

# Programming Project 3 Theory (Part 1)

## *Hartree-Fock theory*

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

$$\hat{H}_e \Psi_K = E_K \Psi_K \quad (1)$$

$$\hat{H}_e = \sum_{A < B} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} - \frac{1}{2} \sum_i \nabla_i^2 - \sum_A \sum_i \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}_i|} + \sum_{i < j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (2)$$

with an optimal balance of accuracy and efficiency for the problem of interest. The most accurate solution possible for a given basis set (cc-pVXZ, DZP, ANO1, etc.) results from expanding the wavefunction

$$\Psi_K = \sum_P \Phi_P C_{PK} \quad (3)$$

in terms of all possible Slater determinants (the  $n$ -electron basis functions  $\{\Phi_P\}$ ) that can be formed from an orthonormal one-electron basis (spin-orbitals  $\{\psi_p\}$ ), and solving for the coefficients as eigenvectors of a matrix representation of  $\hat{H}_e$  in the  $n$ -electron basis,  $(\mathbf{H})_{PQ} = \langle \Phi_P | \hat{H}_e | \Phi_Q \rangle$ . 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  $\{\Phi_P\}$  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 MOs in the one-electron basis <sup>1</sup>, and there are “ $2m$  choose  $n$ ” <sup>2</sup>

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

unique Slater determinants in the  $n$ -electron basis that can be formed from the spin MOs. The upshot is that we typically have to omit some of the Slater determinants in our  $n$ -electron basis in order to make progress toward obtaining a solution in a reasonable amount of time.

As soon as we truncate our Slater determinant expansion (3), 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.

## Defining the problem

It can be shown that the exact ground- and excited-state wavefunctions of a system are stationary points of the Hamiltonian expectation value  $\langle \hat{H}_e \rangle = \langle \Psi | \hat{H}_e | \Psi \rangle$  with  $\Psi$  constrained to be normalized. That is, optimizing  $\langle \Psi | \hat{H}_e | \Psi \rangle$  by varying  $\Psi(1, \dots, n)$  subject only to the constraint  $\langle \Psi | \Psi \rangle = 1$  is equivalent to solving the Schrödinger equation. When we further constrain the form of  $\Psi$ , this is no longer true (i.e. we will generally *not* get an

<sup>1</sup> $m$   $\alpha$ -orbitals and  $m$   $\beta$ -orbitals.

<sup>2</sup>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>

exact eigenfunction of  $\hat{H}_e$ ). However, this *does* generally allow us to get the best approximation to  $\Psi$  subject to the additional constraints.

In order to make the Slater determinant expansion (3) converge with a relatively small number of  $\Phi_P$ 's, we wish to find the best single-determinant approximation to  $\Psi$ . That is, we wish to optimize

$$\langle \Phi | \hat{H}_e | \Phi \rangle \quad \Phi(1, \dots, n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \psi_1(1) & \psi_2(1) & \cdots & \psi_n(1) \\ \psi_1(2) & \psi_2(2) & \cdots & \psi_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(n) & \psi_2(n) & \cdots & \psi_n(n) \end{vmatrix} \quad (4)$$

with respect to variation of the orbitals  $\{\psi_p\}$ , enforcing the normalization constraint by keeping the spin orbitals orthonormal. This orbital optimization procedure is called the *Hartree-Fock method*.

Once we have solved this optimization problem, the expectation value  $\langle \Phi | \hat{H}_e | \Phi \rangle$  is itself a good first approximation to the electronic energy. Indeed, the Hartree-Fock method usually recovers the bulk of the electronic energy.<sup>3</sup> More importantly, however, when we use this new set of Hartree-Fock MOs,  $\{\psi_p\}$ , the FCI expansion tends to converge much more quickly to the true wavefunction. Specifically, when we rewrite the determinant expansion (3) in terms of single  $\{\Phi_i^a\}$ , double  $\{\Phi_{ij}^{ab}\}$ , triple  $\{\Phi_{ijk}^{abc}\}$ , etc. replacements<sup>4</sup> of the orbitals in the Hartree-Fock determinant  $\Phi$  with the remaining orbitals in the basis

$$\Psi = \Phi + \sum_i^a \Phi_i^a c_i^a + \sum_{\substack{a < b \\ i < j}} \Phi_{ij}^{ab} c_{ij}^{ab} + \sum_{\substack{a < b < c \\ i < j < k}} \Phi_{ijk}^{abc} c_{ijk}^{abc} + \dots \quad (5)$$

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

## Derivation

What follows is a derivation of the canonical Hartree-Fock equations, which define the Slater determinant  $\Phi$  that is a stationary point of the Hamiltonian expectation value  $\langle \Phi | \hat{H}_e | \Phi \rangle$ . A discussion of some of the background needed to understand why this derivation works is given in a series of appendices at the end.

