1 Hartree-Fock Theory

The goal of electronic structure theory is to solve the "clamped-nuclei" Schrödinger equation

$$\hat{H}\Psi_k = E_k\Psi_k \qquad \hat{H} = V_{\text{nuc}} + \hat{H}_e = \sum_{a<b}^{\text{nuc.}} \frac{Z_a Z_b}{|\mathbf{R}_a - \mathbf{R}_b|} - \frac{1}{2}\sum_{i}^{\text{elec.}} \nabla_i^2 - \sum_{a}^{\text{nuc.}}\sum_{i}^{\text{elec.}} \frac{Z_a}{|\mathbf{R}_a - \mathbf{r}_i|} + \sum_{i<j}^{\text{elec.}} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{1.1}$$

with an optimal balance of accuracy and efficiency for the problem of interest. The most accurate solution possible for a given atomic orbital (AO) basis set[1] results from expanding the wavefunction

$$\Psi_k = \sum_{\mu} \Phi_\mu c_{\mu k} \tag{1.2}$$

in terms of all possible Slater determinants Φ_μ that can be formed from an orthonormal one-electron basis of spin-orbitals, $\{\psi_p\}$. The expansion coefficients $(\mathbf{c})_k = c_{\mu k}$ are eigenvectors of the matrix $(\mathbf{H})_{\mu\nu} = \langle\Phi_\mu|\hat{H}|\Phi_\nu\rangle$, which is the matrix representation of the Hamiltonian in the determinant basis. This is called the *full configuration-interaction* (FCI) solution.

Any one-electron basis spans the same "function space" as the AO basis set itself, and the full n-electron basis $\{\Phi_\mu\}$ spans the same space of n-electron functions regardless of how one forms spin orbitals from the AO basis set. As a result, one obtains the same FCI solution for any choice of spin-orbitals. In general, however, FCI solutions are completely unfeasible for basis sets of sufficient size to approach the complete basis set limit. One can think of this as a simple counting problem: if there are m functions in the AO basis, then there are $2m$ spin-orbitals in the one-electron basis,[2] and there are "$2m$ choose n" [3]

$$\binom{2m}{n} \equiv \frac{(2m)!}{n!(2m-n)!}$$

unique Slater determinants in the n-electron basis that can be formed from the spin MOs. The upshot is that we usually have to omit some Slater determinants in order to get an answer in a reasonable amount of time.

As soon as we truncate our determinant expansion (1.2), our choice of spin MOs makes a significant difference in the quality of our results. In particular, we need to choose our set of one-electron functions to minimize the number of Slater determinants it takes to "get close to" the exact wavefunction.

1.1 The Hartree-Fock optimization problem

It can be shown that optimizing $\langle\Psi|\hat{H}_e|\Psi\rangle$ by varying Ψ subject to the normalization constraint $\langle\Psi|\Psi\rangle = 1$ is equivalent to solving the Schrödinger equation. When we further constrain the form of Ψ this is no longer true, but it *does* generally allow us to get the best approximation to Ψ for a given approach (or "Ansatz").

In order to make the wavefunction expansion converge with a relatively small number of Φ_μs, we wish to find the best single-determinant approximation to Ψ. That is, we wish to optimize

$$\langle\Phi|\hat{H}_e|\Phi\rangle \qquad\qquad \Phi(1,\ldots,n) = \frac{1}{\sqrt{n!}}\begin{vmatrix} \psi_1(1) & \psi_2(1) & \cdots & \psi_n(1) \\ \psi_1(2) & \psi_2(2) & \cdots & \psi_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(n) & \psi_2(n) & \cdots & \psi_n(n) \end{vmatrix} \tag{1.3}$$

with respect to variation of the orbitals $\{\psi_p\}$, enforcing the normalization constraint by keeping the spin orbitals orthonormal. Note that here the function argument (i) is shorthand for $(\mathbf{r}_i, s_i)$ where $\mathbf{r}_i$ denotes the position of the i^{th} electron and s_i denotes its spin. This optimization problem is the idea behind *Hartree-Fock theory*.

[1] cc-pVXZ, 6-31G, ANO1, etc.

[2] m α-orbitals and m β-orbitals.

