Hubbard model

This simple model of fermions with local interactions provides a testing ground for strongly correlated electron systems. The model, despite its simplicity, cannot be solved exactly. The Hubbard model has been proposed to capture the essential physics of high Tc superconductors.

The model consists of spin-$\frac{1}{2}$ fermions hopping on a lattice, with hopping amplitude $t_{ij}$ between sites $i$ and $j$, plus a local repulsion $U>0$ which penalizes double occupation of sites by spin up and down electrons:

\[ H = - \sum_{ij} \sum_{\sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} - \mu \sum_i n_i \]

Here, $n_{i\sigma} = c_{i\sigma}^\dagger c_{i\sigma}$ is the occupation number of spin-$\sigma$ fermions, $n_i = n_{i\uparrow} + n_{i\downarrow}$, and $\mu$ is a chemical potential which we add in the Grand Canonical ensemble to control the particle density. The interaction term can also be written as

\[ \frac{U}{2} \sum_i n_i (n_i - 1) = \frac{U}{2} \sum_i (n_{i\uparrow} + n_{i\downarrow}) (n_{i\uparrow} + n_{i\downarrow} - 1) = U \sum_i n_{i\uparrow} n_{i\downarrow} \]

where we used the operator identities $n_{i\uparrow}^2 = n_{i\uparrow}$ and $n_{i\downarrow}^2 = n_{i\downarrow}$. This demonstrates that $H$ respects spin-rotational symmetry.

For simplicity, we'll focus on a 2d square lattice in the case where the hopping is restricted to nearest neighbor sites,

\[ t_{ij} = \begin{cases} t & \text{if } i=j \pm \hat{x} \\ 0 & \text{otherwise} \end{cases} \]

where $\hat{x} = \hat{x}$ or $\hat{y}$. Then, the hopping term, including the chemical potential, is

\[ H_0 = -t \sum_{i,\sigma} \sum_{\mu=x,y} c_{i\sigma}^\dagger c_{i\mu \sigma} - \mu \sum_i c_{i\uparrow}^\dagger c_{i\uparrow} \]
This can be diagonalized in plane wave basis, 
\[ c_i = \sum_{k, \sigma} \frac{1}{N} \left( \sqrt{k_y} \right) e^{ik_r} c_{i\sigma} \]
where the crystal momentum has been restricted to the 1st Brillouin zone \( k_x \in \left[ -\frac{\pi}{a}, \frac{\pi}{a} \right] \) (\( a \) is the lattice constant). Then,

\[ H_0 = \sum_{k, \sigma} (E_{k\sigma} - \mu) C_{i\sigma}^\dagger C_{i\sigma} \]
where

\[ E_k = -2t (\cos k_x a + \cos k_y a) \]

Contours of constant energy

Note the saddle points near \((0, \frac{\pi}{a})\) and \((\frac{\pi}{a}, 0)\). This leads to a logarithmic van Hove singularity in the density of states at \( \epsilon = 0 \)

\[ g(\epsilon) = \frac{2}{N} \sum_{k \in BZ} \delta(\epsilon - E_k) \]
\[ g(\epsilon) \sim \ln \frac{1}{|\epsilon|} \text{ as } \epsilon \to 0 \]

When \( U < t \), we can do perturbation theory starting from a metallic state and treating \( U \) as a small parameter. We will focus on the opposite limit, \( U > t \). Then, at half filling (one electron/site), the system is an insulator, and we will treat \( t \) as a perturbation.
For $U \gg t$, we can study the system perturbatively in powers of $t/U$. At half filling, for $t=0$, the ground state contains exactly one electron at every site, but this is a highly degenerate situation due to the spin degree of freedom — the electron at any given site can be in either a spin up or spin down state, leading to $2^N$ ground state configurations.

