# The SCF Workshop Notes 1

Lucas Aebersold                                                                    May 21, 2018

## 1.   Hartree-Fock Self-Consistent Field Theory

In this workshop, we will introduce the theory and implementation the Hartree-Fock Self-Consistent Field Theory (HF-SCF) first with restricted orbitals and closed-shell systems (RHF) and then unrestricted orbitals and open-shelled systems (UHF). First let's take a quick review of what it is we are solving, the Schrödinger equation.

### 1.1.  The electronic structure problem

$$\mathscr{H}\Psi(\mathbf{R}, \mathbf{r}) = E\Psi(\mathbf{R}, \mathbf{r}) \tag{1.1}$$

where

$$\mathbf{R} \rightarrow \text{Nuclear coordinates}$$
$$\mathbf{r} \rightarrow \text{Electronic coordinates}$$

The molecular Hamiltonian is

$$\mathscr{H} = -\sum_A \frac{1}{2m_A}\nabla_A^2(\mathbf{R}) - \frac{1}{2}\sum_i \nabla_i^2(\mathbf{r}) + \sum_A \sum_B \frac{q_A q_B}{r_{AB}} - \sum_A \sum_j \frac{q_A}{r_{Aj}} + \sum_i \sum_j \frac{1}{r_{ij}} \tag{1.2}$$

where, $m$ is the mass, $r$, the radius, $q$ the electric charge. The indices $A$, $B$ represent nuclear indices and the indices $i$, $j$ are electronic indices. We can separate Eq. 1.2 into five distinct terms

$$\mathscr{H} = \underbrace{-\sum_A \frac{1}{2m_A}\nabla_A^2(\mathbf{R})}_{\text{nuclear KE}} \overbrace{- \frac{1}{2}\sum_i \nabla_i^2(\mathbf{r})}^{\text{Electronic KE}} + \underbrace{\sum_A \sum_B \frac{q_A q_B}{r_{AB}}}_{\text{Nuclear repulsion}} \overbrace{- \sum_A \sum_j \frac{q_A}{r_{Aj}}}^{\text{electron-nuclear attraction}} + \underbrace{\sum_i \sum_j \frac{1}{r_{ij}}}_{\text{electron-electron repulsion}}$$

Invoking the Born-Oppenheimer approximation, we treat the nucleus as static, removing the first term. This also separates our wavefunction into a product of nuclear and electronic coordinates. Finally, we also assume atomic units, (the god given units) are

$$\hbar = e = m_e = (4\pi\varepsilon_0)^{-1} = 1$$

For a fixed geometry, term 3, can be evaluated separately, leaving us with the electronic Hamiltonian

$$\mathscr{H} = -\frac{1}{2}\sum_i \nabla_i^2(\mathbf{r}) - \sum_{Aj} \frac{Z_A}{r_{Aj}} + \sum_{ij} \frac{1}{r_{ij}} \tag{1.3}$$

### 1.2.  Theoretical Overview

In order to solve the Schrödinger equation for systems with more than one electron, we must introduce the concept of orbitals. We can represent our wavefunction as a sum of individual orbitals or 'basis functions',

that is a linear combination of basis functions. The most general form of the equation can be written as

$$\psi(\mathbf{r}) = \sum_i c_i \phi_i(\mathbf{r}) \tag{1.4}$$

written in terms of spatial coordinates $\mathbf{r}$. However, a spin coordinate $\omega$ is required as well, making our wavefunction a function of both distance in space and direction of spin $\alpha$ (up) or $\beta$ (down)

$$\Phi(\mathbf{r}, \omega) = \Phi(\mathbf{x}) \qquad \mathbf{x} = \{r, \omega\} \tag{1.5}$$

where our spin coordinates are allowed values of

$$\omega = \begin{cases} \alpha(\omega) & = |\alpha\rangle \\ \beta(\omega) & = |\beta\rangle \end{cases} \tag{1.6}$$

