

Lecture notes for FYS–KJM 4480
Quantum mechanics for many-particle systems

Simen Kvaal

September 26, 2015

Contents

Chapter 1

Fundamental formalism

Suggested reading for this chapter: Raimes [?], sections 1.1–1.3, and Gross/Runge/Heinonen [?], section I.1. Chapter 1 of Szabo/Ostlund [?] contains a nice refresher on mathematical topics, including linear algebra.

Disclaimer: This course is not a mathematics course, so some mathematical details are glossed over. Examples include the fact that the Hamiltonian is rarely a bounded operator, such that there exists Ψ such that $\hat{H}\Psi$ does not make sense in Hilbert space. Often, operators have complicated spectra, including continuous parts. For points in the continuous spectrum, no square-integrable eigenfunction exists. As is usual, we ignore this complication, and basically calculate as if everything were matrices of finite dimension.

1.1 Many-particle systems

1.1.1 Hilbert space and Hamiltonian

We discuss the non-relativistic quantum mechanical description of a system of many particles. For simplicity, we consider N identical particles.

Whereas the classical state of such a system is a point in phase space, the quantum state is a wavefunction depending on all the coordinates:

$$\Psi = \Psi(x_1, x_2, \dots, x_N), \quad (1.1)$$

where x_i is a point in the configuration space X , the space where each particle “lives”. The configuration space¹ for all N particles is thus X^N , and

$$\Psi : X^N \longrightarrow \mathbb{C}. \quad (1.2)$$

Example: The configuration space for an electron is $\mathbb{R}^3 \times \{-\frac{1}{2}, +\frac{1}{2}\}$. A single electron’s configuration is $x = (\vec{r}, s_z)$, where s_z is the projection of the electron spin along some axis. The one-electron wavefunction can thus be considered a *two-component* wavefunction. The N -electron wavefunction is thus a function of N coordinates \vec{r}_i and N spin variables $s_{z,i}$, in total 2^N components.

Example: A nucleon has two discrete degrees of freedom: spin and isospin. Thus, $X = \mathbb{R}^3 \times \{-\frac{1}{2}, +\frac{1}{2}\} \times \{-\frac{1}{2}, +\frac{1}{2}\}$, $x = (\vec{r}, s_z, i_z)$. A single nonrelativistic nucleon thus have a four-component wavefunction, and N nucleons 4^N components.

¹Since the particles are identical, the configuration space is actually the quotient space X^N/S_N , where S_N is the permutation group of N objects. This means that we identify points in X^N that differ only by a permutation. Suppose $X = \mathbb{R}^3$. Then X^N is a flat space. But X^N/S_N is actually a curved space! For low-dimensional systems, $X = \mathbb{R}^1$ or $X = \mathbb{R}^2$, one can show that particle statistics is not confined to only bosons or fermions. See [?].

Remark: Mathematically, X is a *measure space*, which means that a function $\psi : X \rightarrow \mathbb{C}$ can be *integrated* over subsets of X . For subsets of \mathbb{R}^n , the standard measure is Lebesgue measure, which gives an integral slightly more general than the Riemann integral encountered in introductory analysis courses. For discrete sets, the standard measure is *counting measure*, where the integral is simply a sum. See also the small section on finite dimensional spaces further down. This remark is for orientation only. For us, we simply state that we *integrate over continuous degrees of freedom and sum over discrete degrees of freedom*. For $X = \mathbb{R}^d \times S$ with $S = \{s_1, s_2, \dots, s_n\}$ a discrete set, we define

$$\int_X f(x) dx = \sum_{s \in S} \int_{\mathbb{R}^d} f(\vec{r}, s) d^d \vec{r}.$$

The wavefunction has a probabilistic interpretation: $P(x_1, \dots, x_N) = |\Psi(x_1, x_2, \dots, x_N)|^2$ is the probability density for locating all particles at the point $(x_1, \dots, x_N) \in X^N$. Therefore, Ψ must be square integrable, and be in the Hilbert space $L^2(X^N)$,

$$\Psi \in L^2(X^N). \quad (1.3)$$

All physics can be obtained from the state Ψ .

The governing equation in non-relativistic quantum mechanics is the time-dependent Schrödinger equation (TDSE):

$$\hat{H}\Psi(x_1, x_2, \dots, x_N, t) = i\hbar \frac{\partial}{\partial t} \Psi(x_1, x_2, \dots, x_N, t). \quad (1.4)$$

The system Hamiltonian \hat{H} is obtained from its classical counterpart (if such exists) by a procedure called *Weyl quantization* [Add reference!]. If \hat{H} does not explicitly depend on time, the TDSE can be “solved” by instead considering the time-independent Schrödinger equation (TISE),

$$\hat{H}\Psi(x_1, x_2, \dots, x_N) = E\Psi(x_1, x_2, \dots, x_N). \quad (1.5)$$

The reason is well-known: the evolution operator is diagonal in the eigenbasis.

The time-independent Schrödinger equation is the main focus in this course, and we will only scratch the surface. Ψ is a very, very, *very* complicated function. Intuitively, one might think that solving for Ψ is N times as hard as solving for an $N = 1$ wavefunction. However, Ψ is a function of *all* N coordinates. Resolving each coordinate on a grid with, say, K points requires K^N points in total. For $K = 2$ (which is rather coarse) and $N = 40$ (e.g., a ^{40}Ca nucleus), we need $2^{40} \approx 10^{12}$ data points! Describing the correlated motion of N quantum particles is harder than the pioneers of quantum mechanics thought! Literally *thousands* of researchers worldwide are make a living out of devising more or less clever schemes for finding approximate solutions.

1.1.2 More on the space X , and finite dimensional spaces

The space $L^2(X)$ need not be infinite dimensional. Recall that $\psi \in L^2(X)$ if and only of $\psi : X \rightarrow \mathbb{C}$ and

$$\|\psi\|^2 = \int |\psi(x)|^2 dx < +\infty.$$

Let

$$X = \{1, 1, 2, \dots, L\}$$

be a discrete set, such that the integral is a sum. For example, we could discretize space using grid points, or a finite collection of basis functions such as Hermite functions, finite element functions, etc. Then $L^2(X)$ consists of functions $\psi : X \rightarrow \mathbb{C}$, i.e., functions from the integers $1, \dots, L$ to \mathbb{C} . But this is just an ordinary vector in \mathbb{C}^L ! The integral becomes

$$\|\psi\|^2 = \sum_{i=1}^L |\psi_i|^2,$$

which is always finite. Thus, using counting measure on this particular X , $L^2(X) \simeq \mathbb{C}^L$.

1.1.3 The manybody Hamiltonian

Having introduced the wavefunction, we now consider the Hamiltonian. In this course, we shall consider only Hamiltonians on the following generic form:

$$\begin{aligned}\hat{H} &= \sum_{i=1}^N \hat{h}(i) + \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^N \hat{w}(i, j) \\ &= \hat{H}_0 + \hat{W}.\end{aligned}\tag{1.6}$$

where $\hat{h}(i)$ denotes a single-particle operator acting only on the degrees of freedom of particle i , and $\hat{w}(i, j)$ denotes a two-body operator that acts only on the degrees of freedom of the *pair* (i, j) , $i \neq j$.

Of course, one could consider three-body forces as well, and even higher. Such occur in nuclear physics.

Let us take the Hamiltonian of an atom in the Born–Oppenheimer as an example.

The Hamiltonian for a free electron is just its kinetic energy,

$$\hat{t} = \frac{1}{2m_e} p^2 = \frac{1}{2m_e} (-i\hbar \nabla)^2 = -\frac{\hbar^2}{2m_e} \nabla^2.\tag{1.7}$$

If it is moving in an external field, such as the Coulomb field set up by an atomic nucleus of charge $+Ze$ at the location \vec{R} , we obtain the total single-particle Hamiltonian

$$\hat{h} = \hat{t} + \hat{v} = -\frac{\hbar^2}{2m_e} \nabla^2 - \frac{Ze^2}{\|\vec{R} - \vec{r}\|}.\tag{1.8}$$

The Hamiltonian for a system of N electrons, neglecting interactions, becomes

$$\hat{H}_0 = \sum_{i=1}^N \hat{h}(i) = \sum_{i=1}^N \left[-\frac{\hbar^2}{2m_e} \nabla_i^2 - \frac{Ze^2}{|\vec{r}_i - \vec{R}|} \right].\tag{1.9}$$

The electron pair (i, j) interact via the Coulomb force:

$$w(i, j) = \frac{e^2}{|\vec{r}_i - \vec{r}_j|}.\tag{1.10}$$

Thus,

$$\hat{W} = \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}} w(i, j) = \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}} \frac{e^2}{|\vec{r}_i - \vec{r}_j|}.\tag{1.11}$$

1.1.4 Separation of variables

If we neglect the two-body part \hat{W} of the Hamiltonian, we may “solve” the TISE by separation of variables. We do this now as a preliminary step, before we discuss the consequences of the particles being indistinguishable.

We seek an eigenfunction $\Psi \in L^2(X^N)$ to the non-interacting Hamiltonian \hat{H}_0 . Write

$$\Psi(x_1, \dots, x_N) = \psi_1(x_1) \psi_2(x_2) \cdots \psi_N(x_N).\tag{1.12}$$

Plug in to the TISE and divide by Ψ to get

$$\sum_i \psi_i^{-1} [h(i) \psi_i] = E.\tag{1.13}$$

The right hand side is a constant. The left hand side is a sum of functions $f_1 + f_2 + \dots + f_N$, $f_i = f_i(x_i)$. This can only sum to a constant if $f_i(x_i)$ is a constant,

$$\hat{h}\psi_i(x) = \epsilon_i\psi_i(x), \quad (1.14)$$

which is just the TISE for a single particle! Thus, for any collection of N eigenvalues of the single-particle problem, we get a solution of the N particle problem. We obtain that the total eigenfunction is

$$\Psi(x_1, x_2, \dots, x_N) = \psi_{i_1}(x_1)\psi_{i_2}(x_2)\dots\psi_{i_N}(x_N) \quad (1.15)$$

with eigenvalue

$$E = \epsilon_{i_1} + \dots + \epsilon_{i_N}. \quad (1.16)$$

One can also show that the converse is true: any eigenfunction Ψ can be taken on the above form.

1.1.5 Permutations

Are you unfamiliar with permutations? Ask me, and I will add some paragraphs here.

1.1.6 Particle statistics

Our particles are identical, or indistinguishable. There is abundant evidence that all elementary particles must be treated as such. That means that our probability density must be *permutation invariant* in the following sense: let $\sigma \in S_N$ be a permutation of N indices, and let $(x_1, \dots, x_N) \in X^N$ be a configuration of the N particles. Then we must have

$$|\Psi(x_1, x_2, \dots, x_N)|^2 = |\Psi(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(N)})|^2. \quad (1.17)$$

This is equivalent to

$$\Psi(x_1, \dots, x_N) = e^{i\alpha} \Psi(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(N)}) \quad (1.18)$$

for some real α , that may depend on σ . (Clearly, our separation of variables eigenfunctions do not satisfy this!)

Define a linear operator \hat{P}_σ via

$$(\hat{P}_\sigma \Psi)(x_1, \dots, x_N) = \Psi(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(N)}), \quad (1.19)$$

that is, the operator that evaluates Ψ at permuted coordinates. We have reformulated particle indistinguishability as: Ψ is an eigenfunction of \hat{P}_σ for every $\sigma \in S_N$, with eigenvalue possibly depending on σ .

One can show (see the exercises), that either $\hat{P}_\sigma \Psi = \Psi$ for every $\sigma \in S_N$, or $\hat{P}_\sigma \Psi = (-1)^{|\sigma|} \Psi$ for every $\sigma \in S_N$, where $|\sigma|$ is the number of transpositions in σ , and thus $(-1)^{|\sigma|}$ is the sign of the permutation. In the former case, Ψ is “totally symmetric with respect to permutations”, and in the latter case, “totally anti-symmetric”.

It is a *postulate* that particles occurring in quantum theory (in three-dimensional space) is of one of two types: bosons or fermions. Bosons have totally symmetric wavefunctions only, and fermions have totally anti-symmetric wavefunctions only. To cite Leinaas and Myrheim [?], “The physical consequences of this postulate seems to be in good agreement with experimental data.” Wolfgang Pauli proved (using relativistic considerations) that wavefunctions of half-integral *spin* must be anti-symmetric, and wavefunctions of particles with integral spin must be symmetric, connecting the postulate with the intrinsic spin of particles. To this day, no particles with other spin values have been found.

In this course, we focus on fermions. See, e.g., [?] for the general case.

Exercise 1.1. In this exercise, we prove that if $\Psi \in L^2(X^N)$ is an eigenfunction for all \hat{P}_σ , then the eigenvalue is either 1 or $(-1)^{|\sigma|}$.

We introduce transpositions: $\tau \in S_N$ is transposition if it exchanges only a single pair (i, j) , $i \neq j$. Write $\hat{P}_{ij} \equiv \hat{P}_\tau$.

Assume that $\Psi \in L^2(X^N)$ is such that, for all $\sigma \in S_N$,

$$\hat{P}_\sigma \Psi = s_\sigma \Psi, \quad s_\sigma = e^{i\alpha(\sigma)}.$$

Show that $\hat{P}_{ij}^2 = 1$, and find all the possible eigenvalues of \hat{P}_{ij} .

Under the assumption on Ψ , show that if s_{ij} is the eigenvalue of \hat{P}_{ij} ,

$$\hat{P}_{ij} \Psi = s_{ij} \Psi,$$

then, for any other pair (i', j') , the eigenvalue is $s_{ij} = s_{i'j'}$. You will probably need to use the group theoretical properties of permutations.

We have established that the eigenvalue of a transposition is a characteristic of Ψ , let $s = s_{ij}$. Compute the eigenvalue of P_σ for arbitrary σ in terms of s . \triangle

Exercise 1.2. Let

$$\hat{H} = \sum_{i=1}^N \hat{h}(i) + \sum_{(i,j)} \hat{w}(i, j).$$

Show that \hat{H} commutes with P_σ for any permutation $\sigma \in S_N$, i.e., show that for *any* wavefunction $\Psi \in L^2(X^N)$,

$$\hat{H} P_\sigma \Psi = P_\sigma \hat{H} \Psi. \quad (1.20)$$

\triangle

Exercise 1.3. In this exercise, we consider $X = \mathbb{R}^3$, i.e., no spin. Consider each of the below functions.

1. $\Psi(\vec{r}_1, \vec{r}_2) = e^{-\alpha|\vec{r}_1 - \vec{r}_2|}$.
2. $\Psi(\vec{r}_1, \vec{r}_2) = \sin(\vec{e}_z \cdot (\vec{r}_1 - \vec{r}_2))$, where \vec{e}_z is the unit vector in the z -direction.
3. $\Psi(\vec{r}_1, \vec{r}_2, \vec{r}_3) = \sin[\vec{r}_1 \cdot (\vec{r}_2 \times \vec{r}_3)] e^{-|\vec{r}_1|^2} e^{-|\vec{r}_2|^2} e^{-|\vec{r}_3|^2}$

Answer the following questions, per function:

- Is the function totally symmetric?
- Is the function totally antisymmetric?
- Is the function square integrable?

\triangle

1.1.7 Slater determinants

The set of totally antisymmetric wavefunctions $L^2(X^N)_{AS}$ in $L^2(X^N)$ form a *closed subspace* of Hilbert space: it is a linear space which is complete. Thus $L^2(X^N)_{AS}$ is a Hilbert space in its own right, and from our perspective it is the “true” Hilbert space of N identical fermions.

The antisymmetry of a wavefunction of N coordinates is a quite complicated constraint. We are also used to orthonormal bases, and it may seem daunting to come up with such a basis which is also antisymmetric. Slater determinants are the solution.

Exercise 1.4. Prove that $L^2(X^N)_{AS}$ is a linear space. Additionally, if you have the mathematical background, prove that it is a closed subspace using the Hilbert space metric. \triangle

The original space has a tensor product representation:

$$L^2(X^N) = L^2(X) \otimes L^2(X) \otimes \cdots \otimes L^2(X) \quad (N \text{ factors}). \quad (1.21)$$

Here, $L^2(X)$ is the Hilbert space of a single fermion. Let us assume that we have an orthonormal basis (ONB) ϕ_1, ϕ_2, \dots , for this space, such that we can expand any $\psi \in L^2(X)$ as

$$\psi(x) = \sum_{\mu} c_{\mu} \phi_{\mu}(x), \quad (1.22)$$

with

$$\langle \phi_{\mu} | \phi_{\nu} \rangle = \delta_{\mu, \nu} \quad (1.23)$$

and

$$\|\psi\|^2 = \sum_{\mu} |c_{\mu}|^2. \quad (1.24)$$

Thus, $\psi(x)$ is represented by an (infinite) vector $[c_{\mu}] = (c_1, c_2, \dots)$. Because of Eq. (??), we may construct a basis for $L^2(X^N)$ by *tensor products*,

$$\tilde{\Phi}_{\mu_1, \dots, \mu_N}(x_1, \dots, x_N) = \phi_{\mu_1}(x_1) \phi_{\mu_2}(x_2) \cdots \phi_{\mu_N}(x_N). \quad (1.25)$$

Any $\Psi \in L^2(X^N)$ can be written

$$\Psi(x_1, \dots, x_N) = \sum_{\mu_1} \cdots \sum_{\mu_N} c_{\mu_1, \dots, \mu_N} \tilde{\Phi}_{\mu_1, \dots, \mu_N}(x_1, \dots, x_N), \quad (1.26)$$

with

$$\langle \tilde{\Phi}_{\mu_1, \dots, \mu_N} | \tilde{\Phi}_{\nu_1, \dots, \nu_N} \rangle = \delta_{\mu_1, \nu_1} \cdots \delta_{\mu_N, \nu_N}. \quad (1.27)$$

In the $N = 2$ case, we see that $\Psi(x_1, x_2)$ can be represented by an infinite *matrix* $[c_{\mu_1 \mu_2}]$, and in the $N = 3$ case a 3D matrix, and so on.

Remark: Compare this with the separation-of-variables treatment. If the set of eigenfunctions $\psi_i \in L^2(X)$ of \hat{h} is complete, our separation of variables eigenfunctions $\Psi = \psi_1 \psi_2 \cdots \psi_N$ form a complete set too.

Anoter remark: For arbitrary N , the tensor product basis described *can be counted*. For arbitrary N , let us introduce a generic index, a multiindex, $k = (\mu_1, \dots, \mu_N)$. There is a one-to-one mapping between multiindices and the natural numbers $\mathbb{N} = \{0, 1, 2, \dots\}$. Thus, writing $\xi = (x_1, \dots, x_N)$

$$\Psi(\xi) = \sum_k c_k \tilde{\Phi}(\xi), \quad \langle \tilde{\Phi}_k | \tilde{\Phi}_{\ell} \rangle = \delta_{k, \ell} \quad (1.28)$$

all the various N are represented with the same formula. There is nothing special about c being a vector, a matrix, a 3D matrix, etc. They are all fundamentally equivalent, since the basis set can be counted.

Important message so far: a single-particle basis set $\{\phi_\mu\}$ can be used to construct a basis for $L^2(X^N)$.

What about or “actual” Hilbert space, $L^2(X_N)_{AS}$? Can we construct a basis for this using our single-particle basis? Yes, this is the role of *Slater determinants*.

What is the simplest totally antisymmetric wavefunction we can create, starting with some single-particle functions? If we start with $N = 2$, and consider the product $\phi_1(x_1)\phi_2(x_2)$, this is not antisymmetric. But if we consider the linear combination

$$\Phi(x_1, x_2) = \phi_1(x_1)\phi_2(x_2) - \phi_2(x_1)\phi_1(x_2), \quad (1.29)$$

this is antisymmetric if we exchange x_1 and x_2 . Continuing with $N = 3$, we quickly realize that in order to obtain something antisymmetric out of $\phi_1(x_1)\phi_2(x_2)\phi_3(x_3)$, we must take the linear combination

$$\begin{aligned} \Phi(x_1, x_2, x_3) = & \phi_1(x_1)\phi_2(x_2)\phi_3(x_3) - \phi_2(x_1)\phi_1(x_2)\phi_3(x_3) - \phi_1(x_1)\phi_3(x_2)\phi_2(x_3) - \\ & \phi_3(x_1)\phi_2(x_2)\phi_1(x_3) + \phi_2(x_1)\phi_3(x_2)\phi_1(x_3) + \phi_3(x_1)\phi_1(x_2)\phi_2(x_3), \end{aligned} \quad (1.30)$$

each term representing a permutation of the indices (123). There is nothing special about (123) of course, $(\mu_1\mu_2\mu_3)$ also works. Note that if one of these indices are equal, then the whole linear combination is zero as well.

The generalization to N indices is in fact a *determinant*, and we make a definition:

Definition 1.1. Let $\phi_1, \phi_2, \dots, \phi_N$ be arbitrary single-particle functions in $L^2(X)$ (not necessarily orthonormal). The *Slater determinant* defined by these functions is denoted by $[\phi_1\phi_2\cdots\phi_N]$, and is defined via the formula

$$\begin{aligned} [\phi_1, \phi_2, \dots, \phi_N](x_1, \dots, x_N) &= \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(x_1) & \phi_1(x_2) & \cdots & \phi_1(x_N) \\ \phi_2(x_1) & \phi_2(x_2) & \cdots & \phi_2(x_N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N(x_1) & \phi_N(x_2) & \cdots & \phi_N(x_N) \end{vmatrix} \\ &= \frac{1}{\sqrt{N!}} \sum_{\sigma \in S_N} (-1)^{|\sigma|} \prod_{i=1}^N \phi_{\sigma(i)}(x_i) \\ &= \frac{1}{\sqrt{N!}} \sum_{\sigma \in S_N} (-1)^{|\sigma|} \prod_{i=1}^N \phi_i(x_{\sigma(i)}) \end{aligned} \quad (1.31)$$

Note: the $1/\sqrt{N!}$ is there for normalization purposes, see later. The second formula in the definition follows from the theory of matrix determinants.

Exercise 1.5. Show that the two last lines in Eq. (??) are equivalent. This requires some manipulation of permutations. \triangle

Exercise 1.6. Let A be an $N \times N$ matrix. Let $\phi_j, j = 1, \dots, N$ be given single-particle functions, and let $\psi_k, k = 1, \dots, N$ be defined by

$$\psi_k = \sum_j \phi_j A_{jk}. \quad (1.32)$$

Prove that

$$[\psi_1, \psi_2, \dots, \psi_N] = \det(A)[\phi_1, \phi_2, \dots, \phi_N]. \quad (1.33)$$

(Hint: use antisymmetry of Slater determinants with respect to permutations of single-particle functions, and the definition $\det(A) = \sum_{\sigma \in S_N} A_{1\sigma(1)}A_{2\sigma(2)}\cdots A_{N\sigma(N)}$.) \triangle

Exercise 1.7. *NB: This exercise has been updated since it was given as part of Problem set 1 (H2015). The assumption that the indices were sorted was added.* Suppose that $\{\phi_\mu\}$, $\mu = 1, 2, \dots$ are orthonormal. Prove that $\Phi_{\mu_1, \dots, \mu_N} = [\phi_{\mu_1} \phi_{\mu_2}, \dots, \phi_{\mu_N}]$ is normalized,

$$\langle \Phi_{\mu_1 \mu_2 \dots \mu_N} | \Phi_{\mu_1 \mu_2 \dots \mu_N} \rangle = 1.$$

Prove that

$$\langle \Phi_{\mu_1 \mu_2 \dots \mu_N} | \Phi_{\nu_1 \nu_2 \dots \nu_N} \rangle = \delta_{\mu_1 \nu_1} \dots \delta_{\mu_N \nu_N},$$

under the assumption that $\vec{\mu}$ and $\vec{\nu}$ are sorted in increasing order. What do you get for the inner product if the indices are not sorted? \triangle

Observation: Determinant properties imply that permutation of particle indices gives sign change. Permutation of function indices gives sign change:

$$[\phi_1, \dots, \phi_i, \dots, \phi_j, \dots, \phi_N] = -[\phi_1, \dots, \phi_j, \dots, \phi_i, \dots, \phi_N] \quad (1.34)$$

$$[\phi_1, \dots, \phi_N](x_1, \dots, x_i, \dots, x_j, \dots, x_N) = -[\phi_1, \dots, \phi_N](x_1, \dots, x_j, \dots, x_i, \dots, x_N). \quad (1.35)$$

Moreover, two rows equal means that two of the single-particle functions are identical, giving a vanishing determinant.

So the set $\{\mu_1, \mu_2, \dots, \mu_N\}$ with N elements determines the Slater determinant up to a sign change – permuting the elements gives the same determinant.

If two columns in Eq. (??) are identical, the determinant vanishes. Two columns equal mean that we evaluate at some $x_i = x_j$, giving a vanishing determinant. *This is the Pauli exclusion principle.*

Theorem 1.1. *Let $\{\phi_\mu\}$ be an orthonormal basis for $L^2(X)$. Then, any $\Psi \in L^2(X^N)_{AS}$ can be expanded in the Slater determinants*

$$[\phi_{\mu_1}, \phi_{\mu_2}, \dots, \phi_{\mu_N}]. \quad (1.36)$$

Moreover, if we choose an ordering of the indices μ , the Slater determinants satisfying $\mu_1 < \mu_2 < \dots < \mu_N$ form an orthonormal Basis for $L^2(X^N)_{AS}$.