This degeneracy gets lifted for $t>0$, as can be seen by going to $O((t/U)^3)$ in perturbation theory. Let's consider two neighboring sites. When $t=0$, the spin on site 1 is independent of the spin in site 2. For small $t>0$, the electrons can hop from one site to the next to form virtual excited states, but only provided that their spins are different from each other. Whenever the spins are equal, Pauli exclusion prevents these virtual hops.

\[ |\psi_i\rangle = |\uparrow_1 \downarrow_2\rangle \quad |\psi_1\rangle = C_{1\uparrow} C_{2\downarrow} |0\rangle, \quad \epsilon_i = 0 \]

\[ |\psi_m\rangle = |\downarrow_1 \uparrow_2\rangle \quad |\psi_m\rangle = C_{2\uparrow} C_{1\downarrow} |0\rangle, \quad \epsilon_m = U \]

\[ |\psi_f\rangle = C_{1\uparrow} C_{2\uparrow} |\psi_m\rangle = |\psi_1\rangle \quad |\psi_f'\rangle = C_{1\downarrow} C_{2\downarrow} |\psi_m\rangle = -C_{1\downarrow} C_{2\uparrow} |0\rangle \]

$\triangle$ Spin exchange comes with $\otimes$ sign.

\[ \epsilon_f = 0 = \epsilon_f' \]
Of course, we could have done things in the opposite direction and first hop from site 2 to site one. Then, at 2nd order in perturbation theory obtain an effective Hamiltonian acting on the space of singly-occupied sites:

\[
H = 2 \frac{(t+)^2}{U} \left[ |\uparrow \downarrow \rangle \langle \uparrow \downarrow | + |\uparrow \downarrow \rangle \langle \downarrow \uparrow | - |\downarrow \uparrow \rangle \langle \uparrow \downarrow | - |\downarrow \uparrow \rangle \langle \downarrow \uparrow | \right]
\]

\[
+ \varepsilon_i - \varepsilon_m - 2 S_i^+ S_2^\pm + \frac{1}{2}
- S_i^+ S_2^- - S_i^- S_2^+
\]

\[
= \frac{4t^2}{U} \left( S_i^+ S_2^\pm + \frac{1}{2} S_i^+ S_2^- + \frac{1}{2} S_i^- S_2^+ \right) + \text{constant} = \frac{4t^2}{U} \vec{S}_i \cdot \vec{S}_2 + (\text{const})
\]

If we now consider the full square lattice, we obtain

\[
H = J \sum_{\langle ij \rangle} \vec{S}_i \cdot \vec{S}_j, \quad J = \frac{4t^2}{U}
\]

This is the antiferromagnetic Heisenberg model. The source of the antiferromagnetic coupling is that antialigned spins can lower their energy through virtual hops, which are not available to aligned spins due to Pauli exclusion.

One consequence of this result is that, in the strong coupling limit, \( T_c \) is set by \( J \), and goes to zero as \( \frac{t^2}{U} \) in the \( U \rightarrow \infty \) limit. This is in contrast with the Mott-Hubbard gap \( 2\Delta \approx V \).
Antiferromagnet

Consider spin-\(S\) spins on a square lattice, with a nearest neighbor antiferromagnetic interaction

\[
H = J \sum_{\langle ij \rangle} \vec{S}_i \cdot \vec{S}_j = J \sum_{\langle ij \rangle} \left[ S_i^z S_j^z + \frac{1}{2} (S_i^+ S_j^- + S_i^- S_j^+) \right]
\]

"Classical" trial state:

\[
|\Psi_N\rangle = |\uparrow \downarrow \uparrow \downarrow \ldots \uparrow \downarrow \rangle \quad \text{"Néel state"}
\]

\[
\langle S_i^z \rangle = S \cos (\pi, \pi \cdot \vec{r}_i)
\]

\(|\Psi_N\rangle\) is not an eigenstate of \(H\).

Spoiler terms \(S_i^+ S_j^- + \text{h.c.}\)

\[
S_i^+ S_j^- |\ldots \uparrow \downarrow \uparrow \downarrow \ldots \rangle_{ij} = |\ldots \uparrow \uparrow \downarrow \uparrow \ldots \rangle_{ij}
\]

* Do these "quantum fluctuations" destroy the antiferromagnetic order?
* What is the nature of the low energy excitations?

