Lecture notes for FYS–KJM 4480

Quantum mechanics for many-particle systems

Simen Kvaal

October 9, 2015

Contents

Chapter 1

Fundamental formalism

Suggested reading for this chapter: Raimes [1], sections 1.1–1.3, and Gross/Runge/Heinonen [2], section I.1. Chapter 1 of Szabo/Ostlund [3] 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:

$$\Psi = \Psi(x_1, x_2, \cdots, x_N), \tag{1.1}$$

where x_i is a point in the configuration space X, the space where each particle "lives". The configuration space[1] for all N particles is thus X^N, and

$$\Psi : X^N \longrightarrow \mathbb{C}. \tag{1.2}$$

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 has a four-component wavefunction, and N nucleons 4^N components.

[1] Since 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 [4].

Remark: Mathematically, X is a *measure space*, which means that a function $\psi : X \to \mathbb{C}$ 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 $S = \{s_1, s_2, \cdots, s_n\}$ 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: $P(x_1, \cdots, x_N) = |\Psi(x_1, x_2, \cdots, x_N)|^2$ is the probability density for locating all particles at the point $(x_1, \cdots, x_N) \in X^N$. Therefore, Ψ must be square integrable, and be in the Hilbert space $L^2(X^N)$,

$$\Psi \in L^2(X^N). \tag{1.3}$$

All physics can be obtained from the state Ψ.

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

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

The system Hamiltonian $\hat{H}$ is obtained from its classical counterpart (if such exists) by a procedure called *Weyl quantization NB:* 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),

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

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

The time-independent Schödinger 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 $\psi : X \to \mathbb{C}$ and

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

Let

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

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 $\psi : X \to \mathbb{C}$, i.e., functions from the integers $1, \ldots, L$ 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 approximation 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 inter-electronic 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) interacts via the Coulomb force:

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

Thus,

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

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

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

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 $f_1 + f_2 + \cdots f_N$, $f_i = f_i(x_i)$. This can only sum to a constant if $f_i(x_i)$ is a constant,

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

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

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

with eigenvalue

$$E = \epsilon_{i_1} + \cdots + \epsilon_{i_N}. \tag{1.16}$$

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 paragraphhs 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 $(x_1, \cdots, x_N) \in X^N$ be a configuration of the N particles. Then we must have

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

This is equivalent to

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

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

Define a linear operator $\hat{P}_\sigma$ via

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

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

One can show (see the exercises), that either $P_\sigma\Psi = \Psi$ for every $\sigma \in S_N$, or $P_\sigma\Psi = (-1)^{|\sigma|}\Psi$ 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 occuring in quantum theory (in three-dimensional space) are 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 [4], "The physical consequences of this postulate seem 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., [2] 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 $\hat{P}_\sigma$, then the eigenvalue is either 1 or $(-1)^{|\sigma|}$.

We introdice transpositions: $\tau \in S_N$ is transposition if it exhanges 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. △

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 $\Psi \in L^2(X^N)$,

$$\hat{H}P_\sigma\Psi = P_\sigma\hat{H}\Psi. \tag{1.20}$$

△

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 with respect to particle permutations?

Is the function totally antisymmetric with respect to particle permutations?

Is the function square integrable?

△

1.1.7 Slater determinants

The set of totally antisymmetric wavefunctions $L^2(X^N)_{\rm AS}$ in $L^2(X^N)$ form a *closed subspace* of Hilbert space: it is a linear space which is complete. Thus $L^2(X^N)_{\rm AS}$ 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 $L^2(X^N)_{\rm AS}$ is a linear space. Additionally, if you have the mathematical background, prove that it is a closed subspace using the Hilbert space metric. △

The original space has a tensor product representation:

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

Here, $L^2(X)$ is the Hilbert space of a single fermion. Let us assume that we have an orthonormal basis (ONB) $\phi_1, \phi_2, \cdots$, for this space, such that we can expand any $\psi \in L^2(X)$ as

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

with

$$\langle \phi_\mu | \phi_\nu \rangle = \delta_{\mu,\nu} \tag{1.23}$$

and

$$\|\psi\|^2 = \sum_\mu |c_\mu|^2. \tag{1.24}$$

Thus, $\psi(x)$ is represented by an (infinite) vector $[c_\mu] = (c_1, c_2, \cdots)$. Because of Eq. (1.21), we may construct a basis for $L^2(X^N)$ by *tensor products*,

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

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

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

with

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

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.

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

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

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, $L^2(X^N)_{\text{AS}}$? 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

$$\Phi(x_1, x_2) = \phi_1(x_1)\phi_2(x_2) - \phi_2(x_1)\phi_1(x_2), \tag{1.29}$$

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

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

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 combniation is zero as well.

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

Definition 1.1. Let $\phi_1, \phi_2, \ldots, \phi_N$ 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

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

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. (1.31) are equivalent. This requires some manipulation of permutations. △

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

$$\psi_k = \sum_j \phi_j A_{jk}. \tag{1.32}$$

Prove that

$$[\psi_1, \psi_2, \cdots, \psi_N] = \det(A)[\phi_1, \phi_2, \cdots, \phi_N]. \tag{1.33}$$

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

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 $\{\phi_\mu\}$, $\mu = 1, 2, \cdots$ are orthonormal. Prove that $\Phi_{\mu_1,\cdots,\mu_N} = [\phi_{\mu_1}\phi_{\mu_2},\cdots,\phi_{\mu_N}]$ is normalized,

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

Prove that

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

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? △

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

$$[\phi_1,\cdots,\phi_i,\cdots,\phi_j,\cdots,\phi_N] = -[\phi_1,\cdots,\phi_j,\cdots,\phi_i,\cdots,\phi_N] \tag{1.34}$$

$$[\phi_1,\cdots,\phi_N](x_1,\cdots,x_i,\cdots,x_j,\cdots,x_N) = -[\phi_1,\cdots,\phi_N](x_1,\cdots,x_j,\cdots,x_i,\cdots,x_N). \tag{1.35}$$

Moreover, two equal rows (i.e., equal function indices) means that two of the single-particle functions are identical, giving a vanishing determinant. If two *columns* in Eq. (1.31) are identical, the determinant vanishes. Two columns equal mean that we evaluate at some $x_i = x_j$. *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*

$$[\phi_{\mu_1},\phi_{\mu_2},\cdots,\phi_{\mu_N}]. \tag{1.36}$$

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

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

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

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

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

Projecting the two last inequalities onto $\tilde{\Phi}_{\nu_1,\cdots,\nu_N}$ gives

$$c_{\nu_1,\cdots,\nu_j,\cdots,\nu_i,\cdots,\nu_N} = -c_{\nu_1,\cdots,\nu_i,\cdots,\nu_j,\cdots,\nu_N}. \tag{1.39}$$

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

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

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

Note that we can write

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

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

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

This in fact proves that the Slater determinants, when we only use ordered indices, are sufficient to expand any $\Psi L^2(X^N)_{\mathrm{AS}}$. 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!) Thus, the Slater determinants with ordered indices form a basis. □

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. △

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

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

Now,

$$[\phi_1,\cdots,\phi_N] = \sqrt{N!}\mathscr{A}\phi_1(x_1)\cdots\phi_N(x_N). \tag{1.44}$$

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

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

1.2 Second quantization

1.2.1 The creation and annihilation operators

In this section, we introduce the following shorthand:

$$L^2_N \equiv L^2(X^N)_{\mathrm{AS}} \tag{1.45}$$

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^2_N could be formed from an orthonormal basis $\{\phi_\mu\}$ of $L^2(X)$, by computing a set of Slater determinants $\Phi_{\mu_1,\cdots,\mu_N} = [\phi_{\mu_1},\cdots,\phi_{\mu_N}]$, where $\mu_1 < \mu_2 < \cdots < \mu_N$ 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 $[\phi_{\mu_1},\cdots,\phi_{\mu_N}]$ were *functions*, but in quantum mechanics the bra/ket notation is useful. We therefore introduce the ket notation

$$|\psi_1,\cdots,\psi_N\rangle = [\psi_1,\cdots,\psi_N] \tag{1.46}$$

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

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

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

$$\hat{P}_{ij}\,|\mu_1\cdots\mu_i\cdots\mu_j\cdots\mu_N\rangle = -\,|\mu_1\cdots\mu_j\cdots\mu_i\cdots\mu_N\rangle \tag{1.48}$$

and more generally

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

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

$$|\Psi\rangle = \sum_{\vec{\mu}}{}^{\sim}\,|\vec{\mu}\rangle\,\langle\vec{\mu}|\Psi\rangle \tag{1.50}$$

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 (1.31), 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):

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

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:

$$c^\dagger_\nu : L^2_N \to L^2_{N+1} \tag{1.52}$$

$$c_\nu : L^2_N \to L^2_{N-1} \tag{1.53}$$

The operator $c^\dagger_\nu$ is called *a creation operator* and is, roughly defined, by inserting a row and column as described. The operator c_ν is the Hermitian adjoint of $c^\dagger_\nu$, 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^2_0 – 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 $c_\mu^\dagger$ and c_μ.

Definition of the creation operator: For every single-particle index ν, we define the creation operator $c_\nu^\dagger$ acting on the vacuum state by

$$c_\nu^\dagger |-\rangle = |\nu\rangle . \tag{1.54}$$

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

$$c_\nu^\dagger |\mu_1 \cdots \mu_N\rangle \equiv |\nu\mu_1 \cdots \mu_N\rangle . \tag{1.55}$$

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

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

In terms of determinant coordinate expressions as in Eq. (1.31), $c_\nu^\dagger$ inserts a column on the far right with x_{N+1} and inserts a row on the top with the index ν. Finally, the whole expression is renormalized.

[Recall that the basis Slater determinants were the determinants that had *ordered indices*. Assume that $\vec{\mu}$ is ordered. Clearly, $c^\dagger |\vec{\nu}\rangle$ is either zero or equal to $(-1)^j |\mu_1\mu_2 \cdots \mu_j \nu \mu_{j+1} \cdots \mu_N\rangle$, 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

$$c_\nu |-\rangle \equiv 0. \tag{1.57}$$

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

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

In terms of the coordinate determinant expression, this amounts to moving the jth 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 ν. This moving of the jth row may seem like a complication compared to the creation operator, but note that for $c_\nu^\dagger$ we defined its action by *inserting ν on the top*. Moving ν to the $(j+1)$th position will induce a $(-1)^j$. But c_ν removes a row at an in principle arbitrary location.

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

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

△

1.2.2 Anticommutator relations

Recall that the anticommutator of two operators is defined by

$$\{\hat{A}, \hat{B}\} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}. \tag{1.60}$$

In this section, we prove three important anticommutation relations:

$$\{c_{\nu_1}^\dagger, c_{\nu_2}^\dagger\} = 0 \tag{1.61a}$$

$$\{c_{\nu_1}, c_{\nu_2}\} = 0 \tag{1.61b}$$

$$\{c_{\nu_1}, c_{\nu_2}^\dagger\} = \delta_{\nu_1,\nu_2}. \tag{1.61c}$$

Equation (1.65) is called the "fundamental anticommutator".

Let ν_1, ν_2 be a two single-particle indices, and let $N \geq 0$ be arbitrary. By the properties of determinants, it is easy to see, that for any $|\vec{\mu}\rangle \in L^2_N$,

$$c^\dagger_{\nu_1} c^\dagger_{\nu_2} |\vec{\mu}\rangle = -c^\dagger_{\nu_2} c^\dagger_{\nu_1} |\vec{\mu}\rangle . \tag{1.62}$$

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*

$$\{c^\dagger_{\nu_1}, c^\dagger_{\nu_2}\} \equiv c^\dagger_{\nu_1} c^\dagger_{\nu_2} + c^\dagger_{\nu_2} c^\dagger_{\nu_1} = 0. \tag{1.63}$$

Similarly, two annihilation operators anticommute,

$$\{c_{\nu_1}, c_{\nu_2}\} \equiv c_{\nu_1} c_{\nu_2} + c_{\nu_2} c_{\nu_1} = 0. \tag{1.64}$$

We now prove that

$$\{c_{\nu_1}, c^\dagger_{\nu_2}\} \equiv c_{\nu_1} c^\dagger_{\nu_2} + c^\dagger_{\nu_2} c_{\nu_1} = \delta_{\nu_1,\nu_2}. \tag{1.65}$$

Case 1: $\nu_1 = \nu_2 = \nu$. Consider the expression

$$c^\dagger_\nu c_\nu |\vec{\mu}\rangle . \tag{1.66}$$

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

$$c^\dagger_\nu c_\nu |\mu_1 \cdots \mu_N\rangle = c^\dagger_\nu (-1)^{j-1} |\mu_1 \cdots \mu_{j-1}\mu_{j+1}\cdots\mu_N\rangle = (-1)^{j-1} |\mu_j \mu_1 \cdots \mu_{j-1}\mu_{j+1}\cdots\mu_N\rangle = |\mu_1 \cdots \mu_N\rangle . \tag{1.67}$$

We also get

$$c_\nu c^\dagger_\nu |\mu_1 \cdots \mu_N\rangle = c_\nu |\mu_j \mu_1 \cdots \mu_j \cdots \mu_N\rangle = 0. \tag{1.68}$$

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

$$c^\dagger_\nu c_\nu |\mu_1 \cdots \mu_N\rangle = 0. \tag{1.69}$$

On the other hand,

$$c_\nu c^\dagger_\nu |\vec{\mu}\rangle = c_\nu |\nu\vec{\mu}\rangle = (-1)^0 |\vec{\mu}\rangle . \tag{1.70}$$

Case 1 can be summarized as

$$\{c_\nu, c^\dagger_\nu\} = 1 \tag{1.71}$$

as desired.

Case 2: $\nu_1 \neq \nu_2$. Let $\vec{\mu}$ be arbitrary, and consider the expression

$$c^\dagger_{\nu_1} c_{\nu_2} |\vec{\mu}\rangle . \tag{1.72}$$

Case 2a: If either $\nu_1 \in \vec{\mu}$ or $\nu_2 \notin \vec{\mu}$, the expression vanishes. Similarly,

$$c_{\nu_2} c^\dagger_{\nu_1} |\vec{\mu}\rangle = 0. \tag{1.73}$$

Case 2b: $\nu_1 \notin \vec{\mu}$ *and* $\nu_2 \in \vec{\mu}$ ($\nu_2 = \mu_j$):

$$c^\dagger_{\nu_1} c_{\nu_2} |\vec{\mu}\rangle = (-1)^{j-1} c^\dagger_{\nu_1} |\mu_1 \cdots \mu_{j-1}\mu_{j+1}\cdots\mu_N\rangle = |\mu_1 \cdots \mu_{j-1}\nu_1\mu_{j+1}\cdots\mu_N\rangle , \tag{1.74}$$

i.e., $\mu_j = \nu_2$ is replaced by ν_1. On the other hand,

$$c_{\nu_2} c^\dagger_{\nu_1} |\vec{\mu}\rangle = c_{\nu_2} |\nu_1 \mu_1 \cdots \mu_{j-1}\mu_j\mu_{j+1}\cdots\mu_N\rangle = (-1)^j |\nu_1 \mu_1 \cdots \mu_{j-1}\mu_{j+1}\cdots\mu_N\rangle = (-1) |\mu_1 \cdots \mu_{j-1}\nu_1\mu_{j+1}\cdots\mu_N\rangle . \tag{1.75}$$

Summing, we see that Eq. (1.65) 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 $\vec{\nu} = \sigma(\vec{\mu})$, then

$$|\mu_1, \cdots, \mu_N\rangle = (-1)^{|\sigma|} |\mu_{\sigma(1)}, \cdots, \mu_{\sigma(N)}\rangle = (-1)^{|\sigma|} |\nu_1, \cdots, \nu_N\rangle . \tag{1.76}$$

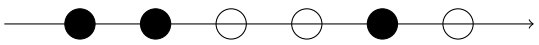
So, the *basis* of Slater determinants can be chosen as those indexed by *sorted* indices.

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 $\nu \notin \vec{\mu}$ 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 ν is set if $\nu \in \vec{\mu}$, i.e., $\nu = \mu_j$ for some $j = 1, \cdots, N$.

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

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

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,

$$|0, 1, 4\rangle = |1_0 1_1 0_2 0_3 1_4\rangle . \tag{1.78}$$

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,

$$|11001\rangle = |0, 1, 4\rangle = -|1, 0, 4\rangle = -|4, 1, 0\rangle = -|0, 4, 1\rangle = +|1, 4, 0\rangle = +|4, 0, 1\rangle , \tag{1.79}$$

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 $\vec{\mu} = \{\mu_1, \cdots, \mu_N\}$ be a given set of occupied orbitals, with occupation number representation

$$|n_0 n_1 n_2 \cdots\rangle$$

Show that:

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

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

△

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 $N = 3$ 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 ↑ and ↓, 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,

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

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

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

The N-body spin operator is

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

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:

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

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

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

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. Only the spin-orbital is a single-particle function in the sense that we use in this text, i.e., a bona fide element in $L^2(X)$, the single-particle Hilbert space. The orbital is an element in $L^2(\mathbb{R}^3)$ and must be adjoined with a spin basis function to become a single-particle function, 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 1.1. In the figure, six spin-orbitals are occupied, and three orbitals are doubly occupied. The illustrated state is

$$c^\dagger_{0,+1} c^\dagger_{0,-1} c^\dagger_{1,+1} c^\dagger_{1,-1} c^\dagger_{2,+1} c^\dagger_{2,-1} |-\rangle . \tag{1.87}$$

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

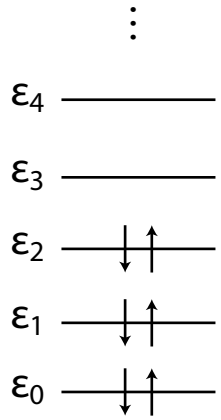


Figure 1.1: Illustration of spin-orbitals and a Slater determinant of 6 electrons

1.2.5 Fock space

The space L_N^2 has the basis consisting of the Slater determinants $|n_0 n_1 n_2 n_3 \cdots\rangle$ with in total N occupied orbitals, or N bits set in the binary representation. The creation operator $c_\mu^\dagger$ 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 1.11

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

$$\mathscr{F} = \bigoplus_{N=0}^{\infty} L_N^2. \tag{1.88}$$

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 single-particle functions.

Thus

$$\langle 0_0 0_1 1_2 0_3 0_4 0_5 0000 \cdots | 1_0 1_1 0_2 0_3 0_4 1_5 0000 \cdots \rangle = 0 \tag{1.89}$$

for example, since the number of particles differ in the two functions.

