_i)}$. Using the complex notation $S^0 = S^0_x + iS^0_y$, $S^1 = S^1_x + iS^1_y$, we can re-write the equation as

$$\begin{aligned} \omega S^0 &= mzJ(1-\varepsilon_{\mathbf{k}})S^1, \\ \omega S^1 &= mzJ(1+\varepsilon_{\mathbf{k}})S^0. \end{aligned} \tag{6.17}$$

It readily follows that $\omega_{\mathbf{k}} = mzJ\sqrt{1-\varepsilon^2_{\mathbf{k}}}$.

6.1.3 Spiral ferromagnet

We work out the elementary excitations for the spiral ferromagnet given by the Hamiltonian

$$\begin{aligned} H_{\mathrm{HDM}} &= -J\sum_i \mathbf{S}_i\cdot(\mathbf{S}_{i+\hat{x}}+\mathbf{S}_{i+\hat{y}}) - K\sum_i\Big(\mathbf{S}_i\times\mathbf{S}_{i+\hat{x}}\cdot\hat{x}+\mathbf{S}_i\times\mathbf{S}_{i+\hat{y}}\cdot\hat{y}\Big) \\ &= -\sum_{\langle ij\rangle}\mathbf{S}_i\cdot\Big(J\mathbf{S}_j + K\mathbf{S}_j\times\hat{e}_{ji}\Big). \end{aligned} \tag{6.18}$$

where we have introduced the unit vector $\hat{e}_{ji}$ extending from site i to site j.

Using the previously derived equation of motion for a ferromagnet, one can write down the equation of motion for the spiral ferromagnet easily

$$\frac{d\mathbf{S}_i}{dt} = \mathbf{S}_i\times\sum_{j\in i}\Big(J\mathbf{S}_j + K\mathbf{S}_j\times\hat{e}_{ji}\Big). \tag{6.19}$$

As before, we use $\mathbf{S}_i = \mathbf{m}_i + \delta\mathbf{S}_i$. The zeroth order term vanishes by default. In fact, the relation

$$\mathbf{m}_i\times\sum_{j\in i}\Big(J\mathbf{m}_j + K\mathbf{m}_j\times\hat{e}_{ji}\Big) = 0 \tag{6.20}$$

defines the classical ground state. The spiral spin state that concerns us is given by $\mathbf{m}_i = \mathbf{m}e^{i\mathbf{k}\cdot\mathbf{r}_i} + c.c.$ where

$$\mathbf{m} = \frac{1}{\sqrt{2}}(\hat{e}_1 - i\hat{e}_2), \quad \hat{e}_1\times\hat{e}_2 = \mathbf{k}. \tag{6.21}$$

Here $\hat{\mathbf{k}} = (1,1)/\sqrt{2}$ is the propagation vector of the spiral spins. One can show that

$$\sum_{j\in i}\Big(J\mathbf{m}_j + K\mathbf{m}_j\times\hat{e}_{ji}\Big) = \alpha\mathbf{m}_i \tag{6.22}$$

with the constant α given by

$$\alpha = 4J\cos Q + 2\sqrt{2}K\sin Q. \tag{6.23}$$

From an independent calculation we know that in two dimensions $\tan Q = K/\sqrt{2}J$, so that

$$\alpha = 4J\cos Q + 2\sqrt{2}K\sin Q = 2\sqrt{2}\sqrt{K^2+2J^2}. \tag{6.24}$$

Then the linearized equation of motion for the spiral ferromagnet becomes

$$\frac{d}{dt}\mathbf{S}_i = \mathbf{m}_i \times \sum_{j\in i}\Big(J\mathbf{S}_j + K\mathbf{S}_j \times \hat{e}_{ji}\Big) + \alpha \mathbf{S}_i \times \mathbf{m}_i. \tag{6.25}$$

We introduce a set of orthogonal vectors explicitly,

$$\begin{aligned}
\mathbf{k} &= \frac{1}{\sqrt{2}}(1,1,0) \\
\mathbf{m}_i &= \Big(-\frac{1}{\sqrt{2}}\cos[\mathbf{Q}\cdot\mathbf{r}_i], \frac{1}{\sqrt{2}}\cos[\mathbf{Q}\cdot\mathbf{r}_i], \sin[\mathbf{Q}\cdot\mathbf{r}_i]\Big) \\
\mathbf{n}_i &= \Big(-\frac{1}{\sqrt{2}}\sin[\mathbf{Q}\cdot\mathbf{r}_i], \frac{1}{\sqrt{2}}\sin[\mathbf{Q}\cdot\mathbf{r}_i], -\cos[\mathbf{Q}\cdot\mathbf{r}_i]\Big).
\end{aligned} \tag{6.26}$$

One can easily verify that the three vectors $(\mathbf{n}_i, \mathbf{m}_i, \mathbf{k})$ form a right-handed orthogonal basis at each site i. And using these orthogonal vectors, the equation (6.25) can be decomposed into three components, $S_i^{\mathbf{k}}$, $S_i^{\mathbf{m}}$, $S_i^{\mathbf{n}}$,where

$$S_i^{\mathbf{k}} = \mathbf{S}_i\cdot\mathbf{k}, \quad S_i^{\mathbf{m}} = \mathbf{S}_i\cdot\mathbf{m}_i, \quad S_i^{\mathbf{n}} = \mathbf{S}_i\cdot\mathbf{n}_i.$$

Inserting the vectors (6.26) into (6.25) and using some vector identities and trigonometric formulas, one immediately obtains $\frac{d}{dt}S_i^{\mathbf{m}_i} = 0$, hence $S^{\mathbf{m}_i} = 0$. The other two equations are

$$\begin{aligned}
\frac{d}{dt}S_i^{\mathbf{k}} &= -\alpha S_i^{\mathbf{n}} - J\cos Q\Big(S^{\mathbf{n}}_{i-\hat{\mathbf{x}}} + S^{\mathbf{n}}_{i+\hat{\mathbf{x}}} + S^{\mathbf{n}}_{i-\hat{\mathbf{y}}} + S^{\mathbf{n}}_{i+\hat{\mathbf{y}}}\Big) \\
&\quad -\frac{K}{\sqrt{2}}\Big[\cos[\mathbf{Q}\cdot\mathbf{r}_i]\Big(S^{\mathbf{k}}_{i+\hat{\mathbf{x}}} - S^{\mathbf{k}}_{i-\hat{\mathbf{x}}} + S^{\mathbf{k}}_{i-\hat{\mathbf{y}}} - S^{\mathbf{k}}_{i+\hat{\mathbf{y}}}\Big) \\
&\quad\quad + \sin Q\Big(S^{\mathbf{n}}_{i-\hat{\mathbf{x}}} + S^{\mathbf{n}}_{i+\hat{\mathbf{x}}} + S^{\mathbf{n}}_{i-\hat{\mathbf{y}}} + S^{\mathbf{n}}_{i+\hat{\mathbf{y}}}\Big)\Big] \\
\frac{d}{dt}S_i^{\mathbf{n}} &= \alpha S_i^{\mathbf{k}} + J\Big(S^{\mathbf{k}}_{i-\hat{\mathbf{x}}} + S^{\mathbf{k}}_{i+\hat{\mathbf{x}}} + S^{\mathbf{k}}_{i-\hat{\mathbf{y}}} + S^{\mathbf{k}}_{i+\hat{\mathbf{y}}}\Big) \\
&\quad -\frac{K}{\sqrt{2}}\Big(\cos[\mathbf{Q}\cdot(\mathbf{r}_i+\hat{\mathbf{x}})]S^{\mathbf{n}}_{i+\hat{\mathbf{x}}} - \cos[\mathbf{Q}\cdot(\mathbf{r}_i-\hat{\mathbf{x}})]S^{\mathbf{n}}_{i-\hat{\mathbf{x}}} \\
&\quad + \cos[\mathbf{Q}\cdot(\mathbf{r}_i-\hat{\mathbf{y}})]S^{\mathbf{n}}_{i-\hat{\mathbf{y}}} - \cos[\mathbf{Q}\cdot(\mathbf{r}_i+\hat{\mathbf{y}})]S^{\mathbf{n}}_{i+\hat{\mathbf{y}}}\Big).
\end{aligned} \tag{6.27}$$

where i in sub and superscript means $\mathbf{r}_i$, so $i\pm\hat{\mathbf{x}} = \mathbf{r}_i \pm \hat{\mathbf{x}}$ and $i \pm \hat{y} = \mathbf{r}_i \pm \hat{\mathbf{y}}$.

6.1.4 Multiple spiral ferromagnet

6.2 Holstein-Primakoff theory

There is a neater way of deriving the same spin wave spectrum we have just derived above, using the method of Holstein-Primakoff (HP) representation of spins. The technique applies somewhat differently for an antiferromagnet as it does for a ferromagnet, and we start off with the simpler case of a ferromagnet.