Spin wave theory

Basic idea — just as in the case of phonons in solids:

(i) Assume classical broken symmetry trial state

(ii) Expand in harmonic fluctuations around the classical spin configuration
Antiferromagnetic spin waves

Classical broken symmetry configuration:

\[ \begin{array}{cccc}
A & B & A & B \\
\uparrow & \downarrow & \uparrow & \downarrow \\
\downarrow & \uparrow & \downarrow & \uparrow \\
B & A & B & A \\
\end{array} \]

"Néel state"

It is convenient to work with a uniform configuration. We therefore rotate the spin coordinates on sites of sub-lattice \( B \) by \( \pi \) around the \( S_x \) axis:

\[ \begin{align*}
\tilde{S}_j^x &= S_j^x, & \tilde{S}_j^y &= -S_j^y, & \tilde{S}_j^z &= -S_j^z, & \forall j \in S.L. \ B
\end{align*} \]

(This unitary transformation is simply a relabeling of the spin axis). On sublattice \( A \), leave spins untouched:

\[ \begin{align*}
\tilde{S}_i^x &= S_i^x, & \tilde{S}_i^y &= S_i^y, & \tilde{S}_i^z &= S_i^z, & \forall i \in S.L. \ A
\end{align*} \]

Now:

\[ \tilde{H} = -J \sum_{\langle i,j \rangle} \tilde{S}_i^z \tilde{S}_j^z + \frac{J}{2} \sum_{\langle i,j \rangle} (\tilde{S}_i^+ \tilde{S}_j^- + \tilde{S}_i^- \tilde{S}_j^+) \]

The Néel state now looks like a ferromagnetic state

\[ \begin{array}{cccc}
\uparrow & \uparrow & \uparrow & \uparrow \\
\uparrow & \uparrow & \uparrow & \uparrow \\
\end{array} \]

But this is not an eigenstate of \( \tilde{H} \) due to the term \( \tilde{S}_i^- \tilde{S}_j^- \).

We can make progress in the limit of large spin \( S \).

In what follows, we will work in this basis but we won't write the tilde (\( \sim \)) over the spin operators anymore.
The basic idea:

In the ground state, in the large $S$ limit,

$$S_i^z = S - \frac{S_i^x + S_i^y}{2S}$$

where we've assumed that $S^x$ & $S^y$ are small fluctuations,

$$\langle S_x^z \rangle, \langle S_y^z \rangle \ll S^2.$$ 

Define operators describing the fluctuations,

$$\hat{x}_i = \frac{S_i^x}{S} \quad \hat{p}_i = \frac{S_i^y}{S}$$

$$[S_i^x, S_i^y] = i S^2 \quad \Rightarrow \quad [\hat{x}_i, \hat{p}_i] = \frac{i}{S} + O\left(\frac{S_i^x + p_i^x}{2}\right)$$

$x_i, p_i$ are canonical conjugates to leading order \( \Rightarrow \frac{1}{S} \leftrightarrow \hbar \)

⇒ The semiclassical expansion is controlled by large $S$.

$$S_i^z = S - \frac{S}{2} \left(x_i^2 + p_i^2\right) + O(x_i^4, p_i^4)$$

Now we can rewrite $H$ in terms of the fluctuations $x_i, p_i$ to quadratic order, and obtain spin-waves by diagonalization \([\text{see, e.g., P.W. Anderson, Phys. Rev. 86, 694 (1952)}]\)

We'll instead use a slightly different approach, which allows for a systematic expansion in $1/S$. 
AF spin waves — Holstein-Primakoff (HP) approach

Exact representation of the spins with HP bosons:

\[ S_i^+ = (\sqrt{2S-n_i}) b_i \quad n_i = b_i^+ b_i \]

\[ S_i^- = b_i^+ (\sqrt{2S-n_i}) \]

\[ S_i^z = S - n_i \]

Using \([b_i, b_j^+] = \delta_{ij}\) it is easy to check that these obey the exact spin algebra.

Our approach will be to substitute these operators into \(\hat{\mathcal{H}}\) and expand the square roots in powers of \(1/S\).