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

A basis for $\mathscr{F}$ is the set of all $|n_0 n_1 \cdots\rangle$ 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 ν be arbitrary. We have that

$$c_\nu |n_0 n_1 \cdots\rangle = (-1)^{\#_\nu} n_\nu |n_0 n_1 \cdots 0_\nu \cdots\rangle , \tag{1.90}$$

and furthermore that

$$c_\nu^\dagger c_\nu |n_0 n_1 \cdots\rangle = (-1)^{2\#_\nu} n_\nu |n_0 n_1 \cdots\rangle . \tag{1.91}$$

Thus,

$$\sum_\nu c_\nu^\dagger c_\nu |n_0 n_1 \cdots\rangle = \sum_\nu n_\nu |n_0 n_1 \cdots\rangle = N |n_0 n_1 \cdots\rangle . \tag{1.92}$$

Therefore, we define

$$\hat{N} \equiv \sum_\nu c_\nu^\dagger c_n u. \tag{1.93}$$

This operator extracts the number of fermions in a state $|\Psi\rangle$ in the sense that for any $|\Psi\rangle \in \mathscr{F}$, $\hat{N} |\Psi\rangle = N |\Psi\rangle$ 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 1.13, 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:

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

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 single-particle functions 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 inifite 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. [Note: This exercise has been updated since it was given as a weekly exercise.]
Let ϕ_μ, $\mu = 1, 2, \cdots, 6$ be given orthonormal single-particle functions.

a) Using the $|\mu_1, \cdots, \mu_N\rangle$ 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 single-particle funcitons. (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?

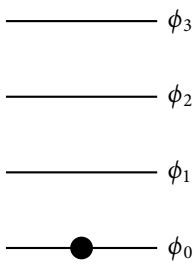
c) Using the occupation number notation $|n_1 n_2 \cdots n_6\rangle$ notation, write down a basis for the same spaces as in exercise a).

c) What is the dimension of the subspace of Fock space you can create with the 6 single-particle functions?

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?

△

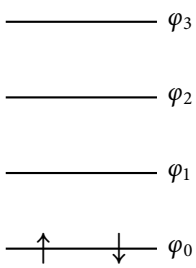
Exercise 1.14 (Note: This exercise has been updated since it was given as a weekly exercise.)**.** Consider the following picture:



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) In a similar fashion as the the above picture, draw a pictures of *all* the distinct Slater determinants you can create using the four single-particle functions. Make sure you consider all possible particle numbers. Caption each picture with the corresponding $|\mu_1\mu_2\cdots\mu_N\rangle$.

We now consider electrons. Consider *4 spin-orbitals* $\varphi_p(\vec{r})$, 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

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

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

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

Write down the matrix of the (single-electron) operator $\hat{h}$ in the spin-orbital basis $\{\phi_{p\sigma}\}$ and find its eigenfunctions. Interpret the spin-orbital diagram in terms of your results. Find the $N = 4$ ground state of $\hat{H}_0$, and draw a picture of it.

△

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:

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

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,

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

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:

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

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

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

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,

$$\hat{W} = \tfrac{1}{4} \sum_{\mu\nu\alpha\beta} \langle \mu\nu|\hat{w}|\alpha\beta\rangle \, c_\mu^\dagger c_\nu^\dagger c_\beta c_\alpha, \tag{1.99}$$

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 notoriouly 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 $\langle \phi_\alpha \phi_\beta|\hat{w}|\phi_\mu \phi_\nu\rangle_{\text{AS}}$ for the anti-symmetric Slater-determinant matrix element (and sometimes we will too), which is equal to:

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

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 [5].

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 $c_\nu^\dagger c_\alpha$ replaces ϕ_α with ϕ_ν (or gives zero of α is not present in $\vec{\mu}$), with no sign change.*

Similarly, $c_{\nu_1}^\dagger c_{\nu_2}^\dagger c_{\alpha_2} c_{\alpha_1}$ replaces α_1 with ν_1, and α_2 with ν_2, or gives zero if one of α_1 or α_2 is not present in $\vec{\mu}$.

Exercise 1.15. Prove the lemma. △

1.3.2 One-body operators

We prove Eq. (1.95) by showing that the actions of the left- and right-hand sides on an arbitrary Slater determinant 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:

$$\begin{aligned}\hat{H}_0 \left|\phi_{\mu_1}, \phi_{\mu_2}, \cdots, \phi_{\mu_N}\right\rangle &= \frac{1}{\sqrt{N!}} \left(\sum_i \hat{h}(i)\right) \sum_\sigma (-1)^{|\sigma|} \hat{P}_\sigma \phi_{\nu_1}(x_1) \cdots \phi_{\nu_N}(x_N) \\ &= \frac{1}{\sqrt{N!}} \sum_\sigma (-1)^{|\sigma|} \hat{P}_\sigma \left(\sum_i \hat{h}(i)\right) \phi_{\nu_1}(x_1) \cdots \phi_{\nu_N}(x_N) \\ &= \left|(\hat{h}\phi_{\nu_1}), \phi_{\nu_2}, \cdots, \phi_{\nu_N}\right\rangle + \left|\phi_{\nu_1}, (\hat{h}\phi_{\nu_2}), \cdots, \phi_{\nu_N}\right\rangle + \cdots + \left|\phi_{\nu_1}, \phi_{\nu_2}, \cdots, (\hat{h}\phi_{\nu_N})\right\rangle\end{aligned} \tag{1.101}$$

Here, we used that $\hat{P}_\sigma$ 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:

$$\psi(x) = \hat{h}\phi_\mu(x) = \sum_\nu \phi_\nu(x) \left\langle \nu|\hat{h}|\mu\right\rangle . \tag{1.102}$$

We insert this expansion:

$$\begin{aligned}\hat{H}_0 \left|\phi_{\mu_1}, \phi_{\mu_2}, \cdots, \phi_{\mu_N}\right\rangle &= \left|(\hat{h}\phi_{\nu_1}), \phi_{\nu_2}, \cdots, \phi_{\nu_N}\right\rangle + \left|\phi_{\nu_1}, (\hat{h}\phi_{\nu_2}), \cdots, \phi_{\nu_N}\right\rangle + \cdots + \left|\phi_{\nu_1}, \phi_{\nu_2}, \cdots, (\hat{h}\phi_{\nu_N})\right\rangle \\ &= \sum_\nu \left\langle \nu|\hat{h}|\mu_1\right\rangle \left|\nu\mu_2, \cdots, \mu_N\right\rangle + \sum_\nu \left\langle \nu|\hat{h}|\mu_2\right\rangle \left|\mu_1\nu\mu_3, \cdots, \mu_N\right\rangle + \cdots + \sum_\nu \left\langle \nu|\hat{h}|\mu_N\right\rangle \left|\mu_1\mu_2, \cdots, \nu\right\rangle\end{aligned} \tag{1.103}$$

Now, we note that

$$\left|\mu_1, \cdots, \mu_{j-1}\nu\mu_{j+1}\cdots\mu_N\right\rangle = c_\nu^\dagger c_{\mu_j} \left|\mu_1, \cdots, \mu_N\right\rangle , \tag{1.104}$$

which we plug in:

$$\begin{aligned}\hat{H}_0 \left|\phi_{\mu_1}, \phi_{\mu_2}, \cdots, \phi_{\mu_N}\right\rangle &= \sum_\nu \left\langle \nu|\hat{h}|\mu_1\right\rangle \left|\nu\mu_2, \cdots, \mu_N\right\rangle + \sum_\nu \left\langle \nu|\hat{h}|\mu_2\right\rangle \left|\mu_1\nu\mu_3, \cdots, \mu_N\right\rangle + \cdots + \sum_\nu \left\langle \nu|\hat{h}|\mu_N\right\rangle \left|\mu_1\mu_2, \cdots, \nu\right\rangle \\ &= \left[\sum_\nu \left\langle \nu|\hat{h}|\mu_1\right\rangle c_\nu^\dagger c_{\mu_1} + \sum_\nu \left\langle \nu|\hat{h}|\mu_2\right\rangle c_\nu^\dagger c_{\mu_2} + \cdots + \sum_\nu \left\langle \nu|\hat{h}|\mu_N\right\rangle c_\nu^\dagger c_{\mu_N}\right] \left|\mu_1, \cdots, \mu_N\right\rangle .\end{aligned} \tag{1.105}$$

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

$$\begin{aligned}\hat{H}_0 \left|\phi_{\mu_1}, \phi_{\mu_2}, \cdots, \phi_{\mu_N}\right\rangle &= \left[\sum_\nu \left\langle \nu|\hat{h}|\mu_1\right\rangle c_\nu^\dagger c_{\mu_1} + \sum_\nu \left\langle \nu|\hat{h}|\mu_2\right\rangle c_\nu^\dagger c_{\mu_2} + \cdots + \sum_\nu \left\langle \nu|\hat{h}|\mu_N\right\rangle c_\nu^\dagger c_{\mu_N}\right] \left|\mu_1, \cdots, \mu_N\right\rangle \\ &= \sum_{\mu\nu} \left\langle \nu|\hat{h}|\mu\right\rangle c_\nu^\dagger c_\mu \left|\mu_1, \cdots, \mu_N\right\rangle .\end{aligned} \tag{1.106}$$

Since $\left|\mu_1, \cdots, \mu_N\right\rangle$ was an arbitrary Slater determinant, we have proven Eq. (1.95).

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,

$$\hat{w}(1,2)\phi_{\mu_1}(x_1)\phi_{\mu_2}(x_2) = \sum_{\nu_1\nu_2} w^{\nu_1\nu_2}_{\mu_1\mu_2}\phi_{\nu_1}(x_1)\phi_{\nu_2}(x_2), \tag{1.107}$$

where the matrix elements $w^{\nu_1\nu_2}_{\mu_1\mu_2}$ are given by the formula (1.98). There is nothing special about the indices $(1,2)$, it may just as well be (i,j). Note also the symmetry property

$$w^{\nu_1\nu_2}_{\mu_1\mu_2} = w^{\nu_2\nu_1}_{\mu_2\mu_1}.$$

As for the one-body case, $\hat{W}$ commutes with $\hat{P}_\sigma$, and we get, using Eq. (1.107),

$$\begin{aligned}
\hat{W}\left|\phi_{\mu_1}\cdots\phi_{\mu_N}\right\rangle &= \frac{1}{\sqrt{N!}}\sum_\sigma (-1)^{|\sigma|}\hat{P}_\sigma\left[\sum_{i<j}\hat{w}(i,j)\phi_{\mu_1}(x_1)\cdots\phi_{\mu_N}(x_N)\right] \\
&= \frac{1}{\sqrt{N!}}\sum_\sigma (-1)^{|\sigma|}\hat{P}_\sigma\left[\sum_{i<j}\hat{w}(i,j)\phi_{\mu_1}(x_1)\cdots\phi_{\mu_N}(x_N)\right] \\
&= \frac{1}{\sqrt{N!}}\sum_\sigma (-1)^{|\sigma|}\hat{P}_\sigma\left[\sum_{i<j}\sum_{\nu_1\nu_2} w^{\nu_1\nu_2}_{\mu_i\mu_j}\phi_{\mu_1}\cdots\phi_{\nu_1}(x_i)\cdots\phi_{\nu_2}(x_j)\cdots\phi_{\mu_N}(x_N)\right] \\
&= \sum_{i<j}\sum_{\nu_1\nu_2} w^{\nu_1\nu_2}_{\mu_i\mu_j}\left|\phi_{\mu_1}\cdots\phi_{\nu_1}\cdots\phi_{\nu_2}\cdots\phi_{\mu_N}\right\rangle \\
&= \sum_{i<j}\sum_{\nu_1\nu_2} w^{\nu_1\nu_2}_{\mu_i\mu_j}c^\dagger_{\nu_1}c^\dagger_{\nu_2}c_{\mu_j}c_{\mu_i}\left|\phi_{\mu_1}\cdots\phi_{\mu_N}\right\rangle.
\end{aligned} \tag{1.108}$$

Here, we used Lemma 1.1 about replacement behaviour of the $c^\dagger c^\dagger cc$ 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:

$$\begin{aligned}
&\sum_{i<j}\sum_{\nu_1\nu_2} w^{\nu_1\nu_2}_{\mu_i\mu_j}c^\dagger_{\nu_1}c^\dagger_{\nu_2}c_{\mu_j}c_{\mu_i}\left|\phi_{\mu_1}\cdots\phi_{\mu_N}\right\rangle = -\sum_{i<j}\sum_{\nu_1\nu_2} w^{\nu_1\nu_2}_{\mu_i\mu_j}c^\dagger_{\nu_1}c^\dagger_{\nu_2}c_{\mu_i}c_{\mu_j}\left|\phi_{\mu_1}\cdots\phi_{\mu_N}\right\rangle \\
&= \sum_{i<j}\sum_{\nu_1\nu_2} w^{\nu_1\nu_2}_{\mu_i\mu_j}c^\dagger_{\nu_2}c^\dagger_{\nu_1}c_{\mu_i}c_{\mu_j}\left|\phi_{\mu_1}\cdots\phi_{\mu_N}\right\rangle = \sum_{j<i}\sum_{\nu_1\nu_2} w^{\nu_1\nu_2}_{\mu_j\mu_i}c^\dagger_{\nu_2}c^\dagger_{\nu_1}c_{\mu_j}c_{\mu_i}\left|\phi_{\mu_1}\cdots\phi_{\mu_N}\right\rangle \\
&= \sum_{j<i}\sum_{\nu_2\nu_1} w^{\nu_2\nu_1}_{\mu_j\mu_i}c^\dagger_{\nu_1}c^\dagger_{\nu_2}c_{\mu_j}c_{\mu_i}\left|\phi_{\mu_1}\cdots\phi_{\mu_N}\right\rangle = \sum_{j<i}\sum_{\nu_2\nu_1} w^{\nu_1\nu_2}_{\mu_i\mu_j}c^\dagger_{\nu_1}c^\dagger_{\nu_2}c_{\mu_j}c_{\mu_i}\left|\phi_{\mu_1}\cdots\phi_{\mu_N}\right\rangle
\end{aligned} \tag{1.109}$$

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

$$\hat{W}\left|\phi_{\mu_1}\cdots\phi_{\mu_N}\right\rangle = \frac{1}{2}\sum_{ij}\sum_{\nu_2\nu_1} w^{\nu_1\nu_2}_{\mu_i\mu_j}c^\dagger_{\nu_1}c^\dagger_{\nu_2}c_{\mu_j}c_{\mu_i}\left|\phi_{\mu_1}\cdots\phi_{\mu_N}\right\rangle. \tag{1.110}$$

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 $c_\alpha\left|\vec{\mu}\right\rangle$ gives zero contributions for such orbitals. Thus, Eq. (1.97) is proven.

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

Exercise 1.16. Prove Eq. (1.99). Start with showing Eq. (1.100). △

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

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

g) Compute the matrix element

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

△

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 1.17, 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

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

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

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

$$|\vec{\nu}\rangle = c_{\nu_j}^{\dagger} c_{\mu_j} |\vec{\mu}\rangle , \quad \nu_j \neq \mu_j. \tag{1.113}$$

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

$$\langle \vec{\mu} | \hat{H}_0 | \vec{\nu} \rangle = h_{\nu_j}^{\mu_j}, \tag{1.114}$$

$$\langle \vec{\mu} | \hat{W} | \vec{\nu} \rangle = \sum_{i} \langle \mu_i \mu_j | \hat{w} | \mu_i \nu_j \rangle_{\mathrm{AS}} . \tag{1.115}$$

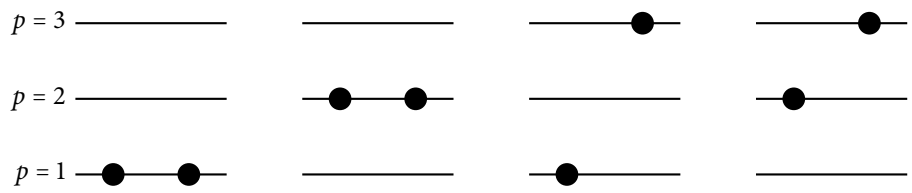


Figure 1.2: Schematic plot of the possible single-particle levels with double degeneracy. The filled circles indicate occupied particle states. The spacing between each level p is constant in this picture. We show some possible two-particle states.

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

$$|\vec{\nu}\rangle = c^\dagger_{\nu_k} c^\dagger_{\nu_j} c_{\mu_j} c_{\mu_j} |\vec{\mu}\rangle\,, \quad j \neq k. \tag{1.116}$$

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

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

$$\langle\vec{\mu}|\hat{W}|\vec{\nu}\rangle = \langle\mu_j\mu_k|\hat{w}|\nu_j\nu_k\rangle_{\rm AS}\,. \tag{1.118}$$

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 1.2. The single-particle states are labelled by the quantum number p and can accomodate 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.

a) How many two-particle Slater determinants can we construct in this space?

b) 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.

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 preceeding 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 $\psi : X \to \mathbb{C}$.

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 $L^2(X^N)_{\mathrm{AS}}$. Furthermore, we defined *Fock space* $\mathscr{F}$,

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

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

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

$$|\Psi_N\rangle = \sum_{\vec{\mu}}^{\sim} |\mu_1\cdots\mu_N\rangle \langle\mu_1\cdots\mu_N|\Psi_N\rangle . \tag{1.119}$$

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* $c^\dagger_\mu$ (these, of course, depend on the basis),

$$|\mu_1\cdots\mu_N\rangle = c^\dagger_{\mu_1} c^\dagger_{\mu_2}\cdots c^\dagger_{\mu_N} |-\rangle . \tag{1.120}$$

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

$$\hat{H}_0 = \sum_{\mu\nu} \langle \mu|\hat{h}|\nu\rangle \, c^\dagger_\mu c_\nu \tag{1.121}$$

$$\hat{W} = \frac{1}{4} \sum_{\substack{\nu_1\nu_2 \\ \mu_1\mu_2}} \langle \mu_1\mu_2|\hat{w}|\nu_1\nu_2\rangle \, c^\dagger_{\mu_1} c^\dagger_{\mu_2} c_{\nu_2} c_{\nu_1}. \tag{1.122}$$

We claimed that these expressions simplify our life a lot.

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

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

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 $\hat{H}\,|\Psi_N\rangle$:

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

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

$$\mathsf{HC} = E\mathsf{C}. \tag{1.125}$$

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. (1.125), we need to compute the matrix elements[2]

$$\mathsf{H}_{0,\vec{\nu},\vec{\mu}} = \langle\vec{\mu}|\hat{H}_0|\vec{\mu}\rangle = \sum_{\mu\nu} \langle\mu|\hat{h}|\nu\rangle\,\langle -|\,c_{\nu_N}c_{\nu_{N-1}}\cdots c_{\nu_1}c^\dagger_\mu c_\nu c^\dagger_{\mu_1}c^\dagger_{\mu_2}\cdots c^\dagger_{\mu_N}\,|-\rangle \tag{1.126}$$

and similarly

$$\mathsf{W}_{\vec{\nu},\vec{\mu}} = \langle\vec{\mu}|\hat{W}|\vec{\mu}\rangle = \frac{1}{4}\sum_{\substack{\alpha_1\alpha_2 \\ \beta_1\beta_2}} \langle\alpha_1\alpha_2|\hat{w}|\beta_1\beta_2\rangle\,\langle -|\,c_{\nu_N}c_{\nu_{N-1}}\cdots c_{\nu_1}c^\dagger_{\alpha_1}c^\dagger_{\alpha_2}c_{\beta_2}c_{\beta_1}c^\dagger_{\mu_1}c^\dagger_{\mu_2}\cdots c^\dagger_{\mu_N}\,|-\rangle \tag{1.127}$$

Notice that we used

$$|\mu_1\cdots\mu_N\rangle = c^\dagger_{\mu_1}\cdots c^\dagger_{\mu_N}\,|-\rangle \tag{1.128}$$

and, by taking the adjoint,

$$\langle\mu_1\cdots\mu_N| = \langle -|\,c_{\mu_N}\cdots c_{\mu_1}. \tag{1.129}$$

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:

$$A = \langle\nu_1\nu_2|c^\dagger_\alpha c_\beta|\mu_1\mu_2\rangle = \langle -|\,c_{\nu_2}c_{\nu_1}c^\dagger_\alpha c_\beta c^\dagger_{\mu_1}c^\dagger_{\mu_2}\,|-\rangle\,. \tag{1.130}$$

[2] On 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* $\hat{H}\,|\vec{\mu}\rangle$.