6.2.1 Ferromagnet

First, I will introduce the representation ($n_i = b_i^\dagger b_i$),

$$\begin{aligned} S_i^+ &= S_i^x + iS_i^y = (2S - n_i)^{1/2} b_i, \\ S_i^- &= S_i^x - iS_i^y = b_i^\dagger (2S - n_i)^{1/2}, \\ S_i^z &= S - n_i, \end{aligned} \tag{6.28}$$

then verify that the new expressions on the r.h.s. of the above equations satisfy all the commutation algebras of ordinary spin operators: $[S_i^\alpha, S_j^\beta] = i\delta_{ij}\epsilon_{\alpha\beta\gamma}S_i^\gamma$, provided the new operator b_i is a canonical boson operator, with the familiar bosonic commutation relations, $[b_i, b_j^\dagger] = \delta_{ij}$, etc.

How this comes about does not concern us so much here; rather we will proceed to apply the brand new technique and see if something good comes out. First rewrite

$$\begin{aligned} \mathbf{S}_i \cdot \mathbf{S}_j &= S_i^z S_j^z + \frac{1}{2}(S_i^+ S_j^- + S_i^- S_j^+) \\ &= (S - n_i)(S - n_j) + \frac{1}{2}(2S - n_i)^{\frac{1}{2}}(2S - n_j)^{\frac{1}{2}}(b_i^\dagger b_j + b_j^\dagger b_i). \end{aligned} \tag{6.29}$$

OK, so it doesn't appear particular illuminating after the substitution. But let's try to appeal to our physical senses to see if some simplifications can occur. We already know the ground state of the Hamiltonian is given by a ferromagnetic arrangement of spins, with all the spins pointing in the $\hat{z}$-direction. If we take the average of the operator S_i^z in the ground state, at zero temperature, we should get $\langle S_i^z \rangle = S$. The same can be said of the r.h.s. of the last of Eq. (6.28) because it is just a mathematically equivalent way of writing down the same quantum-mechanical operator, so we must get $\langle S - n_i \rangle = S$, or $\langle n_i \rangle = 0$! This means that the quantum-mechanical ground state $|0\rangle$ of the ferromagnetic Heisenberg spin model is defined by the condition that $\langle 0|n_i|0\rangle = 0$ for all sites i. An excitation is just a small wiggle of the spins away from the absolutely frozen spin configuration that is the ground state, hence we can suppose that $\langle S_i^z \rangle$ deviates only slightly from its ground state value S, or in the language of bosons, that n_i is some small, but non-zero number. In this limiting circumstance, we can simplify the square roots in Eq. (6.28) by taking $\sqrt{2S - n_i} \approx \sqrt{2S}$, and writing

$$\begin{aligned} \mathbf{S}_i \cdot \mathbf{S}_j &\approx (S - n_i)(S - n_j) + \frac{1}{2} \cdot 2S(b_i^\dagger b_j + b_j^\dagger b_i). \\ &\approx S^2 - S(n_i + n_j) + S(b_i^\dagger b_j + b_j^\dagger b_i). \end{aligned} \tag{6.30}$$

The $n_i n_j$ is a product of two small numbers, so we delete them. The fully HP-transformed Hamiltonian, after the approximations we have just made, will be

$$H_{\mathrm{HP}} \approx -JS \sum_{\langle ij \rangle} (b_i^\dagger b_j + b_j^\dagger b_i - b_i^\dagger b_i - b_j^\dagger b_j). \tag{6.31}$$

Well, now the new Hamiltonian is just a quadratic Hamiltonian involving only a product of two boson operators. Such operators can be diagonalized easily, in this case by going to the Fourier space:

$$b_i = \sum_k e^{ik\cdot r_i} b_k. \tag{6.32}$$

Substitution of the Fourier expression for b_i into Eq. (6.31) immediately gives

$$H = \sum_k \omega_k b_k^\dagger b_k \tag{6.33}$$

where

$$\begin{aligned} \omega_k &= JSz(1-\gamma_k) \\ \gamma_k &= z^{-1}\sum_{j\in i} e^{ik\cdot(r_j - r_i)}. \end{aligned} \tag{6.34}$$

The eigenenergy ω_k agrees with the spin-wave spectrum derived earlier, in Eq. (6.10), using the equation-of-motion theory.

6.2.2 Antiferromagnet

The strategy for dealing with elementary spin excitations of an antiferromagnet is essentially the same, but the technique is more advanced.

First of all, we define a new quantum-mechanical operator $\mathcal{S}_i^y$, $\mathcal{S}_i^z$ related to S_i^y, S_i^z by the relation

$$S_i^y + iS_i^z = e^{i\theta_i}(\mathcal{S}_i^y + i\mathcal{S}_i^z) \tag{6.35}$$

where θ_i is the angle of the classical ground state spins, $\theta_i = \pi(i_x + i_y)$. This represents a rotation of spins, with respect to the x-axis, by the angle that just equals the orientation of the classical ground state spins. The x-component remains unchanged, hence $\mathcal{S}_i^x = S_i^x$. Furthermore, one can easily verify that the new operators $\mathcal{S}_i^\alpha$ satisfy all the commutation rules of the original spins, S_i^z, hence they are as good a representation of the local spin operator as the previous one we used. In terms of $\mathcal{S}_i$, we have the antiferromagnetic Hamiltonian written as

$$H = J\sum_{\langle ij\rangle} \mathbf{S}_i\cdot\mathbf{S}_j \to \frac{J}{2}\sum_{\langle ij\rangle}\left(\mathcal{S}_i^+\mathcal{S}_j^+ + \mathcal{S}_i^-\mathcal{S}_j^-\right) - J\sum_{\langle ij\rangle}\mathcal{S}_i^z\mathcal{S}_j^z. \tag{6.36}$$

Note that we have two raising (lowering) operators instead of one raising and one lowering, as was the case before. The ground state of the new Hamiltonian, Eq. (6.36), is given by the rotation of the antiferromagnetic spins, $\langle S_i^z\rangle = Se^{i\theta_i}$. But this implies that in the new representation, $\langle\mathcal{S}_i^z\rangle = S$ everywhere! Since $\langle S_i^y\rangle = 0$ initially, we have $\langle\mathcal{S}_i^y\rangle = 0$, too. So the new Hamiltonian has a *ferromagnetic* ground state with all the spins pointing in the $+\hat{z}$ direction, and we can make the assumption $\langle\mathcal{S}_i^z\rangle \approx S$, or $\langle n_i\rangle \ll S$ after the HP substitution, for the small-fluctuation analysis.

Under these assumptions the HP-transformed Hamiltonian becomes

$$H \approx JS\sum_{\langle ij\rangle}(b_i^\dagger b_j^\dagger + b_i b_j + n_i + n_j). \tag{6.37}$$

Unlike its ferromagnetic counterpart, Eq. (6.31), this Hamiltonian has a product of two annihilation (creation) operators. As before, we take the Fourier transform and obtain

$$H = JSz \sum_k \left[b_k^\dagger b_k + \frac{\gamma_k}{2} (b_k^\dagger b_{-k}^\dagger + b_k b_{-k}) \right]. \tag{6.38}$$

Whereas doing the Fourier transform was sufficient to yield excitation energies in the previous example of ferromagnetic spins, now the Hamiltonian written in momentum space fails to give eigenenergies. The final trick lies in working out the equation of motion of the boson operators, which gives

$$[H, b_k] = -JSz \left(b_k + \gamma_k b_{-k}^\dagger \right). \tag{6.39}$$

Here it's clear why we could not get the energy from Fourier transform alone. The dynamics of the operator b_k is coupled to that of $b_{-k}^\dagger$, and the dynamics of $b_{-k}^\dagger$ is coupled to that of b_k by the equation

$$[H, b_{-k}^\dagger] = JSz \left(b_{-k}^\dagger + \gamma_k b_k \right). \tag{6.40}$$

Noting that $[H, X] = -idX/dt$ in quantum mechanics, and that for the eigenmodes we can re-write the time derivative as $-iE$, where E is the eigenenergy, the coupled oscillator equation we must solve becomes

$$E_k \begin{pmatrix} b_k \\ b_{-k}^\dagger \end{pmatrix} = JSz \begin{pmatrix} 1 & \gamma_k \\ -\gamma_k & -1 \end{pmatrix} \begin{pmatrix} b_k \\ b_{-k}^\dagger \end{pmatrix}. \tag{6.41}$$

Eigenvalues are readily found as $E_k = \pm JSz\sqrt{1-\gamma_k^2}$. Here I state, without proof, that the negative energies are just artifacts of the theory, and that only the positive branch matters. Hence, the spin waves in the antiferromagnetic background carries the energy given by $JSz\sqrt{1-\gamma_k^2}$ for a wave vector k.