This is justified so long as the spin fluctuations are relatively small

\[ \langle n_i \rangle \ll 2S \quad \text{(or equivalently } \langle S^z \rangle \approx S) \]

The HP representation is tailor made to describe small fluctuations about the \( +S^z \) direction. For a state in which the spins are pointing in a different direction the mapping should be rotated accordingly.
Substituting the HP representation into $\hat{H}$ and expanding the square roots to lowest order, we obtain

$$-JS_i^z S_j^z + \frac{J}{2} (S_i^x S_j^x + S_i^y S_j^y) \simeq -JS^2 + JS(n_i + n_j) + JS (b_i b_j^+ + b_i^+ b_j) + O(s^3)$$

$$\Rightarrow \hat{H} \simeq -\frac{JS^2 N_z}{2} + \hat{H}_{sw}$$

$$\hat{H}_{sw} = JS z \sum_k \left( b_k^+ b_k + \frac{\gamma_k}{z} (b_k b_{-k} + b_{-k}^+ b_k^+) \right)$$

$$\gamma_k = \frac{1}{2} \sum_{\vec{r}_{\text{nearest neighbors}}} e^{i\vec{r}_{\text{nearest neighbors}} \cdot \vec{a}}$$

$$= \frac{1}{2} \left( \cos k_x a + \cos k_y a \right) \quad \text{for square lattice.}$$

Diagonalize using the Bogoliubov transformation:

$$\begin{pmatrix} \alpha_k \\ \alpha_{-k}^+ \end{pmatrix} = \begin{pmatrix} \cosh \theta_k & -\sinh \theta_k \\ -\sinh \theta_k & \cosh \theta_k \end{pmatrix} \begin{pmatrix} b_k \\ b_{-k}^+ \end{pmatrix}$$

where $\theta_k = \theta_{-k}$ is a real number to be determined.

Note that $b_k$ and $b_{-k}^+$ shift momentum by the same amount, as does a linear combination of the two.

The inverse transformation is:

$$\begin{pmatrix} b_k \\ b_{-k}^+ \end{pmatrix} = \begin{pmatrix} \cosh \theta_k & \sinh \theta_k \\ \sinh \theta_k & \cosh \theta_k \end{pmatrix} \begin{pmatrix} \alpha_k \\ \alpha_{-k}^+ \end{pmatrix}$$

$\alpha_k$, $\alpha_{-k}^+$ satisfy bosonic commutation relations, for example:

$$[\alpha_k, \alpha_{-k}^+] = \cosh^2 \theta_k [b_k, b_{-k}^+] + \sinh^2 \theta_k [b_k^+, b_k] = \cosh^2 \theta_k - \sinh^2 \theta_k = 1$$

$$[\alpha_k, \alpha_{-k}] = -\cosh \theta_k \sinh \theta_k [b_k, b_{-k}^+] - \cosh \theta_k \sinh \theta_k [b_k^+, b_k] = 0$$
\[ H_{sw} = JS^2 \sum_k \left[ (\cosh 2\theta_k + \eta_k \sinh 2\theta_k) \alpha_k^\dagger \alpha_k \\
+ \frac{1}{2} (\sin 2\theta_k + \eta_k \cosh 2\theta_k) \left( \alpha_k^\dagger \alpha_{-k}^\dagger + \alpha_{-k} \alpha_k \right) \\
+ \sin^2 \theta_k + \frac{\eta_k}{2} \sinh 2\theta_k \right] \]

We choose \( \theta_k \) so as to make the anomalous terms \( (\alpha^\dagger \alpha^\dagger + \alpha \alpha) \) vanish,

\[ \implies \tanh 2\theta_k = -\eta_k \implies \begin{cases} 
\sinh 2\theta_k = -\frac{\eta_k}{\sqrt{1-\eta_k^2}} \\
\cosh 2\theta_k = \frac{1}{\sqrt{1-\eta_k^2}} 
\end{cases} \]

\[ \cosh 2\theta_k + \eta_k \sinh 2\theta_k = \frac{1-\eta_k^2}{\sqrt{1-\eta_k^2}} = \sqrt{1-\eta_k^2} \]

\[ \sinh^2 \theta_k + \frac{\eta_k}{2} \sinh 2\theta_k = \frac{1}{2} \left( \cosh 2\theta_k - 1 \right) + \frac{\eta_k}{2} \sinh 2\theta_k \]

\[ = \frac{1}{2} \left[ \frac{1-\eta_k^2}{\sqrt{1-\eta_k^2}} - 1 \right] = \frac{1}{2} \sqrt{1-\eta_k^2} - \frac{1}{2} \]

