Lecture 4: Mean-field theory and Hartree-Fock theory

In this lecture we’ll describe a general strategy to approximately solving the many body problem introduced in the previous lecture. This approach falls broadly under the rubric of mean-field theory and is better known, in various contexts, as Hartree-Fock theory, the self-consistent field method, and the Gutzwiller ansatz.

The idea behind mean-field theory is simple: we take as a variational class one that neglects all quantum correlations between particles and apply the variational method. In the application of the variational method one then sees that the influence of all the other particles on a given one are treated in an averaged way. Our treatment of the Helium atom in lecture 2 could be seen as an application of mean-field theory in an embryonic form.

To explain mean-field theory in this lecture we’ll consider a sequence of simplified examples culminating in the study of the Hubbard model.

Example 1: a 1D quantum spin system

Quantum spin systems are simplified models that arise as approximations of systems of electrons moving in the presence of a regular array of binding atoms (see, e.g., Auerbach (1994), chapter 3, for an example derivation).

The degrees of freedom of a quantum spin system are, as the name suggests, quantum spins, localised in a regular array. In this example we only consider an array of spin-$\frac{1}{2}$ degrees of freedom arranged in a regular one-dimensional lattice. A convenient basis for a single spin-$\frac{1}{2}$ degree of freedom is provided by the eigenstates of the $s_z$ spin operator, written $|0\rangle$ and $|1\rangle$. The hilbert space of a single spin-$\frac{1}{2}$ is isomorphic to $\mathbb{C}^2$.

The hilbert space $\mathcal{H}$ for a (one-dimensional) collection of $N$ such spin-$\frac{1}{2}$ degrees of freedom is given by

$$\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2.$$ (1)

A general hamiltonian for a quantum spin system has the form

$$H = \sum_{j=1}^{N-1} h_{j,j+1},$$ (2)

where the operator $h_{j,j+1}$ acts nontrivially only on spins $j$ and $j + 1$. The example that we study here is the transverse Ising model which is written

$$H = -\sum_{j=1}^{N-1} \sigma_j^x \sigma_{j+1}^x + \lambda \sum_{j=1}^{N} \sigma_j^z,$$ (3)

where

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$ (4)
are the Pauli sigma matrices. The first summation in (3) describes an interaction between neighbouring spins that encourages quantum spins to align along the $x$ spin axis. The second summation describes the effect of an external magnetic field of strength $\lambda$ which encourages the spins to line up along the $z$ spin axis. There is an obvious competition between these two terms. Indeed, the interplay between the two terms is sufficiently complex that the model exhibits a great deal of interesting physics, including, a quantum phase transition. The transverse Ising model is actually exactly solvable using a sophisticated map to a fermionic system, but we’ll pretend we don’t know this.

A general state $|\Psi\rangle$ of $N$ quantum spins is written, in the basis of the $s_z = \frac{1}{2} \sigma^z$ operators, as

$$|\Psi\rangle = \sum_{z_1, z_2, \ldots, z_N = 0}^1 \psi(z)|z_1z_2\cdots z_N\rangle.$$  \hspace{1cm} (5)

Note: there are $2^N$ terms in this expansion!

Our objective is to understand the ground state $|\Omega\rangle$ of $H$. If we were to proceed by diagonalising $H$, which is in principle possible, it would take a prohibitive time as $N$ becomes large because $H$ is a $2^N \times 2^N$ matrix (even $N = 12$ is rather difficult on a laptop computer). Thus, if we want to understand such a model as $N$ becomes large we must use another method.

In this example we apply the variational method to $H$ using as our variational class the set of all states with the form

$$\mathcal{V} = \{|\Psi\rangle \mid |\Psi\rangle = |\psi_1\rangle|\psi_2\rangle\cdots|\psi_N\rangle\},$$  \hspace{1cm} (6)

i.e., the set of all product states. The class $\mathcal{V}$ enjoys some important features: (i) it is easy to describe, meaning that it takes only $3N$ real numbers to specify a general member of the class (instead of $2^N$); and (ii) it is easy to calculate, meaning that the expectation value of any reasonable observable quantity in a member of $\mathcal{V}$ is easy to compute. However, the class $\mathcal{V}$ has the considerable downside that no member exhibits any spatial correlations, i.e., suppose $A$ is an observable of the spin at location 1 and $B$ is an observable on the spin at location $j$ (for example, $A = \sigma^z_1$ and $B = \sigma^z_j$), then

$$\langle AB \rangle - \langle A \rangle \langle B \rangle = 0$$  \hspace{1cm} (7)

for all $|\Psi\rangle \in \mathcal{V}$. Despite this drawback the class $\mathcal{V}$, when used in conjunction with the variational method, provides surprisingly good results.

