Mean-field approximation for the repulsive one band Hubbard model

The Hamiltonian is

\[ H = -t \sum_{<ij>} (c_i^\dagger c_j + h.c.) + U \sum_i n_{i\uparrow} n_{i\downarrow} \]

Let us propose a spin-density-wave solution.

We replace \( n_{i\sigma} \) by a mean value. First we write:

\[ n_{i\sigma} = \langle n_{i\sigma} \rangle + (n_{i\sigma} - \langle n_{i\sigma} \rangle) \]

and then we say that \( n_{i\sigma} - \langle n_{i\sigma} \rangle \) is "small". This quantity tells us how much the system deviates from the mean value, which has to be found self-consistently. The interaction becomes

\[ n_{i\uparrow} n_{i\downarrow} = \langle n_{i\uparrow} \rangle \langle n_{i\downarrow} \rangle + (n_{i\uparrow} - \langle n_{i\uparrow} \rangle) \langle n_{i\downarrow} \rangle \]

\[ + (n_{i\downarrow} - \langle n_{i\downarrow} \rangle) \langle n_{i\uparrow} \rangle \]

\[ + (n_{i\uparrow} - \langle n_{i\uparrow} \rangle)(n_{i\downarrow} - \langle n_{i\downarrow} \rangle) \]

This term is of "second order" in the small deviation from the MF Ansatz. Thus, we neglect it. Then:

\[ n_{i\uparrow} n_{i\downarrow} \approx -\langle n_{i\uparrow} \rangle \langle n_{i\downarrow} \rangle + n_{i\uparrow} \langle n_{i\downarrow} \rangle + n_{i\downarrow} \langle n_{i\uparrow} \rangle \]
As Ansatz we propose

\[ \langle \text{Mi}^+ \rangle = \frac{1}{2} (1 + S e^{i \vec{Q}.\vec{r}}) \]
\[ \langle \text{Mi}^- \rangle = \frac{1}{2} (1 - S e^{i \vec{Q}.\vec{r}}) \]

where \( \vec{Q} = (\pi, \pi) \) in 2D. \( S \) is the parameter that we have to find by energy minimization.

\[ e^{i \vec{A}.\vec{r}} = (-1)^{ix + iy} \]

where \( ix, iy \) are the components of the site vector \( \vec{r} \). This factor is 1 for "even" sites, and \((-1)\) for "odd" sites i.e. it divides the square lattice into two sublattices:

\[
\begin{array}{ccc}
-1 & 1 & -1 \\
1 & -1 & 4 \\
-1 & 1 & -1 \\
\end{array}
\]

If \( S = 0 \), then \( \langle \text{Mi}^+ \rangle = \langle \text{Mi}^- \rangle = \frac{1}{2} \) and this is the correct result at \( U = 0 \) and half-filling. In this case translational invariance is unbroken.

If \( S = 1 \), \[ \langle \text{Mi}^+ \rangle = \begin{cases} 1 \text{ for even sites} \\ 0 \text{ for odd sites} \end{cases} \]
\[ \langle \text{Mi}^- \rangle = \begin{cases} 0 \text{ for even sites} \\ 1 \text{ for odd sites} \end{cases} \]
This is correct. It is the result we expect at $U$ large where we have one particle per site at half-filling, and the effective Hamiltonian is the Heisenberg model with antiferromagnetic coupling.

Note that the interaction in the Hubbard model is usually written in a "particle-hole symmetric way" as

$$U \sum_i (\langle m_{i \uparrow} \rangle - \frac{1}{2}) (\langle m_{i \downarrow} \rangle - \frac{1}{2})$$

which, after the MF Ansatz, becomes:

$$U \sum_i (\langle m_{i \uparrow} \rangle \langle m_{i \downarrow} \rangle - \frac{1}{2} (\langle m_{i \uparrow} \rangle + \langle m_{i \downarrow} \rangle) + \frac{1}{4}) \approx$$

$$= U \sum_i \left[ -\langle m_{i \uparrow} \rangle \langle m_{i \downarrow} \rangle + \langle m_{i \uparrow} \rangle + \frac{1}{2} (1 - Se^{i\theta R_i}) + \langle m_{i \downarrow} \rangle \frac{1}{2} (1 + Se^{i\theta R_i}) - \frac{1}{2} (\langle m_{i \uparrow} \rangle + \langle m_{i \downarrow} \rangle) + \frac{1}{4} \right] =$$