Proof. Step 1: Expand Ψ in the tensor product basis.

$$\Psi(x_1, \dots, x_N) = \sum_{\mu_1, \dots, \mu_N} c_{\mu_1, \dots, \mu_N} \tilde{\Phi}_{\mu_1, \dots, \mu_N}(x_1, \dots, x_N). \quad (1.37)$$

Step 2: Show that the coefficients $c_{\vec{\mu}}$ are antisymmetric under permutation. For simplicity, consider a transposition of i with j , $i < j$:

$$\begin{aligned} \hat{P}_{ij} \Psi(x_1, \dots, x_N) &= \sum_{\mu_1, \dots, \mu_N} c_{\mu_1, \dots, \mu_N} \hat{P}_{ij} \tilde{\Phi}_{\mu_1, \dots, \mu_N}(x_1, \dots, x_i, \dots, x_j, \dots, x_N) \\ &= \sum_{\mu_1, \dots, \mu_N} c_{\mu_1, \dots, \mu_N} \tilde{\Phi}_{\mu_1, \dots, \mu_N}(x_1, \dots, x_j, \dots, x_i, \dots, x_N) \\ &= \sum_{\mu_1, \dots, \mu_N} c_{\mu_1, \dots, \mu_N} \tilde{\Phi}_{\mu_1, \dots, \mu_j, \dots, \mu_i, \dots, \mu_N}(x_1, \dots, x_N) \\ &= \sum_{\mu_1, \dots, \mu_N} c_{\mu_1, \dots, \mu_j, \dots, \mu_i, \dots, \mu_N} \tilde{\Phi}_{\mu_1, \dots, \mu_N}(x_1, \dots, x_N) \\ &= - \sum_{\mu_1, \dots, \mu_N} c_{\mu_1, \dots, \mu_N} \tilde{\Phi}_{\mu_1, \dots, \mu_N}(x_1, \dots, x_N) \end{aligned} \quad (1.38)$$

Projecting the two last inequalities onto $\tilde{\Phi}_{v_1, \dots, v_N}$ gives

$$c_{v_1, \dots, v_j, \dots, v_i, \dots, v_N} = -c_{v_1, \dots, v_i, \dots, v_j, \dots, v_N}. \quad (1.39)$$

We decompose an arbitrary $\sigma \in S_N$ into transpositions, and obtain

$$c_{\mu_{\sigma(1)}, \mu_{\sigma(2)}, \dots, \mu_{\sigma(N)}} = (-1)^{|\sigma|} c_{\mu_1, \dots, \mu_N}. \quad (1.40)$$

Step 3: Rearrange summation so that we exhibit Ψ as a linear combination of Slater determinants.

Note that we can write

$$\sum_{\mu_1, \dots, \mu_N} f(\mu_1, \dots, \mu_N) = \sum_{\mu_1 < \mu_2 < \dots < \mu_N} \sum_{\sigma \in S_N} f(\mu_{\sigma(1)}, \dots, \mu_{\sigma(N)}), \quad (1.41)$$

splitting the summation over *ordered* multiindices and permutations of these. We now get

$$\begin{aligned} \Psi &= \sum_{\mu_1 < \dots < \mu_N} \sum_{\sigma} (-1)^{|\sigma|} c_{\mu_1, \dots, \mu_N} \tilde{\Phi}_{\mu_{\sigma(1)}, \mu_{\sigma(2)}, \dots, \mu_{\sigma(N)}} \\ &= \sum_{\mu_1 < \dots < \mu_N} (\sqrt{N!} c_{\mu_1, \dots, \mu_N}) \frac{1}{\sqrt{N!}} \sum_{\sigma} (-1)^{|\sigma|} \tilde{\Phi}_{\mu_{\sigma(1)}, \mu_{\sigma(2)}, \dots, \mu_{\sigma(N)}} \\ &= \sum_{\mu_1 < \dots < \mu_N} (\sqrt{N!} c_{\mu_1, \dots, \mu_N}) [\phi_{\mu_1}, \dots, \phi_{\mu_N}]. \end{aligned} \quad (1.42)$$

This in fact proves that the Slater determinants, when we only use ordered indices, are sufficient to expand any $\Psi L^2(X^N)_{AS}$. Clearly, if we omit one such Slater determinant, not all Ψ can be expanded. (In particular, this omitted Slater determinant cannot be expanded in the rest!) This, the Slater determinants with ordered indices form a basis. \square

Exercise 1.8. How many terms are there in $[\phi_1 \phi_2 \phi_3 \phi_4](x_1, x_2, x_3, x_4)$, when expanded as a linear combination of tensor products? Write down the expansion explicitly. \triangle

Exercise 1.9. In this exercise, we define the antisymmetrization operator \mathcal{A} as

$$\mathcal{A} = \frac{1}{N!} \sum_{\sigma \in S_N} (-1)^{|\sigma|} \hat{P}_{\sigma}. \quad (1.43)$$

Now,

$$[\phi_1, \dots, \phi_N] = \sqrt{N!} \mathcal{A} \phi_1(x_1) \dots \phi_N(x_N). \quad (1.44)$$

An operator U is an orthogonal projector if and only if $U^2 = U$ and $U^\dagger = U$.

Prove that \mathcal{A} is an orthogonal projector from $L^2(X^N)$ onto $L^2(X^N)_{AS}$. \triangle

1.2 Second quantization

1.2.1 The creation and annihilation operators

In this section, we introduce the following shorthand:

$$L_N^2 \equiv L^2(X^N)_{AS} \quad (1.45)$$

since the space X is understood from context, and since we only deal with fermion spaces. We also introduce the bra/ket notation for wavefunctions.

Recall that a basis for L_N^2 could be formed from an orthonormal basis $\{\phi_\mu\}$ of $L^2(X)$, by computing set of Slater determinants $\Phi_{\mu_1, \dots, \mu_N} = [\phi_{\mu_1}, \dots, \phi_{\mu_N}]$, where $\mu_1 < \mu_2 < \dots < \mu_N$ were *ordered*. (If we permute the index set, we get the same function with a possible sign change, so it is not an additional basis function.)

So far we have emphasized that $[\phi_{\mu_1}, \dots, \phi_{\mu_N}]$ were *functions*, but in quantum mechanics the bra/ket notation is useful. We therefore introduce the ket notation

$$|\psi_1, \dots, \psi_N\rangle = [\psi_1, \dots, \psi_N] \quad (1.46)$$

for an arbitrary Slater determinant. When $\{\phi_\mu\}$ is a Single-particle basis, we may choose to suppress all the ϕ 's everywhere, and write

$$|\vec{\mu}\rangle = |\mu_1 \mu_2 \dots \mu_N\rangle, \quad [\phi_{\mu_1}, \dots, \phi_{\mu_N}](x_1, \dots, x_N) = \langle x_1 \dots x_N | \mu_1 \dots \mu_N \rangle \quad (1.47)$$

for a Slater determinant. If $\mu_i = \mu_j$ then $|\vec{\mu}\rangle = 0$ is the zero vector. We recall the antisymmetry properties,

$$\hat{P}_{ij} |\mu_1 \dots \mu_i \dots \mu_j \dots \mu_N\rangle = -|\mu_1 \dots \mu_j \dots \mu_i \dots \mu_N\rangle \quad (1.48)$$

and more generally

$$\hat{P}_\sigma |\mu_1 \dots \mu_N\rangle = (-1)^{|\sigma|} |\mu_{\sigma(1)} \dots \mu_{\sigma(N)}\rangle. \quad (1.49)$$

For any $|\Psi\rangle \in L_N^2$, we have the basis expansion

$$|\Psi\rangle = \sum_{\vec{\mu}} |\vec{\mu}\rangle \langle \vec{\mu} | \Psi \rangle \quad (1.50)$$

connecting with the earlier treatment. The \sim means that we sum *only over ordered sets of indices*. As we saw earlier, the coefficients $\langle \vec{\mu} | \Psi \rangle$ are permutation antisymmetric.

So far, we have used Greek letters μ , ν , etc., as single-particle indices. There is nothing special about this, of course. We will later also use p , q , r , etc.

Looking at the determinant (??), we see that by adding a row containing the index ν , and a column with coordinate x_{N+1} , we obtain an $N + 1$ particle Slater determinant (modulo a constant factor):

$$\langle x_1 \dots x_{N+1} | \nu \mu_1 \mu_2 \dots \mu_N \rangle = \frac{1}{\sqrt{(N+1)!}} \begin{vmatrix} \phi_\nu(x_1) & \phi_\nu(x_2) & \dots & \phi_\nu(x_N) & \phi_\nu(x_{N+1}) \\ \phi_{\mu_1}(x_1) & \phi_{\mu_1}(x_2) & \dots & \phi_{\mu_1}(x_N) & \phi_{\mu_1}(x_{N+1}) \\ \phi_{\mu_2}(x_1) & \phi_{\mu_2}(x_2) & \dots & \phi_{\mu_2}(x_N) & \phi_{\mu_2}(x_{N+1}) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \phi_{\mu_N}(x_1) & \phi_{\mu_N}(x_2) & \dots & \phi_{\mu_N}(x_N) & \phi_{\mu_N}(x_{N+1}) \end{vmatrix} \quad (1.51)$$

Similarly, we can remove a row and column, and obtain an $N - 1$ particle Slater determinant.

This inspires the *creation and annihilation operators*, that map wavefunctions between different particle number spaces:

$$c_\nu^\dagger : L_N^2 \rightarrow L_{N+1}^2 \quad (1.52)$$

$$c_\nu : L_N^2 \rightarrow L_{N-1}^2 \quad (1.53)$$

The operator c_v^\dagger is called a *creation operator* and is, roughly defined, by inserting a row and column as described. The operator c_v is the Hermitian adjoint of c_v^\dagger , and it will be shown that its action on a Slater determinant corresponds to the mentioned removal of a row and column.

We *define* the space L_0^2 – the zero particle space – as a one-dimensional space spanned by the special ket $|- \rangle$, the *vacuum state*. There is nothing mysterious about this, it is just a definition that will be useful later. Note that $|- \rangle \neq 0$.

Recall that a linear operator is fully defined when we specify its action on a basis set. This is how we define c_μ^\dagger and c_μ .

Definition of the creation operator: For every single-particle index v , we define the creation operator c_v^\dagger acting on the vacuum state by

$$c_v^\dagger |- \rangle = |v \rangle. \quad (1.54)$$

Since this is a Slater determinant with a single particle, we have, of course, $\langle x|v \rangle = \phi_v(x)$. For an arbitrary Slater determinant with $N > 0$, we define the action by

$$c_v^\dagger |\mu_1 \cdots \mu_N \rangle \equiv |v \mu_1 \cdots \mu_N \rangle. \quad (1.55)$$

We observe already that if there is a j such that $v = \mu_j$, then $|v \mu_1 \cdots \mu_N \rangle \equiv 0$:

$$\hat{P}_{1j} |v \vec{\mu} \rangle = - |v \vec{\mu} \rangle = |v \vec{\mu} \rangle = 0, \quad v = \mu_j. \quad (1.56)$$

In terms of determinant coordinate expressions as in Eq. (??), c_v^\dagger inserts a column on the far right with x_{N+1} and inserts a row on the top with the index v . Finally, the whole expression is renormalized.

[Recall that the basis determinants were the determinants that had *ordered indices*. Assume that $\vec{\mu}$ is ordered. Clearly, $c^\dagger |\vec{v} \rangle$ is either zero or equal to $(-1)^j |\mu_1 \mu_2 \cdots \mu_j v \mu_{j+1} \cdots \mu_N \rangle$, which is a new basis determinant. Here, j is chosen such that the augmented index set is ordered.]

Let us now consider the *annihilation operator*. There are no particles to remove in the vacuum state, so we set

$$c_v |- \rangle \equiv 0. \quad (1.57)$$

Let $\vec{\mu}$ be a multiindex. If $v = \mu_j$ for some j , we define

$$c_v |\vec{\mu} \rangle \equiv (-1)^{j-1} |\mu_1 \cdots \mu_{j-1} \mu_{j+1} \cdots \mu_N \rangle. \quad (1.58)$$

In terms of the coordinate determinant expression, this amounts to moving the j th row to the top with $j-1$ transpositions, giving the sign factor, and then crossing out the far right column and the first row, now containing the index v . This moving of the j th row may seem like a complication compared to the creation operator, but note that for c_v^\dagger we defined its action by *inserting v on the top*. Moving v to the $(j+1)$ th position will induce a $(-1)^j$. But c_v removes a row at an in principle arbitrary location.

Exercise 1.10. Prove that c_α^\dagger and c_α are Hermitian adjoints of each other, as the notation suggests. Thus, for any $\vec{\mu}$ with N indices, and \vec{v} with $N+1$ indices, show that

$$\langle \vec{\mu} | (c_\alpha |\vec{v} \rangle) = [\langle \vec{v} | (c_\alpha^\dagger |\vec{\mu} \rangle)]^* \quad (1.59)$$

△

1.2.2 Anticommutator relations

Recall that the anticommutator of two operators is defined by

$$\{\hat{A}, \hat{B}\} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}. \quad (1.60)$$

In this section, we prove three important anticommutation relations:

$$\{c_{v_1}^\dagger, c_{v_2}^\dagger\} = 0 \quad (1.61a)$$

$$\{c_{v_1}, c_{v_2}\} = 0 \quad (1.61b)$$

$$\{c_{v_1}, c_{v_2}^\dagger\} = \delta_{v_1, v_2}. \quad (1.61c)$$

Equation (1.61c) is called the “fundamental anticommutator”.

Let v_1, v_2 be two single-particle indices, and let $N \geq 0$ be arbitrary. By the properties of determinants, it is easy to see, that for any $|\vec{\mu}\rangle \in L_N^2$,

$$c_{v_1}^\dagger c_{v_2}^\dagger |\vec{\mu}\rangle = -c_{v_2}^\dagger c_{v_1}^\dagger |\vec{\mu}\rangle. \quad (1.62)$$

Why? The right hand side is obtained by exchanging the two first rows of the determinant on the left hand side.

Since this equation holds for any basis vector, we have shown that the two creation operators *anticommute*

$$\{c_{v_1}^\dagger, c_{v_2}^\dagger\} \equiv c_{v_1}^\dagger c_{v_2}^\dagger + c_{v_2}^\dagger c_{v_1}^\dagger = 0. \quad (1.63)$$

Similarly, two annihilation operators anticommute,

$$\{c_{v_1}, c_{v_2}\} \equiv c_{v_1} c_{v_2} + c_{v_2} c_{v_1} = 0. \quad (1.64)$$

We now prove that

$$\{c_{v_1}, c_{v_2}^\dagger\} \equiv c_{v_1} c_{v_2}^\dagger + c_{v_2}^\dagger c_{v_1} = \delta_{v_1, v_2}. \quad (1.65)$$

Case 1: $v_1 = v_2 = v$. Consider the expression

$$c_v^\dagger c_v |\vec{\mu}\rangle. \quad (1.66)$$

Case 1a: $v = \mu_j$ for some j . We get

$$c_v^\dagger c_v |\mu_1 \cdots \mu_N\rangle = c_v^\dagger (-1)^{j-1} |\mu_1 \cdots \mu_{j-1} \mu_{j+1} \cdots \mu_N\rangle = (-1)^{j-1} |\mu_j \mu_1 \cdots \mu_{j-1} \mu_{j+1} \cdots \mu_N\rangle = |\mu_1 \cdots \mu_N\rangle. \quad (1.67)$$

We also get

$$c_v c_v^\dagger |\mu_1 \cdots \mu_N\rangle = c_v |\mu_j \mu_1 \cdots \mu_{j-1} \mu_{j+1} \cdots \mu_N\rangle = 0. \quad (1.68)$$

Case 1b: $v \notin \vec{\mu}$, v is distinct from all the μ_j . In this case, $c_v |\vec{\mu}\rangle = 0$, so

$$c_v^\dagger c_v |\mu_1 \cdots \mu_N\rangle = 0. \quad (1.69)$$

On the other hand,

$$c_v c_v^\dagger |\vec{\mu}\rangle = c_v |v \vec{\mu}\rangle = (-1)^0 |v \vec{\mu}\rangle. \quad (1.70)$$

Case 1 can be summarised as

$$\{c_v, c_v^\dagger\} = 1 \quad (1.71)$$

as desired.

Case 2: $v_1 \neq v_2$. Let $\vec{\mu}$ be arbitrary, and consider the expression

$$c_{v_1}^\dagger c_{v_2} |\vec{\mu}\rangle. \quad (1.72)$$

Case 2a: If either $v_1 \in \vec{\mu}$ or $v_2 \notin \vec{\mu}$, the expression vanishes. Similarly,

$$c_{v_2} c_{v_1}^\dagger |\vec{\mu}\rangle = 0. \quad (1.73)$$

Case 2b: $v_1 \notin \vec{\mu}$ and $v_2 \in \vec{\mu}$ ($v_2 = \mu_j$):

$$c_{v_1}^\dagger c_{v_2} |\vec{\mu}\rangle = (-1)^{j-1} c_{v_1}^\dagger |\mu_1 \cdots \mu_{j-1} \mu_{j+1} \cdots \mu_N\rangle = |\mu_1 \cdots \mu_{j-1} v_1 \mu_{j+1} \cdots \mu_N\rangle, \quad (1.74)$$

i.e., $\mu_j = v_2$ is replaced by v_1 . On the other hand,

$$c_{v_2} c_{v_1}^\dagger |\vec{\mu}\rangle = c_{v_2} |v_1 \mu_1 \cdots \mu_{j-1} \mu_j \mu_{j+1} \cdots \mu_N\rangle = (-1)^j |v_1 \mu_1 \cdots \mu_{j-1} \mu_{j+1} \cdots \mu_N\rangle = (-1) |\mu_1 \cdots \mu_{j-1} v_1 \mu_{j+1} \cdots \mu_N\rangle. \quad (1.75)$$

Summing, we see that Eq. (??) is proven in general.

1.2.3 Occupation number representation

Consider a given single-particle basis $\{\phi_\mu\}$ and the corresponding basis of Slater determinants $|\vec{\mu}\rangle$.

Given $\vec{\mu}$, we have $N!$ rearrangements of the indices. All of the rearrangements give rise to the same Slater determinant, up to the sign of the permutation. If $\sigma \in S_N$ is the permutation that sorts $\vec{\mu}$ into $\vec{v} = \sigma(\vec{\mu})$, then

$$|\mu_1, \dots, \mu_N\rangle = (-1)^{|\sigma|} |\mu_{\sigma(1)}, \dots, \mu_{\sigma(N)}\rangle = (-1)^{|\sigma|} |v_1, \dots, v_N\rangle. \quad (1.76)$$

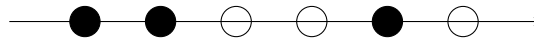
So, the *basis* of Slater determinants can be chosen as those indexed by *sorted* indices.

A sorted set of N indices $\vec{\mu}$ is in 1-1 correspondence with a subset of integers, or equivalently, by a picture of filled/unfilled circles, or *occupied and unoccupied sites*. One may say that the single-particle function ϕ_{μ_j} is *occupied* in $|\vec{\mu}\rangle$, while $v \notin \vec{\mu}$ is *unoccupied*.

A common name for “single-particle function” in chemistry is “orbital”, or “spin-orbital”. We sometimes use the word “orbital” for “single-particle function”.

One can also consider $\vec{\mu}$ as a *binary number* with N bits set: bit number v is set if $v \in \vec{\mu}$, i.e., $v = \mu_j$ for some $j = 1, \dots, N$.

Thus, the Slater determinant $|\mu_1 \mu_2 \mu_3\rangle = |0, 1, 4\rangle$ can be represented by the subset $\{\mu_1, \mu_2, \mu_3\} = \{0, 1, 4\}$, the picture



or the binary number

$$B = 2^{\mu_1} + 2^{\mu_2} + 2^{\mu_3} = 2^0 + 2^1 + 2^4 = 11001_2 = 19. \quad (1.77)$$

The different bits are called *occupation numbers*. The vacuum has no occupied single-particle functions, and is represented by the binary number 0 or the empty set.

We use the notation

$$|n_0 n_1 \cdots n_\mu \cdots\rangle$$

to denote the Slater determinant with occupation numbers $n_\mu \in \{0, 1\}$, and *by definition* we choose the one determinant out of the $N!$ possible that has $\vec{\mu}$ sorted: $\mu_1 < \mu_2 < \cdots < \mu_N$. We have $n_\mu = 1$ if and only if $\mu \in \vec{\mu}$. In the above example,

$$|0, 1, 4\rangle = |1_0 1_1 0_2 0_3 1_4\rangle. \quad (1.78)$$

If no ambiguity can arise, we simply write

$$|11001\rangle$$

etc.

Again, we stress that occupation numbers only represent 1 of the $N!$ Slater determinants possible to construct with μ_1 through μ_N , namely the one where all are sorted. But they still form a basis. In the example,

$$|11001\rangle = |0, 1, 4\rangle = -|1, 0, 4\rangle = -|4, 1, 0\rangle = -|0, 4, 1\rangle = +|1, 4, 0\rangle = +|4, 0, 1\rangle, \quad (1.79)$$

exhausting all possibilities of $N! = 3! = 6$ permutations. All these determinants are clearly linearly dependent.

The following definition can be useful:

Let $\vec{\mu}$ be an index set, and let ν be an arbitrary index. Then $\# \nu$ is the number of μ_j that satisfies $\mu_j < \nu$. Thus $\# \nu$ counts the occupied single-particle functions “before” ν .

Exercise 1.11. Let $\vec{\mu} = \{\mu_1, \dots, \mu_N\}$ be a given set of occupied orbitals, with occupation number representation

$$|n_0 n_1 n_2 \dots\rangle$$

Show that:

$$c_v^\dagger |n_0 n_1 n_2 \dots\rangle = \begin{cases} 0 & \text{if } \nu \text{ is occupied} \\ (-1)^{\# \nu} |n_1 n_2 \dots n_{v-1} 0_\nu n_{v+1} \dots\rangle & \text{if } \nu \text{ is unoccupied} \end{cases} \quad (1.80)$$

$$c_\nu |n_0 n_1 n_2 \dots\rangle = \begin{cases} 0 & \text{if } \nu \text{ is unoccupied} \\ (-1)^{\# \nu} |n_1 n_2 \dots n_{v-1} 1_\nu n_{v+1} \dots\rangle & \text{if } \nu \text{ is occupied} \end{cases} \quad (1.81)$$

△

Exercise 1.12. Let ϕ_i , $i = 1, 2, 3$ be three orthonormal single-particle functions. Consider the determinants $|1, 2, 3\rangle$, $|1, 3, 2\rangle$, $|2, 1, 3\rangle$, $|3, 2, 1\rangle$, $|2, 3, 1\rangle$ and $|3, 1, 2\rangle$.

a) Are there further Slater determinants that can be created using the single-particle orbitals ϕ_i , $i = 1, 2, 3$ only?

b) Write down a basis for the space spanned by the six determinants, i.e., a basis for all the vectors on the form

$$|\Psi\rangle = a_1 |1, 2, 3\rangle + a_2 |1, 3, 2\rangle + a_3 |2, 1, 3\rangle + a_4 |3, 2, 1\rangle + a_5 |2, 3, 1\rangle + a_6 |3, 1, 2\rangle.$$

(Here, a_i are complex numbers.)

△

1.2.4 Spin orbitals and orbital diagrams

Consider a system of electrons. Configuration space is X^N for N electrons, and for a single electron $X = \mathbb{R}^d \times \{+1, -1\}$, so

$$L^2(X) = L^2(\mathbb{R}^3) \otimes \mathbb{C}^2.$$

This means that each $\psi \in L^2(X)$ is a two-component function, one for spin-up and one for spin-down.

The notation for spin can vary. Here, we use +1 for spin up and -1 for spin down, along the z -axis. This is arbitrary, of course. In chemistry, one often uses α for spin up, and β for spin down, as symbols. (This is the notation in Szabo and Ostlund, for instance.) Sometimes one uses arrows \uparrow and \downarrow , or $+\frac{1}{2}$ and $-\frac{1}{2}$.

If $\{\varphi_p(\vec{r})\}$ is an orthonormal basis for $L^2(\mathbb{R}^3)$, the space part, and $\chi_{+1}(\sigma)$ and $\chi_{-1}(\sigma)$ are basis functions for \mathbb{C}^2 , $\sigma \in \{+1, -1\}$, the spin space, we have a basis for $L^2(X)$ via tensor products:

$$\phi_\mu(x) = \phi_{p,\alpha}(\vec{r}, \sigma) = \varphi_p(\vec{r})\chi_\alpha(\sigma).$$

Typically, one chooses

$$\chi_{+1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \chi_{-1} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

That is,

$$\chi_{+1}(+1) = 1, \quad \chi_{+1}(-1) = 0, \quad \chi_{-1}(+1) = 0, \quad \chi_{-1}(-1) = 1.$$

Or, yet another formula,

$$\chi_\alpha(\sigma) = \delta_{\alpha,\sigma}.$$

The operators S_x , S_y and S_z act on spin degrees of freedom only, and their matrices are given by the Pauli matrices:

$$\langle \sigma | S_k | \chi_\alpha \rangle = \frac{1}{2} \hbar \sigma_{k,\alpha\sigma}.$$

Thus,

$$\langle \sigma | S_x | \chi_\alpha \rangle = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (1.82)$$

$$\langle \sigma | S_y | \chi_\alpha \rangle = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad (1.83)$$

$$\langle \sigma | S_z | \chi_\alpha \rangle = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \hbar \alpha \delta_{\alpha\sigma} \quad (1.84)$$

The N -body spin operator is

$$\hat{S}_k = \sum_{i=1}^N S_k(i) \quad (1.85)$$

where $S_k(i)$ acts only on the spin of particle i .

Suppose the Hamiltonian of the electronic system is *independent of spin*, i.e., the Hamiltonian acts only on the degrees of freedom \vec{r}_i , and not σ_i for each particle. Then,

$$[\hat{H}, \hat{S}_z] = 0$$

and we can find a common set of eigenvectors for \hat{H} and \hat{S}_z .

Consider the one-body part \hat{H}_0 of the Hamiltonian, which now is a purely spatial operator:

$$\hat{H}_0 = \sum_i \hat{h}(\vec{r}_i). \quad (1.86)$$

Let \hat{h} , an operator on the space $L^2(\mathbb{R}^3)$, have a complete set of eigenfunctions, $\varphi_p(\vec{r})$,

$$\hat{h}\varphi_p(\vec{r}) = \epsilon_p \varphi_p(\vec{r}).$$

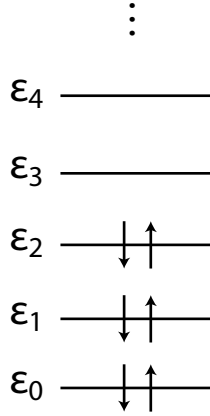


Figure 1.1: Illustration of spin-orbitals and a Slater determinant of 6 electrons

Then, as operator on $L^2(X)$, we have the complete set $\phi_\mu(\vec{r}, \sigma) = \varphi_p(\vec{r})\chi_\alpha(\sigma)$, and we see that the single-particle functions are doubly degenerate:

$$\hat{h}\phi_{p,\alpha}(x) = \epsilon_p\phi_{p,\alpha}(x), \quad \alpha \in \{+1, -1\}.$$

Here, $\mu = (p, \alpha)$ is the combined space/spin quantum numbers.