We are going to consider the case where $N \to \infty$. In this limit the model $H$ is translation invariant. It is fairly reasonable, although not a priori correct (why not?), to assume that the state $|\Psi\rangle \in \mathcal{V}$ minimising

$$\langle \Psi | H | \Psi \rangle$$  \hspace{1cm} (8)

is itself translation invariant. Assuming, regardless, that this is correct we can restrict our variational class to

$$\mathcal{V}_{TI} = \{|\Psi\rangle \mid |\Psi\rangle = \cdots |\psi\rangle|\psi\rangle\cdots \},$$  \hspace{1cm} (9)

2
so that a general member requires only 3 numbers to specify it. Let’s now apply the
variational method to $H$ using the class $\mathcal{V}_{TI}$. The expectation value of the energy of
the system is given by

$$E[|\psi\rangle] = \langle \Psi | H | \Psi \rangle = -\sum_{j=1}^{N-1} \langle \psi | \sigma^x | \psi \rangle^2 + \lambda \sum_{j=1}^{N} \langle \psi | \sigma^z | \psi \rangle.$$  \hspace{1cm} (10)

(We’ve exploited translation invariance of $\mathcal{V}_{TI}$ to drop the subscripts on the pauli
sigma matrices.) Since this expression generically tends to infinity as $N \to \infty$ it is
convenient to focus, rather, on the energy density $e[|\psi\rangle] = E[|\psi\rangle]/N$. In the limit
$N \to \infty$ the energy density becomes

$$e[|\psi\rangle] = -\langle \psi | \sigma^x | \psi \rangle^2 + \lambda \langle \psi | \sigma^z | \psi \rangle.$$  \hspace{1cm} (11)

The variational method now amounts to minimising $e[|\psi\rangle]$ over all normalised states
$|\psi\rangle$ of a single spin-$\frac{1}{2}$ degree of freedom:

$$\min_{|\psi\rangle} -\langle \psi | \sigma^x | \psi \rangle^2 + \lambda \langle \psi | \sigma^z | \psi \rangle.$$  \hspace{1cm} (12)

This minimisation can be done directly, but here we choose a slightly different route.
At this point we exploit the convenient Bloch sphere representation for a general
(mixed) state of a spin-$\frac{1}{2}$ degree of freedom:

$$\rho = \frac{I + r_x \sigma^x + r_y \sigma^y + r_z \sigma^z}{2}.$$  \hspace{1cm} (13)

In order that $\rho$ is a quantum state (i.e., has both eigenvalues $\geq 0$) it is necessary and
sufficient that $\|r\| \leq 1$. If $\rho$ is pure, i.e., $\rho = |\psi\rangle \langle \psi|$, then $\|r\| = 1$ (see, e.g., Nielsen
and Chuang (2000)). Thus, noting that

$$\langle \psi | \sigma^x | \psi \rangle = \text{tr}(\sigma^x | \psi \rangle \langle \psi|) = r_x,$$  \hspace{1cm} (14)

and

$$\langle \psi | \sigma^z | \psi \rangle = \text{tr}(\sigma^z | \psi \rangle \langle \psi|) = r_z$$  \hspace{1cm} (15)

allows us to write our variational problem as the following geometric problem

$$\min_{\|r\|=1} -r_x^2 + \lambda r_z.$$  \hspace{1cm} (16)

Our variational parameters are the three numbers $r_x, r_y,$ and $r_z$, subject to the con-
straint $r_x^2 + r_y^2 + r_z^2 = 1$. Since $r_y$ plays no role in this minimisation we can set it to 0
so as to allow $r_x$ and $r_z$ to vary over the largest domain. Thus our problem becomes

$$\min_{r_x^2 + r_z^2 = 1} -r_x^2 + \lambda r_z.$$  \hspace{1cm} (17)

In polar coordinates $r_x = \cos(\theta)$ and $r_z = \sin(\theta)$ this becomes

$$\min_{\theta} -\cos^2(\theta) + \lambda \sin(\theta).$$  \hspace{1cm} (18)
In the region \( \lambda \in [-2, 2] \) this equation admits extrema at \( \theta = \pi/2, \theta = 3\pi/2 \), and
\[
\theta = \arcsin(-\lambda/2). \tag{19}
\]
Substituting this into \( e[\theta] \) gives us the value
\[
e = \min\{-\lambda, \lambda, -1 - \lambda^2/4\} = -1 - \lambda^2/4. \tag{20}
\]
Outside this region there is only two extrema at \( \theta = \pi/2, \theta = 3\pi/2 \), and the energy density is
\[
e = -|\lambda|. \tag{21}
\]
The point \(|\lambda| = 2\) is special as the energy density behaves nonanalytically and signifies the presence of a quantum phase transition.