[3] The number of unique sets of n marbles that can be drawn from a bag of $2m$ marbles. See `http://en.wikipedia.org/wiki/Combination`

Once we have solved for the Hartree-Fock optimization problem, the expectation value $\langle\Phi|\hat{H}_e|\Phi\rangle$ is itself a good first approximation to the electronic energy. More importantly, however, when we use this new set of Hartree-Fock spin-orbitals, $\{\psi_p\}$, the FCI expansion tends to converge much more quickly to the true wave-function. Specifically, when we rewrite equation 1.2 in terms of single $\{\Phi_i^a\}$, double $\{\Phi_{ij}^{ab}\}$, triple $\{\Phi_{ijk}^{abc}\}$, etc. replacements [4] of the orbitals in the Hartree-Fock determinant Φ with the remaining orbitals in the basis

$$\Psi = \Phi + \sum_{\substack{a \\ i}} \Phi_i^a c_a^i + \sum_{\substack{a<b \\ i<j}} \Phi_{ij}^{ab} c_{ab}^{ij} + \sum_{\substack{a<b<c \\ i<j<k}} \Phi_{ijk}^{abc} c_{abc}^{ijk} + \dots \tag{1.4}$$

the coefficients tend to be very small, and are often virtually negligible for higher than quadruple replacements.

1.2 The Hartree-Fock equations

The electronic Hamiltonian $\hat{H}_e$ contains one- and two-electron operators.

$$\hat{H}_e = \sum_i \hat{h}(i) + \sum_{i<j} \hat{g}(i,j) \qquad \hat{h}(i) \equiv -\frac{1}{2}\nabla_i^2 + \sum_a \frac{Z_a}{|\mathbf{r}_i - \mathbf{R}_a|} \qquad \hat{g}(i,j) \equiv \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{1.5}$$

Its expectation value with respect to a single determinant Φ is given by the first Slater rule

$$\langle\Phi|\hat{H}_e|\Phi\rangle = \sum_i^n \langle\psi_i|\hat{h}|\psi_i\rangle + \frac{1}{2}\sum_{ij}^n \langle\psi_i\psi_j||\psi_i\psi_j\rangle \qquad \langle\psi_p\psi_q||\psi_r\psi_s\rangle \equiv \langle\psi_p\psi_q|\psi_r\psi_s\rangle - \langle\psi_p\psi_q|\psi_s\psi_r\rangle \tag{1.6}$$

where the one- and two-electron integrals are defined as follows.

$$\langle\psi_p|\hat{h}|\psi_q\rangle \equiv \int d(1)\psi_p^*(1)\hat{h}(1)\psi_q(1) \qquad \langle\psi_p\psi_q|\psi_r\psi_s\rangle \equiv \int d(1)d(2)\psi_p^*(1)\psi_q^*(2)\hat{g}(1,2)\psi_r(1)\psi_s(2) \tag{1.7}$$

We wish to optimize equation 1.6 while constraining the orbitals to be normalized and orthogonal.[5]

$$\langle\psi_i|\psi_j\rangle \stackrel{!}{=} \delta_{ij} \tag{1.8}$$

The corresponding Lagrangian functional (see appendix A) is

$$\mathcal{L}[\{\psi_i\},\{\psi_i^*\},\{\epsilon_{ij}\}] = \sum_{i=1}^n \langle\psi_i|\hat{h}|\psi_i\rangle + \frac{1}{2}\sum_{i,j=1}^n \langle\psi_i\psi_j||\psi_i\psi_j\rangle - \sum_{i,j=1}^n \epsilon_{ij}(\langle\psi_i|\psi_j\rangle - \delta_{ij}) \tag{1.9}$$

where $\{\epsilon_{ij}\}$ are our Lagrangian multipliers for the orthonormality constraint. Note that the complex conjugates of the orbitals $\{\psi_i^*\}$ are included as separate arguments of $\mathcal{L}$, since the real and imaginary components of ψ_i can be varied independently.[6]

The stationarity conditions for the Hartree-Fock Lagrangian are (see appendix B)

$$\left.\frac{d\mathcal{L}[\psi_k^* + \varepsilon\eta^*]}{d\varepsilon}\right|_{\varepsilon=0} \stackrel{!}{=} 0 \quad \text{and} \quad \left.\frac{d\mathcal{L}[\psi_k + \varepsilon\eta]}{d\varepsilon}\right|_{\varepsilon=0} \stackrel{!}{=} 0 \quad \text{for all } \eta \qquad\qquad k = 1,\dots,n \tag{1.10}$$

which can be stated in words as follows: For each orbital $\psi_1,\dots,\psi_n$ in the determinant Φ, mixing in a little bit of an arbitrary function $\eta = \eta(\mathbf{r}, s)$ doesn't change the Lagrangian.