In chemistry parlance, $\varphi_p(\vec{r})$ is an *orbital*, while $\phi_\mu(x)$ is a *spin-orbital*.

Each spin-orbital can be occupied by only one electron, but each orbital has room for two – one spin up and one spin down. One typically illustrates the eigenfunctions and the occupations of Slater determinants via a diagram like Figure ???. In the figure, six spin-orbitals are occupied, and three orbitals are doubly occupied. The illustrated state is

$$c_{0,+1}^\dagger c_{0,-1}^\dagger c_{1,+1}^\dagger c_{1,-1}^\dagger c_{2,+1}^\dagger c_{2,-1}^\dagger |-\rangle. \quad (1.87)$$

This is the $N = 6$ -electron ground-state wavefunction of \hat{H}_0 .

1.2.5 Fock space

The space L_N^2 has the basis consisting of the Slater determinants $|n_0 n_1 n_2 n_3 \dots\rangle$ with in total N occupied orbitals, or N bits set in the binary representation. The creation operator c_μ^\dagger inserts a bit in position μ if it is zero, and gives the zero vector if it was already 1. Similarly, the operator c_μ turns a bit off, see Exercise ??

It is natural to consider *Fock space*, the direct sum of all L_N^2 :

$$\mathcal{F} = \bigoplus_{N=0}^{\infty} L_N^2. \quad (1.88)$$

By definition, $\langle \vec{\mu} | \vec{\nu} \rangle = 0$ if $\vec{\nu}$ and $\vec{\mu}$ have different number of particles, i.e., different number of occupied orbitals.

Thus

$$\langle 001000 \dots | 110001010 \dots \rangle = 0 \quad (1.89)$$

for example.

Now, $c_\mu^\dagger : \mathcal{F} \rightarrow \mathcal{F}$ maps entirely inside \mathcal{F} , and similarly with c_μ .

A basis for \mathcal{F} is the set of all $|n_0 n_1 \dots\rangle$ with arbitrary number of orbital occupied.

The binary number representation is quite useful for computer programs involving Slater determinants, as easily can be imagined.

A special operator, the number operator: Let v be arbitrary. We have that

$$c_v |n_0 n_1 \dots\rangle = (-1)^{\#v} n_v |n_0 n_1 \dots 0_v \dots\rangle, \quad (1.90)$$

and furthermore that

$$c_v^\dagger c_v |n_0 n_1 \dots\rangle = (-1)^{2\#v} n_v |n_0 n_1 \dots\rangle. \quad (1.91)$$

Thus,

$$\sum_v c_v^\dagger c_v |n_0 n_1 \dots\rangle = \sum_v n_v |n_0 n_1 \dots\rangle = N |n_0 n_1 \dots\rangle. \quad (1.92)$$

Therefore, we define

$$\hat{N} \equiv \sum_v c_v^\dagger c_v. \quad (1.93)$$

This operator extracts the number of fermions in a state $|\Psi\rangle$ in the sense that for any $|\Psi\rangle \in \mathcal{F}$, $\hat{N} |\Psi\rangle = N |\Psi\rangle$ if and only if $|\Psi\rangle \in L_N^2$.

1.2.6 Truncated bases

For “physical” particles, the Hilbert space is infinite dimensional. But, as we have seen in exercises, especially Exercise ??, we can select *a few* single-particle functions ϕ_μ , and construct Slater determinants out of these. These will be finite in number.

From a mathematical perspective, we can consider these finite single-particle functions to define a single-particle space on their own:

$$V_1 = \text{span}\{\phi_1, \dots, \phi_L\} \subset L^2(X). \quad (1.94)$$

Thus, $\psi \in V_1$ means

$$\psi(x) = \sum_{\mu=1}^L \psi_\mu \phi_\mu(x).$$

Having selected the finite basis, we obtain for different N a Slater determinant basis, spanning $V_N \subset L^2(X^N)_{\text{AS}}$.

Clearly, as we have only L orbitals available, we cannot create more than N particles from vacuum without getting at least one repeated creation operator, i.e., we must have $L \geq N$ to have nonzero dimension. The general dimension is $\dim(V_N) = \binom{L}{N}$.

In computational settings, the truncation of the infinite basis into a finite one is almost universally done. Of, course, we can only numerically diagonalize a finite matrix! But we would still like the basis to be as large as possible to achieve the greatest accuracy. At least intuitively, we expect that as we include more and more single-particle functions, the numerical results will approach the exact result. Under mild assumptions on the basis set and the Hamiltonian under consideration, this is in fact true.

Sometimes, the finite truncation is done after a detailed consideration of the *physics* of the system. This can give considerable physical insight, giving great explanatory power to the second quantized picture. As an example, take the physical explanation of the principles of a *laser*. (See for example https://en.wikipedia.org/wiki/Population_inversion.) Another example is the *Hubbard model* from solid-state physics, see for example https://en.wikipedia.org/wiki/Hubbard_model.

Exercise 1.13. Let ϕ_μ , $\mu = 1, 2, \dots, 6$ be given orthonormal orbitals.

a) Using the $|\mu_1, \dots, \mu_N\rangle$ notation, write down a basis for the finite dimensional subspace of $L^2(X^N)_{\text{AS}}$ for $N = 2$, $N = 3$ and $N = 4$, that you can construct using the given orbitals. (Make sure you include only linearly independent Slater determinants.)

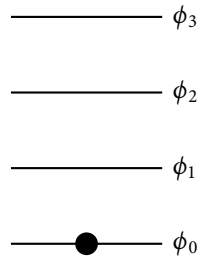
b) Can you construct a Slater determinant for $N = 10$ particles using the given ϕ_μ ?

c) Using the occupation number notation $|n_1 n_2 \dots n_6\rangle$ notation, write down a basis for the same spaces as in exercise a).

d) What is the dimension of the subspace of Fock space you can create with the 6 orbitals?

e) Assume that you have L orbitals instead of just 6. What is the dimension of the N -particle spaces you can build? What is the dimension of the Fock space you can build? \triangle

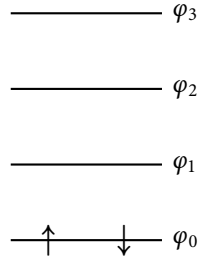
Exercise 1.14. Consider the following picture:



We have four horizontal lines, each representing a single-particle function ϕ_μ . The circle represents an occupied single-particle function, i.e., the Slater determinant $|0\rangle$.

a) Inspired by the above picture, draw a pictures of *all* the distinct Slater determinants you can create using the four single-particle functions. Caption each picture with the corresponding $|\mu_1, \mu_2, \dots\rangle$.

We now consider electrons. Consider *4 orbitals with spin* $\varphi_p(\vec{r})$, i.e., 8 spin-orbitals $\phi_\mu(\vec{r}, \sigma)$. The corresponding diagram for the Slater determinant $|0 \uparrow, 0 \downarrow\rangle$ is:



Each level now can hold 2 electrons, spin up and spin down.

b) Draw all possible 2-electron Slater determinants. Mark those that have total spin projection 0.

c) Consider the one-body operator given by

$$\hat{H}_0 = \sum_p \epsilon_p (c_{p\uparrow}^\dagger c_{p\uparrow} + c_{p\downarrow}^\dagger c_{p\downarrow}).$$

Here, ϵ_p are numbers such that $\epsilon_1 < \epsilon_2 < \dots$. In first quantization,

$$\hat{H}_0 = \sum_{i=1}^N \hat{h}(i).$$

Write down the matrix of \hat{h} in the basis $\phi_{p\sigma}$ and find its eigenfunctions. Interpret the spin-orbital diagram in terms of your results. Find the $N = 2$ ground state of \hat{H}_0 , and draw a picture of it. \triangle

1.3 Representation of operators

1.3.1 What we will prove

In this section, we shall demonstrate the following representation of one-body operators:

$$\hat{H}_0 = \sum_{i=1}^N \hat{h}(i) = \sum_{\mu\nu} \langle \mu | \hat{h} | \nu \rangle c_\mu^\dagger c_\nu. \quad (1.95)$$

Note that the last expression *does not contain N explicitly*. Here, note that $|\mu\rangle$ is a single-particle function – it is the “Slater determinant” $\phi_\mu(x_1)$. The number $\langle \mu | \hat{h} | \nu \rangle$ is the matrix element of the single-particle operator \hat{h} in the given one-particle basis,

$$\langle \mu | \hat{h} | \nu \rangle = \int dx \phi_\mu(x)^* \hat{h} \phi_\nu(x). \quad (1.96)$$

Eq. (??) gives a nice image of how \hat{H}_0 acts on a basis function: each term in the sum manipulates the Slater determinant’s occupied orbitals and weighs it with a matrix element. Simple, and not at all obvious from the “single quantized form”.

We shall also prove the following formula for the two-body operator:

$$\hat{W} = \sum_{(i,j)} \hat{w}(i,j) = \frac{1}{2} \sum_{\mu\nu\alpha\beta} w_{\mu\nu}^{\alpha\beta} c_\mu^\dagger c_\nu^\dagger c_\beta c_\alpha, \quad (1.97)$$

where the ordering of the annihilation operators should be noted. Here,

$$w_{\alpha\beta}^{\mu\nu} = \int dx_1 \int dx_2 \phi_\mu(x_1)^* \phi_\nu(x_2)^* w(x_1, x_2) \phi_\alpha(x_1) \phi_\beta(x_2) \quad (1.98)$$

is a matrix element using tensor product two-body functions, *not* Slater determinants. Using Slater determinant matrix elements we in fact have a similar expansion,

$$\hat{W} = \frac{1}{4} \sum_{\mu\nu\alpha\beta} \langle \mu\nu | \hat{w} | \alpha\beta \rangle c_\mu^\dagger c_\nu^\dagger c_\beta c_\alpha, \quad (1.99)$$

where thus the matrix elements *are antisymmetric*, computed as a matrix element using two-body Slater determinants.

A word of warning: notation for two-body matrix elements is notoriously varying between sources. Some authors use the notation $\langle \phi_\alpha \phi_\beta | \hat{w} | \phi_\mu \phi_\nu \rangle$ for the matrix element $w_{\mu\nu}^{\alpha\beta}$, which is *not* antisymmetric. In our case, the notation clashes with the Slater determinant matrix element, but we will still sin in this respect. Some authors write $\langle \phi_\alpha \phi_\beta | \hat{w} | \phi_\mu \phi_\nu \rangle_{AS}$ for the anti-symmetric Slater-determinant matrix element (and sometimes we will too), which is equal to:

$$\langle \phi_\alpha \phi_\beta | \hat{w} | \phi_\mu \phi_\nu \rangle_{AS} = \langle \alpha\beta | \hat{w} | \mu\nu \rangle = w_{\mu\nu}^{\alpha\beta} - w_{\nu\mu}^{\alpha\beta}. \quad (1.100)$$

This can cause some confusion, as the expansions using tensor products and Slater determinants differ by a factor 2...

The proofs given in this section borrow heavily from [?].

Lemma 1.1. Let $|\mu_1\mu_2\cdots\mu_N\rangle$ be a Slater determinant built from orthonormal single-particle functions ϕ_μ , no particular ordering assumed. The operator $c_v^\dagger c_\alpha$ replaces ϕ_α with ϕ_v (or gives zero if α is not present in $\vec{\mu}$), with no sign change.

Similarly, $c_{v_1}^\dagger c_{v_2}^\dagger c_{\alpha_2} c_{\alpha_1}$ replaces α_1 with v_1 , and α_2 with v_2 , or gives zero if one of α_1 or α_2 is not present in $\vec{\mu}$.

Exercise 1.15. Prove the lemma. \triangle

1.3.2 One-body operators

We prove Eq. (??) by showing that their action on a basis of Slater determinants agree. Let therefore $\{\phi_\mu\}$ be a single-particle basis as usual.

Consider the action of $\hat{H}_0 = \sum_i \hat{h}(i)$ on an arbitrary Slater determinant:

$$\begin{aligned} \hat{H}_0 |\phi_{\mu_1}, \phi_{\mu_2}, \dots, \phi_{\mu_N}\rangle &= \frac{1}{\sqrt{N!}} \left(\sum_i \hat{h}(i) \right) \sum_\sigma (-1)^{|\sigma|} \hat{P}_\sigma \phi_{v_1}(x_1) \cdots \phi_{v_N}(x_N) \\ &= \frac{1}{\sqrt{N!}} \sum_\sigma (-1)^{|\sigma|} \hat{P}_\sigma \left(\sum_i \hat{h}(i) \right) \phi_{v_1}(x_1) \cdots \phi_{v_N}(x_N) \\ &= |(\hat{h}\phi_{v_1}), \phi_{v_2}, \dots, \phi_{v_N}\rangle + |\phi_{v_1}, (\hat{h}\phi_{v_2}), \dots, \phi_{v_N}\rangle + \cdots + |\phi_{v_1}, \phi_{v_2}, \dots, (\hat{h}\phi_{v_N})\rangle \end{aligned} \quad (1.101)$$

Here, we used that \hat{P}_σ commutes with \hat{H}_0 .

Consider the operator \hat{h} acting on a single-particle function ϕ_μ . The result, ψ , can be expanded in the basis:

$$\psi(x) = \hat{h}\phi_\mu(x) = \sum_v \phi_v \langle v | \hat{h} | \mu \rangle. \quad (1.102)$$

We insert this expansion:

$$\begin{aligned} \hat{H}_0 |\phi_{\mu_1}, \phi_{\mu_2}, \dots, \phi_{\mu_N}\rangle &= |(\hat{h}\phi_{v_1}), \phi_{v_2}, \dots, \phi_{v_N}\rangle + |\phi_{v_1}, (\hat{h}\phi_{v_2}), \dots, \phi_{v_N}\rangle + \cdots + |\phi_{v_1}, \phi_{v_2}, \dots, (\hat{h}\phi_{v_N})\rangle \\ &= \sum_v \langle v | \hat{h} | \mu_1 \rangle |\nu\mu_2, \dots, \mu_N\rangle + \sum_v \langle v | \hat{h} | \mu_2 \rangle |\mu_1\nu\mu_3, \dots, \mu_N\rangle + \cdots + \sum_v \langle v | \hat{h} | \mu_N \rangle |\mu_1\mu_2, \dots, \nu\rangle \end{aligned} \quad (1.103)$$

Now, we note that

$$|\mu_1, \dots, \mu_{j-1}\nu\mu_{j+1}, \dots, \mu_N\rangle = c_v^\dagger c_{\mu_j} |\mu_1, \dots, \mu_N\rangle, \quad (1.104)$$

which we plug in:

$$\begin{aligned} \hat{H}_0 |\phi_{\mu_1}, \phi_{\mu_2}, \dots, \phi_{\mu_N}\rangle &= \sum_v \langle v | \hat{h} | \mu_1 \rangle |\nu\mu_2, \dots, \mu_N\rangle + \sum_v \langle v | \hat{h} | \mu_2 \rangle |\mu_1\nu\mu_3, \dots, \mu_N\rangle + \cdots + \sum_v \langle v | \hat{h} | \mu_N \rangle |\mu_1\mu_2, \dots, \nu\rangle \\ &= \left[\sum_v \langle v | \hat{h} | \mu_1 \rangle c_v^\dagger c_{\mu_1} + \sum_v \langle v | \hat{h} | \mu_2 \rangle c_v^\dagger c_{\mu_2} + \cdots + \sum_v \langle v | \hat{h} | \mu_N \rangle c_v^\dagger c_{\mu_N} \right] |\mu_1, \dots, \mu_N\rangle. \end{aligned} \quad (1.105)$$

Finally, we note that $c_\mu |\mu_1 \cdots \mu_N\rangle = 0$ whenever $\mu \notin \vec{\mu}$, so we may extend the summation over μ_j to all of μ , resulting in:

$$\begin{aligned} \hat{H}_0 |\phi_{\mu_1}, \phi_{\mu_2}, \dots, \phi_{\mu_N}\rangle &= \left[\sum_v \langle v | \hat{h} | \mu_1 \rangle c_v^\dagger c_{\mu_1} + \sum_v \langle v | \hat{h} | \mu_2 \rangle c_v^\dagger c_{\mu_2} + \cdots + \sum_v \langle v | \hat{h} | \mu_N \rangle c_v^\dagger c_{\mu_N} \right] |\mu_1, \dots, \mu_N\rangle \\ &= \sum_{\mu v} \langle v | \hat{h} | \mu \rangle c_v^\dagger c_\mu |\mu_1, \dots, \mu_N\rangle. \end{aligned} \quad (1.106)$$

Since $|\mu_1, \dots, \mu_N\rangle$ was an arbitrary Slater determinant, we have proven Eq. (??).

1.3.3 Two-body operators

The operator $\hat{W} = \sum_{i < j} \hat{w}(i, j)$ is a two-body operator. The operator $\hat{w}(1, 2)$ is thus an operator on $L^2(X^2)$ that is completely characterized by its action on a basis: the tensor products $\phi_{\mu_1}(x_1)\phi_{\mu_2}(x_2)$. Thus,

$$\hat{w}(1, 2)\phi_{\mu_1}(x_1)\phi_{\mu_2}(x_2) = \sum_{v_1 v_2} w_{\mu_1 \mu_2}^{v_1 v_2} \phi_{v_1}(x_1)\phi_{v_2}(x_2), \quad (1.107)$$

where the matrix elements $w_{\mu_1 \mu_2}^{v_1 v_2}$ are given by the formula (??). There is nothing special about the indices $(1, 2)$, it may just as well be (i, j) . Note also the symmetry property

$$w_{\mu_1 \mu_2}^{v_1 v_2} = w_{\mu_2 \mu_1}^{v_2 v_1}.$$

As for the one-body case, \hat{W} commutes with \hat{P}_σ , and we get, using Eq. (??),

$$\begin{aligned} \hat{W} |\phi_{\mu_1} \dots \phi_{\mu_N}\rangle &= \frac{1}{\sqrt{N!}} \sum_{\sigma} (-1)^{|\sigma|} \hat{P}_{\sigma} \left[\sum_{i < j} \hat{w}(i, j) \phi_{\mu_1}(x_1) \dots \phi_{\mu_N}(x_N) \right] \\ &= \frac{1}{\sqrt{N!}} \sum_{\sigma} (-1)^{|\sigma|} \hat{P}_{\sigma} \left[\sum_{i < j} \hat{w}(i, j) \phi_{\mu_1}(x_1) \dots \phi_{\mu_N}(x_N) \right] \\ &= \frac{1}{\sqrt{N!}} \sum_{\sigma} (-1)^{|\sigma|} \hat{P}_{\sigma} \left[\sum_{i < j} \sum_{v_1 v_2} w_{\mu_i \mu_j}^{v_1 v_2} \phi_{\mu_1} \dots \phi_{v_1}(x_i) \dots \phi_{v_2}(x_j) \dots \phi_{\mu_N}(x_N) \right] \\ &= \sum_{i < j} \sum_{v_1 v_2} w_{\mu_i \mu_j}^{v_1 v_2} |\phi_{\mu_1} \dots \phi_{v_1} \dots \phi_{v_2} \dots \phi_{\mu_N}\rangle \\ &= \sum_{i < j} \sum_{v_1 v_2} w_{\mu_i \mu_j}^{v_1 v_2} c_{v_1}^{\dagger} c_{v_2}^{\dagger} c_{\mu_j} c_{\mu_i} |\phi_{\mu_1} \dots \phi_{\mu_N}\rangle. \end{aligned} \quad (1.108)$$

Here, we used the Lemma about replacement behaviour of the $c^{\dagger} c^{\dagger} c c$ product. We are currently summing over μ_i and μ_j , such that $i < j$. Including $i = j$ gives zero contribution (why?), and including $j > i$ gives equal contribution:

$$\begin{aligned} \sum_{i < j} \sum_{v_1 v_2} w_{\mu_i \mu_j}^{v_1 v_2} c_{v_1}^{\dagger} c_{v_2}^{\dagger} c_{\mu_j} c_{\mu_i} |\phi_{\mu_1} \dots \phi_{\mu_N}\rangle &= - \sum_{i < j} \sum_{v_1 v_2} w_{\mu_i \mu_j}^{v_1 v_2} c_{v_1}^{\dagger} c_{v_2}^{\dagger} c_{\mu_i} c_{\mu_j} |\phi_{\mu_1} \dots \phi_{\mu_N}\rangle \\ &= \sum_{i < j} \sum_{v_1 v_2} w_{\mu_i \mu_j}^{v_1 v_2} c_{v_2}^{\dagger} c_{v_1}^{\dagger} c_{\mu_i} c_{\mu_j} |\phi_{\mu_1} \dots \phi_{\mu_N}\rangle = \sum_{j < i} \sum_{v_1 v_2} w_{\mu_j \mu_i}^{v_1 v_2} c_{v_2}^{\dagger} c_{v_1}^{\dagger} c_{\mu_j} c_{\mu_i} |\phi_{\mu_1} \dots \phi_{\mu_N}\rangle \\ &= \sum_{j < i} \sum_{v_2 v_1} w_{\mu_j \mu_i}^{v_2 v_1} c_{v_1}^{\dagger} c_{v_2}^{\dagger} c_{\mu_j} c_{\mu_i} |\phi_{\mu_1} \dots \phi_{\mu_N}\rangle = \sum_{j < i} \sum_{v_2 v_1} w_{\mu_i \mu_j}^{v_1 v_2} c_{v_1}^{\dagger} c_{v_2}^{\dagger} c_{\mu_j} c_{\mu_i} |\phi_{\mu_1} \dots \phi_{\mu_N}\rangle \end{aligned} \quad (1.109)$$

Here, we used the anticommutators and symmetry of the matrix elements. Assembling the two contributions,

$$\hat{W} |\phi_{\mu_1} \dots \phi_{\mu_N}\rangle = \frac{1}{2} \sum_{ij} \sum_{v_2 v_1} w_{\mu_i \mu_j}^{v_1 v_2} c_{v_1}^{\dagger} c_{v_2}^{\dagger} c_{\mu_j} c_{\mu_i} |\phi_{\mu_1} \dots \phi_{\mu_N}\rangle. \quad (1.110)$$

We note that the sum over ij is really a sum over two occupied orbitals μ_i and μ_j . We can therefore extend the sum to all unoccupied orbitals as well, since $c_{\alpha} |\vec{\mu}\rangle$ gives zero contributions for such orbitals. Thus, Eq. (??) is proven.

We leave the proof of the antisymmetrized version as an exercise.

Exercise 1.16. Prove Eq. (??). Start with showing Eq. (??).

△

Exercise 1.17. a) Let $\hat{F} = \sum_{i=1}^N \hat{f}(i)$ be a first-quantization operator. Write down the second-quantized form of this operator. Let $\hat{G} = \sum_{i<j} \hat{g}(i, j)$ be a general two-body operator, where $\hat{g}(1, 2) = \hat{g}(2, 1)$. Write down the second-quantized form.

b) Using the fundamental anticommutator relations, compute the matrix element

$$\langle \mu_1 \mu_2 | \hat{F} | \mu_1 \mu_2 \rangle$$

c) Using the fundamental anticommutator relations, compute the matrix element

$$\langle \mu_1 \mu_2 \mu_3 | \hat{F} | \mu_1 \mu_2 \mu_3 \rangle$$

d) Using the fundamental anticommutator relations, compute the matrix element

$$\langle \mu_1 \mu_2 | \hat{G} | \mu_1 \mu_2 \rangle$$

e) Using the fundamental anticommutator relations, compute the matrix element

$$\langle \mu_1 \mu_2 \mu_3 | \hat{G} | \mu_1 \mu_2 \mu_3 \rangle$$

f) Compute the matrix element

$$\langle \mu_1, \mu_2, \dots, \mu_N | \hat{F} | \mu_1, \mu_2, \dots, \mu_N \rangle$$

g) Compute the matrix element

$$\langle \mu_1, \mu_2, \dots, \mu_N | \hat{G} | \mu_1, \mu_2, \dots, \mu_N \rangle$$

△

Exercise 1.18. (Tedious, but very instructive.) In this exercise, we prove the so-called *Slater–Condon rules*: the explicit expressions of matrix elements of one- and two-body operators in a Slater determinant basis.

We do not assume any particular ordering of the occupied single-particle functions considered.