Exercise 1. Calculate the corresponding magnetisation \( \langle \psi|\sigma^z|\psi \rangle \) for the mean-field solution we’ve derived. What happens at \( \lambda = 2\)?

Assignment 1. Carry out a similar analysis as above for the antiferromagnetic Heisenberg model
\[
H = -\sum_{j=1}^{N-1} \sigma_j^x \sigma_{j+1}^x + \sigma_j^y \sigma_{j+1}^y + \sigma_j^z \sigma_{j+1}^z. \tag{22}
\]
First assume that the mean-field solution is translation invariant: what solution do you get in this case. Next try relaxing this assumption by positing that the solution is only 2-periodic:
\[
|\psi \rangle = \cdots |\phi \rangle |\psi \rangle |\phi \rangle |\psi \rangle \cdots . \tag{23}
\]
What value do you get for the energy density in this case?

3 Example 2: spinless fermions on the lattice

In this section we describe the variational principle applied to a class of fermion states known as gaussian or quasi-free. In this case the variational principle is known as Hartree-Fock theory. We consider a second-quantised lattice setting, where the fermion creation and annihilation operators may be given by the finite set
\[
\{a_j, a_k^\dagger\} = \delta_{jk} \mathbb{1}, \quad j, k = 1, 2, \ldots, N. \tag{24}
\]
You can think of \( a_j \) as annihilating a fermion from the single-particle state with wavefunction
\[
\phi_j(x) = \frac{1}{\sqrt{\epsilon}} \chi_{(j\epsilon,(j+1)\epsilon]}(x), \tag{25}
\]
where \( \epsilon \) is the lattice spacing and
\[
\chi_{(a,b]}(x) = 1 \tag{26}
\]
if \( x \in (a,b] \) and zero otherwise. Obviously this is a huge simplification: the operators which annihilate fermions from single-particle states orthogonal to these have been ignored. That such a simplification preserves interesting physical properties of a system of interest is beyond this course but can be found, e.g., in Auerbach (2003).
The model we consider has the second-quantised form
\[ H = -t \sum_{j=1}^{N} (a_j^\dagger a_{j+1} + a_{j+1}^\dagger a_j) + J \sum_{j=1}^{N} a_j^\dagger a_j a_{j+1}^\dagger a_{j+1}, \] (27)
with periodic boundary conditions \( 1 \equiv N + 1 \), and describe fermions hopping on a ring with repulsive interactions between neighbouring sites.

### 3.1 Majorana fermions

In this subsection we follow the paper quant-ph/0404180 closely. It is very much worthwhile reading this paper in full.

The gaussian or quasi-free fermion states are morally analogous to the product states we studied above, and may be defined via several routes (the analogy is that in both cases a system whose state is product/gaussian may be though of as not interacting). Here we define them as all those states arising from a certain closed subset of quadratic physical operations generated by hamiltonians of the form
\[ K = \sum_{j=1}^{N} \epsilon_j a_j^\dagger a_j + H_t + H_t, \] (28)
where
\[ H_t = \sum_{1 \leq j < k \leq N} t_{j,k} a_j^\dagger a_k + \tilde{t}_{k,j} a_j^\dagger a_k; \] (29)
are single-particle, or tunneling, transformations and
\[ H_s = \sum_{1 \leq j < k \leq N} s_{j,k} a_j^\dagger a_k^\dagger + \tilde{s}_{j,k} a_j a_k, \] (30)
which model *squeezing* operations, e.g., an interaction with a bulk \( p \)-wave superconductor where a pair of electrons is swapped against a cooper pair. Both of these generators are quadratic in the fermion operators.