[4] It is typical to use dummy indices i, j, k, l to count over the orbitals in the reference determinant Φ – the "occupied orbitals" – and to use a, b, c, d, to count over the orbitals not contained in Φ – the "unoccupied" or "virtual orbitals." Dummy indices p, q, r, s are generally used to count over the full set of spin-orbitals, whether occupied or not.

[5] The $\stackrel{!}{=}$ sign means "must equal" – these are conditions to be satisfied.

[6] One can explicitly show that for complex variables $\frac{\partial z}{\partial z^*} = 0$. See `https://en.wikipedia.org/wiki/Wirtinger_derivatives#Functions_of_one_complex_variable`.

Separating out the terms in eq (1.9) involving a particular orbital ψ_k, we can write

$$\mathcal{L} = \langle\psi_k|\hat{h}|\psi_k\rangle + \sum_i \langle\psi_k\psi_i||\psi_k\psi_i\rangle - \sum_i \epsilon_{ki}(\langle\psi_k|\psi_i\rangle - \delta_{ki}) - \sum_i \epsilon_{ik}(\langle\psi_i|\psi_k\rangle - \delta_{ik})$$
$$+ \sum_{i\neq k} \langle\psi_i|\hat{h}|\psi_i\rangle + \frac{1}{2}\sum_{i\neq k, j\neq k} \langle\psi_i\psi_j||\psi_i\psi_j\rangle - \sum_{i\neq k, j\neq k} \epsilon_{ij}(\langle\psi_i|\psi_j\rangle - \delta_{ij})$$

using $\langle\psi_k\psi_i||\psi_k\psi_i\rangle = \langle\psi_i\psi_k||\psi_i\psi_k\rangle$, which follows from exchanging integration variables in eq (1.7). The functional directional derivative for varying ψ_k^* along η^* is then

$$\left.\frac{d\mathcal{L}[\psi_k^* + \varepsilon\eta^*]}{d\varepsilon}\right|_{\varepsilon=0} = \frac{d}{d\varepsilon}\left(\langle\psi_k + \varepsilon\eta|\hat{h}|\psi_k\rangle + \sum_i \langle(\psi_k + \varepsilon\eta)\psi_i||\psi_k\psi_i\rangle - \sum_i \epsilon_{ki}\langle\psi_k + \varepsilon\eta|\psi_i\rangle\right)\Bigg|_{\varepsilon=0}$$

where we have dropped ε-independent terms since their derivatives vanish. Evaluating the right-hand side gives

$$\left.\frac{d\mathcal{L}[\psi_k^* + \varepsilon\eta^*]}{d\varepsilon}\right|_{\varepsilon=0} = \langle\eta|\hat{h}|\psi_k\rangle + \sum_i \langle\eta\psi_i||\psi_k\psi_i\rangle - \sum_i \epsilon_{ki}\langle\eta|\psi_i\rangle$$

where the first two terms can be written as follows.[7]

$$\langle\eta|\hat{h}|\psi_k\rangle + \sum_i \langle\eta\psi_i||\psi_k\psi_i\rangle = \int d(1)\eta^*(1)\left(\hat{h}(1) + \sum_i \langle\psi_i(2)|\hat{g}(1,2)(1-\hat{P}(1,2))|\psi_i(2)\rangle\right)\psi_k(1)$$

Here, the *coordinate exchange operator* $\hat{P}(1,2)$ allows us to write $(1-\hat{P}(1,2))\psi_i(2)\psi_k(1)$ as a shorthand for $\psi_i(2)\psi_k(1) - \psi_i(1)\psi_k(2)$. The expression in parentheses constitutes the *Fock operator*.[8]

$$\hat{f}(1) \equiv \hat{h}(1) + \sum_i \langle\psi_i(2)|\hat{g}(1,2)(1-\hat{P}(1,2))|\psi_i(2)\rangle \tag{1.11}$$

Note that it implicitly depends on orbital set, $\hat{f} = \hat{f}[\{\psi_i\}]$.