6.2.3 Spiral ferromagnet

We work out the HP theory of elementary excitations for the spiral ferromagnet given by the Hamiltonian

$$H_{\text{HDM}} = -J \sum_{\mathbf{r}} \mathbf{S}_{\mathbf{r}} \cdot (\mathbf{S}_{\mathbf{r}+\hat{x}} + \mathbf{S}_{\mathbf{r}+\hat{y}}) - K \sum_{\mathbf{r}} \Big(\mathbf{S}_{\mathbf{r}} \times \mathbf{S}_{\mathbf{r}+\hat{x}} \cdot \hat{x} + \mathbf{S}_{\mathbf{r}} \times \mathbf{S}_{\mathbf{r}+\hat{y}} \cdot \hat{y} \Big). \tag{6.42}$$

6.2.4 Multiple spiral ferromagnet

$$\begin{aligned} H_{\text{HDM}} = \quad & -J \sum_{\mathbf{r}} \mathbf{S}_{\mathbf{r}} \cdot (\mathbf{S}_{\mathbf{r}+\hat{x}} + \mathbf{S}_{\mathbf{r}+\hat{y}}) - K \sum_{\mathbf{r}} \Big(\mathbf{S}_{\mathbf{r}} \times \mathbf{S}_{\mathbf{r}+\hat{x}} \cdot \hat{x} + \mathbf{S}_{\mathbf{r}} \times \mathbf{S}_{\mathbf{r}+\hat{y}} \cdot \hat{y} \Big) \\ & + A_1 \sum_{\mathbf{r}} \Big((S_{\mathbf{r}}^x)^4 + (S_{\mathbf{r}}^y)^4 + (S_{\mathbf{r}}^z)^4 \Big) - A_2 \sum_{\mathbf{r}} \Big(S_{\mathbf{r}}^x S_{\mathbf{r}+\hat{x}}^x + S_{\mathbf{r}}^y S_{\mathbf{r}+\hat{y}}^y \Big) \\ & - \mathbf{H} \cdot \sum_{\mathbf{r}} \mathbf{S}_{\mathbf{r}}. \end{aligned} \tag{6.43}$$

6.3 Schwinger bosons

Schwinger found a way to represent the spin operator using a pair of canonical boson operators (b_1, b_2) as

$$\begin{aligned} S^+ &= S^x + iS^y = b_1^\dagger b_2 \\ S^- &= S^x - iS^y = b_2^\dagger b_1 \\ S^z &= \frac{1}{2}(b_1^\dagger b_1 - b_2^\dagger b_2). \end{aligned} \tag{6.44}$$

More concisely one can write $S = \frac{1}{2} b_\alpha^\dagger \sigma_{\alpha\beta} b_\beta$. The spin operators thus defined obey the canonical spin operator relation

$$[S^\alpha, S^\beta] = i\epsilon^{\alpha\beta\gamma} S^\gamma. \tag{6.45}$$

The original spin operator has the additional property $\mathbf{S} \cdot \mathbf{S} = S(S+1)$ which translates into $(b_1^\dagger b_1 + b_2^\dagger b_2)(b_1^\dagger b_1 + b_2^\dagger b_2 + 2)/4$ upon the substitution of Eq. (6.44). To match the original relation, the sum of the two boson occupation numbers must equal $2S$:

$$b_1^\dagger b_1 + b_2^\dagger b_2 = 2S. \tag{6.46}$$

This puts constraint on the the allowed boson Hilbert space of a and b operators.

The spin state with total spin S and $S_z = m$ is represented

$$|Sm\rangle = \frac{(b_1^\dagger)^{S+m}}{\sqrt{(S+m)!}} \frac{(b_2^\dagger)^{S-m}}{\sqrt{(S-m)!}} |0\rangle. \tag{6.47}$$

For spin half the representation becomes $|\uparrow\rangle = b_1^\dagger|0\rangle$, and $|\downarrow\rangle = b_2^\dagger|0\rangle$.

We can also consider the parametrization

$$b_1 = \sqrt{\frac{S}{2}} \cos\frac{\theta}{2}, b_2 = \sqrt{\frac{S}{2}} \sin\frac{\theta}{2} e^{-i\phi} \tag{6.48}$$

from which we deduce $\mathbf{S} = S(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. For a pair of sites i and j we have the relation[2]

$$\sum_\sigma b_{j\sigma}^\dagger b_{i\sigma} = e^{i\omega_{ij}/2} \left| \frac{1 + \Omega_i \cdot \Omega_j}{2} \right|^{1/2}. \tag{6.49}$$

ω_{ij} is the solid angle subtended by $\hat{z}$, Ω_i and Ω_j. Substituting Eq. (6.48) in the action $b_i^\dagger \partial_\tau b_i$ also gives rise to the Berry phase action first discussed by Haldane.

Now let's see how the Heisenberg Hamiltonian looks like in the Schwinger boson representation. Unlike the Holstein-Primakoff theory, we do not need to know *a priori* what the classical ground state looks like. For this reason the Schwinger boson theory is best applied to cases where the ground state does not possess broken symmetry.

It is useful to decompose the spin-spin interaction term $S_i \cdot S_j$ as

[2] Patrick Lee, PRL **63**, 680 (1989)

$$S_i \cdot S_j = S^2 - \frac{1}{2} A_{ij}^+ A_{ij} \tag{6.50}$$

where $A_{ij} = b_{1i} b_{2j} - b_{2i} b_{1j}$ if it is an antiferromagnet, and as

$$S_i \cdot S_j = \frac{1}{2} B_{ij}^+ B_{ij} - S(S+1) = \frac{1}{2} : B_{ij}^+ B_{ij} : -S^2 \tag{6.51}$$

where $B_{ij} = b_{i1}^\dagger b_{j1} + b_{i2}^\dagger b_{j2}$ if it is a ferromagnet. With this substitution the Heisenberg Hamiltonian, which originally expressed how the nearby spins interact, becomes one for two bosons interacting with each other through the $A_{ij}^+ A_{ij}$ term. Not surprisingly, such interaction Hamiltonian is not exactly solvable. Instead one relies on an approximation scheme which renders the Hamiltonian more analytically tractable. We call it the Schwinger boson mean-field theory (SBMFT) which we describe below.

6.3.1 Ferromagnet

6.3.2 Antiferromagnet

Assume that one of the operators in $A_{ij}^+ A_{ij}$ can be replaced by its average, $Q_{ij} = \langle b_{1i} b_{2j} - b_{2i} b_{1j} \rangle$. Then we obtain the mean-field Hamiltonian

$$H_{MF} = -\frac{1}{2} \sum_{\langle ij \rangle} \left(Q_{ij}^* A_{ij} + Q_{ij} A_{ij}^+ \right). \tag{6.52}$$

The new Hamiltonian contains only a pair of boson operators, and like all quantum theories involving only a product of two operators, it is solvable.

Although the theory of diagonalization of the mean-field Hamiltonian can be readily derived, we first introduce a method that will prove useful for treating a variety of lattice situations. In spirit, this is similar to the rotation of the classical spin angles to make all the spins aligned ferromagnetically along a specific direction.

Let's say the putative classical spin average is given by $\langle S_i \rangle = S\Omega_i$ and $\Omega_i = (\sin\theta_i \cos\phi_i, \sin\theta_i \sin\phi_i, \cos\theta_i)$. This will in turn impose the condition on the Schwinger boson average

$$\begin{aligned} \langle b_{i1}^\dagger b_{i2} \rangle &= S e^{i\phi_i} \sin\theta_i \\ \langle b_{i2}^\dagger b_{i1} \rangle &= S e^{-i\phi_i} \sin\theta_i \\ \frac{1}{2} \langle b_{i1}^\dagger b_{i1} - b_{i2}^\dagger b_{i2} \rangle &= S \cos\theta_i. \end{aligned} \tag{6.53}$$

Introduce the unitary matrix U_i

$$U_i = \begin{pmatrix} \cos(\theta_i/2) & -\sin(\theta_i/2) e^{-i\phi_i} \\ \sin(\theta_i/2) e^{i\phi_i} & \cos(\theta_i/2) \end{pmatrix} \tag{6.54}$$

which has the relation

$$U_i^+ (\Omega_i \cdot \sigma) U_i = \sigma^z. \tag{6.55}$$

It is now claimed that the rotation of the Schwinger boson spinor $\psi_i = (b_{i1} b_{i2})$ by $\psi_i \to U_i \psi_i$ yields for the new Schwinger boson operators the average

$$\langle b_{i1}^\dagger b_{i2} \rangle = \langle b_{i2}^\dagger b_{i1} \rangle = 0, \quad \frac{1}{2} \langle b_{i1}^\dagger b_{i1} - b_{i2}^\dagger b_{i2} \rangle = S. \tag{6.56}$$

It also implies that in the rotated basis the classical spin average becomes ferromagnetic, $\langle S_i \rangle = S\hat{z}$.

To make the case more concrete, consider a square-lattice antiferromagnetic for which the classical spin orientation is $\langle S_i \rangle = (-1)^i S\hat{z}$. The unitary matrix (6.54) becomes a unit matrix for A sublattice sites and σ_x for all the B sublattice sites. Therefore, for B lattice sites, (b_{1j}, b_{2j}) becomes $(-b_{2j}, b_{1j})$. The pair amplitude Q_{ij}, $i \in A$, $j \in B$ becomes $b_{1i}b_{1j} + b_{2i}b_{2j}$. Of course it doesn't really matter what we call the A sublattice, so for every pair of sites $\langle ij \rangle$ we have $A_{ij} = \sum_{m=1}^{2} b_{im}b_{jm}$.