If you solved Exercise ??, you solved parts of this exercise.

a) Using the fundamental anticommutator relations, compute $\langle \vec{\mu} | \hat{H}_0 | \vec{\mu} \rangle$ and $\langle \vec{\mu} | \hat{W} | \vec{\mu} \rangle$ and prove that

$$\langle \vec{\mu} | \hat{H}_0 | \vec{\mu} \rangle = \sum_{i=1}^N h_{\mu_i}^{\mu_i}, \quad (1.111)$$

$$\langle \vec{\mu} | \hat{W} | \vec{\mu} \rangle = \sum_{i<j}^N \langle \mu_i \mu_j | \hat{w} | \mu_i \mu_j \rangle_{AS} = \frac{1}{2} \sum_{ij} \langle \mu_i \mu_j | \hat{w} | \mu_i \mu_j \rangle_{AS}. \quad (1.112)$$

b) Let \vec{v} be equal to $\vec{\mu}$, except for one occupied orbital, i.e.,

$$|\vec{v}\rangle = c_{v_j}^\dagger c_{\mu_j} |\vec{\mu}\rangle, \quad v_j \neq \mu_j. \quad (1.113)$$

Using the fundamental anticommutator relations, compute $\langle \vec{\mu} | \hat{H}_0 | \vec{\mu} \rangle$ and $\langle \vec{\mu} | \hat{W} | \vec{\nu} \rangle$, and find

$$\langle \vec{\mu} | \hat{H}_0 | \vec{\nu} \rangle = h_{v_j}^{\mu_j}, \quad (1.114)$$

$$\langle \vec{\mu} | \hat{W} | \vec{\nu} \rangle = \sum_i \langle \mu_i \mu_j | \hat{w} | \mu_i v_j \rangle_{AS}. \quad (1.115)$$

c) Let $\vec{\nu}$ be equal to $\vec{\mu}$, except for two indices, i.e.,

$$|\vec{\nu}\rangle = c_{v_k}^\dagger c_{v_j}^\dagger c_{\mu_j} c_{\mu_k} |\vec{\mu}\rangle, \quad j \neq k. \quad (1.116)$$

Using the fundamental anticommutator relations, compute $\langle \vec{\mu} | \hat{W} | \vec{\nu} \rangle$ and find

$$\langle \vec{\mu} | \hat{H}_0 | \vec{\nu} \rangle = 0, \quad (1.117)$$

$$\langle \vec{\mu} | \hat{W} | \vec{\nu} \rangle = \langle \mu_j \mu_k | \hat{w} | v_j v_k \rangle_{AS}. \quad (1.118)$$

d) Explain that if $\vec{\nu}$ differs from $\vec{\mu}$ in *more* than two occupied functions, then $\langle \vec{\mu} | \hat{W} | \vec{\nu} \rangle = 0$.

△

Exercise 1.19. (This exercise adapted from an exercise by Morten Hjorth-Jensen.)

We will now consider a simple three-level problem, depicted in Figure ???. The single-particle states are labelled by the quantum number p and can accommodate up to two single particles, viz., every single-particle state is doubly degenerate (you could think of this as one state having spin up and the other spin down). We let the spacing between the doubly degenerate single-particle states be constant, with value d . The first state has energy d . There are only three available single-particle states, $p = 1$, $p = 2$ and $p = 3$, as illustrated in the figure.

- How many two-particle Slater determinants can we construct in this space?
- We limit ourselves to a system with only the two lowest single-particle orbits and two particles, $p = 1$ and $p = 2$. We assume that we can write the Hamiltonian as

$$\hat{H} = \hat{H}_0 + \hat{H}_I,$$

and that the onebody part of the Hamiltonian with single-particle operator \hat{h}_0 has the property

$$\hat{h}_0 \psi_{p\sigma} = p \times d \psi_{p\sigma},$$

where we have added a spin quantum number σ . We assume also that the only two-particle states that can exist are those where two particles are in the same state p , as shown by the two possibilities to the left in the figure. The two-particle matrix elements of \hat{H}_I have all a constant value, $-g$. Show then that the Hamiltonian matrix can be written as

$$\begin{pmatrix} 2d - g & -g \\ -g & 4d - g \end{pmatrix},$$

and find the eigenvalues and eigenvectors. What is mixing of the state with two particles in $p = 2$ to the wave function with two-particles in $p = 1$? Discuss your results in terms of a linear combination of Slater determinants.

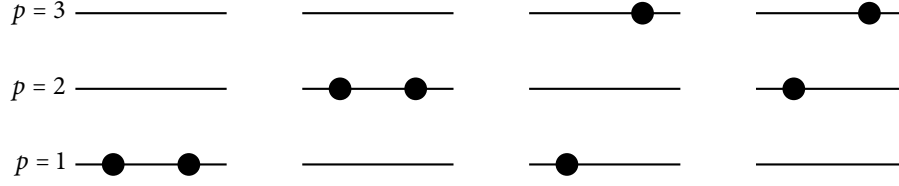


Figure 1.2: Schematic plot of the possible single-particle levels with double degeneracy. The filled circles indicate occupied particle states. The spacing between each level p is constant in this picture. We show some possible two-particle states.

- c) Add the possibility that the two particles can be in the state with $p = 3$ as well and find the Hamiltonian matrix, the eigenvalues and the eigenvectors. We still insist that we only have two-particle states composed of two particles being in the same level p . You can diagonalize numerically your 3×3 matrix.

This simple model catches several birds with a stone. It demonstrates how we can build linear combinations of Slater determinants and interpret these as different admixtures to a given state. It represents also the way we are going to interpret these contributions. The two-particle states above $p = 1$ will be interpreted as excitations from the ground state configuration, $p = 1$ here. The reliability of this ansatz for the ground state, with two particles in $p = 1$, depends on the strength of the interaction g and the single-particle spacing d . Finally, this model is a simple schematic ansatz for studies of pairing correlations and thereby superfluidity/superconductivity in fermionic systems.

△

1.4 Wick's Theorem

Supporting material for this section is Shavitt/Bartlett Ch. 3, Gross/Runge/Heinonen Ch. 19.

1.4.1 A sort of summary and motivation

Let us take a look at what we have so far. In the preceding sections, we introduced a collection of tools for describing (1) many-fermion states in many-particle Hilbert space, and (2) second-quantization language for expressing these states and, importantly, the quantum-mechanical Hamiltonian.

The quantum mechanical Hilbert space for N fermions is defined solely in terms of the *configuration space* X of a single fermion. The Hilbert space of a single particle is $L^2(X)$, the set of square integrable single-particle functions $\psi : X \rightarrow \mathbb{C}$.

Such a space always has an orthonormal basis, say $\{\phi_\mu\}$. Forming *Slater determinants* $|\phi_{\mu_1} \phi_{\mu_2} \cdots \phi_{\mu_N}\rangle$, we obtain totally antisymmetric basis functions for $L^2(X^N)_{AS}$. Furthermore, we defined *Fock space* \mathcal{F} ,

$$\mathcal{F} = \bigoplus_{N=0}^{\infty} L^2(X^N)_{AS}.$$

Inside the Fock space, *every possible wavefunction of a system of some number of fermions exist*.

Given a wavefunction $|\Psi_N\rangle \in \mathcal{F}$, it could be expanded in Slater determinants,

$$|\Psi_N\rangle = \sum_{\vec{\mu}} |\mu_1 \cdots \mu_N\rangle \langle \mu_1 \cdots \mu_N | \Psi_N \rangle. \quad (1.119)$$

Here, the subscript N only indicates that we know the number of particles. The notation $|\mu_1 \cdots \mu_N\rangle$ indicates that a certain single-particle basis $\{\phi_\mu\}$ has been chosen, since we only list the indices μ_j . Each determinant can be constructed from vacuum using *creation operators* c_μ^\dagger (these, of course, depend on the basis),

$$|\mu_1 \cdots \mu_N\rangle = c_{\mu_1}^\dagger c_{\mu_2}^\dagger \cdots c_{\mu_N}^\dagger |-\rangle. \quad (1.120)$$

Finally, we found expressions for one- and two-body operators in terms of creation and annihilation operators:

$$\hat{H}_0 = \sum_{\mu\nu} \langle \mu | \hat{h} | \nu \rangle c_\mu^\dagger c_\nu \quad (1.121)$$

$$\hat{W} = \frac{1}{4} \sum_{\substack{\nu_1 \nu_2 \\ \mu_1 \mu_2}} \langle \mu_1 \mu_2 | \hat{w} | \nu_1 \nu_2 \rangle c_{\mu_1}^\dagger c_{\mu_2}^\dagger c_{\nu_2} c_{\nu_1}. \quad (1.122)$$

We claimed that these expressions simplify our life a lot.

Our life goal in this context is to solve the (time-independent) Schrödinger equation,

$$(\hat{H}_0 + \hat{W}) |\Psi_N\rangle = E |\Psi_N\rangle. \quad (1.123)$$

This expression equates two elements (functions) in Hilbert space. These functions are equal if and only if their basis projections are equal. Thus, we expand $|\Psi_N\rangle$ in the basis, and similarly with the left-hand side $\hat{H} |\Psi_N\rangle$:

$$\sum_{\vec{\mu}} \langle \nu_1 \cdots \nu_N | (\hat{H}_0 + \hat{W}) | \mu_1 \cdots \mu_N \rangle \langle \mu_1 \cdots \mu_N | \Psi_N \rangle = E \langle \nu_1 \cdots \nu_N | \Psi_N \rangle. \quad (1.124)$$

Defining the vector $C_{\vec{\mu}} = \langle \vec{\mu} | \Psi_N \rangle$ and the matrix $H_{\vec{\nu}, \vec{\mu}} = \langle \vec{\nu} | \hat{H} | \vec{\mu} \rangle$, we see that we have a *matrix eigenvalue problem*

$$HC = EC. \quad (1.125)$$

Remark: this equation is (usually) infinite-dimensional, and from a strict mathematical point of view, this must really be carefully defined. But in this course, it is sufficient to think of this as a standard matrix eigenvalue problem.

Ok, so we have a way of describing *vectors* $|\Psi_N\rangle$ and the operators \hat{H}_0 etc. But if we actually want to *solve* Eq. (??), we need to compute the matrix elements²

$$H_{0, \vec{\nu}, \vec{\mu}} = \langle \vec{\mu} | \hat{H}_0 | \vec{\mu} \rangle = \sum_{\mu\nu} \langle \mu | \hat{h} | \nu \rangle \langle - | c_{\nu_N} c_{\nu_{N-1}} \cdots c_{\nu_1} c_\mu^\dagger c_\nu c_{\mu_1}^\dagger c_{\mu_2}^\dagger \cdots c_{\mu_N}^\dagger | - \rangle \quad (1.126)$$

and similarly

$$W_{\vec{\nu}, \vec{\mu}} = \langle \vec{\mu} | \hat{W} | \vec{\mu} \rangle = \frac{1}{4} \sum_{\substack{\alpha_1 \alpha_2 \\ \beta_1 \beta_2}} \langle \alpha_1 \alpha_2 | \hat{w} | \beta_1 \beta_2 \rangle \langle - | c_{\nu_N} c_{\nu_{N-1}} \cdots c_{\nu_1} c_{\alpha_1}^\dagger c_{\alpha_2}^\dagger c_{\beta_2} c_{\beta_1} c_{\mu_1}^\dagger c_{\mu_2}^\dagger \cdots c_{\mu_N}^\dagger | - \rangle \quad (1.127)$$

²On a computer, we need to truncate the basis to obtain a finite-dimensional matrix eigenvalue problem. Only for very small problems will one actually compute the matrix itself, because that is quite expensive. Rather, one computes the *matrix-vector product* $\hat{H} |\vec{\mu}\rangle$.

Notice that we used

$$|\mu_1 \cdots \mu_N\rangle = c_{\mu_1}^\dagger \cdots c_{\mu_N}^\dagger |-\rangle \quad (1.128)$$

and, by taking the adjoint,

$$\langle \mu_1 \cdots \mu_N | = \langle - | c_{\mu_N} \cdots c_{\mu_1}. \quad (1.129)$$

Observe that the order of the annihilation operators is the reverse of the order of the creation operators.

The number $\langle - | c^{(\dagger)} c^{(\dagger)} \cdots c^{(\dagger)} | - \rangle$ is referred to a *vacuum expectation value*, and the problem of computing matrix elements is basically reduced to computing these.

Let us consider an example, and compute a typical vacuum expectation value occuring in the \hat{H}_0 matrix element:

$$A = \langle v_1 v_2 | c_\alpha^\dagger c_\beta | \mu_1 \mu_2 \rangle = \langle - | c_{v_2} c_{v_1} c_\alpha^\dagger c_\beta c_{\mu_1}^\dagger c_{\mu_2}^\dagger | - \rangle. \quad (1.130)$$

Now, how are we going to approach this problem? Recall the anticommutation relations,

$$c_\alpha c_\beta^\dagger + c_\beta^\dagger c_\alpha = \delta_{\alpha\beta} \quad (1.131)$$

$$c_\alpha c_\beta + c_\beta c_\alpha = 0 \quad (1.132)$$

and

$$c_\alpha^\dagger c_\beta^\dagger + c_\beta^\dagger c_\alpha^\dagger = 0. \quad (1.133)$$

So, we can “flip” two creation or annihilation operators adjacent to each other and compensate with a $-$ sign. We can “flip” an annihilation and creation operator by a $-$ sign, *but* we have to “pay a price” in the form of a Kronecker delta, an additional term. However, this additional term has two less creation and annihilation operators.

In this way, we can systematically move the annihilation operators *to the right*, and the creation operators *to the left*, possibly inserting kronecker deltas and generating new terms with fewer operators. But when the annihilation operators are to the right they give zero contribution since $c_\alpha | - \rangle = 0$.

Let us see this in practice, and first remove one pair of creation and annihilation operators:

$$\begin{aligned} A &= \langle - | c_{v_2} c_{v_1} c_\alpha^\dagger c_\beta c_{\mu_1}^\dagger c_{\mu_2}^\dagger | - \rangle = \langle - | c_{v_2} c_{v_1} c_\alpha^\dagger (\delta_{\beta\mu_1} - c_{\mu_1}^\dagger c_\beta) c_{\mu_2}^\dagger | - \rangle \\ &= \delta_{\beta\mu_1} \langle - | c_{v_2} c_{v_1} c_\alpha^\dagger c_{\mu_2}^\dagger | - \rangle - \langle - | c_{v_2} c_{v_1} c_\alpha^\dagger c_{\mu_1}^\dagger c_\beta c_{\mu_2}^\dagger | - \rangle \\ &= \delta_{\beta\mu_1} \langle - | c_{v_2} c_{v_1} c_\alpha^\dagger c_{\mu_2}^\dagger | - \rangle - \langle - | c_{v_2} c_{v_1} c_\alpha^\dagger c_{\mu_1}^\dagger (\delta_{\beta\mu_2} - c_{\mu_2}^\dagger c_\beta) | - \rangle \\ &= \delta_{\beta\mu_1} \langle - | c_{v_2} c_{v_1} c_\alpha^\dagger c_{\mu_2}^\dagger | - \rangle - \delta_{\beta\mu_2} \langle - | c_{v_2} c_{v_1} c_\alpha^\dagger c_{\mu_1}^\dagger | - \rangle. \end{aligned} \quad (1.134)$$

We continue:

$$\begin{aligned} A &= \delta_{\beta\mu_1} \langle - | c_{v_2} (\delta_{v_1\alpha} - c_\alpha^\dagger c_{v_1}) c_{\mu_2}^\dagger | - \rangle - \delta_{\beta\mu_2} \langle - | c_{v_2} (\delta_{v_1\alpha} - c_\alpha^\dagger c_{v_1}) c_{\mu_1}^\dagger | - \rangle \\ &= \delta_{\beta\mu_1} \delta_{v_1\alpha} \langle - | c_{v_2} c_{\mu_2}^\dagger | - \rangle - \delta_{\beta\mu_1} \langle - | c_{v_2} c_\alpha^\dagger c_{v_1} c_{\mu_2}^\dagger | - \rangle - (\mu_1 \leftrightarrow \mu_2). \end{aligned} \quad (1.135)$$

In the last equality, we have indicated that the remaining terms are generated from the previous ones by exchanging μ_1 and μ_2 .

Continuing,

$$A = \delta_{\beta\mu_1} \delta_{v_1\alpha} \langle - | c_{v_2} c_{\mu_2}^\dagger | - \rangle - \delta_{\beta\mu_1} \delta_{v_1\mu_2} \langle - | c_{v_2} c_\alpha^\dagger | - \rangle + \delta_{\beta\mu_1} \langle - | c_{v_2} c_\alpha^\dagger c_{\mu_2}^\dagger c_{v_1} | - \rangle - (\mu_1 \leftrightarrow \mu_2). \quad (1.136)$$

Only the two first terms are non-vanishing, and we note, for example, that $\langle - | c_{v_2} c_{\mu_2}^\dagger | - \rangle = \langle v_2 | \mu_2 \rangle = \delta_{v_2\mu_2}$. (We could also use the anticommutator once more.) This gives:

$$A = \delta_{\beta\mu_1} \delta_{v_1\alpha} \delta_{v_2\mu_2} - \delta_{\beta\mu_1} \delta_{v_1\mu_2} \delta_{v_2\alpha} - (\mu_1 \leftrightarrow \mu_2). \quad (1.137)$$

Yes, our life was made easier by introducing second-quantiation. However, the matrix elements are still quite hard to compute. This is where *Wick's theorem* comes in, by giving a much quicker way of determining the vacuum expectation values.

Observe that the vacuum expectation value is *basis independent*. The value only depends on the anti-commutator relations, and these only depended on the orthonormality of $\{\phi_\mu\}$. So we see that the framework is quite general.

1.4.2 Vacuum expectation values

Consider the computation of a *vacuum expectation value* of a *string of creation and annihilation operators*:

$$\langle - | A_1 A_2 \cdots A_n | - \rangle, \quad (1.138)$$

where each of the A_i are one of the c_μ^\dagger or c_μ . For example, the overlap between two determinants is on this form:

$$\langle \mu_1 \cdots \mu_N | \nu_1 \cdots \nu_N \rangle = \langle - | c_{\mu_N} c_{\mu_{N-1}} \cdots c_{\mu_1} c_{\nu_1}^\dagger c_{\nu_2}^\dagger \cdots c_{\nu_N}^\dagger | - \rangle \quad (1.139)$$

Another example is the matrix elements of an operator on second quantized form, say \hat{H}_0 :

$$\langle \mu_1 \cdots \mu_N | \hat{H}_0 | \nu_1 \cdots \nu_N \rangle = \sum_{\mu\nu} \langle \mu | \hat{h} | \nu \rangle \langle \mu_1 \cdots \mu_N | c_\mu^\dagger c_\nu | \nu_1 \cdots \nu_N \rangle \sum_{\mu\nu} \langle \mu | \hat{h} | \nu \rangle \langle - | c_{\mu_N} \cdots c_{\mu_1} c_\nu^\dagger c_\mu^\dagger \cdots c_{\nu_1}^\dagger | - \rangle. \quad (1.140)$$

The right-hand side is a linear combination of vacuum expectation values. So we see that having a straightforward way to compute Eq. (??) would be of great help.

Wick's Theorem is what we shall need.

1.4.3 Normal ordering and contractions

In this section, we denote a general string of n creation and annihilation operators by

$$A_1 A_2 \cdots A_n, \quad A_i \in \{c_\mu\} \cup \{c_\mu^\dagger\}. \quad (1.141)$$

Our goal is to find a general procedure of computing the vacuum expectation value

$$\langle - | A_1 A_2 \cdots A_n | - \rangle. \quad (1.142)$$

Note that this expectation value only depends on the *orthogonality* of the single-particle functions, not on the functions themselves. I.e., the value of the vacuum expectation value can be computed solely from the anticommutator relations (??).

The procedure we develop is based on *Wick's Theorem*, to be stated and proven. Wick's theorem is based on two fundamental concepts, namely *normal ordering* and *contraction*. The normal-ordered product form of an operator string $A_1 A_2 \cdots A_n$ is defined as

$$N(A_1 A_2 \cdots A_n) \equiv (-1)^{|\sigma|} [\text{creation operators}] \cdot [\text{annihilation operators}] \quad (1.143)$$

Here, $\sigma \in S_n$ denotes a permutation that brings the operator product to the desired order,

$$N(A_1 A_2 \cdots A_n) = (-1)^{|\sigma|} A_{\sigma(1)} A_{\sigma(2)} \cdots A_{\sigma(n)}. \quad (1.144)$$

Note that the string $A_1 \cdots A_n$ and the normal-order product $N(A_1 \cdots A_n)$ is *not* the same operator, since by reordering creation and annihilation operators we neglect the extra terms arising from the Kronecker delta in the anti-commutator relation $\{c_\alpha, c_\beta^\dagger\} = \delta_{\alpha\beta}$. If all individual A_i in fact anticommute, *then* the string and the normal-ordered string are identical as operators, but usually this is not the case.

Remark: The permutation σ in the definition is usually not be unique, but the normal ordered product is unique as operator. Consider for example

$$N(c_1 c_2^\dagger c_3^\dagger c_4) = (-1)^2 c_2^\dagger c_3^\dagger c_1 c_4. \quad (1.145)$$

There are 2×2 possible arrangements of the creation and annihilation operators that conform to the definition of the normal-ordered product. But creation and annihilation operators anticommute among themselves. The permutation sign $(-1)^{|\sigma|}$ in Eq. (??) automatically compensates for this. Thus,

$$N(c_1 c_2^\dagger c_3^\dagger c_4) = c_2^\dagger c_3^\dagger c_1 c_4 = -c_3^\dagger c_2^\dagger c_1 c_4 = c_3^\dagger c_2^\dagger c_4 c_1 = -c_2^\dagger c_3^\dagger c_4 c_1. \quad (1.146)$$

We define the normal order product of *linear combinations* in the obvious way:

$$N(\alpha A_1 \cdots A_n + \beta B_1 \cdots B_m) = \alpha N(A_1 \cdots A_n) + \beta N(B_1 \cdots B_m). \quad (1.147)$$

Mathematical aside for the interested reader: $N(\cdot)$ is now defined as a linear operator on the space of linear combinations of operator strings. The second-quantized formulas for \hat{H}_0 , \hat{W} , etc., are examples of such objects. This space of operators is an example of a C^* -algebra with unity. An algebra is a vector space where multiplication is also defined (the product of two operators is an operator), and roughly speaking the $*$ means that we can form Hermitian adjoints. The operator algebra is said to be generated by the c_μ operators and the unit operator.

A *contraction* between two arbitrary creation and annihilation operators X and Y is a special notation for $\langle -|XY|-\rangle$,

$$\overline{XY} \equiv \langle -|XY|-\rangle. \quad (1.148)$$

Thus, the contraction is a *number*. One can easily show (see exercise ??), that

$$\overline{XY} = XY - N(XY). \quad (1.149)$$

Let us list all the possible contractions:

$$\overline{c_\mu^\dagger c_\nu^\dagger} = \langle -|c_\mu^\dagger c_\nu^\dagger|-\rangle = 0 \quad (1.150a)$$

$$\overline{c_\mu c_\nu} = \langle -|c_\mu c_\nu|-\rangle = 0 \quad (1.150b)$$

$$\overline{c_\mu^\dagger c_\nu} = \langle -|c_\mu^\dagger c_\nu|-\rangle = 0 \quad (1.150c)$$

$$\overline{c_\mu c_\nu^\dagger} = \langle -|c_\mu c_\nu^\dagger|-\rangle = \delta_{\mu\nu}. \quad (1.150d)$$

$$(1.150e)$$

As we see, most contractions are actually zero.

We also define contractions between two operators *inside a normal ordered product*. This is defined by first anticommuting the contracted operators to the front of the product, and then applying Eq. (??). A contraction between two operators at positions x and y (with $x < y$) in the string is thus defined by

$$N(A_1 \cdots \overline{A_x \cdots A_y} \cdots A_n) \equiv (-1)^{x+y+1} N(\overline{A_x A_y} A_1 \cdots \cancel{A_x} \cdots \cancel{A_y} \cdots A_n) = (-1)^{x+y+1} \overline{A_x A_y} N(A_1 \cdots \cancel{A_x} \cdots \cancel{A_y} \cdots A_n) \quad (1.151)$$

Equivalently, let $\sigma \in S_n$ be a permutation such that $\sigma(x) = 1$ and $\sigma(y) = 2$. Then

$$\begin{aligned} N(A_1 \cdots \overline{A_x \cdots A_y} \cdots A_n) &\equiv (-1)^{|\sigma|} N(\overline{A_x A_y} A_{\sigma(3)} \cdots A_{\sigma(n)}) \\ &= (-1)^{|\sigma|} \overline{A_x A_y} N(A_{\sigma(3)} \cdots A_{\sigma(n)}) \end{aligned} \quad (1.152)$$

(This definition also allows the other operators to be permuted among themselves. This is of course perfectly acceptable – the sign of σ accounts for this.)

Thus, the sign factor $(-1)^{x+y+1}$ equals the sign of the permutation σ that brings the two operators to the front. Equivalently, we can count the number c of anticommutations needed, and the sign becomes $(-1)^c$.

Examples:

$$N(c_1 c_2^\dagger \overbrace{c_3 c_4^\dagger}) = \overbrace{c_3 c_4^\dagger} N(c_1 c_2^\dagger) = -\delta_{34} c_1^\dagger c_2 \quad (1.153a)$$

$$N(\overbrace{c_1 c_2^\dagger} c_3 c_4^\dagger) = \overbrace{c_1 c_2^\dagger} N(c_3 c_4^\dagger) = -\delta_{12} c_3^\dagger c_4 \quad (1.153b)$$

$$N(\overbrace{c_1 c_2^\dagger c_3 c_4^\dagger}) = \overbrace{c_1 c_4^\dagger} N(c_2^\dagger c_3) = \delta_{14} c_2^\dagger c_3 \quad (1.153c)$$

$$N(\overbrace{c_1 c_2^\dagger c_3 c_4^\dagger}) = -\overbrace{c_2^\dagger c_4^\dagger} N(c_1 c_3) = 0 \quad (1.153d)$$

In the last example (??), we could tell immediately that the result is zero, since any contraction between two creation operators vanishes.

It is important to realize that the contraction between two operators, with operators between, *only* makes sense when inside the normal-order product operation $N(\dots)$.

We also define normal ordering with *multiple contractions*, say m , leaving $n - 2m$ operators uncontracted. Clearly, there can be at most $\lfloor n/2 \rfloor$ pairs³ if each pair is required to be distinct.

The definition is recursive: each pair of contracted operators is processed in turn according to Eq. (??). This definition is independent of the order of the processing of the pairs. An example: Example:

$$N(\overbrace{c_1 c_2^\dagger c_3 c_4^\dagger c_5 c_6^\dagger}) = (-1)^2 N(\overbrace{c_1 c_4^\dagger c_2^\dagger c_3 c_5 c_6^\dagger}) = (-1)^3 N(\overbrace{c_1 c_4^\dagger c_3 c_6^\dagger} c_2^\dagger c_5) = -\delta_{14} \delta_{36} N(c_2^\dagger c_5). \quad (1.154)$$

Note how the contraction lines cross on the left-hand side.