Using a similar procedure for the variation of ψ_k along η, eq (1.10) evaluates to

$$\int d(1)\eta^*(1)\left(\hat{f}(1)\psi_k(1) - \sum_i \epsilon_{ki}\psi_i(1)\right) \stackrel{!}{=} 0 \quad \text{and} \quad \int d(1)\left(\psi_k^*(1)\hat{f}(1) - \sum_i \psi_i^*(1)\epsilon_{ik}\right)\eta(1) \stackrel{!}{=} 0 \quad \text{for all } \eta$$

which, by the Fundamental Lemma of Calculus of Variations (appendix C), is equivalent to the following

$$\hat{f}(1)\psi_k(1) \stackrel{!}{=} \sum_i \epsilon_{ki}\psi_i(1) \qquad \text{and} \qquad \hat{f}(1)\psi_k^*(1) \stackrel{!}{=} \sum_i \epsilon_{ik}\psi_i^*(1) \tag{1.12}$$

using the Hermitian-ness of the Fock operator, $\langle\psi_k|\hat{f}\eta\rangle = \langle\hat{f}^\dagger\psi_k|\eta\rangle = \langle\hat{f}\psi_k|\eta\rangle$. Subtracting the complex conjugate of the right equation from the left gives

$$\sum_i (\epsilon_{ki} - \epsilon_{ik}^*)\psi_i(1) \stackrel{!}{=} 0$$

which, since the orbitals are linearly independent,[9] implies that $\boldsymbol{\epsilon} = [\epsilon_{ij}]$ forms a Hermitian matrix.

$$\epsilon_{k1} - \epsilon_{1k}^* = \cdots = \epsilon_{kn} - \epsilon_{nk}^* = 0 \tag{1.13}$$

[7] Defining $\langle\psi_p(2)|\hat{g}(1,2)|\psi_q(2)\rangle \equiv \int d(2)\psi_p^*(2)\hat{g}(1,2)\psi_q(2)$

[8] You may see this written as $\hat{f}(1) = \hat{h}(1) + \sum_i(\hat{J}_i(1) - \hat{K}_i(1))$ where $\hat{J}_i(1) \equiv \langle\psi_i(2)|\hat{g}(1,2)|\psi_i(2)\rangle$ and $\hat{K}_i \equiv \langle\psi_i(2)|\hat{g}(1,2)\hat{P}(1,2)|\psi_i(2)\rangle$ are the *Coloumb* and *exchange operators*.

[9] `http://en.wikipedia.org/wiki/Linear_independence\#Definition`

Requiring the multiplier matrix to be Hermitian makes the second condition in eq (1.12) redundant, so that the final *Hartree-Fock equations* can be expressed as follows.

$$\hat{f}\psi_i \stackrel{!}{=} \sum_j \epsilon_{ij}\psi_j \quad \text{and} \quad \boldsymbol{\epsilon} \stackrel{!}{=} \boldsymbol{\epsilon}^\dagger \tag{1.14}$$

To review, these conditions define orbitals which optimize $\langle\Phi|\hat{H}_e|\Phi\rangle$ subject to the constraint $\langle\psi_i|\psi_j\rangle = \delta_{ij}$.

1.3 The canonical Hartree-Fock equations

Appendix D shows that the Hartree-Fock energy and the orthogonality relations are invariant to unitary mixing of the orbitals in Φ. This implies that the solution to the Hartree-Fock optimization problem is not unique, because any unitary transformation of the orbitals in Φ is also a solution. In this section we show how to use this freedom to our advantage, by choosing orbitals which diagonalize the Lagrange multiplier matrix, partially decoupling eq (1.14). These orbitals are known as *canonical Hartree-Fock orbitals*.