Thus,

\[ H_{sw} = -\frac{1}{2} JS^2 N + \sum_k \omega_k \left( \alpha_k^\dagger \alpha_k^\dagger + \frac{1}{2} \right) \]

\[ \omega_k = JS^2 \sqrt{1-\eta_k^2} \]
Long wave length limit \((k \approx 0\) and \(\vec{k} \approx (\pi/a, \pi/a)\)):

\[
\eta_k = \frac{1}{2} \left( \cos k_x a + \cos k_y a \right) \approx \begin{cases} 1 - \frac{1}{4} (\delta k)^2 a^2 & \delta k = \delta \vec{k}^2 \\ -1 + \frac{1}{4} (\delta k)^2 a^2 & \delta k = (\pi/a, \pi/a) + \delta \vec{k}^2 \end{cases}
\]

\[
\omega_k = \int S \pm \sqrt{(1-\eta_k)(1+\eta_k)} \approx JS 2\sqrt{2} a \sqrt{1 \delta \vec{k}^2} \quad \text{square lattice} \quad z=4
\]

This is another example of Goldstone's theorem. However, unlike the ferromagnet where \(\omega_k < k^2\), here \(\omega_k \sim k^2\).

---

**Ground state energy**

The ground state satisfies \(\langle k | GS \rangle = 0\).

It has energy

\[
E = E_{cl} + \Delta E
\]

\[
E_{cl} = -\frac{S^2 J N z}{2} \quad \text{"classical energy"}
\]

\[
\Delta E = \frac{1}{2} \sum_k |J| S^2 z (\sqrt{1-\eta_k^2} - 1) < 0
\]

The quantum correction to the energy is negative.

This makes sense since the "classical" energy was a variational energy in the Neél state. SW theory gives a better variational trial state.
Corrections to the staggered magnetization

Since
\[ \langle S_z \rangle = S - \langle n_i \rangle \]

the reduction in \( \langle S_z \rangle \) due to quantum fluctuations is given by the average HP boson occupation.

\[ \langle n_i \rangle = \langle b_i^+ b_i \rangle = \frac{1}{N} \sum_{k, k'} \langle b_{k}^+ b_{k'} \rangle e^{i(k-k'), r_i} \]

\[ = \frac{1}{N} \sum_{k} \frac{\sinh \theta_k \sinh \theta_{k'}}{\sinh \theta_{k'} - \sinh \theta_k} \langle \alpha_k \alpha_{k'}^+ \rangle e^{i(k-k'), r_i} \quad (\text{since } \langle \alpha_k \alpha_{k'} \rangle = 0) \]

\[ = \frac{1}{N} \sum_{k} \sinh^2 \theta_k \quad \langle \alpha_k \alpha_{k'}^+ \rangle = \delta_{k, k'} \]

\[ = \frac{1}{2N} \sum_{k} \left[ \frac{1}{\sqrt{1-\eta_k^2}} - 1 \right] \]

For two dimensions or higher, this sum converges. This means that, for large enough \( S \), \( \langle S_z \rangle > 0 \). Then, quantum fluctuations renormalize \( \langle S_z \rangle \), but they do not destroy the order.

By contrast, in 1d, the sum is divergent in the thermodynamic limit \( L \to \infty \), when the sum is replaced by an integral,

\[ \langle n_i \rangle = \frac{1}{2} \int \frac{dk}{2\pi} \left[ \frac{1}{\sqrt{1-\eta_k^2}} - 1 \right] \sim \int \frac{dk}{k} \quad \text{at small } k \]

\[ \sim \infty \]

This infrared divergence shows that long range antiferromagnetic order is not possible in 1d, no matter how large is \( S \).