To carry out the mean-field analysis of the transformed Hamiltonian one writes, similar to Eq. (6.57),

$$H_{MF} = -\frac{1}{2} \sum_{\langle ij \rangle} \left(\overline{Q}_{ij} A_{ij} + Q_{ij} A_{ij}^{+} \right). \tag{6.57}$$

with $A_{ij} = b_{i1}b_{j1} + b_{i2}b_{j2}$, $Q_{ij} = \langle A_{ij} \rangle$. We treat the mean-field amplitude as real and uniform, $Q_{ij} = Q$, and obtain

$$\begin{aligned} H_{MF} &= -\frac{Q}{2} \sum_{\langle ij \rangle} \sum_{m=1}^{2} [b_{im}b_{jm}] + h.c. \\ &= -\frac{Q}{4} \left(\sum_{j \in i} e^{ik \cdot (r_j - r_i)} \right) \sum_{km} b_{km}b_{-km} + h.c. \\ &= \sum_{k} \begin{pmatrix} b_{1k}^{\dagger} & b_{2\overline{k}} \end{pmatrix} \begin{pmatrix} 0 & -Q_k \\ -Q_k & 0 \end{pmatrix} \begin{pmatrix} b_{1k} \\ b_{2\overline{k}}^{\dagger} \end{pmatrix} \end{aligned} \tag{6.58}$$

where $Q_k = (Q/4) \sum_{j \in i} e^{ik \cdot (r_j - r_i)}$. Inclusion of the Lagrange multiplier $\sum_i \lambda_i (a_i^{\dagger} a_i + b_i^{\dagger} b_i - 2S)$ renders the mean-field Hamiltonian

$$H_{MF} = \sum_{k} \begin{pmatrix} b_{1k}^{\dagger} & b_{2\overline{k}} \end{pmatrix} \begin{pmatrix} \lambda & -Q_k \\ -Q_k & \lambda \end{pmatrix} \begin{pmatrix} b_{1k} \\ b_{2\overline{k}}^{\dagger} \end{pmatrix}. \tag{6.59}$$

One can bring the Hamiltonian to its diagonalized form by the rotation

$$\begin{pmatrix} b_{1k} \\ b_{2\overline{k}}^{\dagger} \end{pmatrix} = \begin{pmatrix} \cosh\theta_k & \sinh\theta_k \\ \sinh\theta_k & \cosh\theta_k \end{pmatrix} \begin{pmatrix} \gamma_{1k} \\ \gamma_{2k}^{+} \end{pmatrix} \tag{6.60}$$

where the angle θ_k is fixed by

$$\cosh 2\theta_k = \frac{\lambda}{E_k}, \quad \sinh 2\theta_k = \frac{Q_k}{E_k}, \tag{6.61}$$

and the eigenenergy E_k is given by

$$E_k = \sqrt{\lambda^2 - Q_k^2}. \tag{6.62}$$

The minus sign in the energy dispersion (6.62) raises concern for what happens when $\lambda < |Q_k|$. In fact, this never happens since the moment when λ

equals $\min[Q_k]$ is when the bosons begin to condense, and due to the Bose condensation λ is always fixed to maintain $\lambda = \min[Q_k]$.

When boson condensation happens, λ equals Q, and we have

$$E_k = Q\sqrt{1 - (1/4)(\cos k_x + \cos k_y)^2}. \tag{6.63}$$

In fact, this is identical to the spin wave spectrum obtained using the Holstein-Primakoff theory, provided we identify Q with JSz, J =exchange energy, S=spin, z=coordination number. And this is no coincidence. One will find that the same Schwinger boson calculation on the triangular lattice, at the onset of Bose condensation, also has an energy spectrum of the spin waves on the triangular lattice. Schwinger boson theory has the power to capture the physics of both the ordered and the disordered phases of the continuous magnet.

To complete the analysis, the self-consistency relation for Q is derived as

$$Q = \frac{1}{4}\sum_i \tag{6.64}$$

6.3.3 Spiral ferromagnet

6.3.4 Multiple spiral ferromagnet

6.4 Slave boson

In a strong correlation environment there often arises a situation for which each site has a low probability of being occupied by two electrons at once, because of the high energy cost of on-site Coulomb repulsion. Because of this it is useful to formulate the theory using the Hilbert space in which the doubly occupied state is truncated out from the outset. In such a limited Hilbert space consisting of empty ($|0\rangle$), or singly occupied ($|\sigma\rangle$) site, the electron commutation is modified in an odd way. See for example,

$$\{c_\sigma, c_\sigma^+\} = |0\rangle\langle\sigma|\sigma\rangle\langle 0| + |\sigma\rangle\langle 0|0\rangle\langle\sigma| = |0\rangle\langle 0| + |\sigma\rangle\langle\sigma|. \tag{6.65}$$

Normally this should have been one, but due to the limitations of the Hilbert space, the electron creation operator can only connect an empty state to a singly occupied state, but not a singly occupied state to a doubly occupied state. Besides, the result (6.65) equals $1 - |\overline{\sigma}\rangle\langle\overline{\sigma}|$ because we have

$$|0\rangle\langle 0| + \sum_\sigma |\sigma\rangle\langle\sigma| = 1, \tag{6.66}$$

also due to the limited Hilbert space imposed by the no double occupancy constraint. Using a similar idea, we can derive

$$\begin{aligned} \{c_\sigma, c_\sigma^+\} &= 1 - c_{\overline{\sigma}}^+ c_{\overline{\sigma}} \\ \{c_\sigma, c_{\overline{\sigma}}^+\} &= c_{\overline{\sigma}}^+ c_\sigma. \end{aligned} \tag{6.67}$$

On the other hand, commutation relation with the density operator $n = \sum_\sigma |\sigma\rangle\langle\sigma|$ is the same.

$$[n, c_\sigma^+] = c_\sigma^+, \quad [n, c_\sigma] = -c_\sigma. \tag{6.68}$$

The slave-boson approach re-write the electron operator in this restricted Hilbert space as a product of a boson (b_i) and a fermion ($f_{i\sigma}$) operator

$$c_\sigma = b^\dagger f_\sigma. \tag{6.69}$$

Quite remarkably, this prescription recovers all the commutations of the restricted Hilbert space, (6.67)-(6.68), provided we impose the constraint

$$b^\dagger b + \sum_\sigma f_\sigma^+ f_\sigma = 1. \tag{6.70}$$

In this representation, the electron number operator becomes

$$c_\sigma^+ c_\sigma = b b^\dagger f_\sigma^+ f_\sigma = (1 + b^\dagger b) f_\sigma^+ f_\sigma = (2 - \sum_\sigma f_\sigma^+ f_\sigma) f_\sigma^+ f_\sigma = f_\sigma^+ f_\sigma. \tag{6.71}$$

Similarly one can show that the spin operator has the expression

$$S = \frac{1}{2} c_\alpha^+ \sigma_{\alpha\beta} c_\beta = \frac{1}{2} f_\alpha^+ \sigma_{\alpha\beta} f_\beta. \tag{6.72}$$

6.5 Slave-fermion Schwinger boson

The Schwinger boson formalism introduced in the previous section gave a good way to handle the spin dynamics, especially in the incoherent regime before the magnetic long-range order has taken place.

A system where all the dynamics arise from that of the spins is by definition an insulator. On the other hand, a metallic state allows the motion of individual electrons from one site to the next, and this requires the introduction of an electron operator $c_{i\sigma}$ in addition to the spin operator S_i in the Hamiltonian. The Schwinger boson theory by itself does not tell us how one can write down the electron operator in terms of the Schwinger bosons. To do that, one must introduce yet another kind of operator, called the slave-fermion operator denoted f_i. Using this operator, an electron operator gets written as

$$c_\sigma = f^+ b_\sigma. \tag{6.73}$$

Compared to the slave-boson substitution, here the role of boson and fermion is simply reversed. The advantage is that now one obtains an easy generalization of the Schwinger boson theory to treat the conducting electrons. Similar to the slave-boson case, we have

$$\begin{aligned} c_\sigma^+ c_\sigma &= b_\sigma^\dagger b_\sigma \\ c_\alpha^+ \sigma_{\alpha\beta} c_\beta &= b_\alpha^\dagger \sigma_{\alpha\beta} b_\beta \end{aligned} \tag{6.74}$$

The latter equation is precisely the Schwinger boson prescription.

Chapter 7

t-J Model

The "standard model" for the high-T_c cuprates is the t-J model. A popular method used in the solution of the t-J model is first to introduce the slave boson coordinates to rewrite the Hamiltonian, and apply mean-field theory. Other, more sophisticated techniques exist, e.g. DMRG, exact diagonalization, QMC, etc, which I do not talk about here. While the mean-field theory is the least convincing of all the techniques, it is the most versatile and can be applied to all sorts of problems involving t-J Hamiltonian.

The t-J model is defined by the Hamiltonian

$$H = -t\sum_{\langle ij\rangle}(c^\dagger_{j\alpha}c_{i\alpha} + h.c.) + J\sum_{\langle ij\rangle}(\mathbf{S}_i\cdot\mathbf{S}_j - \frac{1}{4}n_in_j) + \frac{V_c}{2}\sum_{i\neq j}\frac{1}{r_{ij}}(n_i-\bar{n})(n_j-\bar{n}), \tag{7.1}$$

with typically the Coulomb term V_c set to zero. Each site i is roughly like an atomic orbital, whose wavefunction hybridizes with its neighbors' to produce itinerancy. Because of the large Coulomb energy of putting two electrons in a given atomic orbital (a few eV), it is usually believed that double occupation of orbitals *never happens*.