In matrix notation, the Hartree-Fock equations can be written as follows.

$$\hat{f}\boldsymbol{\psi} \stackrel{!}{=} \boldsymbol{\epsilon}\boldsymbol{\psi} \quad \text{and} \quad \boldsymbol{\epsilon} = \boldsymbol{\epsilon}^\dagger \qquad\qquad \boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_{11} & \cdots & \epsilon_{1n} \\ \vdots & \ddots & \vdots \\ \epsilon_{n1} & \cdots & \epsilon_{nn} \end{pmatrix}, \quad \boldsymbol{\psi} = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_n \end{pmatrix} \tag{1.15}$$

Since the matrix $\boldsymbol{\epsilon}$ is Hermitian, it can be diagonalized by a unitary transformation $\mathbf{U}$.

$$\boldsymbol{\epsilon} = \mathbf{U}\tilde{\boldsymbol{\epsilon}}\mathbf{U}^\dagger \qquad\qquad \tilde{\boldsymbol{\epsilon}} = \begin{pmatrix} \epsilon_1 & 0 & \cdots & 0 \\ 0 & \epsilon_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \epsilon_n \end{pmatrix} \tag{1.16}$$

Inserting this decomposition into eq (1.15) and multiplying both sides from the left by $\mathbf{U}^\dagger$, we get

$$\hat{f}(\mathbf{U}^\dagger\boldsymbol{\psi}) = \tilde{\boldsymbol{\epsilon}}(\mathbf{U}^\dagger\boldsymbol{\psi})$$

which shows that the problem can be decoupled by using a new set of orbitals $\tilde{\psi}_1, \ldots, \tilde{\psi}_n$, defined as follows.[10]

$$\tilde{\psi}_i = \sum_{j=1}^{n} U_{ji}^* \psi_j \tag{1.17}$$

It can be shown that the Fock operator $\hat{f}$ is invariant to this type of transformation (see appendix D).

Substituting the new orbitals into eq (1.15) and dropping tildes yields the *canonical Hartree-Fock equations*.

$$\hat{f}\psi_i = \epsilon_i\psi_i \qquad i = 1, \ldots, n$$

Since $\boldsymbol{\epsilon}$ is Hermitian, the Lagrangian eigenvalues are real. Note that these equations are not fully decoupled, since $\hat{f}$ still depends on the full orbital set $\{\psi_i\}$. Solving them amounts to solving for the *self-consistent field*

$$\hat{v}(1) \equiv \sum_i \langle\psi_i(2)|\hat{g}(1,2)(1 - \hat{P}(1,2))|\psi_i(2)\rangle = \sum_i (\hat{J}_i(1) - \hat{K}_i(1)) \tag{1.18}$$

in $\hat{f} = \hat{h} + \hat{v}$ that allows all n equations to hold true simultaneously.

[10] In matrix notation this reads $\tilde{\boldsymbol{\psi}} = \mathbf{U}^\dagger\boldsymbol{\psi}$.

A Constrained Optimization

The standard method of optimizing a function subject to a constraint is called Lagrangian optimization. Taking a function of two variables $f(x, y)$ as an example, suppose we want to optimize it subject to a constraint of the form $g(x, y) = c$. In this approach, we define the "Lagrangian function" $\mathcal{L}$ as

$$\mathcal{L}(x, y, \lambda) \equiv f(x, y) - \lambda(g(x, y) - c) \tag{A.1}$$

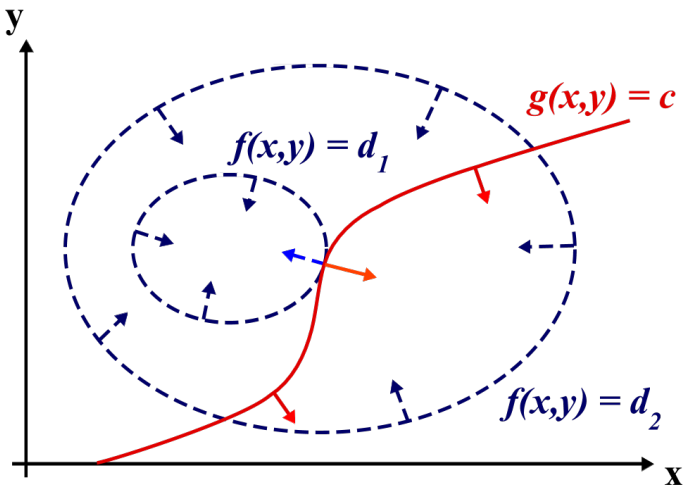
where the parameter λ is called the Lagrange multiplier. The constrained optimization problem can be solved solved by optimizing $\mathcal{L}$ with respect to x, y, and λ. To see why, consider the stationarity conditions for $\mathcal{L}$.