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

$$c_\alpha c_\beta^\dagger + c_\beta^\dagger c_\alpha = \delta_{\alpha\beta} \tag{1.131}$$

$$c_\alpha c_\beta + c_\beta c_\alpha = 0 \tag{1.132}$$

and

$$c_\alpha^\dagger c_\beta^\dagger + c_\beta^\dagger c_\alpha^\dagger = 0. \tag{1.133}$$

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:

$$\begin{aligned} A &= \langle -| c_{\nu_2} c_{\nu_1} c_\alpha^\dagger c_\beta c_{\mu_1}^\dagger c_{\mu_2}^\dagger |-\rangle = \langle -| c_{\nu_2} c_{\nu_1} c_\alpha^\dagger (\delta_{\beta\mu_1} - c_{\mu_1}^\dagger c_\beta) c_{\mu_2}^\dagger |-\rangle \\ &= \delta_{\beta\mu_1} \langle -| c_{\nu_2} c_{\nu_1} c_\alpha^\dagger c_{\mu_2}^\dagger |-\rangle - \langle -| c_{\nu_2} c_{\nu_1} c_\alpha^\dagger c_{\mu_1}^\dagger c_\beta c_{\mu_2}^\dagger |-\rangle \\ &= \delta_{\beta\mu_1} \langle -| c_{\nu_2} c_{\nu_1} c_\alpha^\dagger c_{\mu_2}^\dagger |-\rangle - \langle -| c_{\nu_2} c_{\nu_1} c_\alpha^\dagger c_{\mu_1}^\dagger (\delta_{\beta\mu_2} - c_{\mu_2}^\dagger c_\beta) |-\rangle \\ &= \delta_{\beta\mu_1} \langle -| c_{\nu_2} c_{\nu_1} c_\alpha^\dagger c_{\mu_2}^\dagger |-\rangle - \delta_{\beta\mu_2} \langle -| c_{\nu_2} c_{\nu_1} c_\alpha^\dagger c_{\mu_1}^\dagger |-\rangle . \end{aligned} \tag{1.134}$$

We continue:

$$\begin{aligned} A &= \delta_{\beta\mu_1} \langle -| c_{\nu_2} (\delta_{\nu_1\alpha} - c_\alpha^\dagger c_{\nu_1}) c_{\mu_2}^\dagger |-\rangle - \delta_{\beta\mu_2} \langle -| c_{\nu_2} (\delta_{\nu_1\alpha} - c_\alpha^\dagger c_{\nu_1}) c_{\mu_1}^\dagger |-\rangle \\ &= \delta_{\beta\mu_1}\delta_{\nu_1\alpha} \langle -| c_{\nu_2} c_{\mu_2}^\dagger |-\rangle - \delta_{\beta\mu_1} \langle -| c_{\nu_2} c_\alpha^\dagger c_{\nu_1} c_{\mu_2}^\dagger |-\rangle - (\mu_1 \leftrightarrow \mu_2) . \end{aligned} \tag{1.135}$$

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,

$$A = \delta_{\beta\mu_1}\delta_{\nu_1\alpha} \langle -| c_{\nu_2} c_{\mu_2}^\dagger |-\rangle - \delta_{\beta\mu_1}\delta_{\nu_1\mu_2} \langle -| c_{\nu_2} c_\alpha^\dagger |-\rangle + \delta_{\beta\mu_1} \langle -| c_{\nu_2} c_\alpha^\dagger c_{\mu_2}^\dagger c_{\nu_1} |-\rangle - (\mu_1 \leftrightarrow \mu_2) . \tag{1.136}$$

Only the two first terms are non-vanishing, and we note, for example, that $\langle -| c_{\nu_2} c_{\mu_2}^\dagger |-\rangle = \langle \nu_2 | \mu_2 \rangle = \delta_{\nu_2\mu_2}$. (We could also use the anticommutator once more.) This gives:

$$A = \delta_{\beta\mu_1}\delta_{\nu_1\alpha}\delta_{\nu_2\mu_2} - \delta_{\beta\mu_1}\delta_{\nu_1\mu_2}\delta_{\nu_2\alpha} - (\mu_1 \leftrightarrow \mu_2). \tag{1.137}$$

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 anticommutator 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*:

$$\langle -| A_1 A_2 \cdots A_n |-\rangle , \tag{1.138}$$

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

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

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

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

The right-hand side is a linear combination of vacuum expectation values. So we see that having a straightforward way to compute Eq. (1.138) 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

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

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

$$\langle - | A_1 A_2 \cdots A_n | - \rangle \, . \tag{1.142}$$

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 anticummutator relations (1.61).

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

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

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

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

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

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

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. (1.143) automatically compensates for this. Thus,

$$N(c_1 c_2^\dagger c_3^\dagger c_4) = c_2^\dagger c_3^\dagger c_1 c_4 = -c_3^\dagger c_2^\dagger c_1 c_4 = c_3^\dagger c_2^\dagger c_4 c_1 = -c_2^\dagger c_3^\dagger c_4 c_1. \tag{1.146}$$

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

$$N(\alpha A_1 \cdots A_n + \beta B_1 \cdots B_m) = \alpha N(A_1 \cdots A_n) + \beta N(B_1 \cdots B_n). \tag{1.147}$$

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 to arbitrary creation and annihilation operators X and Y is a special notation for $\langle -|XY|-\rangle$,

$$\overbracket{XY} \equiv \langle -|XY|-\rangle. \tag{1.148}$$

Thus, the contraction is a *number*. One can easily show (see exercise 1.20), that

$$\overbracket{XY} = XY - N(XY). \tag{1.149}$$

Let us list all the possible contractions:

$$\overbracket{c_\mu^\dagger c_\nu^\dagger} = \langle -|c_\mu^\dagger c_\nu^\dagger|-\rangle = 0 \tag{1.150a}$$

$$\overbracket{c_\mu c_\nu} = \langle -|c_\mu c_\nu|-\rangle = 0 \tag{1.150b}$$

$$\overbracket{c_\mu^\dagger c_\nu} = \langle -|c_\mu^\dagger c_\nu|-\rangle = 0 \tag{1.150c}$$

$$\overbracket{c_\mu c_\nu^\dagger} = \langle -|c_\mu c_\nu^\dagger|-\rangle = \delta_{\mu\nu}. \tag{1.150d}$$

$$\tag{1.150e}$$

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. (1.148). A contraction between two operators at positions x and y (with $x < y$) in the string is thus defined by

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

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

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

(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:

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

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

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

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

In the last example (1.153d), 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 $N(\cdots)$.

We also define normal ordering with *multiple contractions*, say m, leaving $n - 2m$ operators uncontracted. Clearly, there can be at most $\lfloor n/2 \rfloor$ pairs[3] if each pair is required to be distinct.

The definition is recursive: each pair of contracted operators is processed in turn according to Eq. (1.151). This definition is independent of the order of the processing of the pairs. An example: Example:

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

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 $i = 1, \cdots, m$. Let $\sigma \in S_n$ be a permutation such that $\sigma(x_1) = 1$, $\sigma(y_1) = 2$, $\sigma(x_2) = 3$, etc. Then

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

Exercise 1.20. Show Eq. (1.149) from Eq. (1.148), by considering the 4 possible cases. △

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

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

△

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,*

$$\begin{aligned} A_1 A_2 \cdots A_n = N(A_1 A_2 \cdots A_n) &+ \sum_{(1)} N\left(A_1 \cdots \overbrace{\cdots \cdots} \cdots A_n\right) + \sum_{(2)} N\left(A_1 \cdots \overbrace{\cdots \overbrace{\cdots}} \cdots A_n\right) \\ &+ \cdots + \sum_{(\lfloor \frac{n}{2} \rfloor)} N\left(\underbrace{A_1 \cdots \overbrace{\cdots \cdots \cdots \cdots \cdots} \cdots A_n}_{\lfloor n/2 \rfloor \text{ contractions}}\right) \end{aligned} \tag{1.157}$$

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 opeators are contracted. If n is odd, there is one uncontracted operator left in each term of the last sum.

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

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,

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

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

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

For even n,

$$\langle -|\, A_1A_2\cdots A_n \,|-\rangle = \sum_{(\lfloor \frac{n}{2} \rfloor)} \underbrace{A_1\cdots\cdots\cdots\cdots\cdots\cdots A_n}_{\text{all contracted}}. \tag{1.160}$$

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

$$\overbrace{c_\alpha c^\dagger_\beta} = \delta_{\alpha\beta}. \tag{1.161}$$

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 $\overbrace{c_\alpha c_\beta}$ or $\overbrace{c^\dagger_\alpha c^\dagger_\beta}$ must be present in Eq. (1.162), 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,

$$\langle -|A_1\cdots A_n|-\rangle = \sum_{\text{all contr.}} \overbrace{A_1A_2A_3A_4}\cdots\overbrace{A_kA_{k+1}A_{k+2}A_{k+3}}\cdots A_n, \tag{1.162}$$

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

$$\overbrace{c_\alpha c^\dagger_\beta}$$

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.) △

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

△

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 (1.157). $\mathcal{P}_1$ 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, $r = 1, \cdots, n$ be creation and annihilation operators. Let B be a creation or annihilation operator. Then,*

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

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,

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

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

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

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. (1.163) 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 $A_1 \cdots A_n$. □

We introduce another lemma, which generalizes Lemma 1.2 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 $S_m = \{1, 2, \cdots, N\} \setminus (\cup_i p_i)$ be the remaining indices when all pairs are removed. Let B a creation or annihilation operator. Then,*

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

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

Proof. let $S = \{1, 2, \cdots, N\}$. 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. (1.166). 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

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

Now,

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

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. (1.166).

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. □

We now prove Wick's Theorem. Assume now that $\mathcal{P}_n$ is true. Multiply Eq. (1.157) from the right by an operator A_{n+1}:

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

Each sum is a sum over m contractions, including the first where we have $m = 0$. We now use Lemma 1.3 and write

$$\sum_{(m)} N\left(\overbrace{A_1A_2A_3A_4}\cdots A_n\right)A_{n+1} = \sum_{(m)} N\left(\overbrace{A_1A_2A_3A_4}\cdots A_nA_{n+1}\right) + \sum_{(m)}\sum_{r} N\left(\overbrace{A_1A_2A_3A_4\cdots A_nA_{n+1}}\right)$$
$$:= X_m + I_m. \tag{1.170}$$

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

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

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

Write

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

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, wee 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,*

$$\langle -|\, \overbrace{A_1A_2A_3\cdots\cdots A_{n-1}A_n}\,|-\rangle = \overbrace{A_{x_1}A_{y_1}}\cdots\overbrace{A_{x_{n/2}}A_{y_{n/2}}}(-1)^s, \tag{1.173}$$

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 annhilation operators must be contracted with a creation operator to the right*.

We now also simplfy 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:

$$\langle \mu_1\cdots\mu_3|\alpha^\dagger\beta|\mu_1\cdots\mu_3\rangle = \langle -|\, \mu_3\mu_2\mu_1\alpha^\dagger\beta\mu_1 1^\dagger\mu_2^\dagger\mu_3^\dagger\,|-\rangle\,. \tag{1.174}$$

Worked example 2:

$$\langle \mu_1\cdots\mu_3|\alpha_1^\dagger\alpha_2^\dagger\beta_2\beta_1|\mu_1\cdots\mu_3\rangle = \langle -|\, \mu_3\mu_2\mu_1\alpha_1^\dagger\alpha_2^\dagger\beta_2\beta_1\mu_1 1^\dagger\mu_2^\dagger\mu_3^\dagger\,|-\rangle\,. \tag{1.175}$$

Exercise 1.24. (Slater–Condon rules revisited)

a) 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?

b) Repeat Exercise 1.18 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

Motivatinal 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 enout, 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 $i, j, k, \cdots \leq N$, and $a, b, c, \cdots > N$. Thus, a summation $\sum_\mu$ is split into $\sum_{i=1}^{N}$ and $\sum_{a=N+1}^{\infty}$.

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

$$b_i = c_i^\dagger, \quad b_a = c_a \tag{1.176}$$

with Hermitian adjoints

$$b_i^\dagger = c_i, \quad b_a^\dagger = c_a^\dagger \tag{1.177}$$

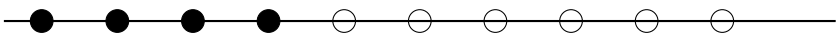
It is an easy exercise to show that the anticommutator relations are preserved:

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

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

$$|\Phi\rangle = |123\cdots N\rangle = c_1^\dagger c_2^\dagger \cdots c_N^\dagger |-\rangle . \tag{1.179}$$

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

$|\Phi\rangle =$ 

Note that

$$b_\mu |\Phi\rangle = 0, \quad \forall \mu. \tag{1.180}$$

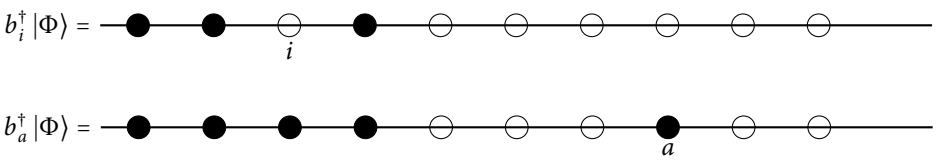
Therefore, $|\Phi\rangle$ has the role of a vacuum for the new operators. It contains zero quasiparticles, since attempting to remove one gives us zero.

Let us create a quasiparticle:

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

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

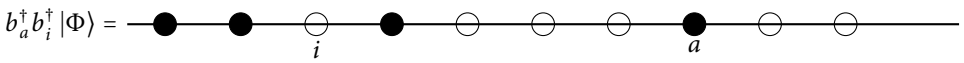
In pictures,



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,



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 pairs[4].

Thus, any wavefunction in with N particles can be written

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

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

$$|\Phi_i^a\rangle = b_a^\dagger b_i^\dagger |\Phi\rangle = c_a^\dagger c_i |\Phi\rangle , \tag{1.184}$$

and

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

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-hol 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 $|ia\rangle_c$, $|\Phi_i^a\rangle$, $|_i^a\rangle$, and others.

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

$$[b_a^\dagger b_i^\dagger,\ b_b^\dagger b_j^\dagger] = 0. \tag{1.186}$$

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

[4] Digression: 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.

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

$$\langle\Phi|c_i^\dagger c_j|\Phi\rangle = \langle\Phi|b_i b_j^\dagger|\Phi\rangle = \overbrace{b_i b_j^\dagger} = \delta_{ij}. \tag{1.187}$$

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

Exercise 1.25. Prove Eq. (1.186) △

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? △

Exercise 1.28. Compute the vacuum expectation value (1.187) ussing the Wick's theorem and the original creation and annihilation operators and compare the method and amount of work.

Repeat for

$$\langle\Phi|c_\alpha^\dagger c_\beta^\dagger c_\delta c_\gamma|\Phi\rangle \tag{1.188}$$

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. △

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

△

1.6 Operators on normal-order form (Not yet lectured)

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 LaTeX.

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

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

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

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

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:

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

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

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

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

$$c_p^\dagger c_q = N(c_p^\dagger c_q) + \overbrace{c_p^\dagger c_q}. \tag{1.193}$$

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

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

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

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

and

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

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

1.6.3 Two-body operators

We continoth an arbitrary two-body operator

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

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

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

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 fiull 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

$$\hat{H} = \hat{H}_{\mathrm{N}} + E_0 = \hat{H}_{\mathrm{N}} + \hat{W}_{\mathrm{N}} + E_0, \tag{1.205}$$

with

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

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

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

Thus, $\hat{H}_{0,\mathrm{N}}$ is the *total* one-quasiparticle operator part of $\hat{H}$, and contains contributions from $\hat{W}$ as well as $\hat{H}_0$, while $\hat{W}_{\mathrm{N}}$ 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 $\hat{H}_{0,\mathrm{N}}$ and $\hat{W}_{\mathrm{N}}$ in terms of quasiparticle operators. This gives a lot of terms, especially in the two-body case. We start with $\hat{H}_{0,\mathrm{N}}$, splitting the sum over pq into four terms:

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

where

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

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

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

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

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

Exercise 1.30. Verify that Eq. (1.209) equals Eq. (1.210). △

Exercise 1.31. Verify that $\hat{W}_{\mathrm{N}}$ in Eq. (1.210) is Hermitian, given that $\hat{W}$ is Hermitian. △

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

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

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

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

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

$$\hat{H} = \hat{F}_{\mathrm{N}} + \hat{G}_{\mathrm{N}} + \hat{X}_{\mathrm{N}} + E_0. \tag{1.213}$$

△

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 $|\Psi\rangle \in L^2_N$ with E a real number such that

$$\hat{H}\,|\Psi\rangle = E\,|\Psi\rangle\,. \tag{2.1}$$

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^2_N has *finite dimension* D, then $\hat{H}$ can be viewed as a *Hermitian matrix*, and we can find a complete set of orthonormal eigenfunctions $|\Psi_k\rangle$, $k = 0, 1, 2, \cdots$ with corresponding eigenvalues E_k, such that

$$\hat{H} = \sum_{k=0}^{D} E_k\,|\Psi_k\rangle\,\langle\Psi_k|\,. \tag{2.2}$$

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*

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

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

$$\hat{H}|\Psi_*\rangle = E_*|\Psi_*\rangle. \tag{2.4}$$

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

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

The stationary point condition can be formulated as

$$f'(0) = 0. \tag{2.6}$$

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

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

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

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

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

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

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

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

Recall that

$$f(\epsilon) = f(0) + \epsilon f'(0) + O(\epsilon^2). \tag{2.11}$$

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

$$A = \frac{aE}{N}, \tag{2.12}$$

that is,

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

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

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

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

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

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

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

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

The proof is complete. □

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,*

$$E_0 = \min \left\{ \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \;\middle|\; 0 \neq |\Psi\rangle \in L^2_N,\ |\langle \Psi | \hat{H} | \Psi \rangle| < +\infty \right\}. \tag{2.18}$$

Theorem 2.2 holds even if Eq. (2.2) 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 dimensons, this is of course not true.