$$= U \sum_i \left[ \frac{1}{4} (1 + Se^{i\theta R_i})(1 - Se^{i\theta R_i}) + \frac{S}{2} e^{i\theta R_i} (\langle m_{i \uparrow} \rangle - \langle m_{i \downarrow} \rangle) + \frac{1}{4} \right]$$

$$= U \sum_i \left[ \frac{S^2}{4} + \frac{S}{2} e^{i\theta R_i} (\langle m_{i \uparrow} \rangle - \langle m_{i \downarrow} \rangle) \right]$$
Then, the MF Hubbard model is

\[ H = -t \sum_{\langle ij \rangle} (c_i^+ c_j + h.c.) + U N S^z \]

- \( \frac{JS}{2} \sum_i e^{iQz_i} (\sigma_i^x - \sigma_i^y) \)

where \( N \) is the total # of sites.

Note that the last term resembles the interaction of a particle in a "staggered" magnetic field \( \vec{H} \) which is of the form:

\[ \sum_i \vec{H}_i \cdot \vec{S}_i = \sum_i \frac{J}{2} e^{iQz_i} \left( \frac{\sigma_i^x - \sigma_i^y}{2} \right) \]

Assume

\( \tau \) points

along \( z \)

The MF Hamiltonian can be handled better in momentum space where the kinetic energy is explicitly diagonal. Note that \( H_{MF} \), being quadratic in \( c^+, c \), must be exactly solvable unlike the full Hamiltonian that contains a "c+c+c+c" complicated interaction.

The kinetic energy becomes:
\[ H_{\text{k.e.}} = -t \sum_{\ell \sigma} (c_\ell^\dagger c_{\ell + \mu \sigma} + c_{\ell + \mu \sigma}^\dagger c_\ell) \]

This is another way to represent nearest neighbors. \( \hat{x}, \hat{y} \) are unit vectors along the axis.

Let us introduce a standard Fourier transformation as:

\[ c_\ell = \frac{1}{\sqrt{N}} \sum_\ell e^{i \ell \cdot \ell} c_{\rho \sigma} \]

Then,

\[ H_{\text{k.e.}} = -t \sum_{\ell, \mu, \sigma} \frac{1}{N} \sum_\rho \rho' \left[ e^{-i \ell \cdot \rho} e^{i \ell \cdot (\rho + \rho')} c_{\rho \sigma}^\dagger c_{\rho' \sigma} + e^{i \ell \cdot \rho} e^{-i \ell \cdot (\rho + \rho')} c_{\rho \sigma} c_{\rho' \sigma} \right] \]

But we know that

\[ \frac{1}{N} \sum_\ell e^{-i \ell \cdot \rho} e^{i \ell \cdot \rho'} = S_{\rho \rho'} \]

Then:

\[ H_{\text{k.e.}} = -t \sum_{\mu, \sigma} \sum_\rho c_{\rho \sigma}^\dagger c_{\rho \sigma} (e^{i \rho \cdot \rho'} + e^{-i \rho \cdot \rho'}) \]
\[ H_{K.E.} = -t \sum_{\phi \sigma} C_{\phi \sigma}^+ C_{\phi \sigma} \sum_{\hat{p}} 2 \cos (\hat{p}.\hat{r}) \]

\[ = \sum_{\phi \sigma} C_{\phi \sigma}^+ C_{\phi \sigma} E_{\phi} \]

where \( E_{\phi} = -2t (\cos \phi_x + \cos \phi_y) \) in 2D.

Since \( \hat{p} \) is either \((1,0)\) or \((0,1)\).

The kinetic energy is indeed diagonal in \( \phi \)-space.