Rather than expressing everything in terms of the non-hermitian operators \( a_j \) and \( a_j^\dagger \) it is convenient to introduce the \( 2N \) hermitian *Majorana* fermion operators
\[ c_{2j-1} = a_j^\dagger + a_j, \] (31)
and
\[ c_{2j} = (-i)(a_j^\dagger - a_j), \] (32)

analogous to the bosonic position and momentum operators. From the anticommutation relations it follows that
\[ \{c_a, c_b\} = 2\delta_{ab}\mathbb{I} \] (33)
for all \( 1 \leq a, b, \leq 2N \). The set of all linear combinations of products of these elements is called the *Clifford algebra* \( Cl_{2N} \). An arbitrary element \( X = Cl_{2N} \) can always be represented as
\[ X = \alpha \mathbb{I} + \sum_{p=1}^{2N} \sum_{1 \leq j_1 < j_2 < \cdots < j_p \leq 2N} \alpha_{j_1,j_2,\ldots,j_p} c_{j_1} c_{j_2} \cdots c_{j_p}, \] (34)
The hamiltonian $K$ takes the form
\[ K = \frac{i}{4} \sum_{a,b=1}^{2N} K_{ab} c_a c_b, \tag{35} \]
where $K_{ab}$ may be an arbitrary antisymmetric real $2N \times 2N$ matrix. Define $V = e^{iK}$, then
\[ V c_a V^\dagger = \sum_{b=1}^{2N} R_{ab} c_b, \tag{36} \]
with $RR^T = I$, and $\det R = 1$. Any rotation in $SO(2N)$ may be implemented with appropriate choice of $K$. (Exercise: prove these statements.)

### 3.2 Grassmann numbers

Grassmann numbers are built using an $n$-dimensional complex vector space: consider a basis
\[ \theta_a, \ 1, \ldots, n \tag{37} \]
of $\mathbb{C}^n$. An example would be simply the column vectors with a 1 in the $j$th place. At the moment all we know is how to add or subtract these elements, i.e., there is no product operation defined on the vector space. We supply a product by defining
\[ \theta_a^2 = 0, \ \forall a \tag{38} \]
and
\[ \theta_a \theta_b + \theta_b \theta_a = 0, \tag{39} \]
an extend it by linearity to an arbitrary element of $\mathbb{C}^2$. No other product relations are imposed. Thus $\theta_a \theta_b$ is not an element of $\mathbb{C}^n$ and the collection of such products provide an additional $n(n-1)/2$ linearly independent elements. Indeed, it is possible to find $2^n$ linearly independent elements in total generated by the above relations. The set of all such elements are called the Grassmann numbers $G_n$. An arbitrary element of $G_n$ may be written as
\[ f = \alpha + \sum_{p=1}^n \sum_{1 \leq j_1 < j_2 < \cdots < j_p \leq n} \alpha_{j_1 j_2 \cdots j_p} \theta_{j_1} \theta_{j_2} \cdots \theta_{j_p}. \tag{40} \]

### 3.3 Gaussian states

Suppose that $X \in \text{Cl}_{2N}$ is some linear combination of products of majorana fermion operators. We can naturally associate a Grassmann number $\omega(X, \theta)$ to such an operator by replacing $c_a$’s with $\theta_a$’s by defining
\[ \omega(X, \theta) \equiv \alpha + \sum_{p=1}^{2N} \sum_{1 \leq j_1 < j_2 < \cdots < j_p \leq 2N} \alpha_{j_1 j_2 \cdots j_p} \theta_{j_1} \theta_{j_2} \cdots \theta_{j_p}. \tag{41} \]
for all
\[ X = \alpha \mathbb{I} + \sum_{p=1}^{2N} \sum_{1 \leq j_1 < j_2 < \cdots < j_p \leq 2N} \alpha_{j_1 j_2 \cdots j_p} c_{j_1} c_{j_2} \cdots c_{j_p}. \] (42)

Warning: this is a map on \( \mathcal{C}_\ell^{2N} \) to \( \mathcal{G}_2^N \) only as linear spaces, the product operation is not preserved by this operation.

Suppose that \( V \) is a transformation implementing the rotation \( R \in SO(2N) \) (see above), and \( X \in \mathcal{C}_\ell^{2N} \) an arbitrary operator. Then
\[ \omega(V XV^\dagger, \theta) = \omega(X, \eta), \quad \eta_a = \sum_{b=1}^{2N} R_{ab} \theta_b. \] (43)

We come now to the main

**Definition 1** A quantum state of \( N \) fermionic modes is Gaussian if and only if its density operator \( \rho \in \mathcal{C}_\ell^{2N} \) has a Gaussian Grassmann representation, i.e.,
\[ \omega(\rho, \theta) = \frac{1}{2^{2N}} e^{\frac{i}{2} \sum_{a,b} \theta_a M_{ab} \theta_b} \] (44)
for some \( 2N \times 2N \) antisymmetric matrix \( M \). The matrix \( M \) is called the correlation matrix of \( \rho \).