We will not prove Theorem 2.2 in its full generality, but we see immediately that it follows from Theorem 2.1: E_0 is a stationary value, and cleary $\mathcal{E}$ 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* $\mathcal{M} \subset L^2_N$, and compute

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

Clearly,

$$E_0 \leq E_0[\mathcal{M}], \tag{2.20}$$

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. (2.19) for the function $|\tilde{\Psi}\rangle \in \mathcal{M}$:

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

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

$$|\Psi_0\rangle \approx |\tilde{\Psi}\rangle + \epsilon |\eta\rangle \tag{2.22}$$

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

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

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

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

The exact ground-state wavefunction is well-known,

$$\psi_0(\vec{r}) = Ce^{-r}, \tag{2.25}$$

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

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

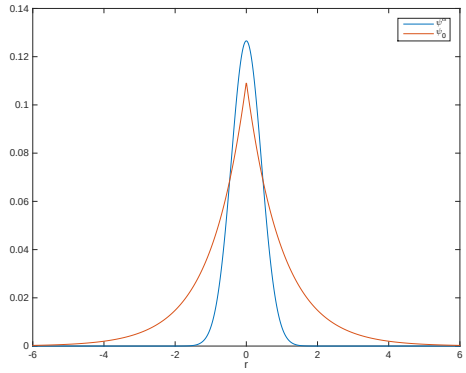


Figure 2.1: Plot of approximate and exact ground-state wavefunction for the Hydrogen example

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,

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

and minimize with respect to α,

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

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. 2.1! 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^2_N defined by a basis set $|\Phi_I\rangle$, $I = 1, 2, \cdots, D$. Then the variational procedure is equivalent to computing the smallest eigenvalue of the matrix

$$\mathsf{H}_{IJ} = \langle\Phi_I|\hat{H}|\Phi_J\rangle\,. \tag{2.29}$$

This is so, because for

$$|\tilde{\Psi}\rangle = \sum_{I=1}^{D} A_I\,|\Phi_I\rangle \tag{2.30}$$

the expectation value becomes

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

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

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

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

Then,

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

Exercise 2.1. Prove Eq. (2.31). △

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,

$$\mathcal{V} = \text{span}\{|\vec{p}\rangle \;\big|\; |\vec{p}\rangle \in \mathcal{S}\} \tag{2.34}$$

Equivalently, any wavefunction in V can be written

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

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

$$\mathcal{V}_{\text{CIS}} = \text{span}\{|\Phi\rangle\,,\ |\Phi_i^a\rangle\ \big|\ i = 1, \cdots N,\ a = N+1, \cdots, L\}, \tag{2.36}$$

and any CIS wavefunction can thus be written

$$|\Psi\rangle = A_0\,|\Phi\rangle + \sum_{ia} A_i^a\,|\Phi_{ia}\rangle\,. \tag{2.37}$$

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

$$\mathcal{V}_{\text{CISD}} = \text{span}\{|\Phi\rangle\,,\ |\Phi_i^a\rangle\ |\Phi_{ij}^{ab}\rangle\big|\ i, j = 1, \cdots N,\ a, b = N+1, \cdots, L\}. \tag{2.38}$$

A wavefunction $|\Psi\rangle \in V_{\text{CISD}}$ can be written

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

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

Sometimes, the doubles term is written

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

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 $|\Phi_{ij}^{ab}\rangle = -|\Phi_{ij}^{ba}\rangle = -|\Phi_{ji}^{ab}\rangle = |\Phi_{ji}^{ba}\rangle$, i.e., we are deliberately over-counting the basis in this expression to keep notation simple.

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

$$\mathcal{V} = \text{span}\{|\Phi_I\rangle\ \big|\ I \in \mathcal{I}\}. \tag{2.41}$$

Our vector expansion becomes

$$|\Psi\rangle = \sum_I A_I\,|\Phi_I\rangle\,, \quad A_I = \langle\Phi_I|\Psi\rangle\,. \tag{2.42}$$

For example, $\mathcal{I} = 1, 2, \cdots, D$ 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 2.4. 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 $\hat{H}_0 = \sum_i^N \hat{h}(i)$. Clearly, the determinant

$$|\Phi\rangle = |123\cdots N\rangle \tag{2.43}$$

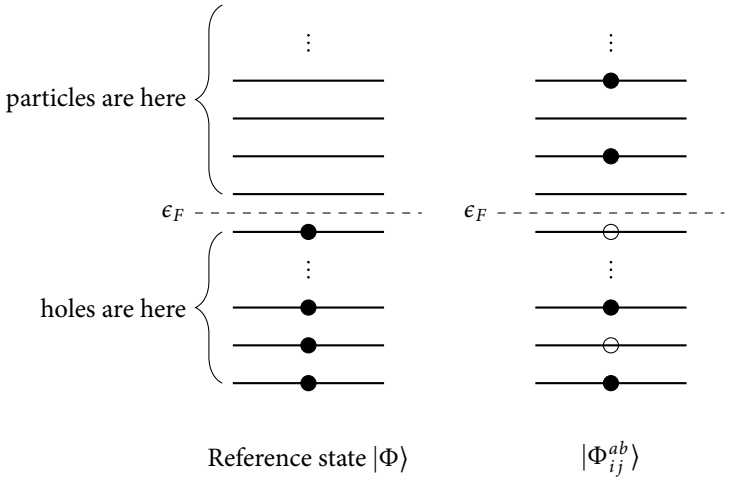


Figure 2.2: Fermi level and quasiparticles. To the left, we have the vacuum state. To the right, we have a doubly excited state, or a two-particle-two-hole-state. Notice how we draw the "Fermi line" between two levels for clarity. In this simple picture, we have assumed that the levels are non-degenerate. 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

$$\hat{H}_0 = \sum_p \epsilon_p c_p^\dagger c_p. \tag{2.44}$$

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,

$$|\Phi\rangle = \left(\prod_{\epsilon_p \leq \epsilon_\mathrm{F}} c_p^\dagger \right) |-\rangle \,, \tag{2.45}$$

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. 2.2

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

$$H_{IJ} = \langle \Phi_I | \hat{H} | \Phi_J \rangle \tag{2.46}$$

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

$$\mathsf{H} = \left(\begin{array}{c|c|c} \langle\Phi|\hat{H}|\Phi\rangle & \langle\Phi|\hat{H}|\Phi_i^a\rangle & \langle\Phi|\hat{H}|\Phi_{ij}^{ab}\rangle \\ \hline \langle\Phi_{i'}^{a'}|\hat{H}|\Phi\rangle & \langle\Phi_{i'}^{a'}|\hat{H}|\Phi_i^a\rangle & \langle\Phi_{i'}^{a'}|\hat{H}|\Phi_{ij}^{ab}\rangle \\ \hline \langle\Phi_{i'j'}^{a'b'}|\hat{H}|\Phi\rangle & \langle\Phi_{i'j'}^{a'b'}|\hat{H}|\Phi_i^a\rangle & \langle\Phi_{i'j'}^{a'b'}|\hat{H}|\Phi_{ij}^{ab}\rangle \end{array}\right) \tag{2.47}$$

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 agorithm 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 partucle 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

$$\vec{A}' = H\vec{A} \tag{2.48}$$

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

$$|\Psi'\rangle = P\hat{H}\,|\Psi\rangle \tag{2.49}$$

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

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

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:

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

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 manipulaing 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 [6], Section 11.8.

2.3.6 Recipe for bit pattern representation.

How can we perform the bitwise operations mentioned above?

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

$$\vec{\mu} = \{1, 5, 6\} \mapsto |\underbrace{010001100\cdots0_2}_{L \text{ bits}}\rangle \mapsto |1 \times 2^1 + 1 \times 2^5 + 1 \times 2^6\rangle = |97_{10}\rangle \,. \tag{2.52}$$

(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. Retun 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 $p = 0, 1, \cdots, L-1$, and thus an occupation number representation of length 8 bits, e.g.,

$$|p = 2, p = 3\rangle = |0_0 0_1 1_2 1_3 0_4 0_5 0_6 0_7\rangle = |00110000\rangle . \tag{2.53}$$

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 2.4.) 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 2.4 and 2.5.) 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 (HF)

2.4.1 The Hartree–Fock equations

Suggested reading for this section: Szabo/Ostlund Ch. 13, Harris/Monkhorst/Freeman Ch. 3, Gross/Runge/Heinonen Ch. 7.

It is highly recommended to read the mathematical supplement in the Appendix, Sec. A.1 on the calculus of variations.

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.*

Thus, our wavefunction manifold $\mathcal{M}$ consists of all possible functions on the form

$$|\Phi\rangle = |\phi_1\phi_2\cdots\phi_N\rangle\,, \quad \langle\phi_i|\phi_j\rangle = \delta_{ij}. \tag{2.54}$$

Note carefully, that a single-particle basis is not given – it is to be found! The expectation value of the Hamiltonian $\hat{H} = \hat{H}_0 + \hat{W}$ now reads (recalling that $\langle\Phi|\Phi\rangle = 1$)

$$\langle\Phi|\hat{H}|\Phi\rangle = \sum_i \langle\phi_i|\hat{h}|\phi_i\rangle + \frac{1}{2}\sum_{ij}\langle\phi_i\phi_j|\hat{w}|\phi_i\phi_j - \phi_j\phi_i\rangle \tag{2.55}$$

as obtained via the Slater–Condon rules, see for example Exercise 1.18. Here,

$$\langle\phi_p\phi_q|\hat{w}|\phi_r\phi_s\rangle \equiv \int dx_1 \int dx_2 \phi_p(x_1)^*\phi_q(x_2)^* w(x_1,x_2)\phi_r(x_1)\phi_s(x_2), \tag{2.56}$$

which satisfies $\langle\phi_p\phi_q|\hat{w}|\phi_r\phi_s\rangle = \langle\phi_q\phi_p|\hat{w}|\phi_s\phi_r\rangle$. The task is now to minimize this energy $\langle\Phi|\hat{H}|\Phi\rangle$ subject to the constraint that the ϕ_i are orthonormalized,

$$\langle\phi_i|\phi_j\rangle = \delta_{ij}. \tag{2.57}$$

When a minimum is found, we denote the solution by $|\Phi_{\mathrm{HF}}\rangle$, the Hartree–Fock state.

The constraints constitute a complication that we want to get rid of. We therefore *Lagrange multipliers*, one for each constraint, giving a Lagrangian functional

$$\begin{aligned}\mathcal{L}[\phi_1, \cdots, \phi_N, \lambda] &= \langle \Phi | \hat{H} | \Phi \rangle - \sum_{ij} \lambda_{ji} (\langle \phi_i | \phi_j \rangle - \delta_{ij}) \\ &= \sum_i \langle \phi_i | \hat{h} | \phi_i \rangle + \frac{1}{2} \sum_{ij} \langle \phi_i \phi_j | \hat{w} | \phi_i \phi_j - \phi_j \phi_i \rangle - \sum_{ij} \lambda_{ji} (\langle \phi_i | \phi_j \rangle - \delta_{ij})\end{aligned} \tag{2.58}$$

Recall, that computing an extremum for the constrained problem is equivalent to an *unconstrained* extremalization of $\mathcal{L}$ with respect to the ϕ_i *and* the Lagrange multipliers (see any text on vector calculus).

Due to symmetry of the constraints, the Lagrange multiplier matrix λ can be assumed to be Hermitian[1].

A word on a special notation. We define a single-particle function

$$\langle \cdot \, \phi_1 | \hat{w} | \phi_2 \phi_3 \rangle \in L_1^2 \tag{2.59}$$

as the function obtained by integrating only over the second particle in the inner product, viz,

$$\langle x_1 | \langle \cdot \, \phi_1 | \hat{w} | \phi_3 \phi_3 \rangle = \langle \cdot \, \phi_1 | \hat{w} | \phi_3 \phi_3 \rangle (x_1) \equiv \int \phi_1(x_2)^* [w(x_1, x_2) \phi_2(x_1) \phi_3(x_2)] \, dx_2. \tag{2.60}$$

The inner product with any single-particle function χ is

$$\langle \chi | \langle \cdot \, \phi_1 | \hat{w} | \phi_2 \phi_3 \rangle = \iint \chi(x_1)^* \phi_1(x_2)^* [w(x_1, x_2) \phi_2(x_1) \phi_3(x_2)] \, dx_1 dx_2 = \langle \chi \phi_1 | \hat{w} | \phi_2 \phi_3 \rangle, \tag{2.61}$$

i.e., the full two-particle integral. Thus, the dot represents an "unused slot" in the two-particle matrix element.

We can expand the function in any orthonormal single-particle basis $\{\chi_p\} \subset L_1^2$,

$$\langle \cdot \, \phi_1 | \hat{w} | \phi_2 \phi_3 \rangle = \sum_p |\chi_p\rangle \langle \chi_p | \langle \cdot \, \phi_1 | \hat{w} | \phi_2 \phi_3 \rangle = \sum_p |\chi_p\rangle \langle \chi_p \phi_1 | \hat{w} | \phi_2 \phi_3 \rangle, \tag{2.62}$$

i.e., a linear combination of to-particle matrix elements. This notation will be useful when 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*

$$\hat{f}(\phi_1, \cdots, \phi_N) |\phi_i\rangle = \epsilon_i |\phi_i\rangle, \tag{2.63}$$

where

$$\hat{f}(\phi_1, \cdots, \phi_N) \equiv \hat{h} + \hat{v}^{direct} - \hat{v}^{exchange}, \tag{2.64}$$

with

$$\hat{v}^{direct} |\psi\rangle \equiv \sum_j \langle \cdot \, \phi_j | \hat{w} | \psi \phi_j \rangle. \tag{2.65}$$

and

$$\hat{v}^{exchange} |\psi\rangle \equiv \sum_j \langle \cdot \, \phi_j | \hat{w} | \phi_j \psi \rangle. \tag{2.66}$$

The equations (2.63) *are referred to as "the Hartree–Fock equations". The operator* $\hat{f}$ *in Eq.* (2.64) *is "the Fock operator", and* $\hat{v}^{direct}$ *and* $\hat{v}^{exchange}$ *are the direct- and exchange potentials, respectively.*

[1]To see this, assume that a_{ji} is a matrix which is not assumed to be Hermitian. Note that the expression $g_{ij} = \langle \phi_i | \phi_j \rangle - \delta_{ij}$ satisfies $g_{ij}^* = g_{ji}$. Thus, $\sum_{ij} a_{ji} g_{ij} = \sum_{ij} a_{ji} g_{ji}^* = \sum_{ij} a_{ij} g_{ij}^* = (\sum_{ij} a_{ij}^* g_{ij})^*$. This gives $\sum_{ij} a_{ji} g_{ij} = \frac{1}{2} \sum_{ij} (a_{ji} + a_{ij}^*) g_{ij}$. Take $\lambda_{ji} = a_{ji} + a_{ij}^*$.

Proof. In the language of Sec. A.1, we need to show that the directional derivative of the Lagrangian vanishes.

We first note that λ_{ji} can be treated separately: $\partial\mathcal{L}/\partial\lambda_{ji} = \langle\phi_i|\phi_j\rangle - \delta_{ij}$, the constraint. These equations are ensured fulfilled in the end by finding solutions ϕ_i that are in fact orthonormal. We thus only compute the directional derivatives with respect to variations of the ϕ_i.

Choose a $k \in \{1, \cdots, N\}$. We are going to compute the directional derivative with respect to changes in the function ϕ_k only, leaving the other fixed. This turns out to be sufficient to find all the equations. Thus, let ϵ be a small real number, and let η be a normalized single-particle function. We write

$$\delta\phi_k = \epsilon\eta.$$

The other functions are fixed, $\delta\phi_i = 0$ for $i \neq k$. Define the function

$$f(\epsilon) = \mathcal{L}(\phi_1, \cdots, \phi_k + \epsilon\eta, \cdots, \phi_N, \lambda), \tag{2.67}$$

To first order in ϵ,

$$f(\epsilon) = f(0) + \epsilon f'(0) + O(\epsilon^2), \tag{2.68}$$

and we for an extremal point of $\mathcal{L}$, we must have that for any η, $f'(0) = 0$. In the language of Sec. A.1, the directional derivative of $\mathcal{L}$ at $\{\phi_i\}_{i=1}^N$ in the direction η (for ϕ_k, the others are fixed) vanishes.

We compute the Taylor expansion of $f(\epsilon)$ by direct computation of the perturbed Lagrangian:

$$\begin{aligned} f(\epsilon) = &\sum_i \langle\phi_i + \delta_{ki}\epsilon\eta|\hat{h}|\phi_i + \delta_{ki}\epsilon\eta\rangle + \frac{1}{2}\sum_{ij}\langle(\phi_i + \delta_{ki}\epsilon\eta)(\phi_j + \delta_{kj}\epsilon\eta)|\hat{w}|(\phi_i + \delta_{ki}\epsilon\eta)(\phi_j + \delta_{kj}\epsilon\eta)\rangle \\ &- \frac{1}{2}\sum_{ij}\langle(\phi_i + \delta_{ki}\epsilon\eta)(\phi_j + \delta_{kj}\epsilon\eta)|\hat{w}|(\phi_j + \delta_{kj}\epsilon\eta)(\phi_i + \delta_{ki}\epsilon\eta)\rangle \\ &- \sum_{ij}\lambda_{ji}(\langle\phi_i + \delta_{ik}\epsilon\eta|\phi_j + \delta_{jk}\epsilon\eta\rangle - \delta_{ij}) \end{aligned} \tag{2.69}$$

We now write out the matrix elements, but keep only terms up to first order in ϵ. This gives

$$\begin{aligned} f(\epsilon) = &\sum_i \langle\phi_i|\hat{h}|\phi_i\rangle + \frac{1}{2}\sum_{ij}\langle\phi_i\phi_j|\hat{w}|\phi_i\phi_j - \phi_j\phi_i\rangle + \epsilon\langle\eta|\hat{h}|\phi_k\rangle + \epsilon\langle\phi_k|\hat{h}|\eta\rangle \\ &+ \frac{1}{2}\epsilon\sum_j\langle\eta\phi_j|\hat{w}|\phi_k\phi_j\rangle + \frac{1}{2}\epsilon\sum_i\langle\phi_i\eta|\hat{w}|\phi_i\phi_k\rangle - \frac{1}{2}\epsilon\sum_j\langle\eta\phi_j|\hat{w}|\phi_j\phi_k\rangle - \frac{1}{2}\epsilon\sum_i\langle\phi_i\eta|\hat{w}|\phi_k\phi_i\rangle \\ &+ \frac{1}{2}\epsilon\sum_j\langle\phi_k\phi_j|\hat{w}|\eta\phi_j\rangle + \frac{1}{2}\epsilon\sum_i\langle\phi_i\phi_k|\hat{w}|\phi_i\eta\rangle - \frac{1}{2}\sum_i\epsilon\langle\phi_i\phi_k|\hat{w}|\eta\phi_i\rangle - \frac{1}{2}\sum_j\epsilon\langle\phi_k\phi_j|\hat{w}|\phi_j\eta\rangle \\ &- \sum_j\lambda_{jk}\epsilon\langle\eta|\phi_j\rangle - \sum_i\lambda_{ki}\epsilon\langle\phi_i|\eta\rangle - \sum_{ij}\lambda_{ji}(\langle\phi_i|\phi_i\rangle - \delta_{ij}) + O(\epsilon^2) \end{aligned} \tag{2.70}$$

We now use the symmetry property of the matrix elements of $\hat{w}$. This gives, for example,

$$\frac{1}{2}\sum_i\langle\phi_i\eta|\hat{w}|\phi_i\phi_k\rangle = \frac{1}{2}\sum_i\langle\eta\phi_i|\hat{w}|\phi_k\phi_i\rangle = \frac{1}{2}\sum_j\langle\eta\phi_j|\hat{w}|\phi_k\phi_j\rangle\,. \tag{2.71}$$