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} \stackrel{!}{=} 0 \qquad \frac{\partial \mathcal{L}}{\partial y} = \frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} \stackrel{!}{=} 0 \qquad \frac{\partial \mathcal{L}}{\partial \lambda} = c - g(x, y) \stackrel{!}{=} 0 \tag{A.2}$$

The last equation is simply the requirement that the constraint $g(x, y) = c$ be satisfied – i.e. that the point (x, y) lies along the contour of $g(x, y)$ specified by $g(x, y) = c$. The first two equations correspond to the requirement that the gradients of the function $f(x, y)$ and the constraint surface $g(x, y)$ be parallel

$$\nabla f = \lambda \nabla g \tag{A.3}$$

which is always true at the point (x, y) of closest approach along the line $g(x, y) = c$ to a minimum or maximum of the function $f(x, y)$. This is best understood visually.



If the gradients were not parallel, we could move along $g(x, y) = c$ to a higher contour of $f(x, y)$ by following the component of ∇f parallel to $g(x, y) = c$.

B Functional Derivatives

A functional is just a function of a function – i.e. some rule F that maps a function f into a number $F[f]$. Definite integrals are a common example. In order to optimize a functional F with respect to its argument f, one needs to take a *functional derivative*.[11] To motivate the definition of a functional derivative, first consider the definition of an ordinary derivative

$$\frac{df(x)}{dx} \equiv \lim_{\varepsilon \to 0} \frac{f(x+\varepsilon) - f(x)}{\varepsilon} \tag{B.1}$$

and note the following identity, which you can verify using $f(x+\varepsilon) = f(x) + \frac{df(x)}{dx}\varepsilon + \mathcal{O}(\varepsilon^2)$.

$$\lim_{\varepsilon \to 0} \frac{f(x+\varepsilon) - f(x)}{\varepsilon} = \left. \frac{df(x+\varepsilon)}{d\varepsilon} \right|_{\varepsilon=0} \tag{B.2}$$

For multivariate functions, we have the concept of a *directional derivative*

$$\mathbf{y} \cdot \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \lim_{\varepsilon \to 0} \frac{f(\mathbf{x}+\varepsilon\mathbf{y}) - f(\mathbf{x})}{\varepsilon} \tag{B.3}$$

which measures the change in $f(\mathbf{x})$ in the direction $\mathbf{y}$. Using equation B.2, the directional derivative can be evaluated as an ordinary scalar derivative with respect to ε.

$$\mathbf{y} \cdot \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \left. \frac{df(\mathbf{x}+\varepsilon\mathbf{y})}{d\varepsilon} \right|_{\varepsilon=0} \tag{B.4}$$

The functional derivative $\frac{\delta F}{\delta f}$ is defined to satisfy an equation analogous to B.3, playing the role of the gradient.

$$\int_{-\infty}^{\infty} dx'\, g(x') \frac{\delta F[f]}{\delta f(x')} \equiv \lim_{\varepsilon \to 0} \frac{F[f+\varepsilon g] - F[f]}{\varepsilon} \tag{B.5}$$

This left-hand side could be called a *functional directional derivative*, giving the change in F upon displacing its argument along the function g. Here, the integral takes the role of the dot product in B.3. Using the same trick as in equation B.4, the functional derivative can be expressed as an ordinary scalar derivative.

$$\int_{-\infty}^{\infty} dx'\, g(x') \frac{\delta F[f]}{\delta f(x')} = \left. \frac{dF[f+\varepsilon g]}{d\varepsilon} \right|_{\varepsilon=0} \tag{B.6}$$

The standard procedure for evaluating the functional derivative is to first evaluate the right-hand side of equation B.6 for an arbitrary g and then infer what $\frac{\delta F[f]}{\delta f(x)}$ must be by comparing to the left-hand side. Equivalently, $g(x')$ can be replaced with a Dirac delta $\delta(x-x')$ in order to arrive at $\frac{\delta F[f]}{\delta f(x)}$ directly.