### Components of $\hat{H}_e$

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

$$\hat{H}_e = V_{\text{Nu}} + \sum_i \hat{h}(i) + \sum_{i < j} \hat{g}(i, j) \quad (6)$$

$$V_{\text{Nu}} \equiv \sum_{A < B} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} \quad (7)$$

$$\hat{h}(i) \equiv -\frac{1}{2} \nabla_i^2 + \sum_A \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} \quad (8)$$

$$\hat{g}(i, j) \equiv \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (9)$$

which can be identified by the number of electron coordinates upon which they act.

<sup>3</sup>Although the importance of the remaining “correlation energy” cannot possibly be underestimated!

<sup>4</sup>It is typical to use dummy indices  $i, j, k, l$  to count over the orbitals in the reference determinant  $\Phi$  – the “occupied orbitals” – and to use  $a, b, c, d$ , to count over the orbitals not contained in  $\Phi$  – the “unoccupied” or “virtual orbitals.” Dummy indices  $p, q, r, s$  are generally used to count over the full set of spin MOs, whether occupied or not.

## Single-determinant expectation value of $\hat{H}_e$

The single-determinant expectation value of the electronic Hamiltonian is given by the first Slater rule

$$\langle \Phi | \hat{H}_e | \Phi \rangle = \sum_i^n \langle \psi_i | \hat{h} | \psi_i \rangle + \frac{1}{2} \sum_{ij}^n \langle \psi_i \psi_j | | \psi_i \psi_j \rangle \quad (10)$$

where the one- and two-electron integrals ( $\langle \psi_p | \hat{h} | \psi_q \rangle$  and  $\langle \psi_p \psi_q | | \psi_r \psi_s \rangle$ , respectively) are defined as follows.

$$\langle \psi_p | \hat{h} | \psi_q \rangle \equiv \int d(1) \psi_p^*(1) \hat{h}(1) \psi_q(1) \quad (11)$$

$$\langle \psi_p \psi_q | | \psi_r \psi_s \rangle \equiv \int d(1, 2) \psi_p^*(1) \psi_q^*(2) \hat{g}(1, 2) \psi_r(1) \psi_s(2) \quad \langle \psi_p \psi_q | | \psi_r \psi_s \rangle \equiv \langle \psi_p \psi_q | \psi_r \psi_s \rangle - \langle \psi_p \psi_q | \psi_s \psi_r \rangle \quad (12)$$

## The Hartree-Fock Lagrangian

We wish to optimize the single-determinant expectation value of  $\hat{H}_e$

$$\langle \Phi | \hat{H}_e | \Phi \rangle = \sum_{i=1}^n \langle \psi_i | \hat{h} | \psi_i \rangle + \frac{1}{2} \sum_{i,j=1}^n \langle \psi_i \psi_j | | \psi_i \psi_j \rangle \quad (13)$$

with respect to variation in of the orbitals  $\{\psi_i\}$ , subject to the constraint that the orbitals remain orthonormal.

$$\langle \psi_i | \psi_j \rangle = \delta_{ij} \quad (14)$$

This means that our Lagrangian functional <sup>5</sup> takes the form

$$\mathcal{L}[\{\psi_i\}, \{\epsilon_{ij}\}] = \sum_{i=1}^n \langle \psi_i | \hat{h} | \psi_i \rangle + \frac{1}{2} \sum_{i,j=1}^n \langle \psi_i \psi_j | | \psi_i \psi_j \rangle - \sum_{i,j=1}^n \epsilon_{ij} (\langle \psi_i | \psi_j \rangle - \delta_{ij}) \quad (15)$$

where  $\{\epsilon_{ij}\}$  are our Lagrangian multipliers for the orthonormality constraint (see appendix 1 for a brief explanation the Lagrangian approach to constrained optimization).