A naive diagonalization of the above Hamiltonian would naturally involve states which have two electrons at a given site, $|\uparrow\downarrow\rangle_i$, which should cost very high energy. In the t-J model this energy is treated as infinity, and the diagonalization takes place within the restricted Hilbert space consisting of an empty site, and single occupied sites of either spin orientation. This sort of problem was confronted sometime in early 80's by using the so-called slave-boson, or auxiliary-boson technique. In a nutshell, one writes an electron operator as a composite of a boson operator b_i and another fermion operator $f_{i\sigma}$ as $c_{i\sigma} = b^\dagger_i f_{i\sigma}$. In this new choice of variables, the no-double-occupancy constraint becomes

$$b^\dagger_i b_i + \sum_\sigma f^\dagger_{i\sigma}f_{i\sigma} = 1 \tag{7.2}$$

for all i. In loose terms it means that the occupation of a given orbital by a vacancy ($b^\dagger_i b_i$) or by a single spin ($f^\dagger_{i\sigma}f_{i\sigma}$) should always add up to one. Double occupancy is then by definition excluded because $\sum_\sigma f^\dagger_{i\sigma}f_{i\sigma} = 1 - b^\dagger_i b_i \leq 1$.

In the path-integral picture the constraint is implemented through the Lagrange multiplier technique, which amounts to adding to the original Lagrangian the term

$$L_{tJ} \to L_{tJ} + i\sum_i (b_i^\dagger b_i + f_i^\dagger f_i - 1). \tag{7.3}$$

The Hamiltonian can be accordingly rewritten as

$$\begin{aligned} H = &-t\sum_{\langle ij\rangle}(b_j b_i^\dagger f_{j\alpha}^\dagger f_{i\alpha} + h.c.) + J\sum_{\langle ij\rangle}(\mathbf{S}_i\cdot\mathbf{S}_j - \frac{1}{4}n_i n_j) \\ &+\frac{V_c}{2}\sum_{i\neq j}\frac{1}{r_{ij}}(n_i-\bar{n})(n_j-\bar{n}) + \sum_i \lambda_i(b_i^\dagger b_i + f_{i\sigma}^\dagger f_{i\sigma} - 1) - \mu(\sum_{i\sigma} f_{i\sigma}^\dagger f_{i\sigma} - N_e). \end{aligned} \tag{7.4}$$

The Lagrange multipliers appear as λ_i and the chemical potential μ is adjusted so as to obtain the desired total electron number N_e. The spin-spin interaction term can be written using the fermion coordinates,

$$S_i = \frac{1}{2} f_{i\alpha}^\dagger \sigma_i^{\alpha\beta} f_{i\beta}. \tag{7.5}$$

In other words,

$$\begin{aligned} 2S_{ix} &= f_{i\uparrow}^\dagger f_{i\downarrow} + f_{i\downarrow}^\dagger f_{i\uparrow} \\ 2S_{iy} &= -i(f_{i\uparrow}^\dagger f_{i\downarrow} - f_{i\downarrow}^\dagger f_{i\uparrow}) \\ 2S_{iz} &= f_{i\uparrow}^\dagger f_{i\uparrow} - f_{i\downarrow}^\dagger f_{i\downarrow}. \end{aligned} \tag{7.6}$$

This fermion representation of spin satisfies all the commutator algebra of the original spin. By substituting Eq. (6) back into Eq. (5) one completes the expression of the t-J model in terms of slave-boson coordinates b_i and $f_{i\sigma}$.

It is horribly difficult to solve a Hamiltonian which contains a quartic interaction term. The J-term is quartic in the fermions, and the t-term is quartic because it contains two b_i's and two $f_{i\sigma}$'s. One can reduce the problem to the fermion-only problem by treating the boson fields b_i as *condensed*, and satisfying the constraint

$$\langle b_i^\dagger b_i\rangle = 1 - \langle f_{i\sigma}^\dagger f_{i\sigma}\rangle. \tag{7.7}$$

In reality one replaces the boson fields in the Hamiltonian by a complex number whose magnitude is determined from $|b_i|^2 = 1 - \langle f_{i\sigma}^\dagger f_{i\sigma}\rangle$. Because the Hamiltonian is invariant under a local phase change $b_i \to e^{i\theta_i} b_i$, $f_{i\sigma} \to e^{\imath\theta_i} f_{i\sigma}$, one can always choose b_i to be real and positive without loss of generality.

The density-density interaction term can be rewritten as

$$\frac{V_c}{2}\sum_{i\neq j}\frac{1}{r_{ij}}(b_i^2 - x)(b_j^2 - x), \tag{7.8}$$

with b_i^2 given as one minus the local electron density.

Having removed the boson degrees of freedom, one can now face the quartic interaction term by mean-field decoupling technique. One groups the quartic term as a product of two bilinear operators, and replace one of them by its expectation values. There are three distinct ways to pair up the fields, hence three different order parameters in the meanfield theory. Keeping track of the Fock and pairing terms in the mean-field decoupling scheme gives

$$\begin{aligned} S_i \cdot S_j &\rightarrow -\frac{3}{8}(\overline{\Delta}^s_{ij} P^s_{ij} + \overline{\chi}^s_{ij} H^s_{ij}) + \frac{1}{8}(\overline{\Delta}^t_{ij} P^t_{ij} + \overline{\chi}^t_{ij} H^t_{ij}) \\ \frac{1}{4} n_i n_j &\rightarrow \frac{1}{8}(\overline{\Delta}^s_{ij} P^s_{ij} - \overline{\chi}^s_{ij} H^s_{ij}) + \frac{1}{8}(\overline{\Delta}^t_{ij} P^t_{ij} - \overline{\chi}^t_{ij} H^t_{ij}) \end{aligned} \tag{7.9}$$

plus their hermitian conjugates. The singlet/triplet operators are defined by

$$\begin{aligned} P^{s/t}_{ij} &= f_{i1} f_{j2} \mp f_{i2} f_{j1}, \\ H^{s/t}_{ij} &= f^\dagger_{j1} f_{i1} \pm f^\dagger_{j2} f_{i2}, \end{aligned} \tag{7.10}$$

and their expectation values by

$$\Delta^{s/t}_{ij} = \langle P^{s/t}_{ij} \rangle, \quad \chi^{s/t}_{ij} = \langle H^{s/t}_{ij} \rangle. \tag{7.11}$$

In actual treatments, the triplet components are assumed to be zero. When I didn't know better, I decoupled the $S_i \cdot S_j$ fully in all three channels, but did the Hartree decoupling only with respect to the $-\frac{1}{4} n_i n_j$ in the J-term[1]. As a result what I had for the mean-field Hamiltonian is

$$\begin{aligned} H_{MF} &= -t \sum_{\langle ij \rangle} (b_j b_i f^\dagger_{j\alpha} f_{i\alpha} + h.c.) - \frac{3J}{8} \sum_{\langle ij \rangle} (\Delta_{ij} P^\dagger_{ij} + \chi_{ij} H^\dagger_{ij} + h.c.) \\ &+ \frac{J}{4} \sum_i M_i \sigma_{iz} - \frac{J}{4} \sum_{\langle ij \rangle} (1 - b_i^2)(1 - b_j^2) \\ &+ \frac{V_c}{2} \sum_{i \neq j} \frac{1}{r_{ij}} (b_i^2 - x)(b_j^2 - x) + \sum_i \lambda_i (b_i^2 + f^\dagger_{i\sigma} f_{i\sigma} - 1) - \mu \sum_i f^\dagger_{i\sigma} f_{i\sigma}, \end{aligned} \tag{7.12}$$

where $P_{ij} = P^s_{ij}$, and $H_{ij} = H^s_{ij}$ and $\Delta_{ij} = \Delta^s_{ij}, K_{ij} = K^s_{ij}$, respectively. σ_{iz} is the Pauli matrix for the local magnetic moment whose axis is chosen in the z-direction, and M_i is the local magnetic field generated by the neighbouring spins, $M_i = \sum_j \langle \sigma_{jz} \rangle$ (j=four nearest neighbours of i). Summation of the spin index σ is implicitly assumed. One can also find treatments where coefficients are $J/2$ for the pairing term, and $J/4$ for the hopping term, because they decoupled $n_i n_j$ according to Eq. (7.9). We will distinguish them as scheme I (sI) and scheme II (sII). The maximal gap value will be equal to $4 \times (3J/8)|\Delta_{ij}|$ (sI), or to $4 \times (4J/8)|\Delta_{ij}|$ (sII).

[1] Also other people decoupled it this way in order to preserve the SU(2) symmetry which came from the Heisenberg model.