For m contractions, the definition is as follows: let (x_i, y_i) be the pairs $x_i < y_i$, for $i = 1, \dots, m$. Let $\sigma \in S_n$ be a permutation such that $\sigma(x_1) = 1$, $\sigma(y_1) = 2$, $\sigma(x_2) = 3$, etc. Then

$$N(\overbrace{A_1 A_2 \dots A_n}^{m \text{ contraction lines}}) = (-1)^{|\sigma|} N(\overbrace{A_{x_1} A_{y_1} \dots A_{x_m} A_{y_m}} \dots A_{\sigma(2m+1)} \dots A_{\sigma(n)}). \quad (1.155)$$

Exercise 1.20. Show Eq. (??) from Eq. (??), by considering the 4 possible cases. △

Exercise 1.21. Prove that, for any permutation $\sigma \in S_n$,

$$N(A_1 A_2 \dots A_n) = (-1)^{|\sigma|} N(A_{\sigma(1)} A_{\sigma(2)} \dots A_{\sigma(n)}). \quad (1.156)$$

△

³ $\lfloor x \rfloor$ is the integer part of x , e.g., $\lfloor 3/2 \rfloor = 1$.

1.4.4 Statement of Wick's Theorem

Wick's theorem states that every string of creation and annihilation operators can be written as a sum of normal-ordered products every possible contraction.

Theorem 1.2 (Wick's Theorem). *Let $A_1 \cdots A_n$ be an operator string of creation and annihilation operators. Then,*

$$A_1 A_2 \cdots A_n = N(A_1 A_2 \cdots A_n) + \sum_{(1)} N \left(A_1 \cdots \overline{} \cdots A_n \right) + \sum_{(2)} N \left(A_1 \cdots \overline{} \cdots A_n \right) + \cdots + \sum_{(\lfloor \frac{n}{2} \rfloor)} N \left(\underbrace{A_1 \cdots \overline{} \cdots A_n}_{[n/2] \text{ contractions}} \right) \quad (1.157)$$

The notation $\sum_{(m)}$ signifies that we sum over all combinations of m contractions.

When n is even, the last sum signifies that we sum over $n/2$ contractions, i.e., all operators are contracted. If n is odd, there is one uncontracted operator left in each term of the last sum.

1.4.5 Vacuum expectation values using Wick's Theorem

Before we start with the proof of Wick's Theorem, we apply it to the evaluation of vacuum expectation values. For any string with at least one factor,

$$\langle -|N(A_1 \cdots A_n)|- \rangle = 0. \quad (1.158)$$

This is so, because in the normal-order product, the annihilation operators are to the right, and the creation operators are on the left. For odd n , therefore, Wick's Theorem gives

$$\langle -|A_1 \cdots A_n|- \rangle = 0 \quad (n \text{ odd number}), \quad (1.159)$$

For even n ,

$$\langle -|A_1 A_2 \cdots A_n|- \rangle = \sum_{(\lfloor \frac{n}{2} \rfloor)} \underbrace{A_1 \cdots \overline{} \cdots A_n}_{\text{all contracted}}. \quad (1.160)$$

where we for brevity omit $N(\cdots)$ since there are no operators left anyway. (Note carefully, that this is abuse of notation!)

The only non-vanishing contractions are

$$\overline{c_\alpha} c_\beta^\dagger = \delta_{\alpha\beta}. \quad (1.161)$$

This reduces the number of contractions we need to consider when evaluating the sum. Moreover, if $A_1 \cdots A_n$ contains a different number of creation and annihilation operators, at least one contraction of the form $\overline{c_\alpha} c_\beta$ or $c_\alpha^\dagger c_\beta^\dagger$ must be present in Eq. (??), in every term, giving a zero expectation value at once.

Finally, one can show that the *sign* of a fully contracted operator product is $(-1)^k$, where k is the number of contraction line crossings. We will not prove this.

Clearly, Wick's theorem provides us with an algebraic method for easy determination of the terms that contribute to the matrix element.

We conclude with a recipe:

Theorem 1.3 (Vacuum expectation values using Wick's Theorem). *Let $A_1 \cdots A_n$ be a string of creation and annihilation operators.*

If n is odd $\langle -|A_1 \cdots A_n| - \rangle = 0$.

Assume n is even. If $A_1 \cdots A_n$ contains a different number of creation operators compared to annihilation operators, $\langle -|A_1 \cdots A_n| - \rangle = 0$.

Finally,

$$\langle -|A_1 \cdots A_n| - \rangle = \sum_{\text{all contr.}} \overbrace{A_1 A_2 A_3 A_4} \cdots \overbrace{A_k A_{k+1} A_{k+2} A_{k+3}} \cdots A_n, \quad (1.162)$$

where the sum runs over all possible combinations of $n/2$ contractions on the form

$$\overbrace{c_\alpha c_\beta}^{\dagger}$$

The sign of each term in the sum is $(-1)^k$, where k is the number of crossings of contraction lines.

Exercise 1.22. (Hard.) Prove the sign rule for the fully contracted terms. (More details will be filled in for this exercise later in the course. Stay tuned.) \triangle

Exercise 1.23. Write out the statement of Wick's Theorem for the following operator strings, and simplify where you can:

1. $c_\beta c_\alpha^\dagger$
2. $c_\alpha^\dagger c_\beta c_\gamma^\dagger c_\delta$
3. $c_\gamma c_\mu^\dagger c_\nu^\dagger c_\alpha c_\beta c_\delta^\dagger$

\triangle

1.4.6 Proof of Wick's Theorem

The proof of Wick's theorem is by induction on the length n of the operator string. In mathematical induction, we prove a statement \mathcal{P}_n for all integers n by first proving it for $n = 1$, and then prove that \mathcal{P}_{n+1} must hold under the assumption that \mathcal{P}_n holds.

Here, the statement \mathcal{P}_n is (??). \mathcal{P}_1 and \mathcal{P}_2 are easily shown to be true (prove it!).

For the rest, it is useful to first prove a lemma.

Lemma 1.2. *Let A_r , $r = 1, \dots, n$ be creation and annihilation operators. Let B be a creation or annihilation operator. Then,*

$$N(A_1 A_2 \cdots A_n) B = \sum_{r=1}^n N(A_1 A_2 \cdots \overbrace{A_r \cdots A_n}^B) + N(A_1 \cdots A_n B). \quad (1.163)$$

Proof. Assume first that B is an annihilation operator. Then all the contractions on the right-hand side vanish. Also, $N(A_1 \cdots A_n)B = N(A_1 \cdots A_n B)$.

Assume next that B is a creation operator, and that all the A_i are annihilation operators. In that case, we can verify that the left- and right-hand sides are equal. The left-hand side is equal to $A_1 \cdots A_n B$ since $A_1 \cdots A_n$ is already a normal-ordered product. We compute $N(A_1 \cdots A_n B) = (-1)^n B A_1 \cdots A_n$. Looking at the left-hand side again,

$$N(A_1 A_2 \cdots A_n)B = A_1 \cdots A_n B = A_1 \cdots A_{n-1} (\overbrace{A_n B} - B A_n), \quad (1.164)$$

since $\{c_\alpha, c_\beta^\dagger\} = \delta_{\alpha,\beta} = c_\alpha \overbrace{c_\beta^\dagger}$. Continuing with the rest of the terms, we get

$$\begin{aligned} N(A_1 A_2 \cdots A_n)B &= A_1 \cdots A_{n-1} \overbrace{A_n B} - A_1 \cdots A_{n-2} \overbrace{A_{n-1} B A_n} + A_1 \cdots A_{n-3} \overbrace{A_{n-2} B A_{n-1} A_n} - \cdots \\ &\quad + (-1)^n B A_1 \cdots A_n \\ &= N(A_1 \cdots A_{n-1} \overbrace{A_n B}) + N(A_1 \cdots A_{n-2} \overbrace{A_{n-1} A_n B}) + N(A_1 \cdots A_{n-3} \overbrace{A_{n-2} A_{n-1} A_n B}) + \cdots \\ &\quad + (-1)^n B A_1 \cdots A_n \end{aligned} \quad (1.165)$$

This proves the case for all A_i annihilation operators, and it remains to prove it when we have creation operators in the mix.

Multiply Eq. (??) from the *left* by a creation operator A_0 . We observe that normal order is preserved on the left hand side since A_0 is a creation operator and can stand to the far right,

$$A_0 N(A_1 \cdots A_n) = N(A_0 \cdots A_n),$$

and similarly $A_0 N(A_1 \cdots A_n)B = N(A_0 \cdots A_n)B$. Also,

$$A_0 \sum_{r=1}^n \sum_{r=1}^n N(A_1 A_2 \cdots \overbrace{A_r \cdots A_n B}) = \sum_{r=0}^n N(A_0 A_1 A_2 \cdots \overbrace{A_r \cdots A_n B}),$$

since $\overbrace{A_0 B} = 0$. Thus, the statement of the lemma is true also when A_0 is a creation operator. Clearly, we can continue, and add as many creation operators we like. Thus, the lemma is true for strings of the form $C_1 \cdots C_k A_{k+1} \cdots A_n$, where C_i are creation operators, and A_i are annihilation operators. By permuting this string, we gain a sign change on all terms, and the terms in the sum over r are reordered, but leaving the sum invariant. Thus, the lemma is proved for arbitrary strings $A_1 \cdots A_n$. \square

We introduce another lemma, which generalizes Lemma ?? to the case where we have an arbitrary number m contractions between the n operators inside the normal order operator.

Lemma 1.3. Suppose $A_1 \cdots A_n$ is a given operator string, and suppose we choose m pairs $p_i = \{x_i, y_i\}$ to contract from this string, with $x_i < y_i$. Let $S_m = \{1, 2, \dots, N\} \setminus (\cup_i p_i)$ be the remaining indices when all pairs are removed. Let B a creation or annihilation operator. Then,

$$N(\overbrace{A_1 A_2 \cdots A_{n-1} A_n} B) = N(\overbrace{A_1 A_2 \cdots A_{n-1} A_n B}) + \sum_{r \in S_m} N(\overbrace{A_1 A_2 \cdots A_r \cdots A_{n-1} A_n B}) \quad (1.166)$$

where the notation indicates that all m pairs are contracted from the A_i s.

Proof. let $S = \{1, 2, \dots, N\}$. The pairs are distinct, which we write mathematically as $p_i \subset S \setminus (\cup_{j=1}^{i-1} p_j)$.

Consider the normal-ordered product of $A_1 \cdots A_n$ with the m given contractions, see the left-hand side of Eq. (??). We perform the pairwise “operator flips” that brings first p_1 to the front, then p_2 , etc. The first pair gives a sign $(-1)^{f_1}$, for f_1 flips. The next pair gives a sign $(-1)^{f_2}$, and so on. (Importantly, f_i depend on the order in which we do the “contraction extractions”.) We arrive at

$$N(\overbrace{A_1 A_2 \cdots A_{n-1} A_n}) = (-1)^{f_1 + f_2 + \cdots + f_m} \overbrace{A_{x_1} A_{y_1}} \cdots \overbrace{A_{x_m} A_{y_m}} N(A_1 \cdots (\text{pairs omitted}) \cdots A_n). \quad (1.167)$$

Now,

$$\begin{aligned} N(\overbrace{A_1 A_2 \cdots A_{n-1} A_n}) B &= (-1)^{f_1 + f_2 + \cdots + f_m} \overbrace{A_{x_1} A_{y_1}} \cdots \overbrace{A_{x_m} A_{y_m}} \left[N(A_1 \cdots (\text{omitted}) \cdots A_n B) \right. \\ &\quad \left. + \sum_{r \in S_m} N(A_1 \cdots (\text{omitted}) \cdots A_r \cdots (\text{omitted}) \cdots A_n B) \right] \end{aligned} \quad (1.168)$$

Consider the first term inside the bracket. We can move the contractions inside again, p_m passing the same operators as when extracted, then p_{m-1} , etc, giving an overall sign change that cancels $(-1)^{f_1 + \cdots + f_m}$. This reproduces the first term on the left-hand side of Eq. (??).

The same is actually true for the second term. Even if we pass a contracted A_r , the “operator flips” count towards the sign, by the definition of $N()$ with contractions.

This completes the proof. \square

We now prove Wick’s Theorem. Assume now that \mathcal{P}_n is true. Multiply Eq. (??) from the right by an operator A_{n+1} :

$$\begin{aligned} A_1 A_2 \cdots A_n A_{n+1} &= N(A_1 A_2 \cdots A_n) A_{n+1} + \sum_{(1)} N(\overbrace{A_1 A_2 \cdots A_n}) A_{n+1} \\ &\quad + \sum_{(2)} N(\overbrace{A_1 A_2 A_3 A_4 \cdots A_n}) A_{n+1} \\ &\quad + \cdots + \sum_{(\lfloor \frac{n}{2} \rfloor)} N(\overbrace{A_1 A_2 A_3 A_4 \cdots A_k A_{k+1} A_{k+2} A_{k+3} \cdots A_n}) A_{n+1} \end{aligned} \quad (1.169)$$

Each sum is a sum over m contractions, including the first where we have $m = 0$. We now use Lemma ?? and write

$$\begin{aligned} \sum_{(m)} N(\overbrace{A_1 A_2 A_3 A_4 \cdots A_n}) A_{n+1} &= \sum_{(m)} N(\overbrace{A_1 A_2 A_3 A_4 \cdots A_n A_{n+1}}) + \sum_{(m)} \sum_r N(\overbrace{A_1 A_2 A_3 A_4 \cdots A_n A_{n+1}}) \\ &:= X_m + I_m. \end{aligned} \quad (1.170)$$

Here, X_m contain all possible m contractions *excluding* A_{n+1} , while I_m contains all possible $m + 1$ contractions *including* A_{n+1} . We now get

$$A_1 \cdots A_{n+1} = X_0 + I_0 + X_1 + I_1 + \cdots X_{\lfloor n/2 \rfloor}. \quad (1.171)$$

Note that $I_{\lfloor n/2 \rfloor} = 0$, since there is no operator left to contract A_{n+1} with after $2\lfloor n/2 \rfloor$ operators have been contracted.

Write

$$A_1 \cdots A_{n+1} = X_0 + (I_0 + X_1) + (I_1 + X_2) + \cdots X_{\lfloor n/2 \rfloor}. \quad (1.172)$$

and note that $(I_m + X_{m+1})$ is the sum over *all possible* $m + 1$ contractions of the string $A_1 \cdots A_{n+1}$. Thus, Wick’s Theorem is proved.

1.4.7 Using Wick's Theorem

In this section, we see some examples of how to use Wick's Theorem to compute vacuum expectation values. First, we state, but do not prove, a theorem regarding the *sign* of a vacuum expectation value of a fully contracted normal-ordered product. The theorem simplifies enormously the work involved in computing the sign of the permutation needed to bring all the contracted pairs to the front.

Theorem 1.4 (Sign rule for vacuum-expectation values). *Let $A_1 \cdots A_n$ be an operator string of creation and annihilation operators, where n is even. Let $n/2$ contractions be assigned, contracting A_{x_i} with A_{y_i} for all $n/2$ pairs of operators, $x_i < y_i$, i.e., we have $m/2$ contractions of the form $\overbrace{A_{x_i} A_{y_i}}$. Then,*

$$\langle - | A_1 A_2 A_3 \cdots A_{n-1} A_n | - \rangle = \overbrace{A_{x_1} A_{y_1}} \cdots \overbrace{A_{x_{n/2}} A_{y_{n/2}}} (-1)^s, \quad (1.173)$$

where s is the number of contraction line crossings on the left-hand side.

Let us compute a few vacuum expectation values with the aid of this rule, and also the simplifications we gain when we know that *all annihilation operators must be contracted with a creation operator to the right*.

We now also simplify the notation a bit, and write, in place of the ordinary creation and annihilation operators,

$$\mu^\dagger \equiv c_\mu^\dagger, \quad \mu = c_\mu.$$

Worked example 1:

$$\langle \mu_1 \cdots \mu_3 | \alpha^\dagger \beta | \mu_1 \cdots \mu_3 \rangle = \langle - | \mu_3 \mu_2 \mu_1 \alpha^\dagger \beta \mu_1^\dagger \mu_2^\dagger \mu_3^\dagger | - \rangle. \quad (1.174)$$

Worked example 2:

$$\langle \mu_1 \cdots \mu_3 | \alpha_1^\dagger \alpha_2^\dagger \beta_2 \beta_1 | \mu_1 \cdots \mu_3 \rangle = \langle - | \mu_3 \mu_2 \mu_1 \alpha_1^\dagger \alpha_2^\dagger \beta_2 \beta_1 \mu_1^\dagger \mu_2^\dagger \mu_3^\dagger | - \rangle. \quad (1.175)$$

Exercise 1.24. (Slater–Condon rules revisited)

- Let $\vec{\mu} = (\mu_1 \cdots \mu_N)$, with $N \geq 2$. Compute the matrix elements $\langle \vec{\mu} | \hat{H}_0 | \vec{\mu} \rangle$ and $\langle \vec{\mu} | \hat{W} | \vec{\mu} \rangle$ using Wick's theorem applied to vacuum expectation values. Do you notice a pattern of which contractions contribute other the rules listed in the main text?
- Repeat Exercise ?? using Wick's Theorem instead of the anticommutator relations to prove the Slater–Condon rules. (Wick's Theorem gives a much less tedious approach.)

△

1.5 Particle-hole formalism

Motivational comments: Often, a single Slater determinant can be a good approximation, for example the Hartree–Fock state, see later. If this approximation is not good enough, one adds a correction on top of that. Therefore it makes sense to develop a convenient way to describe this small correction.

In this section, we will introduce the concept of quasiparticles, or particle-hole formalism.

It is useful to indicate if $\mu \leq N$ or $\mu > N$ in the following. We therefore introduce a rule. A *latin* index $i, j, k, \dots \leq N$, and $a, b, c, \dots > N$. Thus, a summation \sum_μ is split into $\sum_{i=1}^N$ and $\sum_{a=N+1}^\infty$.

We define *quasiparticle creation and annihilation operators* as follows:

$$b_i = c_i^\dagger, \quad b_a = c_a \quad (1.176)$$

with Hermitian adjoints

$$b_i^\dagger = c_i, \quad b_a^\dagger = c_a^\dagger \quad (1.177)$$

It is an easy exercise to show that the anticommutator relations are preserved:

$$\{b_\mu, b_\nu^\dagger\} = \delta_{\mu,\nu}, \quad \{b_\mu, b_\nu\} = 0. \quad (1.178)$$

Let a single-particle basis be given, and consider the Slater determinant

$$|\Phi\rangle = |123\cdots N\rangle = c_1^\dagger c_2^\dagger \cdots c_N^\dagger |-\rangle. \quad (1.179)$$

For the $N = 4$ case, we can draw a picture like this:

$$|\Phi\rangle = \bullet \bullet \bullet \bullet \circ \circ \circ \circ \circ \circ$$

Note that

$$b_\mu |\Phi\rangle = 0, \quad \forall \mu. \quad (1.180)$$

Therefore, $|\Phi\rangle$ has the role of a vacuum for the new operators. It contains zero quasiparticles, since attempting to remove one gives us zero.

Let us create a quasiparticle:

$$b_i^\dagger |\Phi\rangle = c_i |123\cdots N\rangle = (-1)^{i-1} |123\cdots (i-1) (i+1) \cdots N\rangle. \quad (1.181)$$

$$b_a^\dagger |\Phi\rangle = c_a^\dagger |123\cdots N\rangle = (-1)^N |123Na\rangle. \quad (1.182)$$

In pictures,

$$b_i^\dagger |\Phi\rangle = \bullet \bullet \circ_i \bullet \circ \circ \circ \circ \circ \circ$$

$$b_a^\dagger |\Phi\rangle = \bullet \bullet \bullet \bullet \circ \circ \circ \bullet_a \circ \circ$$

Note that $b_i^\dagger |\Phi\rangle$ contains $N - 1$ “real” particles, while $b_a^\dagger |\Phi\rangle$ contains $N + 1$ “real” particles.

The quasiparticles with $\mu = i \leq N$ are called “holes”, while the quasiparticles with $\mu = a > N$ are called “particles”.

Creating a particle-hole pair results in a state with N “real” particles, since $b_a^\dagger b_i^\dagger = c_a^\dagger c_i$ preserves N when acting on a state. Acting on the reference, we get $N - 1$ occupied single-particle functions below N , and 1 occupied single-particle function above N , in pictures,

$$b_a^\dagger b_i^\dagger |\Phi\rangle = \bullet \bullet \circ_i \bullet \circ \circ \circ \bullet_a \circ \circ$$

Clearly, by creating another particle-hole pair with $b_b^\dagger b_i^\dagger$, we get a Slater determinant with two particles and two holes, in total N particles. We are left with $N - 2$ “real” particles below N .

Continuing, it is clear that we can generate *all* the original Slater determinants with N particles by creating up to N particle-hole pairs⁴.

Thus, any wavefunction in with N particles can be written

$$|\Psi_N\rangle = C_0 |\Phi\rangle + \sum_{ia} C_{ia} b_a^\dagger b_i |\Phi\rangle + \frac{1}{2!^2} \sum_{ijab} C_{ijab} b_b^\dagger b_j^\dagger b_a^\dagger b_i^\dagger |\Phi\rangle + \dots, \quad (1.183)$$

where the factor $1/2!^2$ comes from the double counting of the two particle-hole states. The sum extends all the way up to N particle hole pairs.

Se define

$$|\Phi_i^a\rangle = b_a^\dagger b_i^\dagger |\Phi\rangle = c_a^\dagger c_i |\Phi\rangle, \quad (1.184)$$

and

$$|\Phi_{ij}^{ab}\rangle = b_b^\dagger b_j^\dagger b_a^\dagger b_i^\dagger |\Phi\rangle = c_b^\dagger c_j^\dagger c_a^\dagger c_i |\Phi\rangle, \quad (1.185)$$

and so on. The lower indices indicate that they are holes/below N , and the upper indices that they are particles/above N .

In chemistry parlance, a particle-hole pair is called a singles excitation, two particle-hole pairs a doubles excitation, etc. Thus, $|\Phi_i^a\rangle$ is a “singly excited determinant”, $|\Phi_{ij}^{ab}\rangle$ is doubly excited, etc.

There are many different common notations for the particle-hole vacuum: $|\text{vac}\rangle$, $|\Phi\rangle$, $|c\rangle$, etc. Similarly, there are many ways to denote a Slater determinant with m particle-hole pairs, for example $|ia\rangle_c$, $|\Phi_i^a\rangle$, $|i_i^a\rangle$, and others.

We can say that $b_a^\dagger b_i^\dagger$ is a (*particle-hole*) *pair creation operator*. In chemistry language, a *singles excitation*. It is useful to note that

$$[b_a^\dagger b_i^\dagger, b_b^\dagger b_j^\dagger] = 0. \quad (1.186)$$

I.e., $|\Phi_{ij}^{ab}\rangle = |\Phi_{ji}^{ba}\rangle$. We form double excitation operators by products of singles, and so on.

Finally, we note that Wick’s theorem applies equally well to quasiparticles! For example, to compute $\langle\Phi|c_i^\dagger c_j|\Phi\rangle$ we note that $|\Phi\rangle$ is the vacuum and that $c_i^\dagger = b_i$ and $c_j = b_j^\dagger$, so

$$\langle\Phi|c_i^\dagger c_j|\Phi\rangle = \langle\Phi|b_i b_j^\dagger|\Phi\rangle = \overline{b_i b_j^\dagger} = \delta_{ij}. \quad (1.187)$$

Note how the quasiparticles greatly simplified the evaluation of the matrix element, see the exercises.

Exercise 1.25. Prove Eq. (??)

△

Exercise 1.26. Prove the quasiparticle anticommutator relations.

△

Exercise 1.27. If we restrict $a \leq L$, how many *linearly independent* two-particle-two-hole determinants can you create? How many three-particle-three-hole? How would you phrase this in chemistry language? △

⁴Digression: There can be only N hole-particles! I the solution of the Dirac equation, for those who have seen this, the vacuum contains zero electrons. But every time an electron is created, an anti-electron is also created below the “Fermi sea”. There are infinitely many hole-states in Dirac theory.

Exercise 1.28. Compute the vacuum expectation value (??) using the Wick's theorem and the original creation and annihilation operators and compare the method and amount of work.

Repeat for

$$\langle \Phi | c_\alpha^\dagger c_\beta^\dagger c_\delta c_\gamma | \Phi \rangle \quad (1.188)$$

but compute also with quasiparticles, but note that you get several cases, depending if the Greek indices are smaller than or larger than N . Compare with the original formulation. \triangle

Exercise 1.29. Use Wick's Theorem with respect to quasiparticles and write down the following operators as a sum of normal-ordered strings with as few terms as possible (i.e., only include nonvanishing contractions):

a) $b_a b_i^\dagger$

a) $b_i^\dagger b_c b_a^\dagger$

b) $b_i b_j b_a^\dagger b_b^\dagger b_c b_k^\dagger$

c) $b_a b_i b_j b_b^\dagger b_c^\dagger b_d b_k^\dagger$

\triangle

1.6 Operators on normal-order form

1.6.1 The number operator

We need the Hamiltonian and other second-quantized operators on normal-order form, relative to quasiparticle vacuum. I.e., we want the operator to be written such that all quasiparticle annihilation operators are to the right. This is achieved using Wick's Theorem, and results in the original operator obtaining more terms.

This is the task in the current section.

For conformity with much of the literature, we replace Greek indices μ, ν , etc, with p, q , etc. We still reserve i, j , etc for hole indices, and a, b , etc for particles. And let's face it, it is easier to write up in L^AT_EX.

We start with the number operator \hat{N} , as an easy warm-up. First, we rewrite the second-quantized operator using quasiparticle operators:

$$\hat{N} = \sum_p c_p^\dagger c_p = \sum_i c_i^\dagger c_i + \sum_a c_a^\dagger c_a = \sum_i b_i b_i^\dagger + \sum_a b_a^\dagger b_a. \quad (1.189)$$

We now use Wick's theorem, *relative to quasiparticle operators*, to get

$$\begin{aligned} \hat{N} &= \sum_i \left[N(b_i b_i^\dagger) + \overline{b_i b_i^\dagger} \right] + \sum_a \left[N(b_a^\dagger b_a) + \overline{b_a^\dagger b_a} \right] \\ &= \sum_i [-b_i^\dagger b_i + 1] + \sum_a b_a^\dagger b_a \\ &= N - \sum_i b_i^\dagger b_i + \sum_a b_a^\dagger b_a. \end{aligned} \quad (1.190)$$

This is the normal-ordered form of \hat{N} . Interpreting, the last equality counts N minus the number of holes plus the number of particles.