## Hartree-Fock Equations

In order to determine the stationarity condition for the spin MOs, we require that (see 2)

$$\left. \frac{d\mathcal{L}[\psi_k + \epsilon\eta]}{d\epsilon} \right|_{\epsilon=0} \stackrel{!}{=} 0 \quad (16)$$

holds for every orbital  $\psi_k$ . That is, we require that for all of the orbitals  $\psi_1, \dots, \psi_n$  in the Slater determinant  $\Phi$ , adding a little bit of some arbitrary function  $\eta = \eta(\mathbf{r}, m_s)$  doesn't change the Lagrangian.

Separating the terms in (15) involving one of the orbitals  $\psi_k$  from the remaining terms, we can write

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

---

<sup>5</sup>A functional is just a function of a function – i.e. some rule  $F$  that maps a function  $f$  to a number  $F[f]$ . Definite integrals are probably the most common example.

using the fact that  $\langle \psi_k \psi_i | \psi_k \psi_i \rangle = \langle \psi_i \psi_k | \psi_i \psi_k \rangle$ , which can be seen by exchanging dummy variables of integration in the integral (12). The functional derivative for varying  $\psi_k$  is then

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

where we have dropped all terms not involving  $\varepsilon$  (including the  $\delta_{ik}$ s), since their derivatives vanish. Evaluating the derivative and plugging in  $\varepsilon = 0$ , we get

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

The first two terms can be rewritten as <sup>6</sup>

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

where  $\hat{P}(1, 2)$  is an “operator” that swaps the coordinates of the two one-electron functions on its right (prior to integration). <sup>7</sup> The term in parentheses is called the *Fock operator* <sup>8</sup>

$$\hat{f}(1) \equiv \hat{h}(1) + \sum_i \langle \psi_i(2) | \hat{g}(1, 2) (1 - \hat{P}(1, 2)) | \psi_i(2) \rangle \quad (17)$$

which, note, is really an operator functional of the orbitals  $\hat{f} = \hat{f}[\psi_1, \dots, \psi_n]$ .

With this new operator, our stationarity condition takes the form

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

where we see that the second term is the complex conjugate of the first. By the Fundamental Lemma of Calculus of Variations (see appendix 3), this condition is equivalent to requiring

$$\hat{f}(1) \psi_k(1) - \sum_i \epsilon_{ki} \psi_i(1) \stackrel{!}{=} 0 \quad (18)$$

$$\hat{f}(1) \psi_k^*(1) - \sum_i \epsilon_{ik} \psi_i^*(1) \stackrel{!}{=} 0 \quad (19)$$

where we have used the fact that the Fock operator is Hermitian, so that  $\langle \psi_k | \hat{f} \eta \rangle = \langle \hat{f}^\dagger \psi_k | \eta \rangle = \langle \hat{f} \psi_k | \eta \rangle$ . Subtracting the complex conjugate of equation (19) from equation (18) gives

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

<sup>6</sup>Defining  $\langle \psi_p(2) | \hat{g}(1, 2) | \psi_q(2) \rangle \equiv \int d(2) \psi_p^*(2) \hat{g}(1, 2) \psi_q(2)$

<sup>7</sup>This really just a notational trick to allow us to use the short-hand  $\psi_i(1) \psi_j(2) - \psi_j(1) \psi_i(2) \equiv (1 - \hat{P}(1, 2)) \psi_i(1) \psi_j(2)$  which could just as easily be done by instead defining an “index-permuter”  $\hat{P}_{ij}$  to swap the orbital indices.

<sup>8</sup>You have probably seen 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 called the Coloumb and exchange operators.

and, since the orbitals  $\{\psi_i\}$  are linearly independent,<sup>9</sup> this tells us that

$$\epsilon_{k1} - \epsilon_{1k}^* = \dots = \epsilon_{kn} - \epsilon_{nk}^* = 0$$

i.e. the matrix of Lagrange multipliers  $\{\epsilon_{ij}\}$  must be Hermitian.