The fermion part of the Hamiltonian is expressed in the Nambu form by introducing the two spinors

$$\psi_{i\uparrow} = \begin{pmatrix} f_{i\uparrow} \\ f^+_{i\downarrow} \end{pmatrix}, \quad \psi_{i\downarrow} = \begin{pmatrix} f_{i\downarrow} \\ -f^+_{i\uparrow} \end{pmatrix}, \tag{7.13}$$

and the matrices

$$\begin{aligned} T_{ij} &= \begin{pmatrix} -\chi^*_{ij} - (8t/3J)b_i b_j & \Delta_{ij} \\ \Delta^*_{ij} & \chi_{ij} + (8t/3J)b_i b_j \end{pmatrix}, \\ M_{i\sigma} &= \begin{pmatrix} \lambda_i + \sigma m_i & 0 \\ 0 & -\lambda_i + \sigma m_i \end{pmatrix}, \end{aligned} \tag{7.14}$$

as

$$\begin{aligned} H_{MF} &= \frac{3}{8} J \sum_{\langle ij \rangle \sigma} \psi^+_{j\sigma} T_{ij} \psi_{i\sigma} + \frac{1}{2} \sum_{i\sigma} \psi^+_{i\sigma} M_{i\sigma} \psi_{i\sigma} \\ &= \frac{3}{16} J \sum_{i\sigma} \sum_{j \in i} \psi^+_{j\sigma} T_{ij} \psi_{i\sigma} + \frac{1}{2} \sum_{i\sigma} \psi^+_{i\sigma} M_{i\sigma} \psi_{i\sigma}. \end{aligned} \tag{7.15}$$

For actual calculation it is easier if we re-arrange the $\psi_{j\sigma}$ to appear on the right:

$$\begin{aligned} H_{MF} &= \frac{3}{16} J \sum_{i\sigma} \sum_{j \in i} \psi^+_{i\sigma} T_{ji} \psi_{j\sigma} + \frac{1}{2} \sum_{i\sigma} \psi^+_{i\sigma} M_{i\sigma} \psi_{i\sigma}, \\ T_{ji} &= \begin{pmatrix} -\chi_{ij} - (8t/3J)b_i b_j & \Delta_{ij} \\ \Delta^*_{ij} & \chi^*_{ij} + (8t/3J)b_i b_j \end{pmatrix}. \end{aligned} \tag{7.16}$$

From now on we will write $\chi'_{ij} = \chi_{ij} + (8t/3J)b_i b_j$.

The above fermion Hamiltonian can be diagonalized by an appropriate Bogoliubov rotation of the operators. We will first work out the case without magnetism, so that $M_{i\sigma} = M_i = \lambda_i \sigma_z$. The spinors are rotated according to the Bogoliubov matrix (1=spin up, 2=spin down)

$$\begin{aligned} \psi_{i\sigma} &= \sum_n \begin{pmatrix} u_{ni} & -v^*_{ni} \\ v_{ni} & u^*_{ni} \end{pmatrix} \Gamma_{n\sigma}, \\ \Gamma_{n1} &= \begin{pmatrix} \gamma_{i1} \\ \gamma^+_{i2} \end{pmatrix}, \quad \Gamma_{i2} = \begin{pmatrix} \gamma_{i2} \\ -\gamma^+_{i1} \end{pmatrix}. \end{aligned} \tag{7.17}$$

Provided the eigenfunctions (u_{ni}, v_{ni}) obey the equation (taking $J \equiv 1$)

$$\begin{aligned} -\frac{3}{8}\chi'_{ij} u_{nj} + \frac{3}{8}\Delta_{ij} v_{nj} + \lambda_i u_{ni} &= E_n u_{ni} \\ \frac{3}{8}\bar{\Delta}_{ij} u_{nj} + \frac{3}{8}\bar{\chi}'_{ij} v_{nj} - \lambda_i v_{ni} &= E_n v_{ni}, \end{aligned} \tag{7.18}$$

we will have

$$\sum_{j\in i} T_{ji} \begin{pmatrix} u_{nj} & -v^*_{nj} \\ v_{nj} & u^*_{nj} \end{pmatrix} = E_n \begin{pmatrix} u_{ni} & v^*_{ni} \\ v_{ni} & -u^*_{ni} \end{pmatrix}$$
$$\sum_i \sum_{j\in i} \begin{pmatrix} u^*_{ni} & v^*_{ni} \\ -v_{ni} & u_{ni} \end{pmatrix} T_{ji} \begin{pmatrix} u_{nj} & -v^*_{nj} \\ v_{nj} & u^*_{nj} \end{pmatrix} = E_n \sigma_z, \tag{7.19}$$

and the mean field Hamiltonian in diagonal form

$$H_{MF} = \frac{3}{16} \sum_n E_n \Gamma^+_{n\sigma} \sigma_z \Gamma_{n\sigma} = \frac{3}{8} \sum_{n\sigma} E_n \gamma^+_{n\sigma} \gamma_{n\sigma}. \tag{7.20}$$

Self-consistent parameters are calculated according to the formulae ($F(x) \equiv 1/(\exp(x/T)+1)$):

$$\begin{aligned} \chi_{ij} &= \langle \sum_\sigma f^+_{j\sigma} f_{i\sigma} \rangle = \sum_n \Big(u_{ni} u^*_{nj} F(E_n) + v^*_{ni} v_{nj} F(-E_n) \Big) \\ \Delta_{ij} &= \langle \epsilon_{\alpha\beta} f_{i\alpha} f_{j\beta} \rangle = \sum_n \Big(u_{ni} v^*_{nj} F(-E_n) - v^*_{ni} u_{nj} F(E_n) \Big) \\ n_i &= \langle \sum_\sigma f^\dagger_{i\sigma} f_{i\sigma} \rangle = \sum_n \Big(|u_{ni}|^2 F(E_n) + |v_{ni}|^2 F(-E_n) \Big). \end{aligned} \tag{7.21}$$

The sum $\sum_n$ runs over both positive and negative energy sets. The usual identity that (u_{ni}, v_{ni}) of energy $-E$ is equal to $(-v^*_{ni}, u^*_{ni})$ at energy $+E$ still holds and allows the self-consistency equations to simplify to

$$\begin{aligned} \chi_{ij} &= 2 \sum_n u_{ni} u^*_{nj} F(E_n) \\ \Delta_{ij} &= -2 \sum_n v^*_{ni} u_{nj} F(E_n) \\ n_i &= 2 \sum_n |u_{ni}|^2 F(E_n). \end{aligned} \tag{7.22}$$

Now we put the magnetism back in and see how to diagonalize the mean field Hamiltonian. We try a slightly different Bogoliubov rotation than the previous case,

$$\begin{aligned} f_{i1} &= \sum_n (u_{ni} \gamma_{n1} - y^*_{ni} \gamma^+_{n2}), \qquad & f_{i2} &= \sum_n (x_{ni} \gamma_{n2} + v^*_{ni} \gamma^+_{n1}), \\ \psi_{i1} &= \sum_n \begin{pmatrix} u_{ni} & -y^*_{ni} \\ v_{ni} & x^*_{ni} \end{pmatrix} \Gamma_{n1}, \qquad & \psi_{i2} &= \sum_n \begin{pmatrix} x_{ni} & -v^*_{ni} \\ y_{ni} & u^*_{ni} \end{pmatrix} \Gamma_{n2}. \end{aligned} \tag{7.23}$$

This time the correct equations to be satisfied by (u_{ni}, v_{ni}) are

$$\begin{aligned} -\frac{3}{8}\chi'_{ij}u_{nj} + \frac{3}{8}\Delta_{ij}v_{nj} + (m_i+\lambda_i)u_{ni} &= E_n u_{ni} \\ \frac{3}{8}\bar{\Delta}_{ij}u_{nj} + \frac{3}{8}\bar{\chi}'_{ij}v_{nj} + (m_i-\lambda_i)v_{ni} &= E_n v_{ni} \end{aligned} \tag{7.24}$$

with $\chi'_{ij} \equiv \chi_{ij} + (8t/3)b_i b_j$ (sI), or

$$\begin{aligned} -\frac{2}{8}\chi'_{ij}u_{nj} + \frac{4}{8}\Delta_{ij}v_{nj} + (m_i+\lambda_i)u_{ni} &= E_n u_{ni} \\ \frac{4}{8}\bar{\Delta}_{ij}u_{nj} + \frac{2}{8}\bar{\chi}'_{ij}v_{nj} + (m_i-\lambda_i)v_{ni} &= E_n v_{ni} \end{aligned} \tag{7.25}$$

with $\chi'_{ij} \equiv \chi_{ij} + (8t/2)b_i b_j$ (sII). Eigenfunctions (x_{ni}, y_{ni}) are determined from the same equation with $m_i \to -m_i$. Alternatively one can obtain (x_{ni}, y_{ni}) of energy E as $(\bar{v}_{ni}, -\bar{u}_{ni})$ where (u_{ni}, v_{ni}) is an eigenstate of energy $-E$. This statement holds for arbitrary magnetic states: ferromagnetic, antiferromagnetic, and what not.