Let us act with \hat{N} on the quasiparticle vacuum, and observe:

$$\hat{N}|\Phi\rangle = (N - \sum_i b_i^\dagger b_i + \sum_a b_a^\dagger b_a)|\Phi\rangle = N|\Phi\rangle. \quad (1.191)$$

All the terms vanish except the fully contracted term since we have annihilation operators to the right.

Thus, normal-ordered operators can be very useful when we deal with quasiparticles.

1.6.2 One-body operators

We continue with an arbitrary one-body operator

$$\hat{H}_0 = \sum_{pq} h_q^p c_p^\dagger c_q, \quad h_q^p = \langle p|\hat{h}|q\rangle. \quad (1.192)$$

Introducing the quasiparticle operators at this stage leads to four distinct contributions to the operator, corresponding to the different $pq = ij, ia, ai$, and ab contributions. However, it is more convenient to use Wick's Theorem on the above \hat{H}_0 expression without changing the creation- and annihilation operator notation. Thus, beware, when we normal order now, it is *relative to quasiparticles*.

Wick's Theorem gives

$$c_p^\dagger c_q = N(c_p^\dagger c_q) + \overline{c_p^\dagger c_q}. \quad (1.193)$$

Some of the contractions are nonzero, namely, when $pq = ii$. The reader should verify that the rest of the possible contractions vanish identically. Thus,

$$\begin{aligned} \hat{H}_0 &= \sum_{pq} h_q^p N(c_p^\dagger c_q) + \sum_{pq} h_q^p \overline{c_p^\dagger c_q} \\ &= \sum_{pq} h_q^p N(c_p^\dagger c_q) + \sum_i h_i^i \\ &= \hat{H}_0^{(1qp)} + \hat{H}_0^{(0qp)}. \end{aligned} \quad (1.194)$$

Note that \hat{H}_0 is separated into a one-quasiparticle part and a constant zero-quasiparticle part. Explicitly,

$$\hat{H}_0^{(0qp)} = \sum_i h_i^i, \quad (1.195)$$

and

$$\hat{H}_0^{(1qp)} = - \sum_{ij} h_j^i b_j^\dagger b_i + \sum_{ai} h_i^a b_a^\dagger b_i + \sum_{ia} h_a^i b_i b_a + \sum_{ab} h_b^a b_a^\dagger b_b, \quad (1.196)$$

where we have expanded the sum over pq in order to resolve the quasiparticles.

1.6.3 Two-body operators

We continue with an arbitrary two-body operator

$$\hat{W} = \frac{1}{4} \sum_{pqrs} w_{rs}^{pq} c_p^\dagger c_q^\dagger c_s c_r, \quad (1.197)$$

where we assume that w_{rs}^{pq} is anti-symmetrized. (This is at odds with earlier notation, but we do this to save space in the current section.)

Wick's Theorem gives

$$\begin{aligned}
c_p^\dagger c_q^\dagger c_s c_r &= N(c_p^\dagger c_q^\dagger c_s c_r) + N(\overline{c_p^\dagger c_q^\dagger} c_s c_r) + N(\overline{c_p^\dagger c_q^\dagger} c_s c_r) + N(\overline{c_p^\dagger c_q^\dagger} c_s c_r) \\
&\quad + N(c_p^\dagger \overline{c_q^\dagger} c_s c_r) + N(c_p^\dagger \overline{c_q^\dagger} c_s c_r) + N(c_p^\dagger \overline{c_q^\dagger} c_s c_r) \\
&\quad + N(\overline{c_p^\dagger} c_q^\dagger c_s c_r) + N(\overline{c_p^\dagger} c_q^\dagger c_s c_r) + N(\overline{c_p^\dagger} c_q^\dagger c_s c_r) \\
&= N(c_p^\dagger c_q^\dagger c_s c_r) + \overline{c_p^\dagger} c_q^\dagger N(c_s c_r) - \overline{c_p^\dagger} c_s N(c_q^\dagger c_r) + \overline{c_p^\dagger} c_r N(c_q^\dagger c_s) \\
&\quad + \overline{c_q^\dagger} c_s N(c_p^\dagger c_r) - \overline{c_q^\dagger} c_r N(c_p^\dagger c_s) + \overline{c_s} c_r N(c_p^\dagger c_q^\dagger) \\
&\quad + \overline{c_p^\dagger} c_q^\dagger \overline{c_s} c_r - \overline{c_p^\dagger} c_s \overline{c_q^\dagger} c_r + \overline{c_q^\dagger} c_s \overline{c_p^\dagger} c_r
\end{aligned} \tag{1.198}$$

We see immediately, that analogously to the one-body operator, we will get

$$\hat{W} = \hat{W}^{(2qp)} + \hat{W}^{(1qp)} + \hat{W}^{(0qp)}. \tag{1.199}$$

The two-quasiparticle term is

$$\hat{W}^{(2qp)} = \frac{1}{4} \sum_{pqrs} w_{rs}^{pq} N(c_p^\dagger c_q^\dagger c_s c_r). \tag{1.200}$$

The one-quasiparticle term is

$$\hat{W}^{(1qp)} = \sum_{pqi} w_{qi}^{pi} N(c_p^\dagger c_q) \tag{1.201}$$

while the constant term is

$$\hat{W}^{(0qp)} = \frac{1}{2} \sum_{ij} w_{ij}^{ij}. \tag{1.202}$$

1.6.4 Normal-ordered two-body Hamiltonian

Consider the full Hamiltonian on the form

$$\hat{H} = \hat{H}_0 + \hat{W}. \tag{1.203}$$

The Hamiltonian is normal-ordered relative to “real” particles. In terms of quasiparticles, we saw in the previous sections that we could split

$$\hat{H} = \hat{H}_0^{(0qp)} + \hat{W}^{(0qp)} + \hat{H}_0^{(1qp)} + \hat{W}^{(1qp)} + \hat{W}^{(2qp)}, \tag{1.204}$$

separating \hat{H} into zero, one and two-quasiparticle contributions. These were normal-ordered relative to quasiparticles.

It is conventional to write

$$\hat{H} = \hat{H}_N + E_0 = \hat{H}_N + \hat{W}_N + E_0, \tag{1.205}$$

with

$$E_0 = \hat{H}_0^{(0qp)} + \hat{W}^{(0qp)} = \sum_i h_i^i + \frac{1}{2} \sum_{ij} w_{ij}^{ij}, \tag{1.206a}$$

$$\hat{H}_{0,N} = \hat{H}_0^{(1qp)} + \hat{W}^{(1qp)} = \sum_{pq} (h_q^p + \sum_i w_{qi}^{pi}) N(c_p^\dagger c_q), \tag{1.206b}$$

$$\hat{W}_N = \hat{W}^{(2qp)} = \frac{1}{4} \sum_{pqrs} w_{rs}^{pq} N(c_p^\dagger c_q^\dagger c_s c_r). \tag{1.206c}$$

Thus, $\hat{H}_{0,N}$ is the *total* one-quasiparticle operator part of \hat{H} , and contains contributions from \hat{W} as well as \hat{H}_0 , while \hat{W}_N is the total two-quasiparticle operator part of \hat{H} . The subscript N stands for “normal-ordered”.

1.6.5 Full expressions for the normal-ordered Hamiltonian

For completeness, we expand $\hat{H}_{0,N}$ and \hat{W}_N in terms of quasiparticle operators. This gives a lot of terms, especially in the two-body case. We start with $\hat{H}_{0,N}$, splitting the sum over pq into four terms:

$$\begin{aligned}\hat{H}_{0,N} &= \sum_{ij} f_j^i N(c_i^\dagger c_j) + \sum_{ai} f_i^a N(c_a^\dagger c_i) + \sum_{ia} f_a^i N(c_i^\dagger c_a) + \sum_{ab} f_b^a N(c_a^\dagger c_b) \\ &= \sum_{ai} f_i^a b_a^\dagger b_i^\dagger + - \sum_{ij} f_j^i b_j^\dagger b_i + \sum_{ab} f_b^a b_a^\dagger b_b + \sum_{ia} f_a^i b_i b_a +\end{aligned}\quad (1.207)$$

where

$$f_q^p = h_q^p + \sum_j w_{qj}^{pj}. \quad (1.208)$$

Next, we resolve the two-body operator. There are 16 terms:

$$\begin{aligned}\hat{W}_N &= \frac{1}{4} \sum_{ijkl} w_{kl}^{ij} N(c_i^\dagger c_j^\dagger c_l c_k) + \frac{1}{4} \sum_{ijka} w_{ka}^{ij} N(c_i^\dagger c_k^\dagger c_a c_k) + \frac{1}{4} \sum_{ijak} w_{ak}^{ij} N(c_i^\dagger c_j^\dagger c_k c_a) + \frac{1}{4} \sum_{iajk} w_{jk}^{ia} N(c_i^\dagger c_a^\dagger c_k c_j) \\ &+ \frac{1}{4} \sum_{aijk} w_{jk}^{ai} N(c_a^\dagger c_i^\dagger c_k c_j) + \frac{1}{4} \sum_{ijab} w_{ab}^{ij} N(c_i^\dagger c_j^\dagger c_b c_a) + \frac{1}{4} \sum_{iajb} w_{jb}^{ia} N(c_i^\dagger c_a^\dagger c_b c_j) + \frac{1}{4} \sum_{iabj} w_{bj}^{ia} N(c_i^\dagger c_a^\dagger c_j c_b) \\ &+ \frac{1}{4} \sum_{aijb} w_{jb}^{ai} N(c_a^\dagger c_i^\dagger c_b c_j) + \frac{1}{4} \sum_{aibj} w_{bj}^{ai} N(c_a^\dagger c_i^\dagger c_j c_b) + \frac{1}{4} \sum_{abij} w_{ij}^{ab} N(c_a^\dagger c_b^\dagger c_j c_i) + \frac{1}{4} \sum_{iabc} w_{bc}^{ia} N(c_i^\dagger c_a^\dagger c_c c_b) \\ &+ \frac{1}{4} \sum_{aibc} w_{bc}^{ai} N(c_a^\dagger c_i^\dagger c_c c_b) + \frac{1}{4} \sum_{abci} w_{ic}^{ab} N(c_a^\dagger c_b^\dagger c_c c_i) + \frac{1}{4} \sum_{abci} w_{ci}^{ab} N(c_a^\dagger c_b^\dagger c_i c_c) + \frac{1}{4} \sum_{abcd} w_{cd}^{ab} N(c_a^\dagger c_b^\dagger c_d c_c) \\ &+ \frac{1}{4} \sum_{abcd} w_{cd}^{ab} N(c_a^\dagger c_b^\dagger c_d c_c)\end{aligned}\quad (1.209)$$

Some terms are equal, and we rearrange the expression to read:

$$\begin{aligned}\hat{W}_N &= \frac{1}{4} \sum_{abij} w_{ij}^{ab} b_a^\dagger b_b^\dagger b_j^\dagger b_i^\dagger + \frac{1}{2} \sum_{abci} w_{ci}^{ab} b_a^\dagger b_b^\dagger b_i^\dagger b_c + \frac{1}{4} \sum_{aibj} w_{bj}^{ai} b_a^\dagger b_j^\dagger b_b^\dagger b_i + \frac{1}{2} \sum_{aijk} w_{jk}^{ai} b_a^\dagger b_k^\dagger b_j^\dagger b_i \\ &+ \frac{1}{4} \sum_{ijkl} w_{kl}^{ij} b_l^\dagger b_k^\dagger b_i b_j - \frac{1}{4} \sum_{iajb} w_{jb}^{ia} b_a^\dagger b_j^\dagger b_i b_b + \frac{1}{2} \sum_{iabj} w_{bj}^{ia} b_a^\dagger b_j^\dagger b_i b_b + \frac{1}{4} \sum_{abcd} w_{cd}^{ab} b_a^\dagger b_b^\dagger b_d b_c \\ &+ \frac{1}{2} \sum_{ijak} w_{ak}^{ij} b_k^\dagger b_i b_j b_a + \frac{1}{2} \sum_{aibc} w_{bc}^{ai} b_i b_a b_c b_b + \frac{1}{4} \sum_{ijab} w_{ab}^{ij} b_i b_j b_b b_a\end{aligned}\quad (1.210)$$

Exercise 1.30. Verify that Eq. (??) equals Eq. (??). △

Exercise 1.31. Verify that \hat{W}_N in Eq. (??) is Hermitian, given that \hat{W} is Hermitian. △

Exercise 1.32. (Tedious.) Consider a three-body operator

$$\hat{X} = \frac{1}{36} \sum_{pqrstu} x_{stu}^{pqr} c_p^\dagger c_q^\dagger c_r^\dagger c_u c_t c_s, \quad (1.211)$$

where $x_{stu}^{pqr} = \langle pqr | \hat{x}(1, 2, 3) | stu \rangle$ is permutation antisymmetric in the upper and lower indices separately, i.e., it is the matrix of the three-particle operator x . Such operators occur in nuclear physics.

Compute the separation

$$\hat{X} = \hat{X}^{(0qp)} + \hat{X}^{(1qp)} + \hat{X}^{(2qp)} + \hat{X}^{(3qp)}. \quad (1.212)$$

Given a Hamiltonian is given by $\hat{H} = \hat{F} + \hat{G} + \hat{X}$, write down the normal-ordered Hamiltonian, split as

$$\hat{H} = \hat{F}_N + \hat{G}_N + \hat{X}_N + E_0. \quad (1.213)$$

△

Chapter 2

The Standard Methods of approximation

2.1 Introduction

Having dealt with the basic formalism of many-fermion theory, how do we solve the Schrödinger equation approximately? In this section, we discuss the *variational principle*, perhaps *the* most important tool for devising approximate schemes.

We then develop the configuration-interaction method, and then Hartree–Fock theory, and then we combine the two methods.

2.2 The variational principle

Consider the time-independent Schrödinger equation for an N -fermion system, i.e., given our Hamiltonian \hat{H} , find a nonzero $|\Psi\rangle \in L_N^2$ with E a real number such that

$$\hat{H}|\Psi\rangle = E|\Psi\rangle. \quad (2.1)$$

This is an eigenvalue problem for a Hermitian operator \hat{H} over a Hilbert space. The mathematical analysis of this problem is complex. However, if the Hilbert space L_N^2 has *finite dimension* D , then \hat{H} can be viewed as a *Hermitian matrix*, and we can find a complete set of orthonormal eigenfunctions $|\Psi_k\rangle$, $k = 0, 1, 2, \dots$ with corresponding eigenvalues E_k , such that

$$\hat{H} = \sum_{k=0}^D E_k |\Psi_k\rangle \langle \Psi_k|. \quad (2.2)$$

Of course, Hilbert space is usually infinite dimensional, complicating the mathematical analysis of the problem. It may happen that \hat{H} does not even have a ground state, or not even a single eigenvector. However, it turns out, that in most interesting cases the differences are small enough to warrant the assumption that we are dealing with a finite-dimensional problem, or at least that Eq. (2.2) holds with possibly an infinite dimension.

Theorem 2.1 (Variational principle). *Consider the expectation value functional defined by*

$$\mathcal{E}(|\Psi\rangle) \equiv \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle}. \quad (2.3)$$

Let $|\Psi_*\rangle$ be given. Then $E_* = \mathcal{E}(|\Psi_*\rangle)$ is a stationary value of \mathcal{E} with respect to all infinitesimal variations $|\Psi_*\rangle + \epsilon|\eta\rangle$ (with ϵ a small number and $\langle\eta|\eta\rangle = 1$) if and only if

$$\hat{H}|\Psi_*\rangle = E_*|\Psi_*\rangle. \quad (2.4)$$

Proof. Let ϵ be a small real number, $|\Psi\rangle, |\eta\rangle \in L_N^2$ arbitrary vectors, $|\eta\rangle$ normalized. Let $f(\epsilon)$ be defined as

$$f(\epsilon) = \mathcal{E}(|\Psi\rangle + \epsilon|\eta\rangle). \quad (2.5)$$

The stationary point condition can be formulated as

$$f'(0) = 0. \quad (2.6)$$

This condition must hold for all $|\eta\rangle$. Thus, $\epsilon|\eta\rangle$ is an arbitrary infinitesimal variation. Mathematically, $f'(\epsilon)$ is the directional derivative of \mathcal{E} at $|\Psi\rangle$ in the direction $|\eta\rangle$. Then

$$f(\epsilon) = \frac{\langle\Psi|\hat{H}|\Psi\rangle + \epsilon\langle\eta|\hat{H}|\Psi\rangle + \epsilon\langle\Psi|\hat{H}|\eta\rangle + \epsilon^2\langle\eta|\hat{H}|\eta\rangle}{\langle\Psi|\Psi\rangle + \epsilon\langle\eta|\Psi\rangle + \epsilon\langle\Psi|\eta\rangle + \epsilon^2\langle\eta|\eta\rangle}. \quad (2.7)$$

Define $E = \langle\Psi|\hat{H}|\Psi\rangle$, $N = \langle\Psi|\Psi\rangle$. Define $A = \langle\eta|\hat{H}|\Psi\rangle + \langle\Psi|\eta\rangle$, $a = \langle\eta|\Psi\rangle + \langle\Psi|\eta\rangle$.

$$\mathcal{E}(|\Psi\rangle + \epsilon|\eta\rangle) = \frac{E + \epsilon A + O(\epsilon^2)}{N + \epsilon a + O(\epsilon^2)}. \quad (2.8)$$

Using $1/(1+x) = 1-x+O(x^2)$, we expand the denominator to first order in ϵ :

$$\frac{1}{N + \epsilon \frac{a}{N} + O(\epsilon^2)} = \frac{1}{N} \left[1 - \epsilon \frac{a}{N} + O(\epsilon^2) \right]. \quad (2.9)$$

We expand $f(\epsilon)$ to first order in ϵ :

$$\begin{aligned} Nf(\epsilon) &= (E + \epsilon A + O(\epsilon^2)) \left[1 - \epsilon \frac{a}{N} + O(\epsilon^2) \right] \\ &= E + \epsilon \left[A - \frac{aE}{N} \right] + O(\epsilon^2). \end{aligned} \quad (2.10)$$

Recall that

$$f(\epsilon) = f(0) + \epsilon f'(0) + O(\epsilon^2). \quad (2.11)$$

We see that $f'(0) = 0$ if and only if

$$A = \frac{aE}{N}, \quad (2.12)$$

that is,

$$\langle\eta|\hat{H}|\Psi\rangle + \langle\Psi|\hat{H}|\eta\rangle = (\langle\eta|\Psi\rangle + \langle\Psi|\eta\rangle)\mathcal{E}(|\Psi\rangle), \quad (2.13)$$

which must hold for all $|\eta\rangle$. In particular, if it holds for $|\eta\rangle = |u\rangle$ it must also hold for $|\eta\rangle = i|u\rangle$. Plugging these in gives

$$\langle u|\hat{H}|\Psi\rangle + \langle\Psi|\hat{H}|u\rangle = (\langle u|\Psi\rangle + \langle\Psi|u\rangle)\mathcal{E}(|\Psi\rangle), \quad (2.14)$$

$$-i\langle u|\hat{H}|\Psi\rangle + i\langle\Psi|\hat{H}|u\rangle = (-i\langle u|\Psi\rangle + i\langle\Psi|u\rangle)\mathcal{E}(|\Psi\rangle), \quad (2.15)$$

Multiplying the second equation by i and adding the two equations gives

$$\langle u|\hat{H}|\Psi\rangle = \mathcal{E}(|\Psi\rangle)\langle u|\Psi\rangle. \quad (2.16)$$

Since $|u\rangle$ was arbitrary, we must have

$$\hat{H}|\Psi\rangle = \mathcal{E}(|\Psi\rangle)|\Psi\rangle. \quad (2.17)$$

The proof is complete. \square

The *variational principle* in its simplest form states that the ground-state energy E_0 is the minimum of the *expectation value* of the Hamiltonian:

Theorem 2.2 (Variational Principle, Rayleigh–Ritz). *If \hat{H} has a ground state, then the ground-state energy is given by the minimum of the expectation value of \hat{H} , viz,*

$$E_0 = \min \left\{ \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \left| 0 \neq |\Psi\rangle \in L_N^2, |\langle \Psi | \hat{H} | \Psi \rangle| < +\infty \right\}. \quad (2.18)$$

Theorem ?? holds even if Eq. (??) does not hold. It is sufficient that \hat{H} has a lowest eigenvalue. In the infinite dimensional case, we must require that $|\langle \Psi | \hat{H} | \Psi \rangle| < +\infty$, since for most Hamiltonians of interest, there are in fact $|\Psi\rangle$ that has an infinite expectation value. In finite dimensions, this is of course not true.

We will not prove Theorem ?? in its full generality, but we see immediately that it follows from Theorem ??: E_0 is a stationary value, and clearly \mathcal{E} cannot take values *lower* than E_0 . Thus, E_0 must be the minimum.

We now consider *the variational procedure*, a useful method of generating approximate ground-state energies. Suppose we have a *subset of Hilbert space* $\mathcal{M} \subset L_N^2$, and compute

$$E_0[\mathcal{M}] \equiv \inf \left\{ \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \left| 0 \neq |\Psi\rangle \in \mathcal{M}, |\langle \Psi | \hat{H} | \Psi \rangle| < +\infty \right\}. \quad (2.19)$$

Clearly,

$$E_0 \leq E_0[\mathcal{M}], \quad (2.20)$$

since we minimize over a *smaller* set than the full Hilbert space. This *upper bound property* of the variational procedure is very useful, because if we enlarge \mathcal{M} , we will always get a better estimate for E_0 .

Suppose that our variational procedure yields a minimum value in Eq. (??) for the function $|\tilde{\Psi}\rangle \in \mathcal{M}$:

$$E_0[\mathcal{M}] = \mathcal{E}(|\tilde{\Psi}\rangle) = \frac{\langle \tilde{\Psi} | \hat{H} | \tilde{\Psi} \rangle}{\langle \tilde{\Psi} | \tilde{\Psi} \rangle}. \quad (2.21)$$

Suppose also that $|\tilde{\Psi}\rangle$ is fairly close to $|\Psi_0\rangle$, i.e.,

$$|\Psi_0\rangle \approx |\tilde{\Psi}\rangle + \epsilon |\eta\rangle \quad (2.22)$$

Then, from the proof of the variational principle, we expect that

$$\tilde{E}_0 - E_0 = f(\epsilon) - f(0) = [f(0) + \epsilon f'(0) + O(\epsilon^2)] - f(0) = O(\epsilon^2), \quad (2.23)$$

i.e., that the error in the eigenvalue is *quadratic* in the error in the eigenfunction! Thus, the error $E_0[\mathcal{M}] - E_0$ is insensitive to errors in the wavefunction. This explains why the variational procedure is so useful.

Example: The hydrogen atom with Hamiltonian

$$\hat{h} = -\frac{1}{2}\nabla^2 + \frac{1}{r}. \quad (2.24)$$

The exact ground-state wavefunction is well-known,

$$\psi_0(\vec{r}) = C e^{-r}, \quad (2.25)$$

with eigenvalue $E_0 = -1/2$. Here, C is a normalization constant. Let us imagine we did not know ψ_0 , and try a parameterized wavefunction on the form

$$\psi^\alpha(\vec{r}) = (\alpha/\pi)^{3/4} e^{-\alpha r^2/2}. \quad (2.26)$$

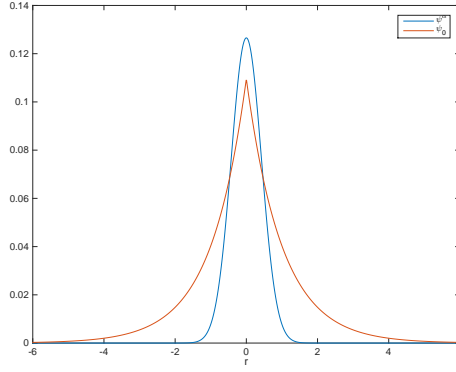


Figure 2.1: Plot of approximate and exact ground-state wavefunction for the Hydrogen example

Thus, $\mathcal{M} = \{|\psi^\alpha\rangle \mid \alpha > 0\}$ is the set of approximate wavefunctions, which all satisfy $\langle\psi^\alpha|\psi^\alpha\rangle = 1$. We can compute the expectation value,

$$\mathcal{E}(|\psi^\alpha\rangle) = \langle\psi^\alpha|\hat{h}|\psi^\alpha\rangle = \frac{3}{4}\alpha - 2\left(\frac{\alpha}{\pi}\right)^{1/2} \quad (2.27)$$

and minimize with respect to α ,

$$E_0[\mathcal{M}] = \inf_{\alpha} \mathcal{E}(|\psi^\alpha\rangle) = \mathcal{E}(|\psi^{\frac{16}{9\pi}}\rangle) = -\frac{4}{3\pi} \approx -0.42. \quad (2.28)$$

This is actually a minimum, obtained at $\alpha = 16/(9\pi)$. Comparing with the exact result, we see that the energies are rather close for such a simple parameterization. The wavefunctions are not that close, see Fig. ??! Note that the exact ground-state is not smooth at $\vec{r} = 0$.

Usually, the set \mathcal{M} contains wavefunction ansätze that are parameterized in some way. In the example, we had a simple Gaussian wavefunction parameterized by the width.