Now, let us Fourier transform the other term that appears in \( H_{mp} \):

\[ -\frac{us}{2} \sum_{l} e^{i\hat{q}\cdot\hat{r}} (\eta_{l\uparrow} - \eta_{l\downarrow}) = \]

\[ = -\frac{us}{2} \sum_{l} e^{i\hat{q}\cdot\hat{r}} (C_{l\uparrow}^+ C_{l\uparrow} - C_{l\downarrow}^+ C_{l\downarrow}) = \]

\[ = -\frac{us}{2} \sum_{l} \sum_{\phi} \sum_{\phi'} e^{i\hat{q}\cdot\hat{r}} e^{-i\hat{p}\cdot\hat{r}} e^{i\hat{p}'\cdot\hat{r}} (C_{\phi\uparrow}^+ C_{\phi'\uparrow} - C_{\phi\downarrow}^+ C_{\phi'\downarrow}) \]

Now we have to use the identity:
\[
\frac{1}{N} \sum_{l} e^{-i(\vec{p} - \vec{p}' - \vec{q}) \cdot \vec{r}} = \sum_{\vec{p}, \vec{p}'} \delta_{\vec{p}, \vec{p} + \vec{q}}
\]

and we get

\[
- \frac{US}{2} \sum_{\vec{q}} \left( C_{\vec{p}' + \vec{q} \uparrow} C_{\vec{p}' \uparrow} - C_{\vec{p}' + \vec{q} \downarrow} C_{\vec{p}' \downarrow} \right)
\]

(Note that \(Q = -Q\) is equivalent)

The full MF Hamiltonian is

\[
H_{\text{MF}} = \frac{US}{4} - \frac{US}{2} \sum_{\vec{p}} \left( C_{\vec{p} + \vec{q} \uparrow} C_{\vec{p} \uparrow} - C_{\vec{p} + \vec{q} \downarrow} C_{\vec{p} \downarrow} \right)
\]

Note that the last term contains terms that change momentum from \(\vec{p}\) to \(\vec{p} + \vec{q}\). This is an apparent violation of \(p\)-conservation. The effect is caused by the MF approximation. The spins are like in a staggered field that allows for \(p\)-non conservation in \(Q\).

\(\vec{p} + \vec{q} \rightarrow \text{stagg. field.}\)
Note also that $H_{NP}$ is "almost" diagonal but not quite yet. Momentum $p$ interacts with itself in the K.E. and with $p+Q$ in the other term. But $p+Q$ interacts with itself and with $p+2Q$ which is $p$. Thus, $p$ and $p+Q$ interact with each other and no other momentum. The Hamiltonian is thus in a simple block form, where each block is

\[
\begin{pmatrix}
C_{p\sigma} & C_{p+Q\sigma} \\
C_{p\sigma}^+ & \begin{pmatrix} E_p & -U S \sigma \\
-U S \sigma & E_{p+Q} \end{pmatrix}
\end{pmatrix}
\]

where $\sigma = +1$ if spin up,

$\sigma = -1$ if spin down.

Note that

\[
E_{p+Q} = -2t \sum_{\mu} \cos (p+Q \cdot \bar{\mu})
\]

\[
= -2t \left[ \cos (p+Q \cdot \bar{x}) + \cos (p+Q \cdot \bar{y}) \right] = +2t \left( \cos p_x + \cos p_y \right)
\]

\[
= -E_p
\]

and the matrix becomes:
\[
\begin{pmatrix}
\varepsilon \rho & -\frac{U S \sigma}{2} \\
-\frac{U S \sigma}{2} & -\varepsilon \rho
\end{pmatrix}
\]

The eigenvalues, \( \varepsilon \rho \), are obtained from:

\[
(\varepsilon \rho - \varepsilon \rho) (-\varepsilon \rho - \varepsilon \rho) - \frac{U^2 S^2 \sigma^2}{4} = 0
\]

\[\varepsilon \rho^2 + \varepsilon \rho^2 = \frac{U^2 S^2 \sigma^2}{4} \]

\[\varepsilon \rho = \pm \sqrt{\varepsilon \rho^2 + \frac{U^2 S^2 \sigma^2}{4}}\]

Note that \( \rho \) is now restricted to half the Brillouin zone. The reason is that the eigenvalues \( \varepsilon \rho \) are obtained after putting in one block \( \rho \) and \( \rho + Q \). At \( S=0 \) which is the noninteracting limit I had \( 2N \) eigenvalues (2 for spin, \( N \) for \# of moments). Now I have \( 2 \times 2 \times \frac{N}{2} = 2N \) also when \( 2 \) is for spin, the second \( 2 \) is for the \( \pm \) values in the equation for \( \varepsilon \rho \) and \( \frac{N}{2} \) are the possible values \( \rho \) takes.
To plot this we use half the Brillouin zone. Consider the 1D case:

At $U=0$ we have the solid line. The dashed line represents the MF solution. This solution at $\mathbf{p}=0$ is

$$E_0 = \pm \sqrt{E_0^2 + \frac{U^2 s^2}{4}}$$

$$= \pm \sqrt{(2t)^2 + \frac{U^2 s^2}{4}}$$

and at $\mathbf{p}=\pi/2$ we know $E_0 = 0$

and

$$E_{\pi/2} = \pm \frac{US}{2}$$

Then, now we have a gap in the spectrum which is the smallest at $\pi/2$. Including negative values for the momenta we have

**Why a gap opens?**

In a state $\uparrow \downarrow \uparrow \downarrow \uparrow \downarrow \cdots$ adding a charge costs zero energy. As single spin!
The total gap is $U_S = 2\Delta$ and the dispersion is
\[ E_p = \pm \sqrt{E_p^2 + \Delta^2} \]
as in the BCS superconductors.

So far we haven't said anything about $S$. We have to find its self-consistent equation. What is the energy of the problem in the MF Ansatz? Well, we have to fill all levels of negative energy which is fine since we are at half-filling and we have enough particles. The energy is:

\[ E = 2 \sum_{\text{MF}} \sum_{\text{spin}} E_p + \frac{UNS^2}{4} \]

\[ \frac{\partial E_{MF}}{\partial S} = 0 = \frac{\partial}{\partial S} \left( 2 \sum_{E_p < 0} \left( -\sum U^2 p^2 + \frac{U S^2}{4} \right) + \frac{UNS^2}{4} \right) \]

\[ = -2 \sum_{E_p < 0} \frac{U^2 S}{2} \frac{2S}{2\sqrt{E_p^2 + \frac{U S^2}{4}}} + \frac{UNS}{4} \frac{2S}{4} \]

\[ = -\frac{U^2 S}{2} \sum_{\text{in the reduced Brillouin zone}} \frac{1}{\sqrt{E_p^2 + \frac{U S^2}{4}}} + \frac{UNS}{2} \]

For each $p$ in the reduced Brillouin zone there is a negative $E_p$.

Let us forget about the trivial $S = 0$ solution.
Then we get:

\[
\frac{1}{N} \sum_{p \in \text{reduced BZ}} \frac{U}{\sqrt{E_p^2 + U^2 s^2}} = 1
\]

This is the "gap equation." Again, very similar to BCS.

This is the "gap equation" in the context of antiferromagnetism.

\[ S \]

\[ \frac{d}{dU} \]

\[ \Delta = \frac{U S}{2} \]

Since \( S \to 1 \) as \( U \to \infty \),

then \( \Delta \approx U/2 \) at large \( U \).
Now, let us calculate the eigenvectors. For $\sigma = +1$, i.e. spin up, the matrix is

\[
\begin{pmatrix}
\mathcal{E}_p & -\Delta \\
-\Delta & -\mathcal{E}_p
\end{pmatrix}
\]

The eigenvector will have the form $\begin{pmatrix} \mu_p \\ -\nu_p \end{pmatrix}$ with $\mu_p^2 + \nu_p^2 = 1$. It means $\mu_p \mathcal{E}_p^+ \rho_\rho > -\nu_p \mathcal{E}_p^+ \rho_\rho$.

As in BCS, let us propose

\[\mu_p^2 = \frac{1}{2} \left( 1 + \frac{\mathcal{E}_p}{|\mathcal{E}_p|} \right)\]

\[\nu_p^2 = \frac{1}{2} \left( 1 - \frac{\mathcal{E}_p}{|\mathcal{E}_p|} \right)\]

The condition of eigenvector is

\[
\begin{pmatrix}
\mathcal{E}_p - \Delta \\
-\Delta - \mathcal{E}_p
\end{pmatrix}
\begin{pmatrix} \mu_p \\ -\nu_p \end{pmatrix} = \mathcal{E}_p \begin{pmatrix} \mu_p \\ -\nu_p \end{pmatrix}
\]

or $\mathcal{E}_p \mu_p + \Delta \nu_p$ should be $\mathcal{E}_p \mu_p$, and $-\Delta \mu_p + \mathcal{E}_p \nu_p$ should be $-\mathcal{E}_p \nu_p$.