Self-consistency equation after considering the relation between (x_{ni}, y_{ni}) and (u_{ni}, v_{ni}) becomes Eq. (7.21) without it being reducible to Eq. (7.22). Additional relation for local magnetic moment is obtained

$$M_i = \langle \sum_\sigma \sigma f^\dagger_{i\sigma} f_{i\sigma} \rangle = \sum_n \Big(|u_{ni}|^2 F(E_n) - |v_{ni}|^2 F(-E_n) \Big). \tag{7.26}$$

Specializing to uniform case: We think of a uniform situation where each eigenstate n is associated with a momentum vector $\mathbf{k}$ and one can write $u_{ni} = \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}_i} u_{\mathbf{k}}$, $v_{ni} = \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}_i} v_{\mathbf{k}}$. For the square lattice one has

$$\begin{aligned} &\chi_{i,i+\hat{x}} = \chi^*_{i+\hat{x},i} = \chi_x, \quad \chi_{i,i+\hat{y}} = \chi^*_{i+\hat{y},i} = \chi_y, \\ &\Delta_{i,i+\hat{x}} = \Delta_{i+\hat{x},i} = \Delta_x, \quad \Delta_{i,i+\hat{y}} = \Delta_{i+\hat{y},i} = \Delta_y \end{aligned} \tag{7.27}$$

with four complex parameters $(\chi_x, \chi_y, \Delta_x, \Delta_y)$. If the local magnetic moment m_i is staggered, a unit cell would have to be doubled. When we consider a uniform magnetization due to, say, a strong external magnetic field, one can take $m_i = m$ and the BdG equation in momentum space reduces to ($\chi'_x = \chi_x + 8t\delta/3$, etc.)

$$\begin{aligned} &-\frac{3}{8}[\chi'_x e^{ik_x} + \overline{\chi}'_x e^{-ik_x} + \chi'_y e^{ik_y} + \overline{\chi}'_y e^{-ik_y}]u_{\mathbf{k}} + \frac{3}{4}[\Delta_x \cos k_x + \Delta_y \cos k_y]v_{\mathbf{k}} + (m+\lambda)u_{\mathbf{k}} = E_{\mathbf{k}} u_{\mathbf{k}} \\ &\frac{3}{4}[\overline{\Delta}_x \cos k_x + \overline{\Delta}_y \cos k_y]u_{\mathbf{k}} + \frac{3}{8}[\chi'_x e^{-ik_x} + \overline{\chi}'_x e^{ik_x} + \chi'_y e^{-ik_y} + \overline{\chi}'_y e^{ik_y}]v_{\mathbf{k}} + (m-\lambda)v_{\mathbf{k}} = E_{\mathbf{k}} v_{\mathbf{k}}. \end{aligned} \tag{7.28}$$

With the definitions

$$A_{\mathbf{k}} = \frac{3}{8}[\chi'_x e^{ik_x} + \overline{\chi}'_x e^{-ik_x} + \chi'_y e^{ik_y} + \overline{\chi}'_y e^{-ik_y}], B_{\mathbf{k}} = \frac{3}{4}[\Delta_x \cos k_x + \Delta_y \cos k_y], \tag{7.29}$$

I can reduce it to a simple-looking form

$$\begin{pmatrix} -A_{\mathbf{k}} + \lambda & B_{\mathbf{k}} \\ B^*_{\mathbf{k}} & A_{-\mathbf{k}} - \lambda \end{pmatrix} \begin{pmatrix} u_{\mathbf{k}} \\ v_{\mathbf{k}} \end{pmatrix} = (E_{\mathbf{k}} - m) \begin{pmatrix} u_{\mathbf{k}} \\ v_{\mathbf{k}} \end{pmatrix}. \tag{7.30}$$

Presumably it is OK to take χ_{ij} to be real in this case, so that $A_{\mathbf{k}} = (3/4)[\chi'_x \cos k_x + \chi'_y \cos k_y] = A_{-\mathbf{k}}$. The eigenenergies follow as $E_{\mathbf{k}} = m \pm \varepsilon_{\mathbf{k}}$, $\varepsilon_{\mathbf{k}} = \sqrt{(\lambda - A_{\mathbf{k}})^2 + |B_{\mathbf{k}}|^2}$. The average quantities are

$$\begin{aligned}
\chi_x &= \sum_{E_{\mathbf{k}}} \Big(|u_{\mathbf{k}}|^2 e^{-ik_x} F(E_{\mathbf{k}}) + |v_{\mathbf{k}}|^2 e^{ik_x} F(-E_{\mathbf{k}}) \Big) \\
\chi_y &= \sum_{E_{\mathbf{k}}} \Big(|u_{\mathbf{k}}|^2 e^{-ik_y} F(E_{\mathbf{k}}) + |v_{\mathbf{k}}|^2 e^{ik_y} F(-E_{\mathbf{k}}) \Big) \\
\Delta_x &= \sum_{E_{\mathbf{k}}} u_{\mathbf{k}} v^*_{\mathbf{k}} \Big(e^{-ik_x} F(-E_{\mathbf{k}}) - e^{ik_x} F(E_{\mathbf{k}}) \Big) \\
\Delta_y &= \sum_{E_{\mathbf{k}}} u_{\mathbf{k}} v^*_{\mathbf{k}} \Big(e^{-ik_y} F(-E_{\mathbf{k}}) - e^{ik_y} F(E_{\mathbf{k}}) \Big) \\
n &= \sum_{E_{\mathbf{k}}} \Big(|u_{\mathbf{k}}|^2 F(E_{\mathbf{k}}) + |v_{\mathbf{k}}|^2 F(-E_{\mathbf{k}}) \Big).
\end{aligned} \tag{7.31}$$

For the positive energy branch, $E_{\mathbf{k}} = m + \varepsilon_{\mathbf{k}}$, we obtain

$$|u_{\mathbf{k}}|^2 = \frac{1}{2}\left(1 + \frac{\lambda - A_{\mathbf{k}}}{\varepsilon_{\mathbf{k}}}\right), |v_{\mathbf{k}}|^2 = \frac{1}{2}\left(1 - \frac{\lambda - A_{\mathbf{k}}}{\varepsilon_{\mathbf{k}}}\right), u_{\mathbf{k}} v^*_{\mathbf{k}} = \frac{1}{2}\frac{B_{\mathbf{k}}}{\varepsilon_{\mathbf{k}}}. \tag{7.32}$$

For the negative energy branch, $E_{\mathbf{k}} = m - \varepsilon_{\mathbf{k}}$, we obtain

$$|u_{\mathbf{k}}|^2 = \frac{1}{2}\left(1 - \frac{\lambda - A_{\mathbf{k}}}{\varepsilon_{\mathbf{k}}}\right), |v_{\mathbf{k}}|^2 = \frac{1}{2}\left(1 + \frac{\lambda - A_{\mathbf{k}}}{\varepsilon_{\mathbf{k}}}\right), u_{\mathbf{k}} v^*_{\mathbf{k}} = -\frac{1}{2}\frac{B_{\mathbf{k}}}{\varepsilon_{\mathbf{k}}}. \tag{7.33}$$

Chapter 8

Exact Spin Hamiltonians

8.1 AKLT States

The AKLT (Affleck, Kennedy, Lieb, Tasaki) Hamiltonian is given by

$$H^{\mathrm{AKLT}} = \sum_{\langle ij \rangle} \mathcal{P}_{2S}(ij), \tag{8.1}$$

where $\mathcal{P}_{2S}$ projects the spin operators $\boldsymbol{J}_{ij} = \boldsymbol{S}_i + \boldsymbol{S}_j$ for nearest-neighbor pair of adjacent spins $\langle ij \rangle$ onto a subspace of magnitude $2S$. There are two constraints on the spin operators in writing down the AKLT states,

1. $\boldsymbol{S}_i^2 = S(S+1)$ for any spin $\boldsymbol{S}_i$.
2. The value of S must satisfy $z = 2S$, where z is the coordination number.

The projection operator is given by

$$\mathcal{P}_{2S}(ij) = \frac{\boldsymbol{J}_{ij}^2}{2S(2S+1)} \times \frac{\boldsymbol{J}_{ij}^2 - 1 \cdot 2}{2S(2S+1) - 1 \cdot 2} \times \cdots \times \frac{\boldsymbol{J}_{ij}^2 - (2S-1) \cdot 2S}{2S(2S+1) - 2S(2S-1)} \ . \tag{8.2}$$

Any pair of spins with a net spin less than $2S$ is projected out, *i.e.* $\mathcal{P}_{2S}(ij)|J < 2S\rangle = 0$. For $J = 2S$, which is the maximal value attainable for the sum of two spin-S operators, we get $\mathcal{P}_{2S}(ij)|2S\rangle = |2S\rangle$. For $S = 1$,

$$\mathcal{P}_2(ij) = \frac{\boldsymbol{J}_{ij}^2(\boldsymbol{J}_{ij}^2 - 2)}{2 \cdot 3 \cdot 4} = \frac{1}{24}\boldsymbol{J}_{ij}^2(\boldsymbol{J}_{ij}^2 - 2) \tag{8.3}$$

and since $\boldsymbol{J}_{ij}^2 = (\boldsymbol{S}_i + \boldsymbol{S}_j)^2 = 2S(S+1) + 2\boldsymbol{S}_i \cdot \boldsymbol{S}_j$,

$$\mathcal{P}_2(ij) = \boldsymbol{S}_i \cdot \boldsymbol{S}_j + \frac{1}{3}(\boldsymbol{S}_i \cdot \boldsymbol{S}_j)^2 + \frac{2}{3} \ . \tag{8.4}$$

Hence the AKLT Hamiltonian for a one-dimensional spin-1 chain is written as

$$H^{\mathrm{AKLT}} = \sum_{\langle ij \rangle} \left[\boldsymbol{S}_i \cdot \boldsymbol{S}_j + \frac{1}{3}(\boldsymbol{S}_i \cdot \boldsymbol{S}_j)^2 \right] \ . \tag{8.5}$$