2.2.1 The Cauchy interlace theorem and linear models

Suppose that the set \mathcal{M} is a linear space, i.e., a subspace \mathcal{V} of L_N^2 defined by a basis set $|\Phi_I\rangle$, $I = 1, 2, \dots, D$. Then the variational procedure is equivalent to computing the smallest eigenvalue of the matrix

$$H_{IJ} = \langle\Phi_I|\hat{H}|\Phi_J\rangle. \quad (2.29)$$

This is so, because for

$$|\tilde{\Psi}\rangle = \sum_{I=1}^D A_I |\Phi_I\rangle \quad (2.30)$$

the expectation value becomes

$$\mathcal{E}(|\tilde{\Psi}\rangle) = \frac{\mathbf{A}^H \mathbf{H} \mathbf{A}}{\mathbf{A}^H \mathbf{A}}, \quad (2.31)$$

which is simply the expectation value functional for the quantum system with the Hamiltonian H and wavefunction A , and we can apply the variational principle to this functional.

It is a fact, that under very mild assumptions on $\{|\Phi_I\rangle\}$ and \hat{H} , the eigenvalues of the matrix H converge to the eigenvalues of \hat{H} , even in the infinite dimensional case.

For the finite-dimensional case, the Cauchy interlace theorem states that for a linear model as here described, *all* the eigenvalues of H actually approximate eigenvalues of the full Hamiltonian H from above. For a general nonlinear model \mathcal{M} , we cannot say this. In general *only* the ground-state energy is approximated.

The theorem implies that *truncating a single-particle basis* or *truncating a Slater determinant basis* makes sense.

We will not prove the theorem.

Theorem 2.3 (Cauchy Interlace Theorem). *Let \mathcal{V}_1 and \mathcal{V}_2 be linear spaces, of dimension D_1 and D_2 , respectively. Let $\mathcal{V}_1 \subset \mathcal{V}_2$ be a subspace.*

Let $\{|\Phi_I\rangle\}_{I=1}^{D_2}$ be an orthonormal basis for \mathcal{V}_2 , such that $\{|\Phi_I\rangle\}_{I=1}^{D_1}$ is a basis for \mathcal{V}_1 .

Let $\hat{H} : \mathcal{V}_2 \rightarrow \mathcal{V}_2$ be a Hermitian operator with matrix $H_2 \in \mathbb{C}^{D_2 \times D_2}$, $H_{IJ} = \langle \Phi_I | \hat{H} | \Phi_J \rangle$.

Let H_1 be the projection of \hat{H} onto \mathcal{V}_1 , i.e., the matrix H_1 of this operator is equal to the upper left $D_1 \times D_1$ block of the $D_2 \times D_2$ matrix H_2 .

Let $E_k^{(i)}$ be the D_i eigenvalues of H_i , arranged such that

$$E_k^{(i)} \leq E_{k+1}^{(i)} \quad \forall k. \quad (2.32)$$

Then,

$$E_k^{(2)} \leq E_k^{(1)} \leq E_{k+\delta}^{(2)}, \quad \delta = D_2 - D_1. \quad (2.33)$$

Exercise 2.1. Prove Eq. (??). △

2.3 The Configuration-interaction method (CI)

2.3.1 General description

We now describe an approach to manybody theory called *configuration-interaction theory* (CI). It basically entails truncating both the single-particle basis and the resulting Slater determinant basis according to certain rules.

Let an orthonormal single-particle basis $\{\phi_p\}$ be given, with associated creation operators c_p^\dagger , and corresponding Slater determinants $|\vec{p}\rangle$. Suppose we expand an N -fermion wavefunction in the Slater determinant basis, but *truncate* the expansion, including only a finite subset \mathcal{S} of Slater determinants. The determinants then span a D -dimensional subspace of L_N^2 ,

$$\mathcal{V} = \text{span}\{|\vec{p}\rangle \mid |\vec{p}\rangle \in \mathcal{S}\} \quad (2.34)$$

Equivalently, any wavefunction in V can be written

$$|\Psi\rangle = \sum_{\vec{p} \in \mathcal{S}} A_{\vec{p}} |\vec{p}\rangle, \quad A_{\vec{p}} = \langle \vec{p} | \Psi \rangle. \quad (2.35)$$

The set \mathcal{S} may of course be chosen in many different ways. One typical choice is the set of all possible Slater determinants generated by the first L single-particle functions ϕ_0 through ϕ_{L-1} . This gives a space of dimension $\binom{L}{N}$, and is called the *full configuration-interaction space* (FCI space).

Another typical approach is to have a *reference determinant* $|\Phi\rangle$ and consider particle-hole states on top of that, or excitations in chemistry language.

For example, the one-particle-one-hole space (CI singles, CIS) wavefunction is given by the choice

$$\mathcal{V}_{\text{CIS}} = \text{span}\{|\Phi\rangle, |\Phi_i^a\rangle \mid i = 1, \dots, N, a = N+1, \dots, L\}, \quad (2.36)$$

and any CIS wavefunction can thus be written

$$|\Psi\rangle = A_0 |\Phi\rangle + \sum_{ia} A_i^a |\Phi_{ia}^a\rangle. \quad (2.37)$$

Furthermore, CI singles-and-doubles (CISD) is defined by the space

$$\mathcal{V}_{\text{CISD}} = \text{span}\{|\Phi\rangle, |\Phi_i^a\rangle, |\Phi_{ij}^{ab}\rangle \mid i, j = 1, \dots, N, a, b = N+1, \dots, L\}. \quad (2.38)$$

A wavefunction $|\Psi\rangle \in \mathcal{V}_{\text{CISD}}$ can be written

$$|\Psi\rangle = A_0 |\Phi\rangle + \sum_{ia} A_i^a |\Phi_i^a\rangle + \sum_{i<j} \sum_{a<b} A_{ij}^{ab} |\Phi_{ij}^{ab}\rangle. \quad (2.39)$$

Configuration-interaction singles-doubles-and-triples (CISDT), etc, are defined similarly.

Sometimes, the doubles term is written

$$\sum_{i<j} \sum_{a<b} A_{ij}^{ab} |\Phi_{ij}^{ab}\rangle = \frac{1}{4} \sum_{ij} \sum_{ab} A_{ij}^{ab} |\Phi_{ij}^{ab}\rangle. \quad (2.40)$$

The coefficients satisfy $A_{ij}^{ab} = -A_{ji}^{ab} = -A_{ij}^{ba} = A_{ji}^{ba}$, and the factor 1/4 comes from the fact that $|\Phi_{ij}^{ab}\rangle = -|\Phi_{ji}^{ba}\rangle = -|\Phi_{ji}^{ab}\rangle = |\Phi_{ij}^{ba}\rangle$, i.e., we are deliberately over-counting the basis in this expression to keep notation simple.

Exercise 2.2. Compute the dimension of \mathcal{V}_{CIS} , $\mathcal{V}_{\text{CISD}}$, etc. △

Clearly, indexing the Slater determinants using the vector \vec{p} directly can be cumbersome. Using a different notation, we let $I \in \mathcal{I}$ be an index that enumerates the basis determinants, and write

$$\mathcal{V} = \text{span}\{|\Phi_I\rangle \mid I \in \mathcal{I}\}. \quad (2.41)$$

Our vector expansion becomes

$$|\Psi\rangle = \sum_I A_I |\Phi_I\rangle, \quad A_I = \langle \Phi_I | \Psi \rangle. \quad (2.42)$$

For example, $\mathcal{I} = 1, 2, \dots, D$ is a possibility, with some way of choosing an I for every \vec{p} we are interested in. Or $I = (a, i)$, $I = (ab, ij)$, etc, enumerates the CIS, CISD, etc, hierarchy of spaces.

How do we choose the single-particle functions and the reference state in CI theory? The most common choice in chemistry is to employ a basis of *Hartree-Fock spin-orbitals*. This is the topic of Section ???. A more general picture is as follows: if $\hat{H} = \hat{H}_0 + \hat{W}$, it is also possible to consider \hat{W} a perturbation of \hat{H}_0 , assuming that the eigenstates and eigenvalues of \hat{H}_0 are good approximations to those of the full \hat{H} . (This is also true for the Hartree-Fock paradigm to be considered later.)

Let therefore $\{\phi_p\}$ be a complete set of eigenfunctions for the single-particle operator \hat{h} with eigenvalues ϵ_p arranged in increasing order. Then, the Slater determinants $|\vec{p}\rangle$ are eigenstates of the one-body Hamiltonian $\hat{H}_0 = \sum_i \hat{h}(i)$. Clearly, the determinant

$$|\Phi\rangle = |123 \dots N\rangle \quad (2.43)$$

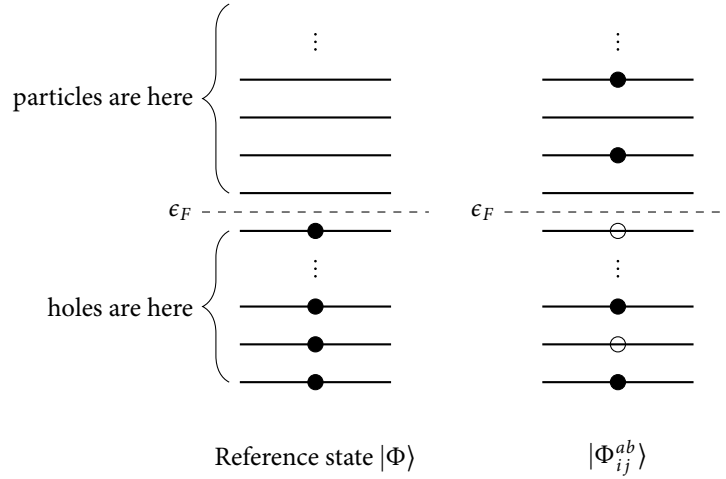


Figure 2.2: Fermi level and quasiparticles. To the left, we have the vacuum state. To the right, we have a doubly excited state, or a two-particle-two-hole-state. Notice how we draw the “Fermi line” between two levels for clarity. In this simple picture, we have assumed that the levels are non-degenerate. If we had spin present, we could fit two particles per level, and so on.

is the ground-state wavefunction of \hat{H}_0 , whose second quantized expression is

$$\hat{H}_0 = \sum_p \epsilon_p c_p^\dagger c_p. \quad (2.44)$$

Note, that if the eigenvalues of \hat{h} are degenerate, then this wavefunction may or may not be unique.

In this picture, the truncated CI scheme as outlined above is a natural approach, since it is reasonable to assume that singles, doubles, etc, will systematically improve upon the “zero-order” wavefunction $|\Phi\rangle$.

In the context of a reference function $|\Phi\rangle$ defined in terms of a zero-order Hamiltonian, such as \hat{H}_0 , it is common to define the *fermi level* ϵ as the energy of the occupied orbital with the highest energy, ϵ_F , assuming that all degenerate levels are included. With this terminology,

$$|\Phi\rangle = \left(\prod_{\epsilon_p \leq \epsilon_F} c_p^\dagger \right) |-\rangle, \quad (2.45)$$

for example. Moreover, we say that a *hole* is “below the Fermi level” and a *particle* is “above the Fermi level”. An excitation excites a fermion from below the Fermi level to above the Fermi level. Thus, the index N is replaced by the one-body *energy* of that level, ϵ_F . See Fig. ??

Notice that the truncated CI scheme favors the description of the ground-state wavefunction.

2.3.2 Matrix elements of the CI method

Having established the parametrization of the approximate wavefunction, a linear space \mathcal{V} , we turn to the variational principle, which tells us (together with the Cauchy Interlace Theorem that) that the matrix of the Hamiltonian \hat{H} with respect to the chosen basis is the central object. Diagonalizing this matrix gives us approximations to the ground-state energy and in total D eigenvalues of the full system.

Thus, in the CI method, we need to diagonalize the matrix $H = [H_{IJ}]$ given by

$$H_{IJ} = \langle \Phi_I | \hat{H} | \Phi_J \rangle \quad (2.46)$$

If we look at the CISD case, the matrix then obtains a block form:

$$H = \begin{pmatrix} \langle \Phi | \hat{H} | \Phi \rangle & \langle \Phi | \hat{H} | \Phi_i^a \rangle & \langle \Phi | \hat{H} | \Phi_{ij}^{ab} \rangle \\ \langle \Phi_{i'}^{a'} | \hat{H} | \Phi \rangle & \langle \Phi_{i'}^{a'} | \hat{H} | \Phi_i^a \rangle & \langle \Phi_{i'}^{a'} | \hat{H} | \Phi_{ij}^{ab} \rangle \\ \langle \Phi_{i'j'}^{a'b'} | \hat{H} | \Phi \rangle & \langle \Phi_{i'j'}^{a'b'} | \hat{H} | \Phi_i^a \rangle & \langle \Phi_{i'j'}^{a'b'} | \hat{H} | \Phi_{ij}^{ab} \rangle \end{pmatrix} \quad (2.47)$$

2.3.3 Computer implementation of CI methods

In chemistry, *speed* and *reliability* are crucial factors. Computations are performed by non-specialists using highly optimized codes like DALTON, MOLPRO, or GAUSSIAN.

We will not try and compete with such codes, of course, but instead indicate how various methods may be implemented.

2.3.4 Naive CI

The simplest approach, which we here call “naive CI”, is to

1. Write down a list of all the Slater determinants in the desired basis,

$$I \mapsto |\Phi_I\rangle.$$

2. Compute all the matrix elements H_{IJ} and store them in computer memory as a big $D \times D$ matrix. This can be done using, say, the Slater–Condon rules (see Exercise ??) that are basically formulae for the matrix elements given in terms of the occupied single-particle functions in $|\Phi_I\rangle$ and $|\Phi_J\rangle$.
3. Use a diagonalization algorithm to find, say, the ground-state energy or other eigenvalues of the matrix.

The biggest problem with this approach, is that the dimension D of the CI space grows pretty fast. The matrix is, in principle, a table with D^2 elements. For FCI, D grows like $\binom{L}{N}$, which very quickly is prohibitive. For CIS, it grows only like $N(L - N)$, but CIS is not that fancy. For CISD, the dimension grows like $N^2(L - N)^2$. We see that the spaces in any case become huge for moderate particle numbers and numbers L of single-particle functions.

2.3.5 Direct CI

More common than “naive CI” is direct CI. For systems of interest, the matrix size grows so quickly that storing the matrix H in memory is out of question. Moreover, diagonalization of dense matrices scales as D^3 , quickly becoming too expensive for practical calculations.

Luckily, we have *iterative algorithms* such as the Lanczos algorithm. These rely only on the *matrix-vector product*. Nowhere is the actual value of H_{IJ} needed, only the action on a vector A_I , i.e., the algorithm needs to compute

$$\vec{A}' = H\vec{A} \quad (2.48)$$

for some input vector \vec{A} . I.e., we must have an algorithm to compute

$$|\Psi'\rangle = P\hat{H}|\Psi\rangle \quad (2.49)$$

where $P = \sum_I |\Phi_I\rangle \langle \Phi_I|$ is the projection operator onto our chosen basis, i.e., we throw away the part of $\hat{H}|\Psi\rangle$ which is not describable in terms of our basis.

It is useful to represent $|\Phi_I\rangle$ in terms of its occupation number vector, a bit string $B = B[I]$. These are integers, and we need a table of these in computer memory. Since our $|\Phi_I\rangle$ must be linearly independent,

there is a one-to-one correspondence between the $B[I]$'s and the I 's, i.e., we can *invert* the table to obtain $I = I[B]$, given B . We write $|B\rangle = |\Phi_{I[B]}\rangle$ for brevity, and we stress that now B is an integer written on binary form.

The central observation is now that, for any string of creation and annihilation operators

$$C_1 C_2 \cdots C_n |B\rangle = \begin{cases} 0 & \text{or} \\ (-1)^s |B'\rangle \end{cases}. \quad (2.50)$$

The result can be found by *manipulating the bits of B and keeping track of the resulting sign*. When B' has been found, the corresponding index I' can be found by searching the bit pattern table. Thus, let us write:

$$\begin{aligned} |\Psi'\rangle &= P \hat{H} |\Psi\rangle = \sum_{B'} |B'\rangle \langle B'| \hat{H} \sum_B A_B |B\rangle \\ &= \sum_B A_B \sum_{B'} |B'\rangle \langle B'| \left(\sum_{pq} h_q^p c_p^\dagger c_q + \frac{1}{4} \sum_{pqrs} w_{rs}^{pq} c_p^\dagger c_q^\dagger c_s c_r \right) |B\rangle \\ &= |\Psi'\rangle = \sum_B A_B \left(\sum_{pq} h_q^p \langle B'| c_p^\dagger c_q |B\rangle + \frac{1}{4} \sum_{pqrs} w_{rs}^{pq} \langle B'| c_p^\dagger c_q^\dagger c_s c_r |B\rangle \right) |B'\rangle \end{aligned} \quad (2.51)$$

This gives us the following algorithm for computing the \hat{H}_0 contribution to $|\Psi'\rangle$ (the \hat{W} part is similar):

1. Initialize $A'_{I'} = 0$ for all I' .
2. Loop over I :
 - (a) Fetch $B = B[I]$.
 - (b) Loop over p, q .
 - i. Compute $c_p^\dagger c_q |B\rangle = 0$ or $(-1)^s |B'\rangle$ by manipulating the bits in B .
 - ii. If the result is nonzero, compute I' such that $B[I'] = B'$ by searching the bit pattern table.
 - iii. If the pattern is found, update $A'_{I'} \leftarrow A'_{I'} + A_I h_q^p (-1)^s$.

Of course, this algorithm is just a sketch. There are many ways to improve it.

How does one search for the index I' in step 2/b/ii? One way is to ensure that the table of bit patterns (integers) are sorted, and then use *binary search*. This requires on average $O(D \log D)$ operations, and since we need to do this $O(D)$ times, this slows down our program drastically. One can also use a *hash map* (e.g., the C++ STL class `std::map<int, int>` can be used). This is no faster.

A *much* faster approach can be taken using *graphical methods*. It is actually possible to find a formula for the inverse map. This formula is $O(1)$, dramatically reducing the computer work for direct CI. For more information on this technique, see Helgaker/Jørgensen/Olsen [?], Section 11.8.

2.3.6 Recipe for bit pattern representation.

How can we perform the bitwise operations mentioned above?

Each Slater determinant $|\mu_1, \dots, \mu_N\rangle$ is, via the occupation numbers, mapped to the bit pattern $|n_0 n_1 n_2 \dots n_L\rangle$ where each $n_\mu \in \{0, 1\}$. We identify the bit pattern with the integer $B[\mu \dots \mu_N]$ it encodes. Thus,

$$\vec{\mu} = \{1, 5, 6\} \mapsto \underbrace{|010001100 \dots 0_2\rangle}_{L \text{ bits}} \mapsto |1 \times 2^1 + 1 \times 2^5 + 1 \times 2^6\rangle = |97_{10}\rangle. \quad (2.52)$$

(But who is thinking in terms of base-10 numbers these days anyway?) All integers between 0 and 2^{L-1} encode all possible Fock space basis functions. A basis for N -fermion space is composed of all the integers whose bit patterns have precisely N bits in total.

Annihilation operator: $c_p |B\rangle$ is either 0 or $(-1)^k |B'\rangle$ for some k and B' . We have the following algorithm:

1. If bit p is not set, return the zero result.
2. Else, compute k as the number of bits set *before* p .
3. Erase bit p to obtain B' .
4. Return the sign $(-1)^k$ and B' .

Creation operator: $c_p^\dagger |B\rangle$ is either 0 or $(-1)^k |B'\rangle$ for some k and B' . We have the following algorithm:

1. If bit p is set, return the zero result.
2. Else, compute k as the number of bits set *before* p .
3. Light bit p to obtain B' .
4. Return the sign $(-1)^k$ and B' .

The product $c_p^\dagger c_q |B\rangle$ can be computed by repeating the above algorithms, and similarly with *any* string of creation and annihilation operators.

Exercise 2.3. We are given $L = 8$ orbitals, numbered $p = 0, 1, \dots, L - 1$, and thus an occupation number representation of length 8 bits, e.g.,

$$|p = 2, p = 3\rangle = |0_0 0_1 1_2 1_3 0_4 0_5 0_6 0_7\rangle = |00110000\rangle. \quad (2.53)$$

Write down the result of the following expressions, on occupation number form. Remember the sign factor:

- a) $c_1^\dagger |01100000\rangle$
- b) $c_5 |01000101\rangle$
- c) $c_4^\dagger c_1^\dagger |01101000\rangle$
- d) $c_1^\dagger |01100000\rangle$
- e) $c_6 |11111111\rangle$
- f) $c_6^\dagger |01111001\rangle$
- g) $c_1^\dagger c_2^\dagger c_3^\dagger |00000000\rangle$
- h) $c_4^\dagger c_4 |11101000\rangle$

△

Exercise 2.4. Write a program that generates all possible bit patterns of length L with N bits set and writes them to screen.

Check that you have the correct number of patterns, $\binom{L}{N}$.

△

Exercise 2.5. (continues exercise ??.) Write a program that correctly creates/annihilates particles from a bit pattern representation $|B\rangle$ of a Slater determinant, returning the proper sign.

△

Exercise 2.6. (continues exercises ?? and ??.) Write a program that, given h_q^p and w_{rs}^{pq} (antisymmetrized or otherwise) as input arrays, computes $\hat{H}|B\rangle$ using direct CI.

△

2.4 Hartree–Fock theory

One of the earliest and most successful approximation methods for many-fermion systems was the *Hartree–Fock method* (HF method). In Hartree–Fock theory we parametrize our wavefunction as a single Slater determinant. However, *the single-particle functions are the unknowns to be determined by the variational procedure.*

Suggested reading for this section: Szabo/Ostlund Ch. 13, Harris/Monkhorst/Freeman Ch. 3, Gross/Runge/Heinonen Ch. 7.

Thus, our wavefunction manifold \mathcal{M} consists of all possible functions on the form

$$|\Phi\rangle = |\phi_1\phi_2\cdots\phi_N\rangle, \quad \langle\phi_i|\phi_j\rangle = \delta_{ij}. \quad (2.54)$$

Note carefully, that a single-particle basis is not given – it is to be found! The task is now to minimize $\langle\Phi|\hat{H}|\Phi\rangle$ (recalling that $\langle\Phi|\Phi\rangle = 1$) under the constraint that the ϕ_i are orthonormalized. When a minimum is found, we denote the solution by $|\Phi_{\text{HF}}\rangle$, the Hartree–Fock state, and the Hartree–Fock energy is correspondingly

$$\begin{aligned} E_{\text{HF}} &\equiv \langle\Phi_{\text{HF}}|\hat{H}|\Phi_{\text{HF}}\rangle = \min_{\phi_1, \dots, \phi_N} \langle\Phi|\hat{H}|\Phi\rangle \\ &= \sum_i \langle\phi_i|\hat{h}|\phi_i\rangle + \frac{1}{2} \sum_{ij} \langle\phi_i\phi_j|\hat{w}|\phi_i\phi_j - \phi_j\phi_i\rangle. \end{aligned} \quad (2.55)$$

as obtained via the Slater–Condon rules, see for example Exercise ??. Here, we use the convention that

$$\langle\phi_p\phi_q|\hat{w}|\phi_r\phi_s\rangle \equiv \int dx_1 \int dx_2 \phi_p(x_1)^* \phi_q(x_2)^* w(x_1, x_2) \phi_r(x_1) \phi_s(x_2), \quad (2.56)$$

which is not antisymmetric. Thus, in total, the matrix element occurring in Eq. (??) is anti-symmetric. We use this notation here to emphasize the dependence of the expectation value on the single-particle functions.

We introduce Lagrange multipliers for the constraints, giving a Lagrangian

$$\begin{aligned} \mathcal{L}(\phi_1, \dots, \phi_N, \lambda) &= \langle\Phi|\hat{H}|\Phi\rangle - \sum_{ij} \lambda_{ij} (\langle\phi_i|\phi_j\rangle - \delta_{ij}) \\ &= \sum_i \langle\phi_i|\hat{h}|\phi_i\rangle + \frac{1}{2} \sum_{ij} \langle\phi_i\phi_j|\hat{w}|\phi_i\phi_j - \phi_j\phi_i\rangle - \sum_{ij} \lambda_{ij} (\langle\phi_i|\phi_j\rangle - \delta_{ij}) \end{aligned} \quad (2.57)$$

Due to symmetry, the Lagrange multiplier matrix λ can be assumed to be Hermitian. Computing a stationary point for the constrained problem is equivalent to an *unconstrained* optimization of \mathcal{L} with respect to the ϕ_i and the Lagrange multipliers.

A word on a special notation: We are going to compute the variation of expressions like $\langle \phi_i \phi_j | \hat{w} | \phi_i \phi_j \rangle$ with respect to ϕ_i . This variation will be a single-particle function. For an arbitrary (possibly unsymmetrized) two-particle function $|U\rangle \in L_1^2 \otimes L_1^2$, and a single-particle function $|v\rangle \in L_1^2$, we define

$$\langle \cdot | v | U \rangle \in L_1^2 \quad (2.58)$$

as the single-particle function obtained by integrating only over the second particle, viz,

$$\langle x_1 | \langle \cdot | v | U \rangle \equiv \int v(x_2)^* U(x_1, x_2) dx_2. \quad (2.59)$$

In other words, given a single-particle basis $\{\chi_\mu\} \subset L_1^2$, the expansion of $\langle \cdot | v | U \rangle$ is given by

$$\langle \cdot | v | U \rangle = \sum_\mu |\chi_\mu\rangle \langle \chi_\mu | v | U \rangle, \quad \langle \chi_\mu | v | U \rangle = \int \int \chi_\mu(x_1)^* v(x_2)^* U(x_1, x_2) dx_1 dx_2. \quad (2.60)$$

We now state and prove our result:

Theorem 2.4 (Hartree–Fock equations). *The single-particle functions of the Hartree–Fock state $|\Phi_{HF}\rangle$ satisfy the nonlinear eigenvalue problem*

$$\hat{f}(\phi_1, \dots, \phi_N) |\phi_i\rangle = \epsilon_i |\phi_i\rangle, \quad (2.61)$$

where

$$\hat{f}(\phi_1, \dots, \phi_N) = \hat{h} + \hat{v}^{direct} - \hat{v}^{exchange}, \quad (2.62)$$

with

$$\hat{v}^{direct} |\psi\rangle = \sum_j \langle \cdot | \phi_j | \hat{w} | \psi \phi_j \rangle. \quad (2.63)$$

and

$$\hat{v}^{exchange} |\psi\rangle = \sum_j \langle \cdot | \phi_j | \hat{w} | \phi_j \psi \rangle. \quad (2.64)$$

The equations (??) are referred to as “the Hartree–Fock equations”. The operator \hat{f} in Eq. (??) is “the Fock operator”, and \hat{v}^{direct} and $\hat{v}^{exchange}$ are the direct- and exchange potential, respectively.