This gives us a new set of conditions for stationarity of  $\langle \Phi | \hat{H}_e | \Phi \rangle$  subject to the constraint  $\langle \psi_i | \psi_j \rangle = \delta_{ij}$

$$\hat{f}\psi_i \stackrel{!}{=} \sum_j \epsilon_{ij}\psi_j \quad \epsilon_{ij} \stackrel{!}{=} \epsilon_{ji}^* \quad (20)$$

which is one form of the *Hartree-Fock equations*. We can show that this set of equations can be partially decoupled<sup>10</sup> by diagonalizing the multiplier matrix to give a series of eigenvalue equations  $\hat{f}\psi_i = \epsilon_i\psi_i$ , known as the *canonical Hartree-Fock equations*.

## Canonical Hartree-Fock Equations

Equation (20) can be written in matrix form as

$$\hat{f}\psi \stackrel{!}{=} \epsilon\psi \quad \epsilon = \epsilon^\dagger \quad (21)$$

where we have defined a matrix of Lagrange multipliers and a vector of orbitals.

$$\epsilon = \begin{pmatrix} \epsilon_{11} & \dots & \epsilon_{1n} \\ \vdots & \ddots & \vdots \\ \epsilon_{n1} & \dots & \epsilon_{nn} \end{pmatrix} \quad \psi = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_n \end{pmatrix} \quad (22)$$

Since  $\epsilon$  is a Hermitian matrix, it can be diagonalized by a unitary<sup>11</sup> transformation  $\mathbf{U}$ .

$$\epsilon = \mathbf{U}\tilde{\epsilon}\mathbf{U}^\dagger \quad (23)$$

Inserting this decomposition into equation (21) and multiplying both sides from the left by  $\mathbf{U}^\dagger$ , we get

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

which shows that the problem can be put into canonical form by using a new set of orbitals  $\tilde{\psi} = \mathbf{U}^\dagger\psi$  formed from  $\psi_1, \dots, \psi_n$  as

$$\tilde{\psi}_i = \sum_{j=1}^n U_{ji}^* \psi_j \quad (24)$$

We can show that this “re-mixing” of the orbitals in  $\Phi$  leaves the Fock operator  $\hat{f}$  and the expectation value  $\langle \Phi | \hat{H}_e | \Phi \rangle$  unchanged, while also preserving orbital orthonormality (see 4).

Dropping the tildes, the *canonical Hartree-Fock equations* can be written as

$$\hat{f}\psi_i = \epsilon_i\psi_i \quad i = 1, \dots, n$$

where  $\epsilon_i$  is an element of the diagonal multiplier matrix and the requirement that  $\epsilon = \epsilon^\dagger$  now simply means that each  $\epsilon_i$  is real,  $\epsilon_i = \epsilon_i^*$ . Note that these equations are not fully decoupled, since  $\hat{f}$  depends on the full set  $\{\psi_i\}$ . Solving these equations amounts to solving for the *self-consistent field*

$$\hat{v}(1) \equiv \sum_i \langle \psi_i(2) | \hat{g}(1,2)(1 - \hat{P}(1,2)) | \psi_i(2) \rangle = \sum_i (\hat{J}_i(1) - \hat{K}_i(1)) \quad (25)$$

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

<sup>9</sup>[http://en.wikipedia.org/wiki/Linear\\_independence#Definition](http://en.wikipedia.org/wiki/Linear_independence#Definition)

<sup>10</sup>We can never fully decouple them, because the Fock operator depends on all of the orbitals.

<sup>11</sup>[http://en.wikipedia.org/wiki/Unitary\\_matrix](http://en.wikipedia.org/wiki/Unitary_matrix)

# 1 Appendix: Constrained Optimization

The standard method of optimizing a function or functional 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