This gives a simplification of $f(\epsilon)$, and we regroup:

$$\begin{aligned} f(\epsilon) = &\sum_i \langle \phi_i | \hat{h} | \phi_i \rangle + \frac{1}{2} \sum_{ij} \langle \phi_i \phi_j | \hat{w} | \phi_i \phi_j - \phi_j \phi_i \rangle - \sum_{ij} \lambda_{ji} (\langle \phi_i | \phi_i \rangle - \delta_{ij}) \\ &+ \epsilon \langle \eta | \hat{h} | \phi_k \rangle + \epsilon \langle \phi_k | \hat{h} | \eta \rangle + \epsilon \sum_j \langle \eta \phi_j | \hat{w} | \phi_k \phi_j \rangle - \epsilon \sum_j \langle \eta \phi_j | \hat{w} | \phi_j \phi_k \rangle \\ &+ \epsilon \sum_j \langle \phi_k \phi_j | \hat{w} | \eta \phi_j \rangle - \sum_j \epsilon \langle \phi_j \phi_k | \hat{w} | \eta \phi_j \rangle - \epsilon \sum_j \lambda_{jk} \langle \eta | \phi_j \rangle - \epsilon \sum_j \lambda_{kj} \langle \phi_j | \eta \rangle + O(\epsilon^2) \end{aligned} \tag{2.72}$$

We recognize that the zeroth order term is just $f(0) = \mathcal{L}(\phi_1, \cdots, \phi_N, \lambda)$. We read off $f'(\epsilon)$, and obtain the directional derivative, and hence the equation

$$\begin{aligned} 0 = &\langle \eta | \hat{h} | \phi_k \rangle + \langle \phi_k | \hat{h} | \eta \rangle + \sum_j \langle \eta \phi_j | \hat{w} | \phi_k \phi_j \rangle - \sum_j \langle \eta \phi_j | \hat{w} | \phi_j \phi_k \rangle \\ &+ \sum_j \langle \phi_k \phi_j | \hat{w} | \eta \phi_j \rangle - \sum_j \langle \phi_j \phi_k | \hat{w} | \eta \phi_j \rangle - \sum_j \lambda_{kj} \langle \phi_j | \eta \rangle - \sum_j \lambda_{jk} \langle \eta | \phi_j \rangle , \end{aligned} \tag{2.73}$$

which must be valid for all choices of the function η. In particular we can also insert $i\eta$, giving, after dividing the result by i,

$$\begin{aligned} 0 = &- \langle \eta | \hat{h} | \phi_k \rangle + \langle \phi_k | \hat{h} | \eta \rangle - \sum_j \langle \eta \phi_j | \hat{w} | \phi_k \phi_j \rangle + \sum_j \langle \eta \phi_j | \hat{w} | \phi_j \phi_k \rangle \\ &+ \sum_j \langle \phi_k \phi_j | \hat{w} | \eta \phi_j \rangle - \sum_j \langle \phi_j \phi_k | \hat{w} | \eta \phi_j \rangle - \sum_j \lambda_{kj} \langle \phi_j | \eta \rangle + \sum_j \lambda_{jk} \langle \eta | \phi_j \rangle , \end{aligned} \tag{2.74}$$

Subtracting the two equations gives us

$$0 = \langle \eta | \hat{h} | \phi_k \rangle + \sum_j \langle \eta \phi_j | \hat{w} | \phi_k \phi_j \rangle - \sum_j \langle \eta \phi_j | \hat{w} | \phi_j \phi_k \rangle - \sum_j \lambda_{jk} \langle \eta | \phi_j \rangle , \quad \forall \eta. \tag{2.75}$$

Let $\{\chi_p\}$ be a complete orthonormal basis for the single-particle space L_1^2. Inserting $\eta = \chi_p$ in Eq. (2.75), we obtain

$$\begin{aligned} 0 &= \sum_p |\chi_p\rangle \left\{ \langle \chi_p | \hat{h} | \phi_k \rangle + \sum_j \langle \chi_p \phi_j | \hat{w} | \phi_k \phi_j \rangle - \sum_j \langle \chi_p \phi_j | \hat{w} | \phi_j \phi_k \rangle - \sum_j \lambda_{jk} \langle \chi_p | \phi_j \rangle \right\} \\ &= \hat{h} | \phi_k \rangle + \sum_j \sum_p |\chi_p\rangle \langle \chi_p \phi_j | \hat{w} | \phi_k \phi_j \rangle - \sum_j \sum_p |\chi_p\rangle \langle \chi_p \phi_j | \hat{w} | \phi_j \phi_k \rangle - \sum_j \lambda_{jk} | \phi_j \rangle . \end{aligned} \tag{2.76}$$

Here, we used

$$1 = \sum_p |\chi_p\rangle \langle \chi_p| . \tag{2.77}$$

We use Eq. (2.62), to get

$$0 = \hat{h} | \phi_k \rangle + \sum_j \langle \cdot \phi_j | \hat{w} | \phi_k \phi_j \rangle - \sum_j \langle \cdot \eta \phi_j | \hat{w} | \phi_j \phi_k \rangle - \sum_j \lambda_{jk} | \phi_j \rangle . \tag{2.78}$$

We now get rid of λ, replacing it with a diagonal matrix with diagonal elements ϵ_k (not to be confused with the small parameter ϵ above, which we now are done with.)

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

$$\tilde{\phi}_k = \sum_j \phi_j U_{jk} \tag{2.79}$$

with U a unitary matrix, then $|\tilde{\Phi}\rangle = \det(U)\,|\Phi\rangle$, i.e., the same state, and clearly the energy must be the same too.

As argued, $\lambda_{ij} = \lambda^*_{ji}$ can be assumed Hermitian. Select therefore U such that $\lambda = UEU^H$, with $E_{jk} = \delta_{jk}\epsilon_k$ the elements of a diagonal matrix (the eigenvalues of λ):

$$\lambda_{ji} = \sum_\ell U_{j\ell}\epsilon_\ell U^*_{i\ell}. \tag{2.80}$$

Let $|r_i\rangle$ be the right-hand side of Eq. (2.78), and consider

$$\sum_k |r_k\rangle\, U_{ki} = 0. \tag{2.81}$$

Since U is unitary, Eq. (2.78) is satisfied for all k if and only if Eq. (2.81) is satisfied for all i. Computing the sum in Eq. (2.81) (see Exercise 2.9) we obtain

$$\hat{h}\,|\tilde{\phi}_i\rangle + \sum_j \left[\langle\,\cdot\,\tilde{\phi}_j|\hat{w}|\tilde{\phi}_i\tilde{\phi}_j\rangle - \langle\,\cdot\,\tilde{\phi}_j|\hat{w}|\tilde{\phi}_j\tilde{\phi}_i\rangle\right] - \epsilon_i\,|\tilde{\phi}_i\rangle = 0. \tag{2.82}$$

This must hold for all $i = 1, \cdots, N$ simultaneously.

With the definitions of $\hat{v}^{\text{direct}}$ and $\hat{v}^{\text{exchange}}$ in the theorem formulation, we are finished. □

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 $\hat{v}^{\text{direct}}$ 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

$$\hat{w}_{\text{Coul}}(x_1, x_2) = \frac{1}{|\vec{r}_1 - \vec{r}_2|}, \quad x_i = (\vec{r}_i, \sigma). \tag{2.83}$$

The operator $\hat{v}^{\text{direct}}$ 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

$$\begin{aligned}\langle x_1|\,(\hat{v}^{\text{direct}}\,|\psi\rangle) &= \sum_j \langle x_1|\,\langle\,\cdot\,\phi_j|\hat{w}|\psi\phi_j\rangle = \sum_j \int \phi^*_j(x_2) w(x_1, x_2)\phi_j(x_2)\psi(x_1)\,dx_2 \\ &= \left[\int \sum_j |\phi_j(x_2)|^2 w(x_1, x_2)\,dx_2\right]\psi(x_1) \equiv v^{\text{direct}}(x_1)\psi(x_1).\end{aligned} \tag{2.84}$$

Thus, $\hat{v}^{\text{direct}}$ 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

$$\langle x_1|\,(\hat{v}^{\text{exchange}}\,|\psi\rangle) = \sum_j \langle x_1|\,\langle\,\cdot\,\phi_j|\hat{w}|\phi_j\psi\rangle = \sum_j \int \phi^*_j(x_2) w(x_1, x_2)\psi(x_2)\phi(x_1)\,dx_2. \tag{2.85}$$

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

$$\gamma(x, x') = \sum_j \phi_j(x)\phi_j(x')^*, \tag{2.86}$$

we can express

$$\langle x_1 | \hat{v}^{\text{direct}} | \psi \rangle = \psi(x_1) \int \gamma(x_2, x_2) w(x_1, x_2) \, dx_2. \tag{2.87}$$

$$\langle x_1 | \hat{v}^{\text{exchange}} | \psi \rangle = \int \gamma(x_1, x_2) w(x_1, x_2) \psi(x_2) \, dx_2. \tag{2.88}$$

The reduced density matrix γ will turn out to be a useful concept in Hartree–Fock theory.

In the proof of the HF equations, we first found an equation whose solutions were not eigenfunctions, Eq. (2.78). However, by forming a particular linear combination, the equation was brought on eigenvalue form, Eq. (2.63). 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 diagonal form of the HF equations are referred to as *the canonical HF equations*, while the non-diagonal form is *non-canonical*.

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 when these have been found, $\hat{f} = \hat{f}(\phi_1, \cdots, \phi_N)$ is a fixed Hermitian operator (see Exercise 2.7), and we can in principle[2] find a complete basis of eigenvectors of $\hat{f}$,

$$\{\phi_p\} = \{\phi_i\} \cup \{\phi_a\}. \tag{2.89}$$

This particular orthonormal basis is often taken as basis for proper manybody treatments, such as CI calculations, perturbation theory, and coupled-cluster (CC) theory (these are topics we return to later). It is referred to as the *canonical basis*, and Eq. (2.91) is an extenion of Eq. (2.63) to include the extra single-particle functions ϕ_a.

This rather central, that we write it up as a definition that we can refer to later:

Definition 2.1 (Canonical HF equations, HF basis)**.** For a given two-body Hamiltonian

$$\hat{H} = \sum_{i=1}^{N} \hat{h}(i) + \sum_{i<j}^{N} \hat{w}(i, j), \tag{2.90}$$

The equation

$$\hat{f}(\phi_1, \cdots, \phi_N) |\phi_p\rangle = \epsilon_p |\phi_p\rangle . \tag{2.91}$$

with the Fock operator

$$\hat{f}(\phi_1, \cdots, \phi_N) = \hat{h} + \hat{v}^{\text{direct}} - \hat{v}^{\text{exchange}}, \tag{2.92}$$

is referred to as the *canonical Hartree–Fock equations*, and the solutions are called the *canonical single-particle functions*.

The first N HF single-particle functions ϕ_i are often called *occupied*, while the rest, ϕ_a, are often called *virtual* single-particle functions.

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* (ϕ_i, ϵ_i)*,* $i = 1, \cdots, N$*, to the canonical Hartree–Fock equations have been found. Then, the Hartree–Fock energy is given by*

$$E_{HF} = \sum_i \epsilon_i - \frac{1}{2} \sum_{ij} \langle \phi_i \phi_i | \hat{w} | \phi_i \phi_j - \phi_j \phi_i \rangle . \tag{2.93}$$

[2] It happens that $\hat{f}$ has a continuous spectrum, so our statement must really be limited to finite-dimensional one-particle spaces for strict validity.

Proof. Multiply the HF equation from the left by $\langle\phi_i|$ and sum over i to obtain

$$\sum_i \epsilon_i = \sum_i \langle\phi_i|\hat{h}|\phi_i\rangle + \sum_{ij} \langle\phi_i\phi_j|\hat{w}|\phi_i\phi_j - \phi_j\phi_i\rangle\,. \tag{2.94}$$

We see that the interaction is double counted compared to Eq. (2.55), and we are finished. □

Exercise 2.7. Suppose the HF single-particle functions have been found, so that the Fock operator $\hat{f}$ is a fixed operator. Prove that it is Hermitian, i.e., for any two single-particle functions $\psi(x)$ and $\psi'(x)$,

$$\langle\psi|\hat{f}|\psi'\rangle = [\langle\psi'|\hat{f}|\psi\rangle]^*.$$

△

Exercise 2.8. 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

$$\tilde{\phi}_i = \sum_j \phi_j U_{ji}. \tag{2.95}$$

Show that

$$\gamma(x,x') = \sum_j \tilde{\phi}_j(x)\tilde{\phi}_j(x')^*. \tag{2.96}$$

What can you conclude about $\hat{v}^{\text{direct}}$ and $\hat{v}^{\text{exchange}}$, which are functions of γ? △

Exercise 2.9. In this exercise, we fill in the details between Eq. (2.81) and Eq. (2.82) in the proof of Theorem 2.4.

a) Verify that

$$\sum_k U_{ki}\hat{h}\,|\phi_k\rangle = \hat{h}\,|\tilde{\phi}_i\rangle\,. \tag{2.97}$$

b) Next, show that

$$\sum_k U_{ki} \sum_j \lambda_{jk}\,|\phi_j\rangle = \epsilon_i\,|\tilde{\phi}_i\rangle\,. \tag{2.98}$$

c) As an intermediate calculation, verify that

$$|\phi_i\rangle = \sum_k U^H_{ki}\,|\tilde{\phi}_k\rangle = \sum_k U^*_{ik}\,|\tilde{\phi}_k\rangle\,. \tag{2.99}$$

d) Show that

$$\sum_k U_{ki} \sum_j \langle\,\cdot\,\phi_j|\hat{w}|\phi_k\phi_j\rangle = \sum_j \langle\,\cdot\,\tilde{\phi}_j|\hat{w}|\tilde{\phi}_i\tilde{\phi}_j\rangle\,. \tag{2.100}$$

You may do the transformations of the various ϕ_ℓ into $\tilde{\phi}_\ell$ using c), or use Exercise 2.8.

e) Show that

$$\sum_k U_{ki} \sum_j \langle\,\cdot\,\phi_j|\hat{w}|\phi_j\phi_k\rangle = \sum_j \langle\,\cdot\,\tilde{\phi}_j|\hat{w}|\tilde{\phi}_j\tilde{\phi}_i\rangle\,. \tag{2.101}$$

f) Gather the results of a), b), d), and e), to show that Eq. (2.81) becomes Eq. (2.82).

△

2.4.2 The Hartree–Fock equations in a given basis: the Roothan–Hall equations

How do we solve the HF equations (2.63)? 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. However, we do *not* assume that it is orthonormal. Thus, we have a *possibly non-diagonal overlap matrix S of size* $L \times L$,

$$S_{pq} \equiv \langle \chi_p | \chi_q \rangle \,. \tag{2.102}$$

and we must have that S^{-1} exists since the ϕ_p form a basis.

Such basis functions are common in quantum chemistry, where a non-orthogonal basis of *Gaussian functions centered on the atoms* is typically employed. See for example Szabo/Ostlund or Helgaker/Jørgensen/Olsen for details. For now, we just keep this remark as a motivation for not assuming orthogonality. In nuclear physics or solid state physics, orthogonal functions χ_p are more typical.

We expand our HF functions as

$$|\phi_p\rangle = \sum_q |\chi_q\rangle \, U_{qp}, \tag{2.103}$$

where U is in general not a unitary matrix, since the basis is not orthogonal. (However, we have $U^H S U = I$, the identity matrix, see Exercise 2.10.) We notice that the *columns* of U 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

$$\gamma(x,x') = \sum_i \langle x|\phi_i\rangle \langle \phi_i|x'\rangle = \sum_{pq}\sum_i U_{qi} \, |\chi_q\rangle \langle \chi_p| \, U^*_{pi} = \sum_{pq}(\sum_i U_{qi}U^*_{pi}) \langle x|\chi_q\rangle \langle \chi_p|x'\rangle = \sum_{pq}(U_{1:N}U^H_{1:N})_{qp} \langle x|\chi_q\rangle \langle \chi_p|x'\rangle \,, \tag{2.104}$$

and it makes sense to define

$$D = U_{1:N}U^H_{1:N} = \sum_i u_i u_i^H, \tag{2.105}$$

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 canonical HF equations (2.91) can be written

$$F(D)U = SU\epsilon, \tag{2.106}$$

where

$$F_{pq} = \langle \chi_p | \hat{f}(\phi_1, \cdots, \phi_N) | \chi_q \rangle \tag{2.107}$$

are the matrix elements of the Fock operator in the fixed basis, and where $\epsilon = \mathrm{diag}(\epsilon_1, \cdots, \epsilon_L)$ is a diagonal matrix. Equation (2.91) is a nonlinear generalized eigenvalue problem.

Let us look at the matrix elements of f,

$$F_{qp} = \langle \chi_q | \hat{f} | \chi_p \rangle = \langle \chi_q | \hat{h} | \chi_p \rangle + \langle \chi_q | \hat{v}^{\mathrm{direct}} | \chi_q \rangle - \langle \chi_q | \hat{v}^{\mathrm{exchange}} | \chi_p \rangle \,. \tag{2.108}$$

The direct term is

$$\begin{aligned} \langle \chi_q | \hat{v}^{\mathrm{direct}} | \chi_p \rangle &= \sum_j \langle \chi_q \phi_j | \hat{w} | \chi_p \phi_j \rangle = \sum_{p'q'j} U_{jq'} U^*_{jp'} \langle \chi_q \chi_{q'} | \hat{w} | \chi_p \chi_{p'} \rangle \\ &= \sum_{p'q'} D_{q'p'} \langle \chi_q \chi_{q'} | \hat{w} | \chi_p \chi_{p'} \rangle \,. \end{aligned} \tag{2.109}$$

Correspondingly,

$$\begin{aligned} \langle \chi_q | \hat{v}^{\mathrm{exchange}} | \chi_p \rangle &= \sum_j \langle \chi_q \phi_j | \hat{w} | \phi_j \chi_p \rangle = \sum_{p'q'j} U_{jq'} U^*_{jp'} \langle \chi_q \chi_{q'} | \hat{w} | \chi_{p'} \chi_p \rangle \\ &= \sum_{p'q'} D_{q'p'} \langle \chi_q \chi_{q'} | \hat{w} | \chi_{p'} \chi_p \rangle \,. \end{aligned} \tag{2.110}$$

We obtain

$$F_{qp} = \langle \chi_q | \hat{h} | \chi_p \rangle + \sum_{p'q'} D_{p'q'} (\langle \chi_q \chi_{q'} | \hat{w} | \chi_p \chi_{p'} \rangle - \langle \chi_q \chi_{q'} | \hat{w} | \chi_{p'} \chi_p \rangle). \tag{2.111}$$

Note that we have expressed F_{qp} in terms of *non-antisymmetric* matrix elements of $\hat{w}$.

Thus, projecting the LHS of the canonical HF equations onto the basis gives

$$\langle \chi_q | \hat{f} | \phi_p \rangle = \sum_{q'} \langle \chi_q | \hat{f} | \chi_{q'} \rangle U_{q'p} = \sum_{q'} F_{qq'} U_{q'p}, \quad \forall q, p. \tag{2.112}$$

The right-hand side gives the projection

$$\langle \chi_q | \phi_p \rangle \epsilon_p = \sum_{q'} \langle \chi_q | \chi_{q'} \rangle U_{q'p} \epsilon_p = \sum_{q'} S_{qq'} U_{q'p} \epsilon_i, \quad \forall q, p. \tag{2.113}$$

Gathering, we find

$$F(D)U = SU\epsilon, \tag{2.114}$$

and we are finished. This equation is called the Roothan–Hall equation.