Proof. Since for the Hartree–Fock state the energy \mathcal{E} is a minimum, we need to show that the energy has zero variation under all the possible variations of the ϕ_i that preserve orthonormality $\langle \phi_i | \phi_j \rangle = \delta_{ij}$. In other words that the Lagrangian \mathcal{L} defined in Eq. (??) has a zero variation under *unconstrained* variation, allowing also those that destroy orthonormality. We also need to identify the Lagrange multipliers, and they will turn out to be related to the eigenvalues ϵ_i .

As in the proof for the variational principle, we can treat ϕ_i and ϕ_i^* as independent variables, varying only the bra states. Let ϵ be a small real number, and let η_i be arbitrary normalized single-particle functions. We consider a variation $\delta\phi_i = \epsilon\eta_i$, and define the function

$$f(\epsilon) = \mathcal{L}(\phi_i + \epsilon\eta_i, \lambda). \quad (2.65)$$

The variational equations are equivalent to $f'(0) = 0$ for all choices of all the η_i .

We compute

$$f(\epsilon) = \mathcal{E}(\phi_1, \dots, \phi_N) + \epsilon f'(0) + O(\epsilon^2) \quad (2.66)$$

where $f'(0)$ is obtained as

$$f'(0) = \sum_i \langle \eta_i | \hat{h} | \phi_i \rangle + \frac{1}{2} \sum_{ij} [\langle \eta_i \phi_j | \hat{w} | \phi_i \phi_j \rangle - \langle \eta_i \phi_j | \hat{w} | \phi_j \phi_i \rangle + \langle \phi_i \eta_j | \hat{w} | \phi_i \phi_j \rangle - \langle \phi_i \eta_j | \hat{w} | \phi_j \phi_i \rangle] - \sum_{ij} \lambda_{ij} \langle \eta_i | \phi_j \rangle. \quad (2.67)$$

By symmetry,

$$f'(0) = \sum_i \langle \eta_i | \hat{h} | \phi_i \rangle + \sum_{ij} [\langle \eta_i \phi_j | \hat{w} | \phi_i \phi_j \rangle - \langle \eta_i \phi_j | \hat{w} | \phi_j \phi_i \rangle] - \sum_{ij} \lambda_{ij} \langle \eta_i | \phi_j \rangle. \quad (2.68)$$

Equating $f'(0) = 0$ gives us the stationary conditions, which must hold for *all* η_i choices. In particular, setting all $\eta_j = 0$, $j \neq i$, and $\eta_i = \chi_\mu$, a single-particle basis function, we obtain the equation:

$$\langle \chi_\mu | \hat{h} | \phi_i \rangle + \sum_j [\langle \chi_\mu \phi_j | \hat{w} | \phi_i \phi_j \rangle - \langle \chi_\mu \phi_j | \hat{w} | \phi_j \phi_i \rangle] - \sum_j \lambda_{ij} \langle \chi_\mu | \phi_j \rangle = 0 \quad (2.69)$$

We recognize this as the single-particle equation

$$\hat{h} | \phi_i \rangle + \sum_j [\langle \cdot | \phi_j | \hat{w} | \phi_i \phi_j \rangle - \langle \cdot | \phi_j | \hat{w} | \phi_j \phi_i \rangle] - \sum_j \lambda_{ij} | \phi_j \rangle. \quad (2.70)$$

The orbitals are orthonormal. The determinant $|\Phi\rangle$ is invariant (up to an irrelevant phase) under a unitary mixing of the single-particle functions, i.e, if we let

$$\tilde{\phi}_k = \sum_j \phi_j U_{jk}^* \quad (2.71)$$

with U a unitary matrix, then $|\tilde{\Phi}\rangle = \det(U)^* |\Phi\rangle$, i.e., the same state, and clearly the energy is invariant.

Select U such that $U^H \lambda U = E$ is diagonal, i.e., $\lambda = U E U^H$,

$$\lambda_{ij} = \sum_k U_{ik} E_k U_{jk}^*. \quad (2.72)$$

Inserted into the equation (??), we get

$$\hat{h} | \tilde{\phi}_i \rangle + \sum_j [\langle \cdot | \tilde{\phi}_j | \hat{w} | \tilde{\phi}_i \tilde{\phi}_j \rangle - \langle \cdot | \tilde{\phi}_j | \hat{w} | \tilde{\phi}_j \tilde{\phi}_i \rangle] = E_i | \tilde{\phi}_i \rangle. \quad (2.73)$$

This must hold for all $i = 1, \dots, N$ simultaneously.

With the definitions of \hat{v}^{direct} and $\hat{v}^{\text{exchange}}$ in the theorem formulation, we are finished. \square

Note that the theorem does not guarantee that the solutions to the HF equations correspond to a the actual HF solution, i.e., a global minimum, or even a local minimum. It could well be a saddle point. Indeed, it has been found that the standard algorithms for the HF equations sometimes give local minima **NB: insert citation.**

Let us consider the unfamiliar operators \hat{v}^{direct} and $\hat{v}^{\text{exchange}}$ in some detail. To this end, suppose that the two-body operator is a local potential $\hat{w}(x_1, x_2)$, such as the Coulomb potential

$$\hat{w}_{\text{Coul}}(x_1, x_2) = \frac{1}{|\vec{r}_1 - \vec{r}_2|}, \quad x_i = (\vec{r}_i, \sigma). \quad (2.74)$$

The operator \hat{v}^{direct} is a one-body operator. When acting on a one-body function $|\psi\rangle$ it produces a new one-body function, which at x_1 takes the value

$$\begin{aligned} \langle x_1 | (\hat{v}^{\text{direct}} |\psi\rangle) &= \sum_j \langle x_1 | \langle \cdot | \phi_j | \hat{w} | \psi \phi_j \rangle = \sum_j \int \phi_j^*(x_2) w(x_1, x_2) \phi_j(x_2) \psi(x_1) dx_2 \\ &= \left[\int \sum_j |\phi_j(x_2)|^2 w(x_1, x_2) dx_2 \right] \psi(x_1) \equiv v^{\text{direct}}(x_1) \psi(x_1). \end{aligned} \quad (2.75)$$

Thus, \hat{v}^{direct} is a *local potential*, given by a sort of average of $w(x_2, x_1)$ over x_2 , weighted by $\rho(x) \equiv \sum_j |\phi_j(x)|^2$, giving a “mean-field potential”.

The operator $\hat{v}^{\text{exchange}}$ is, however, *non-local*: the value $\langle x_1 | (\hat{v}^{\text{exchange}} |\psi\rangle)$ depends on $\psi(x_2)$ in every point x_2 . To see this, we compute

$$\langle x_1 | (\hat{v}^{\text{exchange}} |\psi\rangle) = \sum_j \langle x_1 | (\cdot \phi_j | \hat{w} | \phi_j \psi) = \sum_j \int \phi_j^*(x_2) w(x_1, x_2) \psi(x_2) \phi_j(x_1) dx_2. \quad (2.76)$$

The operator $\hat{v}^{\text{exchange}}$ is still *linear* when acting on $|\psi\rangle$, it is just not interpretable as a local potential.

If we introduce the *reduced one-particle density matrix* $\gamma(x_1, x_2)$ as

$$\gamma(x, x') = \sum_j \phi_j(x) \phi_j(x')^*, \quad (2.77)$$

we can express

$$\langle x_1 | \hat{v}^{\text{direct}} |\psi\rangle = \psi(x_1) \int \gamma(x_2, x_2) w(x_1, x_2) dx_2. \quad (2.78)$$

$$\langle x_1 | \hat{v}^{\text{exchange}} |\psi\rangle = \int \gamma(x_2, x_1) w(x_1, x_2) \psi(x_2) dx_2. \quad (2.79)$$

In the proof of the HF equations, we first found an equation whose solutions were not eigenfunctions, Eq. (??). However, by forming a particular linear combination, the equation was brought on eigenvalue form, Eq. (??). We realized that the HF single-particle functions were not unique; any unitary transformation among the orbitals produces the *same* $|\Phi_{\text{HF}}\rangle$.

The HF equations are a set of eigenvalue equations that are nonlinear in the eigenvectors. Thus, the equations need to be solved *self-consistently*. The fermions experience an averaged interaction from the other electrons – hence, we often call HF theory for *mean-field theory*.

We only used the N first eigenvectors of \hat{f} to construct our HF wavefunction. But clearly, when these have been found, $\hat{f} = \hat{f}(\phi_1, \dots, \phi_N)$ is “just like” any other one-body operator found so far, and we can in principle¹ find a complete basis

$$\{\phi_p\} = \{\phi_i\} \cup \{\phi_a\}, \quad (2.80)$$

satisfying

$$\hat{f} |\phi_p\rangle = \epsilon_p |\phi_p\rangle. \quad (2.81)$$

This particular orthonormal basis is often taken as basis for CI treatments.

We now show an interesting relation for the Hartree–Fock energy. It is tempting to assume that $E_{\text{HF}} = \sum_i \epsilon_i$. However, this is not the case.

Theorem 2.5 (Energy expression for Hartree–Fock). *Assume that a solution to the canonical Hartree–Fock equations have been found. The Hartree–Fock energy is given by*

$$E_{\text{HF}} = \sum_i \epsilon_i - \frac{1}{2} \sum_{ij} \langle \phi_i \phi_i | \hat{w} | \phi_i \phi_j - \phi_j \phi_i \rangle. \quad (2.82)$$

Proof. Multiply the HF equation from the left by $\langle \phi_i |$ and sum over i to obtain

$$\sum_i \epsilon_i = \sum_i \langle \phi_i | \hat{h} | \phi_i \rangle + \sum_{ij} \langle \phi_i \phi_j | \hat{w} | \phi_i \phi_j - \phi_j \phi_i \rangle. \quad (2.83)$$

We see that the interaction is double counted compared to Eq. (??), and we are finished. \square

¹It happens that \hat{f} has a continuous spectrum, so our statement must really be limited to finite-dimensional one-particle spaces for strict validity.

Exercise 2.7. We show that the reduced one-particle density matrix is the same for canonical and non-canonical orbitals: Let U be a unitary matrix and define

$$\tilde{\phi}_j = \sum_i \phi_i U_{ji}. \quad (2.84)$$

Show that

$$\gamma(x, x') = \sum_j \tilde{\phi}_j(x) \tilde{\phi}_j^*(x'). \quad (2.85)$$

What can you conclude about \hat{v}^{direct} and $\hat{v}^{\text{exchange}}$? \triangle

2.4.1 The Hartree–Fock equations in a given basis

How do we solve the HF equations (??)? In this section, we reformulate the HF equations relative to a fixed basis, $\{\chi_p\}_{p=1}^L$. For practical reasons, of course, the basis must have a finite size L . We expand our HF functions as

$$|\phi_p\rangle = \sum_q |\chi_q\rangle U_{qp}, \quad (2.86)$$

where U is a unitary matrix. We notice that the *columns* are the basis expansions of each ϕ_p . We write u_p for column number p , $|\phi_p\rangle = \sum_q |\chi_q\rangle (u_p)_q$.

The reduced density matrix becomes

$$\gamma = \sum_i |\phi_i\rangle \langle \phi_i| = \sum_{pq} \sum_i U_{qi} |\chi_q\rangle \langle \chi_p| U_{pi}^* = \sum_{pq} \left(\sum_i U_{qi} U_{pi}^* \right) |\chi_q\rangle \langle \chi_p| = \sum_{pq} (U_{1:N} U_{1:N}^\dagger)_{qp} |\chi_q\rangle \langle \chi_p|, \quad (2.87)$$

and it makes sense to define

$$D = U_{1:N} U_{1:N}^H = \sum_i u_i u_i^H, \quad (2.88)$$

which we interpret as the reduced density matrix relative to the given basis $\{\chi_p\}$, depending on the N first columns of U only.

We now demonstrate how the HF equations (??) can be written

$$F(D)u_p = \epsilon_p u_p, \quad (2.89)$$

where

$$F_{pq} = \langle \chi_p | \hat{f}(\gamma) | \chi_q \rangle \quad (2.90)$$

are the matrix elements of the Fock operator in the fixed basis.

Let us look at the matrix elements of f ,

$$F_{qp} = \langle \chi_q | \hat{f} | \chi_p \rangle = \langle \chi_q | \hat{h} | \chi_p \rangle + \langle \chi_q | \hat{v}^{\text{direct}} | \chi_q \rangle - \langle \chi_q | \hat{v}^{\text{exchange}} | \chi_p \rangle. \quad (2.91)$$

The direct term is

$$\begin{aligned} \langle \chi_q | \hat{v}^{\text{direct}} | \chi_p \rangle &= \sum_j \langle \chi_q \phi_j | \hat{w} | \chi_p \phi_j \rangle = \sum_{p'q'j} U_{jq'} U_{jp'}^* \langle \chi_q \chi_{q'} | \hat{w} | \chi_p \chi_{p'} \rangle \\ &= \sum_{p'q'} D_{q'p'} \langle \chi_q \chi_{q'} | \hat{w} | \chi_p \chi_{p'} \rangle. \end{aligned} \quad (2.92)$$

Correspondingly,

$$\begin{aligned}\langle \chi_q | \hat{V}^{\text{exchange}} | \chi_p \rangle &= \sum_j \langle \chi_q \phi_j | \hat{w} | \phi_j \chi_p \rangle = \sum_{p'q'j} U_{jq'} U_{jp'}^* \langle \chi_q \chi_{q'} | \hat{w} | \chi_{p'} \chi_p \rangle \\ &= \sum_{p'q'} D_{q'p'} \langle \chi_q \chi_{q'} | \hat{w} | \chi_{p'} \chi_p \rangle.\end{aligned}\quad (2.93)$$

We obtain

$$F_{qp} = \langle \chi_q | \hat{h} | \chi_p \rangle + \sum_{p'q'} D_{p'q'} (\langle \chi_q \chi_{q'} | \hat{w} | \chi_p \chi_{p'} \rangle - \langle \chi_q \chi_{q'} | \hat{w} | \chi_{p'} \chi_p \rangle). \quad (2.94)$$

Note that we have expressed F_{qp} in terms of *non-antisymmetric* matrix elements of \hat{w} .

2.4.2 Self-consistent field iteration

How do we find self-consistent solutions of Eq. (??)? The standard approach is by self-consistent field iterations (SCF iterations), Finding hopefully better and better approximations $u_i^{(k)}$, $k = 1, 2, 3, \dots$, to the canonical HF functions, starting from a well-selected initial guess $u_i^{(0)}$.

Let $D^{(k)} = \sum_i u_i^{(k)} (u_i^{(k)})^H$ be the k 'th iteration's density matrix. Then, the basic SCF iteration is to compute a complete set of orthonormal vectors

$$F(D^{(k)}) u_p^{(k+1)} = \epsilon_p^{(k+1)} u_p^{(k+1)} \quad (2.95)$$

by numerical diagonalization, sorting the eigenvalues $\epsilon_p^{(k+1)}$ in ascending order. Then, $p = 1, \dots, N$ gives the next approximation to the HF eigenpairs (ϕ_i, ϵ_i) , while the next $L - N$ form additional basis vectors to be used in, say, CI calculation.

If the SCF iteration converges, it often converges to a solution that corresponds to the true HF minimum wavefunction. Sometimes it does not converge to the true solution, but is still useful. Sometimes it does not converge at all, and one needs to “fix” the SCF iteration.

In fact, the basic SCF iteration has very problematic convergence properties. The most common scheme today is the so-called *direct inversion in the iterative subspace* iteration (DIIS), but this is out of scope for the present course. Read more about DIIS in Helgaker/Jørgensen/Olsen [?], and see also <https://en.wikipedia.org/wiki/DIIS>.

Exercise 2.8. NB: To be completed Change-of-basis and second quantization.

...

△

Exercise 2.9. NB: To be completed Second-quantization reformulation of HF equations that gives Brillouin's Theorem directly as the stationary condition.

△

2.4.3 The fluctuation potential

We have now established the canonical Hartree–Fock orbitals, that can be used as a basis just like any other orthonormal basis. Each canonical orbital ϕ_p is associated with a creation operator c_p^\dagger , and in terms of the *original* basis $\{\chi_p\}$ we have for a two-body operator

$$\langle pq|\hat{w}|rs\rangle = \langle \phi_p\phi_q|\hat{w}|\phi_r\phi_s\rangle_{\text{AS}} = \sum_{p'q'r's'} U_{p'p}^* U_{q'q}^* U_{r'r} U_{s's} \langle \chi_{p'}\chi_{q'}|\hat{w}|\chi_{r'}\chi_{s'}\rangle_{\text{AS}} \quad (2.96)$$

and

$$\langle p|\hat{h}|q\rangle = h_q^p = \langle \phi_p|\hat{h}|\phi_q\rangle = \sum_{p'q'} U_{p'p}^* U_{q'q} \langle \chi_{p'}|\hat{h}|\chi_{q'}\rangle, \quad (2.97)$$

and similarly for any one-body operator.

It is convenient to write the Hamiltonian on the following form

$$\hat{H} = \hat{H}_0 + \hat{W} = \hat{F} + \hat{U}, \quad (2.98)$$

where the *Fock operator* is given by

$$\hat{F} = \sum_{i=1}^N \hat{f}(i) = \hat{H}_0 + \hat{V}^{\text{direct}} - \hat{V}^{\text{exchange}}, \quad (2.99)$$

and where the *fluctuation potential* is given by

$$\hat{U} = \hat{W} - \hat{V}^{\text{direct}} + \hat{V}^{\text{exchange}}. \quad (2.100)$$

Here,

$$\hat{V}^{\text{direct}} = \sum_i \hat{v}^{\text{direct}}(i), \quad \hat{V}^{\text{exchange}} = \sum_i \hat{v}^{\text{exchange}}(i). \quad (2.101)$$

The fluctuation potential is so named, because if one considers the HF solution as a reference $|\Phi\rangle$ (and now we drop the “HF” subscript), “most” of the interactions between the particles in $|\Phi\rangle$ are described by the Fock operator, and \hat{U} should be “small” causing the exact wavefunction $|\Psi\rangle = |\Phi\rangle + \delta|\Psi\rangle$ to describe “small fluctuations” on top of $|\Phi\rangle$.

Indeed, the Slater determinant $|\Phi\rangle$ is the ground-state eigenfunction of \hat{F} with eigenvalue $\sum_i \epsilon_i$. Moreover, the Slater determinants $|\phi_{p_1}\cdots\phi_{p_N}\rangle$ diagonalize \hat{F} with eigenvalues $\sum_i \epsilon_{p_i}$.

An expression for the direct potential operator matrix element is

$$\langle \phi_q|\hat{V}^{\text{direct}}|\phi_p\rangle = \sum_i \langle \phi_i\phi_q|\hat{w}|\phi_i\phi_p\rangle, \quad (2.102)$$

with non-antisymmetric matrix elements. Thus,

$$\hat{V}^{\text{direct}} = \sum_{pq} \sum_i \langle \phi_i\phi_q|\hat{w}|\phi_i\phi_p\rangle c_q^\dagger c_p. \quad (2.103)$$

Similarly, for the the exchange potential we get

$$\hat{V}^{\text{exchange}} = \sum_{pq} \sum_i \langle \phi_i\phi_q|\hat{w}|\phi_p\phi_i\rangle c_q^\dagger c_p. \quad (2.104)$$

This results in (using antisymmetrized matrix elements (??))

$$\hat{F} = \hat{H}_0 + \sum_{pq} \sum_i \langle qi|\hat{w}|pi\rangle c_q^\dagger c_p \quad (2.105)$$

$$\hat{U} = \hat{W} - \sum_{pq} \sum_i \langle qi|\hat{w}|pi\rangle c_q^\dagger c_p. \quad (2.106)$$

Having dealt with the second-quantized form of the Hartree–Fock partitioned Hamiltonian, let us turn to the Slater determinants. Since the c_p^\dagger are creation operators for the canonical HF orbitals, a basis of Slater determinants can be taken to be the $|p_1 \cdots p_N\rangle$, with $p_1 < p_2 < \cdots < p_N$ as before. Alternatively, we can use the quasiparticle picture, and let the HF function be the reference,

$$|\Phi\rangle = c_1^\dagger \cdots c_N^\dagger |-\rangle, \quad (2.107)$$

all other Slater determinant basis functions can be written

$$|\Phi_i^a\rangle = c_a^\dagger c_i |\Phi_{\text{HF}}\rangle = b_a^\dagger b_i^\dagger |\Phi_{\text{HF}}\rangle, \quad (2.108)$$

$$|\Phi_{ij}^{ab}\rangle = c_b^\dagger c_j c_a^\dagger c_i |\Phi_{\text{HF}}\rangle = b_b^\dagger b_j^\dagger b_a^\dagger b_i^\dagger |\Phi_{\text{HF}}\rangle, \quad (2.109)$$

etc, where we have introduced the quasiparticle creation- and annihilation operators.

In quantum chemistry, it is *standard* to start with the HF orbitals and perform corrections on top of that, such as CISD, giving rise to the term “post-Hartree–Fock methods”.

What is special about the HF reference, is of course that it is chosen to be optimize a certain aspect of the basis, namely that the reference state has minimal energy. This has an important consequence, namely *Brillouin’s Theorem*:

Theorem 2.6 (Brillouin’s Theorem). *Using HF single-particle basis,*

$$\langle \Phi_i^a | \hat{H} | \Phi \rangle = 0, \quad \forall i, a. \quad (2.110)$$

Proof. The Fock matrix is diagonal in the HF orbitals,

$$f_q^p = h_q^p + \sum_j \langle p j | \hat{w} | q j \rangle = \delta_{pq} \epsilon_q. \quad (2.111)$$

In particular, the off-diagonal elements vanish,

$$f_i^a = h_i^a + \sum_j \langle a j | \hat{w} | i j \rangle = 0. \quad (2.112)$$

But this is precisely (see the Slater–Condon rules from Exercise ??) the expression for $\langle \Phi_i^a | \hat{H} | \Phi \rangle$, which therefore must vanish for all i, a . □

2.4.4 Normal-ordered Hamiltonian in HF basis

Recall that for a two-body Hamiltonian $\hat{H} = \hat{H}_0 + \hat{W}$, the normal-ordered Hamiltonian (with respect to quasiparticles) was

$$\hat{H} = E_0 + \hat{H}_{0,\text{N}} + \hat{W}_{\text{N}}, \quad (2.113)$$

with

$$E_0 = \sum_i h_i^i + \frac{1}{2} \sum_{ij} \langle ij | \hat{w} | ij \rangle \quad (2.114)$$

$$\hat{H}_{0,\text{N}} = \sum_{pq} (h_q^p + \sum_j \langle p j | \hat{w} | q j \rangle) N(c_p^\dagger c_q), \quad (2.115)$$

$$\hat{W}_{\text{N}} = N(\hat{W}) = \frac{1}{4} \sum_{pqrs} \langle pq | \hat{w} | rs \rangle N(c_p^\dagger c_q^\dagger c_s c_r). \quad (2.116)$$

Each of the operators with subscript “N” is thus normal-ordered with respect to quasiparticle vacuum, thereby simplifying many formulas and manipulations.

Suppose now our single-particle basis is the HF basis. Looking carefully at the above equations, and recalling the operator $N(\cdot)$ is defined for linear combinations of strings, we recognize that

$$E_0 = E_{\text{HF}}, \quad \hat{H}_{0,N} = N(\hat{F}), \quad \text{and} \quad \hat{W}_N = N(\hat{U}). \quad (2.117)$$

Thus, using HF orbitals, the normal-ordered Hamiltonian takes on a particularly simple form:

$$\hat{H} = \hat{F} + \hat{U} = E_{\text{HF}} + N(\hat{F}) + N(\hat{U}), \quad (2.118)$$

where we recall that the normal-ordering operator is relative to quasiparticle vacuum. Here, the quasiparticle reference is the HF state $|\Psi_{\text{HF}}\rangle = |\Phi\rangle$. Recall, that the normal-ordering operator is defined linear combinations of strings,

$$N(\hat{F}) = N\left(\sum_{pq} f_q^p c_p^\dagger c_q\right) = \sum_{pq} f_q^p N(c_p^\dagger c_q). \quad (2.119)$$

But beware! In general, $N(\hat{H}_0) \neq \hat{H}_{0,N}$! The operator $\hat{H}_{0,N}$ depends on the whole Hamiltonian, i.e., also the two-body interaction. It is just that in *the particular case of the HF partitioning* of the Hamiltonian, $N(\hat{F}) = \hat{F}_N$.

We now also use the fact that \hat{F} is diagonal in the HF basis,

$$\hat{F} = \sum_p \epsilon_p c_p^\dagger c_p. \quad (2.120)$$

This gives a considerable simplification, since

$$N(\hat{F}) = \sum_p \epsilon_p N(c_p^\dagger c_p) = \sum_a \epsilon_a b_a^\dagger b_a - \sum_i \epsilon_i b_i^\dagger b_i. \quad (2.121)$$

Exercise 2.10. Set up the CISD formalism using L Hartree–Fock orbitals. Use the normal-ordered Hamiltonian. Compute the matrix elements $\langle \Phi | \hat{H} | \Phi \rangle$, $\langle \Phi_i^a | \hat{H} | \Phi \rangle$, $\langle \Phi_{ij}^{ab} | \hat{H} | \Phi \rangle$, $\langle \Phi_i^a | \hat{H} | \Phi_k^c \rangle$, $\langle \Phi_i^a | \hat{H} | \Phi_{kl}^{cd} \rangle$, and $\langle \Phi_{ij}^{ab} | \hat{H} | \Phi_{kl}^{cd} \rangle$. Use Wick's Theorem for quasiparticle operators to achieve this. (One could also use the Slater–Condon rules, but this exercise is about quasiparticles and normal-ordered operators.) \triangle