The correlation matrix for a Gaussian state \( \rho \) can be found via
\[ M_{ab} = \frac{i}{2} \text{tr}(\rho [c_a, c_b]) = \begin{cases} \text{tr}(\rho c_a c_b), & a \neq b \\ 0, & a = b. \end{cases} \] (45)

The correlation matrix completely characterises \( \rho \) via Wick’s theorem because the expectation value of any higher-order monomial of fermion operators may be computed using the formula
\[ \text{tr}(\rho^{i p} c_{a_1} c_{a_2} \cdots c_{a_{2p}}) = \text{Pf}(M|_{a_1 a_2 \cdots a_{2p}}), \] (46)
with \( 1 \leq a_1 < \cdots < a_{2p} \leq 2N \), \( \text{Pf}(\cdot) \) denotes the Pfaffian, and \( M|_{a_1 a_2 \cdots a_{2p}} \) denotes the \( 2p \times 2p \) submatrix of \( M \) with the indicated rows and columns. The only case we’re really going to use is
\[ \text{tr}(\rho c_1 c_2 c_3 c_4) = -M_{12} M_{34} + M_{13} M_{24} - M_{14} M_{23}. \] (47)

Any real \( 2N \times 2N \) antisymmetric matrix \( M \) can be converted into a block diagonal form by an appropriate choice of rotation \( R \in SO(2N) \) via
\[ R^T M R = \bigoplus_{j=1}^{N} \begin{pmatrix} 0 & \lambda_j \\ -\lambda_j & 0 \end{pmatrix}. \] (48)

The absolute values \( |\lambda_j|, j = 1, 2, \ldots, N \) are the Williamson eigenvalues of \( M \). It follows that any Gaussian state may be transformed via \( R \in SO(2N) \) into a product form
\[ \rho(\lambda_1, \ldots, \lambda_N) = \frac{1}{2^N} \prod_{j=1}^{N} (\mathbb{I} + i \lambda_j c_{2j-1} c_{2j}). \] (49)

In order the \( \rho \) be a legal quantum state it is necessary that \( |\lambda_j| \leq 1, j = 1, 2, \ldots, N \), which is the same as saying that the eigenvalues of \( M^T M \) must all lie in \([-1, 1]\).
3.4 Generalised Hartree-Fock theory

We finally come to the formulation of generalised Hartree-Fock theory. We follow, in part, the paper arXiv:1005.5284.

By transforming our original fermion operators $a_j$ to the Majorana representation $c_a$ our original hamiltonian takes the form

$$H = i \sum_{ab} T_{ab} c_a c_b + \sum_{abde} U_{abde} c_a c_b c_d c_e,$$

(50)

with $T$ antisymmetric. Exercise: what is the exact form of $T$ and $U$ in our case?

Let’s now apply the variational principle to $H$ using as our variational class the set of all Gaussian states, both mixed and pure. Thus we aim to solve the optimisation problem

$$\inf_{\rho \text{gaussian}} \text{tr}(\rho H).$$

(51)

This is greatly simplified by noticing that

$$i \text{tr}(\rho c_a c_b) = M_{ab}$$

(52)

and

$$\text{tr}(\rho c_a c_b c_d c_e) = -M_{ab} M_{de} + M_{ad} M_{be} - M_{ae} M_{bd}$$

(53)

so that

$$\text{tr}(\rho H) = \sum_{ab} T_{ab} M_{ab} + \sum_{abde} U_{abde} (-M_{ab} M_{de} + M_{ad} M_{be} - M_{ae} M_{bd}).$$

(54)

Notice what a huge simplification this is: to specify our state we need only specify the $n(n-1)/2$ numbers defining the upper triangular portion of $M$, and the energy is a function purely of these numbers. Generalised Hartree-Fock theory is then to carry out the minimisation

$$\inf_{M, M^T \leq 1} \sum_{ab} T_{ab} M_{ab} + \sum_{abde} U_{abde} (-M_{ab} M_{de} + M_{ad} M_{be} - M_{ae} M_{bd}).$$

(55)

This is far from trivial for arbitrary $T$ and $U$, and we must take recourse, in general to numerical methods gradient descent methods. However, we have made a huge saving because this problem can at least be stored in a computer’s memory for large $N$, in contrast to the situation where non-Gaussian states are considered. Additionally, symmetries may allow us to compute the objective function efficiently.