A similar reasoning leads to the AKLT Hamiltonian in two dimensional square and triangular lattices. For a two-dimensional square lattice with spin-2 the Hamiltonian is

$$H_{\square}^{\rm AKLT} = \sum_{\langle ij\rangle} (\boldsymbol{S}_i \cdot \boldsymbol{S}_j)(\boldsymbol{S}_i \cdot \boldsymbol{S}_j + 3)(\boldsymbol{S}_i \cdot \boldsymbol{S}_j + 5)(\boldsymbol{S}_i \cdot \boldsymbol{S}_j + 6) \tag{8.6}$$

The Hamiltonian of a two-dimensional triangular lattice with spin-3 is

$$\begin{aligned} H_{\triangle}^{\rm AKLT} = \sum_{\langle ij\rangle} &(\boldsymbol{S}_i \cdot \boldsymbol{S}_j + 12)(\boldsymbol{S}_i \cdot \boldsymbol{S}_j + 11)(\boldsymbol{S}_i \cdot \boldsymbol{S}_j + 9) \times \\ &(\boldsymbol{S}_i \cdot \boldsymbol{S}_j + 6)(\boldsymbol{S}_i \cdot \boldsymbol{S}_j + 2)(\boldsymbol{S}_i \cdot \boldsymbol{S}_j - 3) \end{aligned} \tag{8.7}$$

By construction, the AKLT Hamiltonian is positive semi-definite and therefore the ground state can be at most the zero energy eigenstate. Such a state can be constructed in the Schwinger boson representation as

$$|{\rm AKLT}\rangle = \prod_{\langle ij\rangle} (a_i^\dagger b_j^\dagger - b_i^\dagger a_j^\dagger)|0\rangle \ . \tag{8.8}$$

Each $\langle ij\rangle$ bond is covered by a dimer, a spin singlet, made up of two spin-1/2 constituents. The AKLT state has, for each $\langle ij\rangle$ bond, a maximum $J_{ij,z}$ value of $2S-1$. Then we must have J of that bond less than $2S$, and $\mathcal{P}_{2S}(ij)|{\rm AKLT}\rangle = 0$. To see that $J_{ij,z}$ does not exceed $2S-1$, we take the square lattice as an example.

From Eq. (8.8), it is clear that $a_i^\dagger b_j^\dagger - b_i^\dagger a_j^\dagger$ for the $\langle ij\rangle$ bond gives $J_{ij,z} = 0$. According to Fig. **??**, bonds $\langle i2\rangle$, $\langle i3\rangle$, and $\langle i4\rangle$ contribute $a_i^\dagger b_2^\dagger$, $a_i^\dagger b_3^\dagger$, and $a_i^\dagger b_4^\dagger$, respectively and bonds $\langle j5\rangle$, $\langle j6\rangle$, and $\langle j7\rangle$ contribute $a_j^\dagger b_5^\dagger$, $a_j^\dagger b_6^\dagger$, and $a_j^\dagger b_7^\dagger$, respectively. We have up to six up $S = 1/2$ spins or $(a_i^\dagger)^3 (a_j^\dagger)^3$ for the $\langle ij\rangle$ bonds. So $J_{ij,z}^{\rm max} = 3 = 2(z-1)\cdot\frac{1}{2} = z - 1 = 2S - 1$. Since the wavefunction does not possess any $J = 2S$ component for $\langle ij\rangle$ bonds, one must have $J_{2S}^2(ij)|{\rm AKLT}\rangle = 0$.

8.2 Majumdar-Ghosh states

The Majumdar-Ghosh (MG) model offers an exactly solvable model of spin-1/2 chains.

$$H_{\rm MG} = \frac{4J}{3} \sum_{i=1}^{N} (\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1} + \frac{1}{2} \boldsymbol{S}_i \cdot \boldsymbol{S}_{i+2}) + \frac{NJ}{2}, \quad \boldsymbol{S}_{N+1} = \boldsymbol{S}_1 \ . \tag{8.9}$$

$H_{\rm MG}$ possesses two degenerate ground states one of which is given by

$$|{\rm MG}\rangle = \prod_{i=1}^{N/2} \frac{1}{\sqrt{2}} \Big((a_{2i-1}^\dagger b_{2i}^\dagger - b_{2i-1}^\dagger a_{2i}^\dagger \Big) |0\rangle, \tag{8.10}$$

and the other is a translation of this state by one lattice unit. The proof is as follows: construct the total spin operator for a triad of adjacent spins,

$$\begin{aligned} \boldsymbol{J}_i &= \boldsymbol{S}_{i-1} + \boldsymbol{S}_i + \boldsymbol{S}_{i+1} \\ \boldsymbol{J}_i^2 &= J(J+1), \quad J = \frac{1}{2}, \ \frac{3}{2} \ . \end{aligned} \tag{8.11}$$

Since one is adding up three $S = 1/2$ operators, the Hilbert space for the triad of spins must be decomposed into either $J = 3/2$ or $J = 1/2$ sectors. The appropriate projection operator is constructed as

$$\frac{1}{3}\left(\boldsymbol{J}_i^2 - \frac{3}{4}\right) \equiv \mathcal{P}_{3/2}(i-1,i,i+1) \ . \tag{8.12}$$

If $\{\boldsymbol{S}_{i-1},\ \boldsymbol{S}_i,\ \boldsymbol{S}_{i+1}\}$ form a $J = 1/2$ state, then $\mathcal{P}_{3/2}(i-1,i,i+1)\,|J = 1/2\rangle = 0$, but if $J = 3/2$ then $\mathcal{P}_{3/2}(i-1,i,i+1)\,|J = 3/2\rangle = |J = 3/2\rangle$. $H_{\rm MG}$ is a sum of projectors, $H_{\rm MG} = J\sum_i \mathcal{P}_{3/2}(i-1,i,i+1)$. Due to the preceding argument about the allowed values of J, each average $\langle\mathcal{P}_{3/2}(i-1,i,i+1)\rangle$ must be non-negative, and we can anticipate that the zero-energy state, if it exists, is the ground state.

Expanding

$$\begin{aligned}\mathcal{P}_{3/2}(i-1,i,i+1) &= \frac{1}{3}\left[(\boldsymbol{S}_{i-1}+\boldsymbol{S}_i+\boldsymbol{S}_{i+1})^2 - \frac{3}{4}\right] \\ &= \frac{1}{3}\left[\frac{3}{2} + 2(\boldsymbol{S}_{i-1}\cdot\boldsymbol{S}_i + \boldsymbol{S}_i\cdot\boldsymbol{S}_{i+1} + \boldsymbol{S}_{i-1}\cdot\boldsymbol{S}_{i+1})\right] \\ &= \frac{1}{2} + \frac{2}{3}(\boldsymbol{S}_{i-1}\cdot\boldsymbol{S}_i + \boldsymbol{S}_i\cdot\boldsymbol{S}_{i+1} + \boldsymbol{S}_{i-1}\cdot\boldsymbol{S}_{i+1}) \ . \end{aligned} \tag{8.13}$$

Hence,

$$H_{\rm MG} = J\sum_{i=1}^{N}\mathcal{P}_{3/2}(i-1,i,i+1) = \frac{NJ}{2} + \frac{4}{3}J\sum_i\left(\boldsymbol{S}_i\cdot\boldsymbol{S}_{i+1} + \frac{1}{2}\sum_i \boldsymbol{S}_i\cdot\boldsymbol{S}_{i+2}\right). \tag{8.14}$$

Two of the three sites in the $|{\rm MG}\rangle$ states are already bound into a singlet. The portion of the wave function featuring the three sites $i-1, i, i+1$ are linear combinations of

$$(a_{i-1}^\dagger b_i^\dagger - b_{i-1}^\dagger a_i^\dagger)a_{i+1}^\dagger \quad \& \quad (a_{i-1}^\dagger b_i^\dagger - b_{i-1}^\dagger a_i^\dagger)b_{i+1}^\dagger. \tag{8.15}$$

The singlet combination $(a_{i-1}^\dagger b_i^\dagger - b_{i-1}^\dagger a_i^\dagger)$ remains invariant under an arbitrary SU(2) rotation of spins, and $a_{i+1}^\dagger$ and $b_{i+1}^\dagger$ rotates into linear combinations of each other. So, no matter what the basis to write down the Schwinger bosons, one always finds $J_z = 1/2$ for the triad, and thus $J = 1/2$. If no triad possesses $J = 3/2$ component, $\mathcal{P}_{3/2}(i-1,i,i+1)\,|d\rangle_\pm = 0$ for all i, so $H_{\rm MG}\,|d\rangle_\pm = 0$. Since $H_{\rm MG}$ is a sum of positive semi-definite operators, the zero-energy eigen state is the ground state. The spin-spin correlation for $|{\rm MG}\rangle$ is

$$\langle\boldsymbol{S}_i\cdot\boldsymbol{S}_j\rangle = \begin{cases} 3/4, & i = j \\ -3/4, & |i-j| = 1 \\ 0, & |i-j| > 1 \end{cases} \tag{8.16}$$