Note, the Hamiltonian in nonrelativistic quantum mechanics does not depend on spin, but spin must be included for some answers to be sensible. Basically, spin is derived from Dirac's relativistic work, and just nicely placed back into the nonrelativistic Hamiltonian. Like pretty much everything else in quantum mechanics, spin coordinates are orthonormal, i.e. $\langle \omega_i | \omega_j \rangle = \delta_{\alpha\beta}$. What happens if we 'swap' the positions of the electrons? For right now, we will simply assert that electrons, which are fermions, are antisymmetric with respect to exchange, that is

$$\psi(x_1, \ldots x_i, \ldots x_j, \ldots x_n) = -\psi(x_1, \ldots x_j, \ldots x_i, \ldots x_n) \tag{1.7}$$

More plainly stated, if we put electron 1 into electron 2's orbital and electron 2 into electron 1's orbital, the wavefunction will change its 'sign'. Lastly, we will also distinguish between the spatial and spin orbitals

$$\phi(\mathbf{r}) = \text{spatial orbital} \tag{1.8}$$

$$\chi(\mathbf{x}) = \begin{cases} \phi(\mathbf{r})\alpha(\omega) \\ \phi(\mathbf{r})\beta(\omega) \end{cases} = \text{spin orbital} \tag{1.9}$$

where the spin orbital is the product of the spatial orbital and the spin coordinate. What this comes down to is we must use determinants because they have the property that if we rows or columns, we change the sign.

## 1.3. The Hartree-Fock approximation for a two electron system

Now we will focus on the Hartree-Fock approximation for a two electron system. First, let us introduce a more compact notation for the Slater determinant, where it is written as a ket of the diagonal elements $\chi_k(i)$,

$$|\Psi_0\rangle = |\chi_1 \chi_2 \cdots \chi_N\rangle \tag{1.10}$$

The ground state energy is

$$E_0 = \langle \Psi_0 | \mathcal{H} | \Psi_0 \rangle \tag{1.11}$$

We need to find the $|\Psi_0\rangle$ that minimizes $E_0$. The electronic Hamiltonian from Eq. (1.3) can be rewritten as

$$\mathcal{H} = \sum_i \left( -\frac{1}{2}\nabla_i^2 - \sum_A \frac{Z_A}{r_{iA}} \right) + \sum_{ij} \frac{1}{r_{ij}} \tag{1.12}$$

Where $i$ represents one electron, $A$ represents the nuclei and $i, j$ represent pairs of electrons.

$$\mathcal{H} = \underbrace{\sum_i \left( -\frac{1}{2}\nabla_i^2 - \sum_A \frac{Z_A}{r_{iA}} \right)}_{\text{one-electron operator}} + \underbrace{\sum_{ij} \frac{1}{r_{ij}}}_{\text{two-electron operator}}$$

To solve this we will break it up into two parts, the one-electron part, which includes the kinetic and potential energy terms

$$h(i) = -\frac{1}{2}\nabla_i^2 - \sum_A \frac{Z_A}{r_{iA}} = T + V \tag{1.13}$$

and write the full Hamiltonian in terms of the one-electron part and the two-electron part

$$\mathcal{H} = \sum_i h(i) + \sum_{ij} \frac{1}{r_{ij}} \tag{1.14}$$

First we'll consider a system with only two electrons, expanding equation (1.14) out would then look like

$$\mathcal{H} = \overset{1}{\overbrace{h(1)}} + \overset{2}{\overbrace{h(2)}} + \overset{3}{\overbrace{\frac{1}{r_{ij}}}} \tag{1.15}$$

Solving for term (1), we get

$$\langle \psi_0 | h(1) | \psi_0 \rangle = \int \left[ 2^{-1/2} \left( \chi_1(1)\chi_2(2) - \chi_2(1)\chi_1(2) \right) \right]^* h(1) \left[ 2^{-1/2} \left( \chi_1(1)\chi_2(2) - \chi_2(1)\chi_1(2) \right) \right] d\mathbf{r}_1 d\mathbf{r}_2$$

$$= \frac{1}{2} \int \left[ \chi_1^*(1)\chi_2^*(2)h(r_1)\chi_1(1)\chi_2(2) + \chi_2^*(1)\chi_1^*(2)h(r_1)\chi_2(1)\chi_1(2) \right.$$
$$\left. - \chi_1^*(1)\chi_2^*(2)h(r_1)\chi_2(1)\chi_1(2) - \chi_2^*(1)\chi_1^*(2)h(r_1)\chi_1(1)\chi_2(2) \right] d\mathbf{r}_1 d\mathbf{r}_2$$