Using eq. B.6 and the lemma in appendix C, we find that the stationarity condition for a functional

$$\frac{\delta F[f]}{\delta f} \stackrel{!}{=} 0 \tag{B.7}$$

is equivalent to the following condition.

$$\left. \frac{dF[f+\varepsilon g]}{d\varepsilon} \right|_{\varepsilon=0} \stackrel{!}{=} 0 \qquad \text{for all } g(x) \tag{B.8}$$

[11] `http://en.wikipedia.org/wiki/Functional_derivative`

C Fundamental Lemma of Calculus of Variations

The *Fundamental Lemma of Calculus of Variations*[12] says that, for continuous functions, the condition

$$\int_{-\infty}^{\infty} dx f(x)\eta(x) = 0 \qquad \text{for all } \eta(x) \tag{C.1}$$

holds only when $f(x) = 0$ for all x. We can see this by considering the case $\eta(x) = f(x)$. Since $f(x)^2$ is nonnegative everywhere, the integral yields a positive number whenever $f(x) \neq 0$ on a finite range of x values.

[12] http://en.wikipedia.org/wiki/Fundamental_lemma_of_calculus_of_variations

D Unitary Invariances for Hartree-Fock Orbitals

Orthonormality. By definition, unitary transformations preserve overlaps. This can be verified as follows

$$\langle\tilde{\psi}_i|\tilde{\psi}_j\rangle = \sum_{kl} U_{ki}U_{lj}^*\langle\psi_k|\psi_l\rangle = \sum_{kl} U_{ki}U_{lj}^*\delta_{kl} = \sum_k U_{ki}U_{kj}^* = \delta_{ij}$$

using $\sum_k U_{ki}U_{kj}^* = (\mathbf{U}\mathbf{U}^\dagger)_{ji} = (\mathbf{1})_{ji} = \delta_{ji}$.

Fock operator. Only the Coulomb and exchange parts of the Fock operator depend on the orbital set. For the Coulomb part, we have

$$\sum_i \langle\tilde{\psi}_i(2)|\hat{g}(1,2)|\tilde{\psi}_i(2)\rangle = \sum_{ijk} U_{ji}U_{ki}^*\langle\psi_j(2)|\hat{g}(1,2)|\psi_k(2)\rangle = \sum_{jk}\delta_{jk}\langle\psi_j(2)|\hat{g}(1,2)|\psi_k(2)\rangle = \sum_j \langle\psi_j(2)|\hat{g}(1,2)|\psi_j(2)\rangle$$

using the fact that $\sum_i U_{ji}U_{ki}^* = \delta_{jk}$. For the exchange part, we have the same thing with a $\hat{P}(1,2)$ sandwiched in there.

Hamiltonian expectation value. The vector notation $\boldsymbol{\psi}$ for our orbitals allows us to express Φ and $\tilde{\Phi}$ as

$$\Phi(1,\ldots,n) = \tfrac{1}{\sqrt{n!}}|\boldsymbol{\psi}(1)\cdots\boldsymbol{\psi}(n)| \qquad\qquad \tilde{\Phi}(1,\ldots,n) = \tfrac{1}{\sqrt{n!}}|\tilde{\boldsymbol{\psi}}(1)\cdots\tilde{\boldsymbol{\psi}}(n)|$$

which, noting that the matrix $\left(\tilde{\boldsymbol{\psi}}(1)\ \cdots\ \tilde{\boldsymbol{\psi}}(n)\right)$ is simply

$$\left(\tilde{\boldsymbol{\psi}}(1)\ \cdots\ \tilde{\boldsymbol{\psi}}(n)\right) = \left(\mathbf{U}^\dagger\boldsymbol{\psi}(1)\ \cdots\ \mathbf{U}^\dagger\boldsymbol{\psi}(n)\right) = \mathbf{U}^\dagger\left(\boldsymbol{\psi}(1)\ \cdots\ \boldsymbol{\psi}(n)\right)$$

implies $\tilde{\Phi} = \det(\mathbf{U}^\dagger)\Phi = \det(\mathbf{U})^*\Phi$. Therefore, $\tilde{\Phi}$ and Φ have the same energy expectation values.

$$\langle\tilde{\Phi}|\hat{H}_e|\tilde{\Phi}\rangle = \det(\mathbf{U}\mathbf{U}^\dagger)\langle\Phi|\hat{H}_e|\Phi\rangle = \langle\Phi|\hat{H}_e|\Phi\rangle \tag{D.1}$$