But $\mathcal{E}_p \mu_p + \Delta \nu_p = \mu_p \left( \mathcal{E}_p + \Delta \frac{\mathcal{E}_p}{|\mathcal{E}_p|^2} \right) = \mu_p \left( \mathcal{E}_p + \Delta \sqrt{\frac{|\mathcal{E}_p - \mathcal{E}_p^0|^2}{|\mathcal{E}_p|^2 - \mathcal{E}_p^0}} \right) = \mu_p \left( \mathcal{E}_p + \Delta \sqrt{\frac{|\mathcal{E}_p|^2 - |\mathcal{E}_p - \mathcal{E}_p^0|^2}{\Delta^2}} \right) = \mu_p \left[ \mathcal{E}_p + \left( \mathcal{E}_p - \mathcal{E}_p^0 \right) \right] = \mu_p |\mathcal{E}_p| \text{ O.K.}
\]

and $-\Delta \mu_p + \mathcal{E}_p \nu_p = (-\nu_p) \left( + \Delta \frac{\mu_p}{\nu_p} - \mathcal{E}_p \right) = (\nu_p) \left( \Delta \frac{\mathcal{E}_p + \mathcal{E}_p}{|\mathcal{E}_p + \mathcal{E}_p|^2} - \mathcal{E}_p \right) = \nu_p \left( \Delta \frac{\mathcal{E}_p + \mathcal{E}_p}{|\mathcal{E}_p + \mathcal{E}_p|^2} - \mathcal{E}_p \right)$.
\[
\left(\nabla_p \right) \left[ \Delta \sqrt{\frac{E_p^2 + E_p}{E_p^2 - E_p^2}} - E_p \right] = \left(\nabla_p \right) \left[ \Delta \sqrt{\frac{E_p^2 + E_p^2}{\Delta^2}} - E_p \right] = (\nabla_p \mid E_p) \tag{14}
\]

Then \((\mu_p)\) is indeed eigenvector with eigenvalue \(\sqrt{E_p^2 + \Delta^2} = E_p\).

The other one will be \((\nu_p)\) which is orthogonal to the previous one. There is no need to verify that it is since just showing the orthogonality to the previous eigenvector is here enough. Indeed

\[
(\mu_p - \nu_p)(\nu_p) = 0.
\]

The operators \(\hat{\sigma}^{\pm}\) corresponding to eigenvectors are

\[
\begin{align*}
\hat{\sigma}^+ &= \mu_p \hat{C}_{p+1} - \nu_p \hat{C}_{p+1}^+ \\
\hat{\sigma}^- &= \nu_p \hat{C}_{p+1}^+ + \mu_p \hat{C}_{p+1} \\
\hat{\sigma}^0 &= \mu_p \hat{C}_{p+1} + \nu_p \hat{C}_{p+1}^+
\end{align*}
\]

On schematically

\[
\begin{pmatrix}
\hat{\sigma}^+
\hat{\sigma}^-
\hat{\sigma}^0
\end{pmatrix} = \begin{pmatrix}
\nu_p & -\mu_p \\
\mu_p & \nu_p \\
\end{pmatrix} \begin{pmatrix}
\hat{C}_{p+1} \\
\hat{C}_{p+1}^+
\end{pmatrix}
\]

Corresponding to block \(p\).

The Hamiltonian is now:

\[
\text{Eigenvectors are mixture of 2 types of electrons!}
\]

\[
\Leftarrow \quad \text{Eigenvalue } |E_p|
\]

\[
\Leftarrow \quad \text{ } -E_p
\]

Little date game is true for \(\Delta^2 \neq 0\).

Then, if \(\mu_p\) with \(p=0\) is dominated by \(\nu_p\) for \(\Delta^2 \neq 0\).