Which can be written in Dirac notation as

$$\langle \psi_0 | h(1) | \psi_0 \rangle = \frac{1}{2} \left[ \langle 1|h|1 \rangle + \langle 2|h|2 \rangle \right] \tag{1.16}$$

The equation for electron 2 will follow similarly only with the indices switched, giving us

$$\langle \psi_0 | h(2) | \psi_0 \rangle = \frac{1}{2} \left[ \langle 2|h|2 \rangle + \langle 1|h|1 \rangle \right] \tag{1.17}$$

The two-electron part can be written as

$$\langle \psi_0 | \frac{1}{r_{12}} | \psi_0 \rangle = \frac{1}{2} \int \left[ \chi_1^*(1)\chi_2^*(2)\frac{1}{r_{12}}\chi_1(1)\chi_2(2) + \chi_2^*(1)\chi_1^*(2)\frac{1}{r_{12}}\chi_2(1)\chi_1(2) \right.$$
$$\left. - \chi_1^*(1)\chi_2^*(2)\frac{1}{r_{12}}\chi_2(1)\chi_1(2) - \chi_2^*(1)\chi_1^*(2)\frac{1}{r_{12}}\chi_1(1)\chi_2(2) \right] d\mathbf{r}_1 d\mathbf{r}_2 \tag{1.18}$$

The two-electron Slater determinant can now be written as

$$\langle \psi_0 | \frac{1}{r_{12}} | \psi_0 \rangle = \langle 12|12 \rangle - \langle 12|21 \rangle = \mathcal{J}_{12} - \mathcal{K}_{12}$$

where we have introduced two terms, $\mathcal{J}$ is the Coulomb integral (the electronic repulsion from two clouds of charge) and $\mathcal{K}$ is the exchange integral, the correction to the Coulomb integral. This takes into account two particles of the same spin cannot be right on top of each other, removing it from the integral. I must mention there are two notations for the two-electron integral $\langle ij|kl \rangle$, physicists' notation which is

$$\langle ij|kl \rangle = \int \chi_i^*(1)\chi_j^*(2)\frac{1}{r_{12}}\chi_k(1)\chi_l(2)\, d\mathbf{r}_1 d\mathbf{r}_2 \tag{1.19}$$

Then there's also chemists' notation,

$$[ij|kl] = \int \chi_i^*(2)\chi_j(2)\frac{1}{r_{12}}\chi_k^*(1)\chi_l(1)\, d\mathbf{r}_1 d\mathbf{r}_2 \tag{1.20}$$

Here electron 2 appears completely on the bra side, complex-conjugate first, and electron 1 appears completely on the ket side. This notation, is not as physically intuitive as physicist's notation, but when coding these equations, it makes much more sense to use it. Interestingly, S&O comment on it being unfortunate there are two types of notation, but never comment on why chemists' notation exists, and use physicists' notation the rest of the chapter. However, they then proceed to use chemists notation in Appendix A, which is the discussion and derivation of how to integrate Gaussian type orbitals and in Appendix B, the SCF code. Again they give no reason as to the switch. The reason is creating indexing loops for symmetry are much easier in chemists notation, hence nearly all codes use chemists' notation.