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

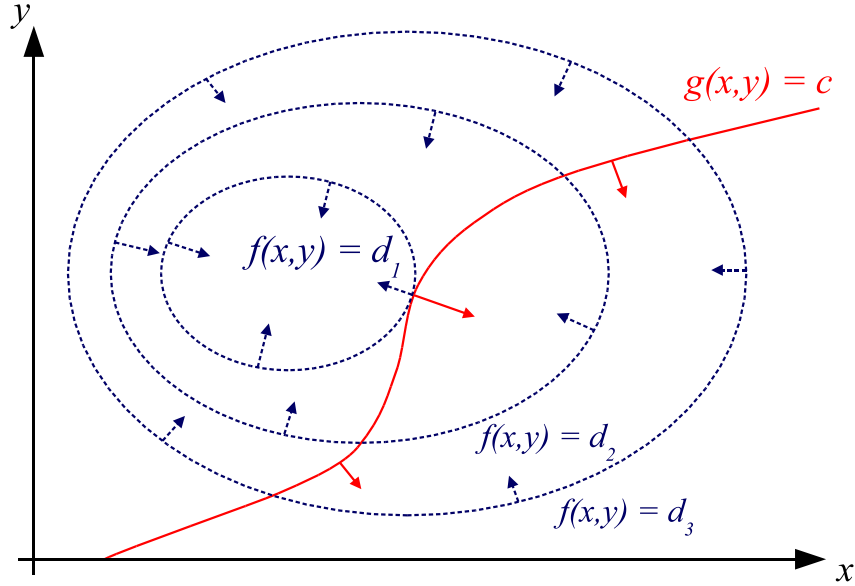
where the parameter  $\lambda$  is called the Lagrange multiplier. The constrained optimization of  $f(x, y)$  can be achieved by determining a stationary point of  $\mathcal{L}$  in all of its arguments.<sup>12</sup>

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial x} &= \frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} \stackrel{!}{=} 0 \\ \frac{\partial \mathcal{L}}{\partial y} &= \frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} \stackrel{!}{=} 0 \\ \frac{\partial \mathcal{L}}{\partial \lambda} &= c - g(x, y) \stackrel{!}{=} 0\end{aligned}$$

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 surface  $f(x, y)$  and the constraint surface  $g(x, y)$  be parallel

$$\nabla f = \lambda \nabla g \tag{26}$$

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 surface  $f(x, y)$ . 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  $\nabla f$  parallel to  $g(x, y) = c$ .

<sup>12</sup>The  $\stackrel{!}{=}$  sign means “must equal” – these are the conditions to be satisfied.

## 2 Appendix: Optimizing Functionals

In order to optimize  $\langle \Phi | \hat{H}_e | \Phi \rangle$  with respect to variation of the spin MOs, we need to learn how to take a derivative of a functional with respect to a function (a *functional derivative* <sup>13</sup>). First, consider the definition of the derivative of a simple 1D function  $f(x)$ .

$$\frac{df(x)}{dx} \equiv \lim_{\epsilon \rightarrow 0} \frac{f(x + \epsilon) - f(x)}{\epsilon}$$

which is simply looking at how  $f(x)$  changes when we add “a little bit”  $\epsilon$  of the number 1. The stationarity condition for a function is that it doesn’t change when we add “a little bit” to its argument.

$$\lim_{\epsilon \rightarrow 0} \frac{f(x + \epsilon) - f(x)}{\epsilon} \stackrel{!}{=} 0 \quad (27)$$

The stationarity condition of a functional is similar. To optimize a functional  $F[f]$ , we require that it doesn’t change when we add “a little bit”  $\epsilon$  of an arbitrary function  $g$ .

$$\lim_{\epsilon \rightarrow 0} \frac{F[f + \epsilon g] - F[f]}{\epsilon} \stackrel{!}{=} 0 \quad (28)$$

Note, however, that  $F[f]$  must be stationary with respect to *any* form for  $g$  – otherwise we haven’t really optimized the thing.

One can (you should) show that that a single derivative may also be written as

$$\frac{df(x)}{dx} = \left. \frac{df(x + \epsilon)}{d\epsilon} \right|_{\epsilon=0} \quad (29)$$

by expanding  $f$  on the right side in a Taylor series and evaluating the derivative. This latter form of derivative readily generalizes to functionals.

$$\lim_{\epsilon \rightarrow 0} \frac{F[f + \epsilon g] - F[f]}{\epsilon} = \left. \frac{dF[f + \epsilon g]}{d\epsilon} \right|_{\epsilon=0} \quad (30)$$

and is particularly convenient because it allows us to deal with a plain ol’ derivative rather than something more complicated. Hence, we can obtain stationarity conditions for functionals by requiring that

$$\left. \frac{dF[f + \epsilon g]}{d\epsilon} \right|_{\epsilon=0} \stackrel{!}{=} 0 \quad (31)$$

for *any* possible choice of  $g$ .