In terms of each column, i.e., each ϕ_p,

$$F(D)u_p = \epsilon_p S u_p. \tag{2.115}$$

Exercise 2.10. Prove that $U^H S U = I$ (the identity matrix) by using $\langle \phi_p | \phi_q \rangle = \delta_{pq}$ and

$$|\phi_p\rangle \sum_q S_{qp} |\chi_q\rangle, \quad S_{qp} = \langle \chi_q | \chi_p \rangle. \tag{2.116}$$

△

2.4.3 Self-consistent field iteration

How do we find self-consistent solutions of Eq. (2.115)? The standard approach is by self-consistent field iteractions (SCF iterations), Finding hopefully better and better approximations $u_i^{(k)}$, $k = 1, 2, 3, \cdots$, 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

$$F(D^{(k)}) u_p^{(k+1)} = \epsilon_p^{(k+1)} S u_p^{(k+1)} \tag{2.117}$$

by numerical diagonalization, sorting the eigenvalues $\epsilon_p^{(k+1)}$ in ascending order. Then, $p = 1, \cdots, N$ gives the next approximation to the HF eigenpairs (ϕ_i, ϵ_i), while the next $L - N$ form the additional canonical functions.

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 aboud DIIS in Helgaker/Jørgensen/Olsen [6], and see also `https://en.wikipedia.org/wiki/DIIS`.

2.4.4 Basis expansions in HF single-particle functions

We have now established the canonical Hartree–Fock single-particle functions, which can be used as a basis just like any other orthonormal basis. Each canonical 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

$$\langle pq|\hat{w}|rs\rangle_{\text{AS}} = \langle \phi_p\phi_q|\hat{w}|\phi_r\phi_s\rangle_{\text{AS}} = \sum_{p'q'r's'} U^*_{p'p}U^*_{q'q}U_{r'r}U_{s's}\,\langle \chi_{p'}\chi_{q'}|\hat{w}|\chi_{r'}\chi_{s'}\rangle_{\text{AS}} \tag{2.118}$$

and

$$\langle p|\hat{h}|q\rangle = h^p_q = \langle \phi_p|\hat{h}|\phi_q\rangle = \sum_{p'q'} U^*_{p'p}U_{q'q}\,\langle \chi_{p'}|\hat{h}|\chi_{q'}\rangle\,, \tag{2.119}$$

and similarly for any one-body operator. In a situation where HF single-particle functions are used in, say, a CI program, the matrix elements $\langle \chi_{p'}\chi_{q'}|\hat{w}|\chi_{r'}\chi_{s'}\rangle_{(\text{AS})}$ and $\langle \chi_{p'}|\hat{h}|\chi_{q'}\rangle$ will be produced by external codes. This is especially true in chemistry, where the computation of matrix elements is a business on its own.

In quantum chemistry, it is *standard* to start with the HF single-particle functions and perform corrections on top of that, such as CISD, giving rise to the term "post-Hartree–Fock methods".

It is convenient to write the Hamiltonian on the following form

$$\hat{H} = \hat{H}_0 + \hat{W} = \hat{F} + \hat{U}, \tag{2.120}$$

where the second-quantized Fock operator is given by

$$\hat{F} = \sum_{i=1}^{N}\hat{f}(i) = \hat{H}_0 + \hat{V}^{\text{direct}} - \hat{V}^{\text{exchange}}, \tag{2.121}$$

and where the *fluctuation potential* is given by

$$\hat{U} = \hat{W} - \hat{V}^{\text{direct}} + \hat{V}^{\text{exchange}}. \tag{2.122}$$

Here,

$$\hat{V}^{\text{direct}} = \sum_i \hat{v}^{\text{direct}}(i), \quad \hat{V}^{\text{exchange}} = \sum_i \hat{v}^{\text{exchange}}(i). \tag{2.123}$$

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": after all, we have chosen the HF state such that it contains as much of the interaction energy as possible, by minimizing the energy over all possible determinants. Thus, the exact wavefunction $|\Psi\rangle = |\Phi\rangle + \delta\,|\Psi\rangle$ consists of "small fluctutations" on top of $|\Phi\rangle$ caused by $\hat{U}$.

An expression for the direct potential operator matrix element is

$$\langle \phi_q|\,\hat{v}^{\text{direct}}\,|\phi_p\rangle = \sum_i \langle \phi_i\phi_q|\hat{w}|\phi_i\phi_p\rangle\,, \tag{2.124}$$

with non-antisymmetric matrix elements. Thus,

$$\hat{V}^{\text{direct}} = \sum_{pq}\sum_i \langle \phi_i\phi_q|\hat{w}|\phi_i\phi_p\rangle\, c_q^\dagger c_p. \tag{2.125}$$

Similarly, for the the exchange potential we get

$$\hat{V}^{\text{exchange}} = \sum_{pq}\sum_i \langle \phi_i\phi_q|\hat{w}|\phi_p\phi_i\rangle\, c_q^\dagger c_p. \tag{2.126}$$

This results in (using antisymmetrized matrix elements (2.118))

$$\hat{F} = \hat{H}_0 + \sum_{pq}\sum_{i} \langle qi|\hat{w}|pi\rangle_{\mathrm{AS}}\, c_q^\dagger c_p \tag{2.127}$$

$$\hat{U} = \hat{W} - \sum_{pq}\sum_{i} \langle qi|\hat{w}|pi\rangle_{\mathrm{AS}}\, c_q^\dagger c_p. \tag{2.128}$$

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 single-particle functionss, a basis of Slater determinants can be taken to be the $|p_1 \cdots p_N\rangle$, with $p_1 < p_2 < \cdots p_N$. Alternatively, we can use the quasiparticle picture, and let the HF function be the reference,

$$|\Phi\rangle = c_1^\dagger \cdots c_N^\dagger\, |-\rangle\,. \tag{2.129}$$

All other Slater determinant basis functions can be written

$$|\Phi_i^a\rangle = c_a^\dagger c_i\, |\Phi_{\mathrm{HF}}\rangle = b_a^\dagger b_i^\dagger\, |\Phi\rangle\,, \tag{2.130}$$

$$|\Phi_{ij}^{ab}\rangle = c_b^\dagger c_j c_a^\dagger c_i\, |\Phi\rangle = b_b^\dagger b_j^\dagger b_a^\dagger b_i^\dagger\, |\Phi\rangle\,, \tag{2.131}$$

etc, where we have introduced the quasiparticle creation- and annihilation operators.

All the determinants $|p_1, \cdots, p_N\rangle$ are eigenfunctions of $\hat{F}$,

$$\hat{F}\, |p_1, \cdots, p_N\rangle = \left(\sum_i \epsilon_{p_i}\right) |p_1, \cdots, p_N\rangle\,, \tag{2.132}$$

and in particular the HF function $|\Phi\rangle$ is the "ground-state" of $\hat{F}$.

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 a reformulation in terms of second-quantization, namely *Brillouin's Theorem*:

Theorem 2.6 (Brillouin's Theorem)**.** *Let an orthonormal single-particle basis $\{\phi_p\}$ be given, and and assume that these satisfy the canonical HF equations. Then,*

$$\langle \Phi_i^a|\hat{H}|\Phi\rangle = 0, \qquad \forall\, i, a. \tag{2.133}$$

Proof. Assume that the HF equations are satisfied. Since $\hat{f}$ is Hermitian, the single-particle basis functions are orthonormal. The Fock matrix becomes diagonal,

$$f_q^p = h_q^p + \sum_j \langle pj|\hat{w}|qj\rangle = \delta_{pq}\epsilon_q. \tag{2.134}$$

In particular,

$$f_i^a = h_i^q + \sum_j \langle aj|\hat{w}|ij\rangle = 0. \tag{2.135}$$

But this is precisely (see the Slater–Condon rules from Exercise 1.18) the expression for $\langle \Phi_i^a|\, \hat{H}\, |\Phi\rangle$, which therefore must vanish for all i, a. □

The converse of Brilloin's theorem is also true, in the sense that $f_i^a = 0$ is equivalent to the *non-canonical* HF equations. Recall that the HF state is the same for the non-canonical and canonical single-particle functions.

Theorem 2.7 (Converse of Brillouin's Theorem)**.** *Let a single-particle basis be given. This basis satisfies*

$$\langle \Phi_i^a | \hat{H} | \Phi \rangle = 0, \qquad \forall i, a \tag{2.136}$$

if and only if the non-canonical HF equations are satified for the occupied ϕ_i*,* $i = 1, \cdots, N$*.*

Proof. Since $f_a^i = (f_i^a)^* = 0$,

$$\hat{f} \, |\phi_i\rangle = \sum_p \langle \phi_p | \hat{f} | \phi_i \rangle \, |\phi_p\rangle = \sum_j \langle \phi_j | \hat{f} | \phi_i \rangle \, |\phi_j\rangle \,, \tag{2.137}$$

This is implies $\hat{f} \, |\phi_i\rangle = \sum_j \lambda_{ji} \, |\phi_j\rangle$ with $\lambda_{ji} = f_i^j$, which are the non-canonical HF equations. Conversely, assume that the non-canonical HF equations are satisfied by the ϕ_i,

$$\hat{f} \, |\phi_i\rangle = \sum_j \lambda_{ji} \, |\phi_j\rangle \,. \tag{2.138}$$

Forming the inner product with ϕ_j, we otain $f_i^j = \lambda_{ji}$, and Eq. (2.137) is satisfied. □

Because of Brillouin's Theorem, a configuration-interaction treatment win only singles (CIS) yields no correction over the HF treatment alone, and we have to go to doubles.

2.4.5 Restricted Hartree–Fock for electronic systems (RHF)

[There is an unfortunate overlap between the notation for spin functions χ_α and the basis functions χ_p in the previous section. Hopefully no confusion arises.]

We now discuss the *restricted Hartree–Fock (RHF) method* for electronic systems. Supporting material: Szabo and Ostlund.

Motivation:

Consider N electrons, which we assume to a first approximation do not interact among themselves, i.e., we neglect the inter-electron repulsion operator given by

$$\hat{w}(\vec{r}_1, \vec{r}_2) = \frac{1}{|\vec{r}_1 - \vec{r}_2|}, \tag{2.139}$$

in suitable units. The electrons are thus described by a one-body Hamiltonian $\hat{H}_0 = \sum_i \hat{h}(i)$,

$$\hat{h}(\vec{r}) = -\frac{1}{2}\nabla^2 + v(\vec{r}), \tag{2.140}$$

where $v(\vec{r})$ is an external electrostatic potential, such as the one set up by an atomic nucleus. The operator $\hat{h}$ does not couple to electron spin, so that the single-particle eigenfunctions of $\hat{h}$ separate,

$$\phi_\mu(\vec{r}, \sigma) = \varphi_p(\vec{r}) \chi_\alpha(\sigma), \quad \mu = (p, \sigma), \tag{2.141}$$

where $\alpha = \pm 1/2$ is the value of the projection of the electron spin along the z-axis. Also, $\sigma = \pm 1/2$, and $\langle \chi_\alpha | \chi_\beta \rangle = \delta_{\alpha\beta}$. The eigenvalue problem of $\hat{h}(\vec{r})$ becomes

$$\hat{h} \varphi_p(\vec{r}) \chi_\alpha(\sigma) = e_p \varphi_p(\vec{r}) \chi_\alpha(\sigma), \quad \sigma = \pm 1/2. \tag{2.142}$$

where the eigenvalue e_p is seen to be doubly degenerate due to spin. The N-electron ground-state of $\hat{H}_0$ is now given by the Slater determinant with the N first eigensolutios $\phi_{(p,\sigma)}$ occupied. Assuming N even, we get

$$|\Phi\rangle = |\phi_{1,\frac{1}{2}} \phi_{1,-\frac{1}{2}} \cdots \phi_{\frac{N}{2},\frac{1}{2}} \phi_{\frac{N}{2},-\frac{1}{2}}\rangle \,. \tag{2.143}$$

(If N is odd, the ground-state is doubly degenerate, with an electron occupying $\phi_{\lfloor \frac{N}{2} \rfloor + 1, \alpha}$, for $\alpha = +1/2$ or $\alpha = -1/2$.) A common notation is

$$|\Phi_{\mathrm{RHF}}\rangle = |\varphi_1 \bar{\varphi}_1 \varphi_2 \bar{\varphi}_2 \cdots \varphi_{N/2} \bar{\varphi}_{N/2}\rangle \tag{2.144}$$

with the understanding that φ_p represents $\phi_{p,+1/2}$ and $\bar{\varphi}_p$ represents $\phi_{p,-1/2}$.

The idea of RHF is to assume that the exact ground-state has a similar structure. Thus, we do not optimize all the N single-particle functions freely, we assume that they form a set of doubly occupied orbitals. In RHF we therefore compute the HF single-particle functions by minimizing the energy under the assumption that $|\Phi\rangle$ is on the form (2.144).

The HF energy is simplified because of the special case of single-particle functions on factorized spin-orbital form. Consider for example the matrix element

$$\langle \phi_{p,\alpha} | \hat{h} | \phi_{q,\beta} \rangle = \langle \chi_\alpha | \chi_\beta \rangle \int \varphi_p(\vec{r})^* \hat{h}(\vec{r}) \varphi_q(\vec{r}) \, d\vec{r} \equiv \delta_{\alpha\beta} (\varphi_p | \hat{h} | \varphi_q), \tag{2.145}$$

where we have introduced a special notation for the spatial matrix element. Similarly,

$$\langle \phi_{p\alpha} \phi_{q\beta} | \hat{w} | \phi_{r\gamma} \phi_{s\delta} \rangle = \langle \chi_\alpha | \chi_\gamma \rangle \langle \chi_\beta | \chi_\delta \rangle \iint \varphi_p(\vec{r}_1)^* \varphi_q(\vec{r}_2)^* \hat{w}(\vec{r}_1, \vec{r}_2) \varphi_r(\vec{r}_1) \varphi_s(\vec{r}_2) \, d\vec{r}_1 d\vec{r}_2 \equiv \delta_{\alpha\gamma} \delta_{\beta\delta} (\varphi_p \varphi_q | \hat{w} | \varphi_r \varphi_s), \tag{2.146}$$

where we also introduce a special notation to be used in the sequel.

We use Eqs. (2.145–2.146) and compute the energy of $|\Phi\rangle$:

$$\begin{aligned} \langle \Phi | \hat{H} | \Phi \rangle &= \sum_\alpha \sum_{i=1}^{N/2} \langle \phi_{i\alpha} | \hat{h} | \phi_{i\alpha} \rangle + \frac{1}{2} \sum_\alpha \sum_{i=1}^{N/2} \sum_\beta \sum_{j=1}^{N/2} \langle \phi_{i\alpha} \phi_{j\beta} | \hat{w} | \phi_{i\alpha} \phi_{j\beta} - \phi_{j\beta} \phi_{i\alpha} \rangle \\ &= 2 \sum_{i=1}^{N/2} (\varphi_i | \hat{h} | \varphi_i) + 2 \sum_{ij}^{N/2} (\varphi_i \varphi_j | \hat{w} | \varphi_i \varphi_j) - \sum_{ij}^{N/2} (\varphi_i \varphi_j | \hat{w} | \varphi_j \varphi_i) \end{aligned} \tag{2.147}$$

Observe the factor 2 in front of the two first terms.

The RHF state is obtained by minimizing the energy with respect to orthonormal orbitals φ_i, $i = 1, \cdots, N/2$. We obtain the restricted HF equationn.

Theorem 2.8 (Restricted Hartree–Fock equations)**.** *The orbitals of the minimizing RHF state $|\Phi_{RHF}\rangle$ satisfies the RHF equations:*

$$\hat{f}(\gamma) \varphi_i(\vec{r}) = \epsilon_i \varphi_i(\vec{r}), \quad i = 1, \cdots, N/2, \tag{2.148}$$

where the (RHF) Fock operator is given by

$$\hat{f}(\gamma) = \tag{2.149}$$

and where the reduced denity matrix is

$$\gamma(\vec{r}, \vec{r}') = 2 \sum_i \varphi_i(\vec{r}) \varphi_i(\vec{r})^*. \tag{2.150}$$

The RHF energy is

$$E_{RHF} = 2 \sum_{i=1}^{N/2} \epsilon_i - 2 \sum_{ij}^{N/2} (\varphi_i \varphi_j | \hat{w} | \varphi_i \varphi_j) + \sum_{ij}^{N/2} (\varphi_i \varphi_j | \hat{w} | \varphi_j \varphi_i) \tag{2.151}$$

Proof. (Optional reading.) Optimization of RHF energy, and RHF equations: Introducing Lagrange multipliers for the orthonormality constraints, we obtain a Lagrangian

$$\mathcal{L}[\varphi_1, \cdots, \varphi_{N/2}, \lambda] = 2\sum_i (\varphi_i|\hat{h}|\varphi_i) + 2\sum_{ij}(\varphi_i\varphi_j|\hat{w}|\varphi_i\varphi_j) - \sum_{ij}(\varphi_i\varphi_j|\hat{w}|\varphi_j\varphi_i) - 2\sum_{ij}\lambda_{ji}[(\varphi_i|\varphi_j) - \delta_{ij}], \tag{2.152}$$

where we have introduced Lagrange multipliers for the orthonormality constraints. The factor 2 in front of the constraint term is for convenience.

A procedure similar to the derivation of the HF equations (see Exercise 2.11) gives:

$$\hat{h}\varphi_i + 2\sum_j(\,\cdot\,\varphi_j|\hat{w}|\varphi_i\varphi_j) - \sum_j(\,\cdot\,\varphi_j|\hat{w}|\varphi_j\varphi_i) - \sum_j \lambda_{ji}\varphi_j = 0. \tag{2.153}$$

A unitary transformation similar to the one for the HF equations allow us to replace λ by a diagonal matrix, finally obtaining

$$[\hat{h} + \hat{v}^{\text{Coulomb}} - \hat{v}^{\text{exchange}}]\varphi_i(\vec{r}) = \epsilon_i\varphi_i(\vec{r}), \quad i = 1, \cdots, N/2, \tag{2.154}$$

with

$$\hat{v}^{\text{Coulomb}}(\vec{r}) = \int \gamma(\vec{r}', \vec{r}')\frac{1}{|\vec{r} - \vec{r}'|}\, d\vec{r}' \tag{2.155}$$

being a local potential, and where

$$[\hat{v}^{\text{exchange}}\psi](\vec{r}) = \frac{1}{2}\int \gamma(\vec{r}', \vec{r})\frac{1}{|\vec{r} - \vec{r}'|}\psi(\vec{r}')\, d\vec{r}' \tag{2.156}$$

is a non-local potential. The reduced density matrix is

$$\gamma(\vec{r}, \vec{r}') \equiv 2\sum_{j=1}^{N/2}\varphi_j(\vec{r})\varphi_j(\vec{r}')^* \tag{2.157}$$

The proof of Eq. (2.151) is obtained by taking the inner product of Eq. (2.148) with φ_i and summing over i, then multiplying with 2. □

Exercise 2.11. In this exercise, we prove Theorem 2.8 (To be filled in.) △

2.4.6 Unrestricted Hartree–Fock for electronic systems (UHF)

Supporting material: Szabo and Ostlund.