**Table 1.** Helpful table of common quantum chemistry expressions and their meaning.

| Expression | Description |
|---|---|
| $\mathbf{x}_i$ | The space and spin coordinates of particle $i$ |
| $\mathbf{r}_i$ | The spatial coordinates of particle $i$ |
| $\phi_i$ | The $i$-th one-particle basis function (orbital), can also be written as $i$ |
| | Usually denotes a spin-orbital obtained from an HF procedure. |
| $\chi_i$ | The $i$-th one-particle basis function (orbital) |
| | Usually denotes atomic spin-orbital |
| $\langle i\|h\|j\rangle$ | One-electron integral in physicists' notation (spin orbitals) |
| | $= \int \phi_i^*(\mathbf{x}_1)h(\mathbf{r}_1)\phi_j(\mathbf{x}_1)\,d\mathbf{x}_1$ |
| $[i\|h\|j]$ | One-electron integral in chemists' notation (spin orbitals) |
| | $= \langle i\|h\|j\rangle$ |
| $\langle ij\|kl\rangle \quad =$ | $\int \phi_i^*(\mathbf{x}_1)\phi_j^*(\mathbf{x}_2)r_{12}^{-1}\phi_k(\mathbf{x}_1)\phi_l(\mathbf{x}_2)\,d\mathbf{x}_1 d\mathbf{x}_2$ |
| | Two-electron integral for spin orbitals in physicists' notation |
| $[ij\|kl] \quad =$ | $\int \phi_i^*(\mathbf{x}_1)\phi_j(\mathbf{x}_1)r_{12}^{-1}\phi_k(\mathbf{x}_2)\phi_l^*(\mathbf{x}_2)\,d\mathbf{x}_1 d\mathbf{x}_2$ |
| | Two-electron integral for spin orbitals in chemists' notation |
| $\langle ij\|\|kl\rangle \quad =$ | $\langle ij\|kl\rangle - \langle ij\|lk\rangle$ |
| | Antisymmetrized two-electron integral |
| $(i\|h\|j)$ | One-electron integral in chemists' notation (spatial orbitals) |
| $(ij\|kl) \quad =$ | $\int \phi_i^*(\mathbf{r}_1)\phi_j(\mathbf{r}_1)r_{12}^{-1}\phi_k(\mathbf{r}_2)\phi_l^*(\mathbf{r}_2)\,d\mathbf{r}_1 d\mathbf{r}_2$ |
| | A two-electron integral in chemists' notation (spatial orbitals) |

## 1.4. The Roothaan Equations

We kind of skip a few steps here, and will dive straight into discussion of the When we eliminate spin, the calculation of our molecular orbitals then involves the spatial integro-differential equation

$$f(r_1)\psi_i(r_1) = \varepsilon_i \psi_i(r_1) \tag{1.21}$$

However, there is a problem with Eq. (1.21) in that no practical procedures exist for getting numerical solutions for molecules. However, Roothaan would come up with a method that when a set of known spatial basis functions is introduced, a differential equation can be converted to a set of algebraic equations and solved via matrix methods.

So, we introduce a set of $K$ known basis functions $\{\phi_\mu(r)|\mu = 1, 2, \ldots, K\}$ and expand the unknown molecular orbitals in the linear expansion

$$\psi_i = \sum_{\mu=1}^{K} C_{\mu i}\phi_\mu \qquad i = 1, 2, \ldots, K \tag{1.22}$$

If the set of basis functions $\{\phi_\mu\}$ was complete, this would be an exact expansion, unfortunately we are also

restricted to a finite set of $K$ basis functions.

$$f(1) \sum_\nu C_{\nu i} \phi_\nu(1) = \varepsilon_i \sum_\nu C_{\nu i} \phi_\nu(1) \tag{1.23}$$

Multiplying $\phi_\mu^*(1)$ on the left and integrating, we turn the integro-differential equation into a matrix equation.

$$\sum_\nu C_{\nu i} \int \phi_\mu^*(1) f(1) \phi_\nu(1) \, d\mathbf{r}_1 = \varepsilon_i \sum_\nu C_{\nu i} \int \phi_\mu^*(1) \phi_\nu(1) \, d\mathbf{r}_1 \tag{1.24}$$

We'll define two matrices

1. The *overlap matrix* **S** has elements

$$S_{\mu\nu} = \int \phi_\mu^*(1) \phi_\nu(1) \, d\mathbf{r}_1 \tag{1.25}$$

   and is a $K \times K$ Hermitian (usually real and symmetric) matrix. The basis functions $\{\phi_\mu\}$ are not usually orthogonal to each other and hence elements will have an overlap of $0 \leq S_{\mu\nu} \geq 1$, namely the off diagonal elements, while the diagonal elements are always one.
   Fun facts about the overlap matrix

   (a) The sign of the off-diagonal elements depends on the relative sign of the two basis functions.
   (b) If two off-diagonal elements approach unity, i.e. they completely overlap, the basis functions approach linear dependence
   (c) **S** is Hermitian it can be diagonalized by a unitary matrix
   (d) The eigenvalues of **S** are necessarily positive numbers $\rightarrow$ positive-definite matrix.
   (e) Linear dependence in the basis set is associated with eigenvalues of **S** approaching zero.