But if \(\Delta^2 = 0\), then \(p=0\) and \(p\) has a negative energy and both can never be at the same amount of time.
\[
\begin{pmatrix}
\hat{t}^{(+)} & \hat{t}^{(-)} \\
\hat{y}^{(+)} & \hat{y}^{(-)}
\end{pmatrix}
\begin{pmatrix}
|E_p| & 0 \\
0 & -|E_p|
\end{pmatrix}
\begin{pmatrix}
\delta^{(4)} \\
\delta^{(-4)}
\end{pmatrix} = |
\begin{pmatrix}
\delta^{(4)} \\
\delta^{(-4)}
\end{pmatrix}
\end{pmatrix}
\]

\[
= |E_p| \begin{pmatrix}
\delta^{(4)} \\
\delta^{(-4)}
\end{pmatrix}
\begin{pmatrix}
\delta^{(+)} \\
\delta^{(-)}
\end{pmatrix}
\]

where \( |E_p| = + \sqrt{\epsilon_p^2 + \Delta^2} \).

\[
\hat{H} = \sum_{p \neq \pm} |E_p| \begin{pmatrix}
\delta^{(+)} \\
\delta^{(-)}
\end{pmatrix}
\begin{pmatrix}
+\delta^{(+)} \\
-\delta^{(-)}
\end{pmatrix}
\]

Now, for spin down we simply change \(-\Delta \rightarrow \Delta\) in the Hamiltonian block. In such a case we propose \((\mu_p)\) and we get

\[
\begin{pmatrix}
\epsilon_p & \Delta \\
\Delta & -\epsilon_p
\end{pmatrix}
\begin{pmatrix}
\mu_p \\
\nu_p
\end{pmatrix} = E_p
\begin{pmatrix}
\mu_p \\
\nu_p
\end{pmatrix}
\]

or \( \epsilon_p \mu_p + \Delta \nu_p \) should be \( E_p \mu_p \),

\( \Delta \mu_p - \epsilon_p \nu_p \) should be \( E_p \nu_p \).

\[
\epsilon_p \mu_p + \Delta \nu_p = \mu_p \left( \epsilon_p + \frac{\Delta \nu_p}{\mu_p} \right) = \mu_p \left( \epsilon_p + \Delta \frac{|E_p| - \epsilon_p}{|E_p| + \epsilon_p} \right) = \\
= \mu_p \left( \epsilon_p + \Delta \frac{\sqrt{(|E_p| - \epsilon_p)^2}}{|E_p| - \epsilon_p} \right) = \mu_p \left( \epsilon_p + \Delta \frac{|E_p| - \epsilon_p}{\Delta^2} \right) = \\
= \mu_p |E_p|.
\]

An orthogonal vector is \( (\nu_p \quad -\mu_p) \)
and thus the eigenoperators are:

\[
\hat{J}^{(+)}_p \psi = \lambda_p \hat{C}_p \psi + \nu_p \hat{C}_p+Q \psi
\]

\[
\hat{J}^{(-)}_p \psi = \nu_p \hat{C}_p \psi - \lambda_p \hat{C}_p+Q \psi
\]

\[\text{eigenvalue } \epsilon_p \]

\[\text{eigenvalue } \pm \epsilon_p \]

The ground state is then:

\[
\prod_{p \in \text{reduced BZ}} \hat{J}^{(-)}_p \hat{J}^{(+)\dagger}_p 10 > =
\]

\[
\prod_{p \in \text{reduced BZ}} (\nu_p \hat{C}_p\hat{C}_p^+ + \lambda_p \hat{C}_p\hat{C}_p^+Q) (\nu_p \hat{C}_p\hat{C}_p^+ - \lambda_p \hat{C}_p\hat{C}_p^+Q) 10 >
\]

The state \(\hat{C}_p\hat{C}_p^+ 10 >\) has charge \(\nu_p\):

\[
' = \hat{C}_p\hat{C}_p^+ 10 > \quad ' = \nu_p
\]

\[
' = \hat{C}_p\hat{C}_p^+ 10 > \quad ' = \nu_p
\]

\[
' = \hat{C}_p\hat{C}_p^+ 10 > \quad ' = \nu_p
\]

If \(\Delta = 0\), \(\nu_p = 1\) in the reduced BZ and \(\lambda_p = 0\) "" "" "" "" "" ""

If \(\Delta \neq 0\) but small, then \(\nu_p \gg \lambda_p\) and the state \(\hat{C}_p\hat{C}_p^+ 10 >\) dominates although the rest have a nonzero chance specifically near the boundary of the reduced BZ, where \(\epsilon_p = 0\).