The RHF model is usually a good approximation, but fails in some circumstances. The *unrestricted* Hartree–Fock model is an intermediate between the general HF model and the restricted HF model. In RHF space orbital i for both spins were required to be identical. In UHF we allow them to be different,

$$\phi_{i,\alpha}(\vec{r}, \sigma) = \varphi_i^\alpha(\vec{r})\chi_\alpha(\sigma). \tag{2.158}$$

Thus, the orbital carries a spin-index as well as a space index, compare with the RHF model. The UHF state can be written

$$|\Phi_{\text{UHF}}\rangle = |\varphi_1^{1/2}\bar{\varphi}_1^{-1/2}\varphi_2^{1/2}\bar{\varphi}_2^{-1/2}\cdots\varphi_{N/2}^{1/2}\bar{\varphi}_{N/2}^{-1/2}\rangle\,, \tag{2.159}$$

compare with Eq. (**??**) Notice that the spin-orbitals are still orthogonal for different spins. Notice also that the general HF model is more general than UHF: there, each spin-orbital was not required to separate into a product of space and spin functions.

The UHF energy expectation value is (see Exercise 2.12)

$$E_{\text{UHF}} = \sum_{\alpha} \sum_{i=1}^{N/2} (\varphi_i^\alpha | \hat{h} | \varphi_i^\alpha) + \frac{1}{2} \sum_{\alpha} \sum_{\beta} \sum_{ij}^{N/2} (\varphi_i^\alpha \varphi_j^\beta | \hat{w} | \varphi_i^\alpha \varphi_j^\beta) - \frac{1}{2} \sum_{\alpha} \sum_{ij}^{N/2} (\varphi_i^\alpha \varphi_j^\alpha | \hat{w} | \varphi_j^\alpha \varphi_i^\alpha). \tag{2.160}$$

The variational UHF equations become

$$\hat{h} \varphi_i^\alpha(\vec{r}) + \sum_{\beta} \sum_{j} (\cdot \, \varphi_j^\beta | \hat{w} | \varphi_i^\alpha \varphi_j^\beta) - \sum_{j} (\cdot \, \varphi_j^\alpha | \hat{w} | \varphi_j^\alpha \varphi_i^\alpha) = \epsilon_i^\alpha \varphi_i^\alpha(\vec{r}), \tag{2.161}$$

where we note that each spin-orbital is not doubly degenerate anymore. We introduce the UHF Coulomb potential,

$$v^{\text{Coulomb}}(\vec{r}) = \int \sum_{j\beta} |\varphi_j^\beta(\vec{r}')|^2 \frac{1}{|\vec{r} - \vec{r}'|} \, d\vec{r}', \tag{2.162}$$

and the UHF exchange potential operator

$$[\hat{v}^{\alpha,\text{exchange}} \psi](\vec{r}) = \int \sum_{j} \varphi_j^\alpha(\vec{r}) \varphi_j^\alpha(\vec{r}')^* \psi(\vec{r}) \frac{1}{|\vec{r} - \vec{r}'|} \, d\vec{r}', \tag{2.163}$$

to obtain

$$[\hat{h} + \hat{v}^{\text{Coulomb}} - \hat{v}^{\alpha,\text{exchange}}] \phi_i^\alpha(\vec{r}) = \epsilon_i^\alpha \phi_i^\alpha(\vec{r}). \tag{2.164}$$

Exercise 2.12. Prove Eq. (2.160), by showing

$$\langle \Phi_{\text{UHF}} | \hat{H} | \Phi_{\text{UHF}} \rangle = E_{\text{UHF}}. \tag{2.165}$$

△

2.4.7 The symmetry dilemma

2.4.8 Hartree–Fock for the electron gas

2.4.9 Normal-ordered Hamiltonian in HF basis (Not yet lectured)

Recall that for a two-body Hamiltonian $\hat{H} = \hat{H}_0 + \hat{W}$, the normal-ordered Hamiltonian (with respect to quasiparticles) was

$$\hat{H} = E_0 + \hat{H}_{0,\text{N}} + \hat{W}_{\text{N}}, \tag{2.166}$$

with

$$E_0 = \sum_i h_i^i + \frac{1}{2} \sum_{ij} \langle ij | \hat{w} | ij \rangle \tag{2.167}$$

$$\hat{H}_{0,\text{N}} = \sum_{pq} (h_q^p + \sum_j \langle pj | \hat{w} | qj \rangle) N(c_p^\dagger c_q), \tag{2.168}$$

$$\hat{W}_{\text{N}} = N(\hat{W}) = \frac{1}{4} \sum_{pqrs} \langle pq | \hat{w} | rs \rangle \, N(c_p^\dagger c_q^\dagger c_s c_r). \tag{2.169}$$

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

$$E_0 = E_{\rm HF}, \quad \hat{H}_{0,\rm N} = N(\hat{F}), \quad \text{and} \quad \hat{W}_{\rm N} = N(\hat{U}). \tag{2.170}$$

Thus, using HF orbitals, the normal-ordered Hamiltonian takes on a particularly simple form:

$$\hat{H} = \hat{F} + \hat{U} = E_{\rm HF} + N(\hat{F}) + N(\hat{U}), \tag{2.171}$$

where we recall that the normal-ordering operator is relative to quasiparticle vacuum. Here, the quasiparticle reference is the HF state $|\Psi_{\rm HF}\rangle = |\Phi\rangle$. Recall, that the normal-orering operator is defined linear combinations of strings,

$$N(\hat{F}) = N\left(\sum_{pq} f_q^p c_p^\dagger c_q\right) = \sum_{pq} f_q^p N(c_p^\dagger c_q). \tag{2.172}$$

But beware! In general, $N(\hat{H}_0) \neq \hat{H}_{0,\rm N}$! The operator $\hat{H}_{0,\rm N}$ depends on the whole Hamiltonian, i.e., also the two-body interaction. It is just that in *in the particular case of the HF partitioning* of the Hamiltonian, $N(\hat{F}) = \hat{F}_{\rm N}$.

We now also use the fact that $\hat{F}$ *is diagonal in the HF basis*,

$$\hat{F} = \sum_p \epsilon_p c_p^\dagger c_p. \tag{2.173}$$

This gives a considerable simplification, since

$$N(\hat{F}) = \sum_p \epsilon_p N(c_p^\dagger c_p) = \sum_a \epsilon_a b_a^\dagger b_c - \sum_i \epsilon_i b_i^\dagger b_i. \tag{2.174}$$

Exercise 2.13. 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.) △

2.5 Perturbation theory for the ground-state (PT)

2.5.1 Rayleigh–Schrödinger perturbation theory (RSPT)

Perturbation theory is a powerful method for systematic improvement of a model wavefunction. We can for the moment "forget" everything we know about second quantization, Slater determinants, quasiparticles, etc: PT is a generic theory applicable to all matrix problems.

Supporting material: Szabo and Ostlund; Bartlett and Shavitt; Helgaker, Jørgensen and Olsen.

Suppose we have a Hamiltonian H for which we seek eigenfunctions and eigenvalues,

$$\hat{H}\,|\Psi_k\rangle = E_k\,|\Psi_k\rangle\,. \tag{2.175}$$

The idea is to partition the Hamiltonian into a part that we can "solve" and a perturbation $\hat{V}$,

$$\hat{H} = \hat{H}_0 + \hat{V}. \tag{2.176}$$

The operator $\hat{H}_0$ is "solved", in the sense that we know its complete set of eigenfunctions,

$$\hat{H}_0 \left|\Phi_k\right\rangle = \epsilon_k \left|\Phi_k\right\rangle . \tag{2.177}$$

The set $\{|\Phi_k\rangle\}$ is an orthonormal basis for Hilbert space, and we should, in principle, be able to express the exact eigenvectors and (and therefore the eigenvalues) in terms of the this basis and $\hat{V}$.

In perturbation theory, we seek such an expression in terms of *power series in the perturbation* $\hat{V}$. We introduce an *order parameter* λ and write

$$\hat{H}_\lambda = \hat{H}_0 + \lambda \hat{V}, \tag{2.178}$$

i.e., $\hat{H} = \hat{H}_1$ is the full Hamiltonian. It is not unreasonable to assume that the eigenvalues and eigenvectors of $\hat{H}_\lambda$ become smooth functions of λ, at least for λ sufficiently small.

The Schrödinger equation for $\hat{H}(\lambda)$ reads

$$\hat{H}_\lambda \left|\Psi_k(\lambda)\right\rangle = E_k(\lambda) \left|\Psi_k(\lambda)\right\rangle . \tag{2.179}$$

We now assume that we can expand the eigenvectors and eigenvalues in *power series* around $\lambda = 0$.

$$\left|\Psi_k(\lambda)\right\rangle = \sum_{n=0}^{\infty} \left|\Psi_k^{(n)}\right\rangle \lambda^n \tag{2.180a}$$

$$E_k(\lambda) = \sum_{n=0}^{\infty} E_k^{(n)} \lambda^n . \tag{2.180b}$$

The unperturbed problem is obtained at $\lambda = 0$: $\left|\Psi_k(0)\right\rangle = \left|\Phi_k\right\rangle = \left|\Psi_k^{(0)}\right\rangle$ and $E_k(0) = \epsilon_k$, and the full problem at $\lambda = 1$: $\left|\Psi_k(1)\right\rangle = \left|\Psi_k\right\rangle$, and $E_k(1) = E_k$.

We now seek the perturbation corrections $E_k^{(n)}$ and $\left|\Psi_k^{(n)}\right\rangle$. This is done by plugging Eqs. (2.180) into the Schrödinger equation. This gives

$$(\hat{H}_0 + \lambda \hat{V}) \sum_{n=0}^{\infty} \left|\Psi_k^{(n)}\right\rangle \lambda^n = \left(\sum_{n=0}^{\infty} E_k^{(n)} \lambda^n \right) \sum_{m=0}^{\infty} \left|\Psi_k^{(m)}\right\rangle \lambda^m . \tag{2.181}$$

For this equation to hold *for all* λ, it must hold order-by-order. The λ^0-part of the equation is simply Eq. (2.177). The n'th order equation is

$$\hat{H}_0 \left|\Psi_k^{(n)}\right\rangle + \hat{V} \left|\Psi_k^{(n-1)}\right\rangle = \sum_{j=0}^{n} E_k^{(j)} \left|\Psi_k^{(n-j)}\right\rangle , \quad n > 0. \tag{2.182}$$

The solution $\left|\Psi_k(\lambda)\right\rangle$ to the Schrödinger equation is not unique. By scaling it we obtain a new solution. Thus, in order to write $\left|\Psi_k(\lambda)\right\rangle$ as a smooth function of λ, we need to select one particular normalization for each λ. We obtain particularly simple expressions using *intermediate normalization*:

$$\left\langle \Phi_k | \Psi_k(\lambda) \right\rangle = 1. \tag{2.183}$$

Inserting the power series for $\left|\Psi_k(\lambda)\right\rangle$ we obtain

$$1 = \left\langle \Phi_k | \Psi_k(\lambda) \right\rangle = 1 + \lambda \left|\Psi_k^{(1)}\right\rangle + \lambda^2 \left|\Psi_k^{(2)}\right\rangle + \cdots, \tag{2.184}$$

Since this expression is to hold for all λ, it must hold order-by-order, which gives

$$\left\langle \Phi_k | \Psi_k^{(n)} \right\rangle = 0, \quad \forall n \geq 1, \tag{2.185}$$

i.e., *all* the higher-order corrections are orthogonal to the unperturbed vector $|\Phi_k\rangle$.

We now use Eq. (2.185) and project Eq. (2.182) onto $|\Phi_k\rangle$ to obtain

$$\langle\Phi_k|\hat{V}|\Psi_k^{(n-1)}\rangle = E_k^{(n)}, \tag{2.186}$$

which is an expression for the n-th order energy perturbation in terms of the $n-1$-th order correction in the wavefunction. In particular,

$$E_k^{(1)} = \langle\Phi_k|\hat{V}|\Phi_k\rangle\,. \tag{2.187}$$

If we can find an expression for $|\Psi_k^{(n)}\rangle$ in terms of $|\Psi_k^{(j)}\rangle$, $j < n$, then we have a recursive procedure for determining all the perturbation corrections.

To this end, rearrange Eq. (2.182) as

$$(\epsilon_k - \hat{H}_0)\,|\Psi_k^{(n)}\rangle = \hat{V}\,|\Psi_k^{(n-1)}\rangle - \sum_{j=0}^{n-1} E_k^{(n-j)}\,|\Psi_k^{(j)}\rangle\,. \tag{2.188}$$

On the right-hand side we only have wavefunction corrections of order less than n. We also know that the $E_k^{(n)}$, which occurs on the right-hand side, is a function of $|\Psi_k^{(n-1)}\rangle$, so if we can somehow invert $\epsilon_k - \hat{H}_0$ then we have an expression for $|\Psi_k^{(n)}\rangle$ in terms of lower-order corrections only.

Let $\hat{P} = |\Phi_k\rangle\,|\Phi_k\rangle$, the projection operator onto the unperturbed eigenvector. Let $\hat{Q} = 1 - \hat{P}$, which is then the projector onto the subspace spanned by all the other $|\Phi_j\rangle$, $j \neq k$:

$$\hat{Q} = \sum_{j\neq k} |\Phi_j\rangle\,|\Phi_j\rangle\,. \tag{2.189}$$

We know that for $n > 0$,

$$|\Psi_k^{(n)}\rangle = \hat{P}\,|\Psi_k^{(n)}\rangle + \hat{Q}\,|\Psi_k^{(n)}\rangle = \hat{Q}\,|\Psi_k^{(n)}\rangle \tag{2.190}$$

by the intermediate normalization. Moreover,

$$[\hat{H}_0, \hat{Q}] = 0, \tag{2.191}$$

since the $|\Phi_j\rangle$ are eigenfunctions of $\hat{H}_0$. Acting on Eq. (2.188) with $\hat{Q}$ we then obtain

$$(\epsilon_k - \hat{H}_0)\hat{Q}\,|\Psi_k^{(n)}\rangle = \hat{Q}\hat{V}\,|\Psi_k^{(n-1)}\rangle - \sum_{j=1}^{n-1} E_k^{(n-j)}\hat{Q}\,|\Psi_k^{(j)}\rangle\,, \tag{2.192}$$

where we remark that the $j = 0$-term from the sum on the right-hand side is eliminated. We have

$$\epsilon_k - \hat{H}_0 = \sum_{j\neq k}(\epsilon_k - \epsilon_j)\,|\Phi_j\rangle\,\langle\Phi_j| = \hat{Q}(\epsilon_k - \hat{H}_0)\hat{Q}, \tag{2.193}$$

i.e., the operator acts only within the space orthogonal to $|\Phi_k\rangle$. Define the operator

$$\hat{R} = \hat{Q}\hat{R}\hat{Q} = \sum_{j\neq k} \frac{1}{\epsilon_j - \epsilon_k}\,|\Phi_j\rangle\,\langle\Phi_j|\,. \tag{2.194}$$

It is important to note that we must *here assume that the unperturbed eigenvalue ϵ_k is non-degenerate*. Otherwise there are infinite terms in the sum. Now, for every $|u\rangle = |Q\rangle\,|u\rangle$ (such as $|\Psi_k^{(n)}\rangle$) we have

$$\hat{R}(\epsilon_k - \hat{H}_0)\,|u\rangle = |u\rangle\,. \tag{2.195}$$

The operator $\hat{R}$ is called a *pseudoinverse*, and since $\hat{R} = \hat{Q}\hat{R}\hat{Q}$ it is common to write

$$\hat{R} = \frac{\hat{Q}}{\epsilon_k - \hat{H}_0}, \tag{2.196}$$

even though the fraction notation for matrices and operators is in general not sound. Acting with $\hat{R}$ on Eq. (2.188), we obtain

$$|\Psi_k^{(n)}\rangle = \frac{\hat{Q}}{\epsilon_k - \hat{H}_0}\left[\hat{V}\,|\Psi_k^{(n-1)}\rangle - \sum_{j=1}^{n-1} E_k^{(n-j)}\,|\Psi_k^{(j)}\rangle\right]. \tag{2.197}$$

We summarize as a theorem:

Theorem 2.9 (Non-degenerate Rayleigh–Schrödinger Perturbation Theory)**.** *Let $\hat{H} = \hat{H}_0 + \lambda\hat{V}$ be given, and assume that*

$$\hat{H}_0\,|\Phi_k\rangle = \epsilon_k\,|\Phi_k\rangle \tag{2.198}$$

where the eigenvectors for a complete basis. Let

$$(\hat{H}_0 + \lambda\hat{V})\,|\Psi_k(\lambda)\rangle = E_k(\lambda)\,|\Psi_k(\lambda)\rangle\,, \quad \langle\Phi_k|\Psi_k(\lambda)\rangle = 1, \tag{2.199}$$

for a given k, and assume that ϵ_k is a non-degenerate eigenvalue for $\hat{H}_0$. Assume furthermore, that $E_k(\lambda)$ and $|\Psi_k(\lambda)\rangle$ are analytic in a neighborhood of $\lambda = 0$,

$$E_k(\lambda) = \sum_{n=0}^{\infty} E_k^{(n)}\lambda^n \tag{2.200}$$

$$|\Psi_k(\lambda)\rangle = \sum_{n=0}^{\infty} |\Psi_k^{(n)}\rangle\,\lambda^n. \tag{2.201}$$

Then the n-th order corrections are given recursively in terms of the $j < n$-th order corrections via the formulae

$$E_k^{(n)} = \langle\Phi_k|\hat{V}|\Psi_k^{(n-1)}\rangle \tag{2.202}$$

$$|\Psi_k^{(n)}\rangle = \frac{\hat{Q}}{\epsilon_k - \hat{H}_0}\left[\hat{V}\,|\Psi_k^{(n-1)}\rangle - \sum_{j=1}^{n-1} E_k^{(n-j)}\,|\Psi_k^{(j)}\rangle\right]. \tag{2.203}$$

2.5.2 Low-order RSPT

Theorem 2.9 gives a recursive procedure for the n-th order corrections of the energies and wavefunctions. We now consider the explicit expressions up to $n = 3$.

For notational simplicity, we omit the subscript k in the following, and write $\epsilon \equiv \epsilon_k$ for the unperturbed energy, $|\Phi\rangle =\equiv |\Phi_k\rangle$ for the unperturbed wavefunction, etc.