---

<sup>13</sup>[http://en.wikipedia.org/wiki/Functional\\_derivative](http://en.wikipedia.org/wiki/Functional_derivative)

### 3 Appendix: Fundamental Lemma of Calculus of Variations

The *Fundamental Lemma of Calculus of Variations*<sup>14</sup> allows us to deal with situations like

$$\int_{-\infty}^{\infty} dx f(x) \eta(x) = 0 \quad \text{for all } \eta(x) \quad (32)$$

and tells us that this can only be true if the function  $f(x)$  vanishes everywhere – that is, if  $f(x) = 0$ . We can see that this is true by considering the case  $\eta(x) = f(x)$ . Since  $f(x)^2$  is nonnegative everywhere, the integral can only be zero in this case if  $f(x) = 0$ .

**Complex functions.** Consider an analogous situation for two complex functions  $f$  and  $g$  that satisfy

$$\int_{-\infty}^{\infty} dx \eta^*(x) f(x) + \int_{-\infty}^{\infty} dx \eta(x) g(x) = 0 \quad \text{for all } \eta(x). \quad (33)$$

We can show that this can only be satisfied if both  $f(x)$  and  $g(x)$  separately vanish everywhere. Since  $\eta(x)$  is arbitrary, this equation must hold for some other function  $\gamma(x) = i\eta(x)$ .

$$\int_{-\infty}^{\infty} dx \gamma^*(x) f(x) + \int_{-\infty}^{\infty} dx \gamma(x) g(x) = 0 \quad (34)$$

$$\int_{-\infty}^{\infty} dx [i\eta(x)]^* f(x) + \int_{-\infty}^{\infty} dx [i\eta(x)] g(x) = 0 \quad (35)$$

$$-i \int_{-\infty}^{\infty} dx \eta^*(x) f(x) + i \int_{-\infty}^{\infty} dx \eta(x) g(x) = 0 \quad (36)$$

Adding  $i$  times equation (36) to equation (33) yields:

$$\int_{-\infty}^{\infty} dx \eta^*(x) f(x) = 0 \quad \text{for all } \eta(x). \quad (37)$$

Subtracting  $i$  times equation (36) to equation (33) yields:

$$\int_{-\infty}^{\infty} dx \eta(x) g(x) = 0 \quad \text{for all } \eta(x). \quad (38)$$

Again by the Fundamental Lemma of Calculus of Variations, equations (37) and (38) can only hold true together if  $f(x) = 0$  and  $g(x) = 0$ .<sup>15</sup>

<sup>14</sup>[http://en.wikipedia.org/wiki/Fundamental\\_lemma\\_of\\_calculus\\_of\\_variations](http://en.wikipedia.org/wiki/Fundamental_lemma_of_calculus_of_variations)

<sup>15</sup>Since we could always choose  $\eta(x) = f(x)$  or  $\eta(x) = g^*(x)$ , giving integrands  $|f(x)|^2$  and  $|g(x)|^2$  that are nonnegative everywhere.



## 4 Appendix: Unitary Invariances for Hartree-Fock Orbitals

**orbital orthonormality** Unitary transformations actually always preserve orthonormality by definition, but we can show it:

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

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

**Fock operator** The orbitals enter only into the Coulomb and exchange part of the Fock operator. 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 elegantly express  $\Phi$  and  $\tilde{\Phi}$  as follows.

$$\Phi(1, \dots, n) = \frac{1}{\sqrt{n!}} |\boldsymbol{\psi}(1) \cdots \boldsymbol{\psi}(n)| \quad \tilde{\Phi}(1, \dots, n) = \frac{1}{\sqrt{n!}} |\tilde{\boldsymbol{\psi}}(1) \cdots \tilde{\boldsymbol{\psi}}(n)|$$

But note that the matrix  $(\tilde{\boldsymbol{\psi}}(1) \cdots \tilde{\boldsymbol{\psi}}(n))$  is simply

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

and so  $\tilde{\Phi} = \det(\mathbf{U}^\dagger) \Phi = \det(\mathbf{U})^* \Phi$ , which means

$$\langle \tilde{\Phi} | \hat{H}_e | \tilde{\Phi} \rangle = \det(\mathbf{U} \mathbf{U}^\dagger) \langle \Phi | \hat{H}_e | \Phi \rangle = \langle \Phi | \hat{H}_e | \Phi \rangle \quad (39)$$