The first-order correction to the energy is simple,

$$E^{(1)} = \langle\Phi|\hat{V}|\Phi\rangle\,. \tag{2.204}$$

For $E^{(2)}$ we need the first-order wavefuncton correction,

$$|\Psi^{(1)}\rangle = \frac{\hat{Q}}{\epsilon - \hat{H}_0}\hat{V}\,|\Phi\rangle\,, \tag{2.205}$$

which then gives

$$E^{(2)} = \langle\Phi|\hat{V}\frac{\hat{Q}}{\epsilon - \hat{H}_0}\hat{V}|\Phi\rangle\,. \tag{2.206}$$

For $E^{(3)}$ we need the second-order wavefunction correction,

$$|\Psi^{(2)}\rangle = \frac{\hat{Q}}{\epsilon - \hat{H}_0}[\hat{V} - \langle\Phi|\hat{V}|\Phi\rangle]\frac{\hat{Q}}{\epsilon - \hat{H}_0}\hat{V}|\Phi\rangle\,. \tag{2.207}$$

This gives

$$\begin{aligned} E^{(3)} &= \langle\Phi|\hat{V}\frac{\hat{Q}}{\epsilon - \hat{H}_0}[\hat{V} - \langle\Phi|\hat{V}|\Phi\rangle]\frac{\hat{Q}}{\epsilon - \hat{H}_0}\hat{V}|\Phi\rangle \\ &= \langle\Phi|\hat{V}\frac{\hat{Q}}{\epsilon - \hat{H}_0}\hat{V}\frac{\hat{Q}}{\epsilon - \hat{H}_0}\hat{V}|\Phi\rangle - \langle\Phi|\hat{V}|\Phi\rangle\langle\Phi|\hat{V}\frac{\hat{Q}}{\epsilon - \hat{H}_0}\frac{\hat{Q}}{\epsilon - \hat{H}_0}\hat{V}|\Phi\rangle \end{aligned} \tag{2.208}$$

Comparing $E^{(1)}$, $E^{(2)}$, and $E^{(3)}$, we see that the third-order term differ from the pattern. When we discuss diagrams, the second term of $E^{(3)}$ is referred to as an unlinked term. Higher order corrections obtain more and more such terms. [They can be systematically generated from the first term $\langle\Phi|\hat{V}\hat{R}\cdots\hat{R}\hat{V}|\Phi\rangle$ a procedure called the "bracketing procedure", see Paldus and Cizek. This is, however, out of scope for the current course.]

Exercise 2.14. Prove Eq. (2.207) △

Exercise 2.15. Derive the fourth and fift order perturbation theory corrections to the energy. △

2.5.3 A two-state example

It is instructive to consider a two-state example, since we can diagonalize it exactly and obtain closed-form expressions. The behavior of the perturbation series can then be considered.

Let

$$H_0 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad V = \begin{pmatrix} 0 & \epsilon \\ \epsilon & 0 \end{pmatrix}, \tag{2.209}$$

so that

$$H(\lambda) = \begin{pmatrix} -1 & \lambda\epsilon \\ \lambda\epsilon & 1 \end{pmatrix}. \tag{2.210}$$

The exact eigenvalues of $H(\lambda)$ are given by the roots of the polynomial

$$\det(H(\lambda) - eI) = (-1-e)(1-e) - (\lambda\epsilon)^2 = -(1+e)(1-e) - (\lambda\epsilon)^2. \tag{2.211}$$

Solving, we find the two roots

$$e_\pm = \pm[1 + (\lambda\epsilon)^2]^{1/2}. \tag{2.212}$$

As functions of λ, see Fig. 2.3. Note well, that our calculations are also true for *complex* λ, only that $e_\pm$ are no longer real, but complex roots in general.

In RSPT, we would seek the Taylor series of, say, $e_-(\lambda)$ around $\lambda = 0$. Since we have a closed-form expression we can compute this series. The first three terms are

$$e_-(\lambda) = e_-(0) + \lambda e_-'(0) + \frac{1}{2}\lambda^2 e_-''(0) + O(\lambda^3). \tag{2.213}$$

Explicit evaluation of the derivatives:

$$e'_-(\lambda) = -[1 + (\lambda\epsilon)^2]^{-1/2}\lambda\epsilon^2 \tag{2.214}$$

$$e''_-(\lambda) = [1 + (\lambda\epsilon)^2]^{-3/2}(\lambda\epsilon^2)^2 - [1 + (\lambda\epsilon)^2]^{-1/2}\epsilon^2 \tag{2.215}$$

We obtain the Taylor series (with a few extra terms obtained by computer algebra)

$$e_-(\lambda) = -1 - \frac{1}{2}(\lambda\epsilon)^2 + \frac{1}{8}(\lambda\epsilon)^4 - \frac{1}{26}(\lambda\epsilon)^6 + \frac{5}{128}(\lambda\epsilon)^8 + O(\lambda^{10}). \tag{2.216}$$

A natural question arises: does the series *converge*? Does it converge for our desired parameter value $\lambda = 1$? Well, our function has a *branch-point singularity* since it is a square-root function. The branch-point singularity arises when the two roots coincide in the complex plane, at $\lambda = \pm i/\epsilon$. At these points the eigenvalue functions are no longer analytic. The Taylor series only converges in a disc around $\lambda = 0$ that does not contain the singularity. Thus, the Taylor series will only converge within the circle $|\lambda| < 1/|\epsilon|$, i.e., it will converge for $\lambda = 1$ only if $|\epsilon| < 1$.

Thus, we see directly that the strength of the perturbation may affect the convergence properties of the perturbation series.

The points $\lambda = \pm i/\epsilon$ are called *avoided crossings* since, if the parameter λ is real, it "narrowly misses" the branch-point and hence an exact crossing. Often, in eigenvalue plots, one can see the function behaviour $\pm[a + (\epsilon\lambda)^2]^{1/2}$, indicating an avoided crossing and hence a singluarity located approximately at this λ-value.

For an n-state problem, each of the n eigenvalues may collide with $n-1$ eigenvalues (again for complex λ in general), giving quite a lot of possible branch points, and thus many singularities. Determining whether the RSPT series converges is thus a virtually impossible task for many-body calculations. Still, a few terms may still give a good approximation.

2.5.4 Møller–Plesset perturbation theory (MPPT)

Hartree–Fock + PT.

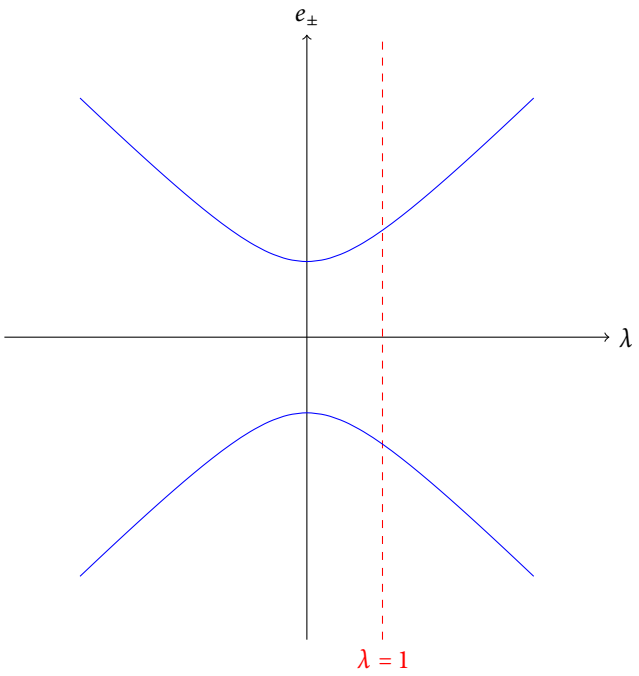


Figure 2.3: The eigenvalues of a two-state problem, as function of the perturbation parameter λ. Here, $\epsilon = 1$.

Chapter 3

Feynman diagrams for Rayleigh–Schrödinger perturbation theory

Chapter 4

Coupled-cluster theory (CC)

Appendix A

Mathematical supplement

A.1 Calculus of variations

A.1.1 Functionals

In the calculus of variations, we compute the extrema of a possibly nonlinear *function of a function*. Such objects are often called *functionals*. Thus, a functional $F[u]$ takes some *function* u and produces a *number*. One can think of F depending on infinitude of function values $u(x)$. In the case of the energy expectation value, the N-body wavefunction $|\Psi\rangle$ is mapped to the number

$$\mathcal{E}[|\Psi\rangle] = \langle\Psi|\hat{H}|\Psi\rangle \,/\, \langle\Psi|\Psi\rangle \,.$$

Suppose we expand the wavefunction in a basis, say, a Slater determinant basis,

$$|\Psi\rangle = \sum_I A_I \,|\Phi_I\rangle \,.$$

Then, $\mathcal{E}$ becomes a function of the vector $\vec{A}$, a possibly infinite set of coefficients. This may be an easier way to think of a functional: a function that depends on K variables, where K may be infinite.

A functional can also depend on more than one function. In Hartree–Fock theory, the energy functional depends on N single-particle functions ϕ_i, $i = 1, \cdots, N$. Moreover, the Hartree–Fock Lagrangian function that we *actually* optimize is a functional that also depends on a matrix $\lambda = [\lambda_{ij}]$ of Lagrange multipliers, $\mathcal{L} = \mathcal{L}[\phi_1, \cdots, \phi_N, \lambda]$. Given expansions of the ϕ_i as $\phi_i(x) = \sum_p \chi_p(x) U_{ip}$, we see that $\mathcal{L}$ becomes a function of the matrix U and the matrix λ. Thus, functionals are not too different from ordinary functions of a vectors.

How do we go about computing the extrema of a functional? A function of a single real variable has an intuitive notion of a local extremum, and most readers probably have an intuitive notion of extrema of two-variable functions as well. But if we go to higher dimensions (or infinite dimensions!) it becomes more complicated.

We will therefore introduce the concept of a *directional derivative* in a rather informal way. This is very handy, and allows us to read off the condition for an extremum in a straight-forward manner. This framework is called *the calculus of variations*, since we are computing the "variation in $F[u]$" with respect to arbitrary "variations δu of the function u".

A.1.2 Functions of one real variable

Consider first a simple function $F : I \to \mathbb{R}$, $I \subset \mathbb{R}$ being an interval. Suppose $x_0 \in I$. Assuming that F can be differentiated at leat twice, we can compute a second-order Taylor expansion around x_0, viz,

$$F(x_0 + \epsilon) \approx F(x_0) + \epsilon F'(x_0) + \frac{1}{2}\epsilon^2 F''(x_0). \tag{A.1}$$

The error in this approximation vanishes as $\epsilon \to 0$.

The condition for an extremum at x_0 is $F'(x_0) = 0$. The second-order term tells us the nature of the extremum: if $F''(x_0) > 0$ then x_0 is a local minimum. If $F''(x_0) < 0$ then x_0 is a local maximum. Finally, if $F''(x_0) = 0$, we cannot determine right away if we have a maximum or minimum. We may have neither, as for $F(x) = x^3$, where $x_0 = 0$ is a saddle point. A minimum and a saddle point is illustrated in Fig. A.1.

A.1.3 Functions of two real variables

Consider yourself in a landscape of mountains and valleys. The elevation is $F(x, y)$. You are trying to find, say, a local minimum (x_0, y_0) of elevation. On a map, a local minimum will show up as successively smaller closed curves of equal elevation, see Fig. A.2. (The same is true for a maximum, and a saddle point

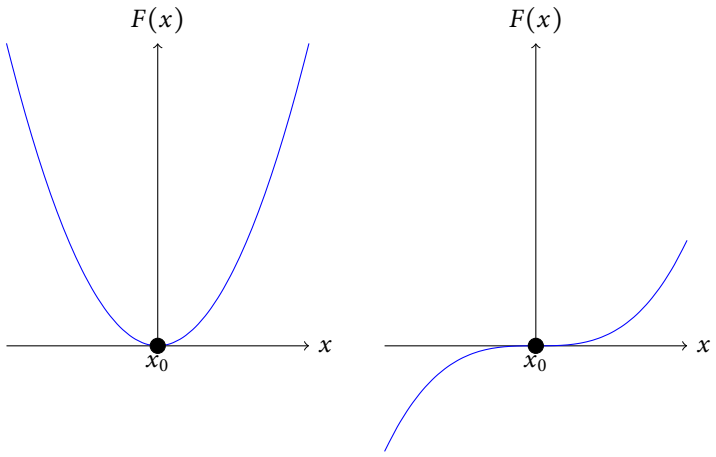


Figure A.1: Simple functions of one real variables with a local minimum ($F''(x_0) > 0$) (left) and a saddle point ($F''(x_0) = 0$) (right)..

is a crossing of lines of equal elevation.) We now observe, that if you move in a direction $\eta = (\delta x, \delta y) \neq 0$ from the local minium, you *will always walk uphill*, that is, the function

$$f(\epsilon) = F(x_0 + \epsilon\delta x, y_0 + \epsilon\delta y)$$

has a local minimum at $\epsilon = 0$, irrespective of η. If you were standing on a mountaintop (a local maximum) you would always walk downhill, and $f(\epsilon)$ would always have a local maximum at $\epsilon = 0$.

Finally, if you are standing between two mountaintops to the east and west, and looking down at valleys to the south and north, you are standing on a saddle point. You are walking downhill if you go north or south, but uphill if you go east or west: $f(\epsilon)$ has a local minimum for some η, and a maximum for other η.

We see that, at least intuitively, we can determine wheter F has a local extremum at (x_0, y_0) by studying the behaviour of $f(\epsilon)$, for all possible choices of η. We now prove this claim:

Let us compute the Taylor expansion of $f(\epsilon)$:

$$\begin{aligned} f(\epsilon) &\approx f(0) + \epsilon f'(0) + \tfrac{1}{2}\epsilon^2 f''(0) \\ &= F(x_0, y_0) + \epsilon \nabla F(x_0, y_0)^T \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} + \frac{1}{2}\epsilon^2 (\delta x \ \delta y) H(x_0, y_0) \begin{pmatrix} \delta x \\ \delta y \end{pmatrix}. \end{aligned} \tag{A.2}$$

We used the chain rule, and introduced the gradient and the Hessian matrix H, given by

$$\nabla F(x_0, y_0) = \begin{pmatrix} \frac{\partial F(x_0,y_0)}{\partial x} \\ \frac{\partial F(x_0,y_0)}{\partial y} \end{pmatrix} \tag{A.3}$$

and

$$H(x_0, y_0) = \begin{pmatrix} \frac{\partial^2 F(x_0,y_0)}{\partial x^2} & \frac{\partial^2 F(x_0,y_0)}{\partial x \partial y} \\ \frac{\partial^2 F(x_0,y_0)}{\partial y \partial x} & \frac{\partial^2 F(x_0,y_0)}{\partial y^2} \end{pmatrix}. \tag{A.4}$$

Now, F has an extremum at (x_0, y_0) if and only if $\nabla F(x_0, y_0) = 0$, while $f(\epsilon)$ has an extremum at $\epsilon = 0$ if and only if the second term in Eq. (A.2) vanishes. But if $\nabla F(x_0, y_0)^T \eta = 0$ for all $\eta \neq 0$, then clearly $\nabla F(x_0, y_0) = 0$ and vice versa. QED.

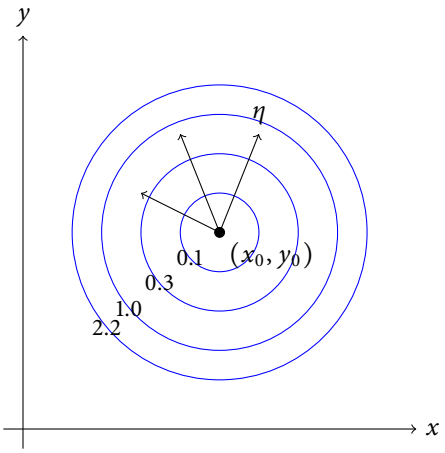


Figure A.2: The condition for a local minimum (x_0, y_0) for a function $F(x, y)$: in all directions $\eta \neq 0$ you walk uphill from (x_0, y_0).

We introduce the *directional derivative* of F at (x_0, y_0) in the direction $\eta = (\delta x, \delta y)$,

$$F'(x_0, y_0; \eta) \equiv \frac{d}{d\epsilon} F(x_0 + \epsilon\delta x + y_0 + \epsilon\delta y)\Big|_{\epsilon=0} \tag{A.5}$$

which is precisely the second term in Eq. (A.2),

$$f(\epsilon) \approx F(x_0, y_0) + \epsilon F'(x_0, y_0; \eta) + \frac{1}{2}\epsilon^2 \eta^T H(x_0, y_0)\eta. \tag{A.6}$$

Thus, the extremum condition is equivalent to $F'(x_0, y_0; \eta) = 0$ for all $\eta \neq 0$.

What about the nature of the extremum? If

$$\eta^T H(x_0, y_0)\eta > 0 \tag{A.7}$$

for all possible directions η, we have a local minimum. This is precisely the condition that $H(x_0, y_0)$ is a positive definite matrix. Since $H(x_0, y_0)$ is a symmetric matrix, this is equivalent to all the eigenvalues being positive. Thus, $f(\epsilon)$ must have a local minimum at $\epsilon = 0$ for every $\eta \neq 0$.

Similarly, if $H(x_0, y_0)$ is negative definite,

$$\eta^T H(x_0, y_0)\eta < 0, \quad \forall \eta \tag{A.8}$$

then we have a local maximum. However, ff $H(x_0, y_0)$ is neither positive nor negative definite, we cannot say whether we have a maximum or minimum. We may in fact have a saddle point, as in the case of standing between mountains and valleys.

A.1.4 Extremalization of a functional

The concept of the directional derivative is of course valid for more than two dimensions. For a function $F : \mathbb{R}^n \to \mathbb{R}$, the localization of an extremum can be formulated as: find $x_0 \in \mathbb{R}^n$ such that the directional derivative vanishes for every nonzero $\eta \in \mathbb{R}^n$:

$$F'(x_0; \eta) = \frac{d}{d\epsilon} F(x_0 + \epsilon\eta)\Big|_{\epsilon=0} = 0, \quad \forall \eta \in \mathbb{R}^n, \eta \neq 0. \tag{A.9}$$

This condition is equivalent to $\nabla F(x_0)^T = 0$.

Turning to a *functional* $F[u]$ for some function u, or set of functions, the directional derivative in the direction of the *function* η is in principle straightforward:

$$F'[u;\eta] = \frac{d}{d\epsilon} F[u + \epsilon\eta]\Big|_{\epsilon=0}. \tag{A.10}$$

Computing $F[u + \epsilon\eta]$ as a series in ϵ is usually straightforward, allowing an expression for $F'[u;\eta]$ to be read off. Typically, this leads to a *differential equation*: the variational principle gave us the Schrödinger equation, while extremalization of the Hartree–Fock energy gave us the Hartree–Fock equations.

The term "calculus of variations" is historical, and comes from the idea that we are "computing infinitesimal variations $\delta F[u]$ in the functional under infinitesimal variations δu of the function" in all possible ways, i.e., a different way of saying that we are computing directional derivatives.

Bibliography

[1] S. Raimes. *Many-electron theory*. North-Holland, Amsterdam, Netherlands, 1972.

[2] E.K.U. Gross, E. Runge, and O. Heinonen. *Many-Particle Theory*. Adam Hilger, Bristol, Philadelphia and New York, 1991.

[3] A. Szabo and N.S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover, New York, USA, 1996.

[4] J.M. Leinaas and J. Myrheim. On the theory of identical particles. *Il Nuovo Cimento*, 37:1, 1977.

[5] F.E. Harris, H.J. Monkhorst, and D.L. Freeman. *Algebraic and Diagrammatic Methods in Many-Fermion Theory*. Oxford, New Work, 1992.

[6] R. Helgaker, P. Jørgensen, and Olsen. J. *Molecular Electronic-Structure Theory*. Wiley, Chichester, UK, 2002.