2. The *Fock matrix* **F** has elements

$$F_{\mu\nu} = \int \phi_\mu^*(1) f(1) \phi_\nu(1) \, d\mathbf{r}_1 \tag{1.26}$$

   The Fock matrix **F** is the matrix representation of the Fock operator with the set of basis functions $\{\phi_\mu\}$. With definitions of **F** and **S** we can write the integrated HF equation as

$$\sum_\nu F_{\mu\nu} C_{\nu i} = \varepsilon_i \sum_\nu S_{\mu\nu} C_{\nu i} \qquad i = 1, 2, \ldots, K \tag{1.27}$$

These are the notorious *Roothaan equations*, which can be written as a seemingly harmless single matrix equation

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon} \tag{1.28}$$

where **C** is the matrix of orbital coefficients

$$\mathbf{C} = \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1K} \\ C_{21} & C_{22} & \cdots & C_{2K} \\ \vdots & \vdots & & \vdots \\ C_{K1} & C_{K2} & \cdots & C_{KK} \end{pmatrix} \tag{1.29}$$

and $\boldsymbol{\varepsilon}$ is a diagonal matrix of the orbital energies $\varepsilon_i$

$$
\boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 & & & \\ & \varepsilon_2 & & \\ & & \ddots & \\ & & & \varepsilon_K \end{pmatrix}
\tag{1.30}
$$

Note that from (1.22) and (1.29) we know the columns of $\mathbf{C}$ will describe the molecular orbitals, i.e. the coefficients describing $\psi_1$ are the first column of $\mathbf{C}$, those describing $\psi_2$ are in the second column of $\mathbf{C}$, etc.

### 1.4.1 The Fock Matrix

The Fock operator allows us to build a new set of orbitals iteratively and approach the exact solution

$$
f(1) = h(1) + \sum_a^{N/2} 2\mathcal{J}_a(1) + \sum_a \mathcal{K}_a(1)
\tag{1.31}
$$

here $a$ runs over all occupied orbitals. The Coulomb and exchange operators represent the repulsion/correlation from all the electrons

$$
f\chi_a = \varepsilon_a \chi_a
\tag{1.32}
$$

Diagonalizing this would give us a set of orbitals that minimizes the energy. We first need a form for the Hamiltonian matrix $\mathbf{H}$, which can be written as

$$
H_{\mu\nu}^{\text{core}} + \sum_a^{N/2} 2\,(\mu\nu|aa) - (\mu a|a\nu)
\tag{1.33}
$$

where $\mu$ and $\nu$ operate over the basis functions. This is the core-Hamiltonian

$$
H_{\mu\nu}^{\text{core}} = \int \phi_\mu^*(1) h(1) \phi_\nu(1)\, d\mathbf{r}_1
\tag{1.34}
$$

The elements of the core Hamiltonian matrix are integrals involving the one-electron operator $h(1)$, describing the kinetic energy and nuclear attraction of an electron, i.e.

$$
h(1) = -\frac{1}{2}\nabla_i^2 - \sum_A \frac{Z_A}{|\mathbf{r}_1 - \mathbf{R}_A|}
\tag{1.35}
$$

Calculating the elements of the core-Hamiltonian matrix thus involves the kinetic energy integrals

$$
T_{\mu\nu} = \int \phi_\mu^*(1)\left(-\frac{1}{2}\nabla_1^2\right)\phi_\nu(1)\, d\mathbf{r}_1
\tag{1.36}
$$

and the nuclear attraction integrals

$$
V_{\mu\nu}^{\text{nucl}} = \int \phi_\mu^*(1)\left(-\sum_A \frac{Z_A}{|\mathbf{r}_1 - \mathbf{R}_A|}\right)\phi_\nu(1)\, d\mathbf{r}_1
\tag{1.37}
$$

where

$$H_{\mu\nu}^{\text{core}} = T_{\mu\nu} + V_{\mu\nu}^{\text{nucl}} \tag{1.38}$$

To return to expression (1.33), we can represent the Fock matrix $\mathbf{F}$ as

$$F_{\mu\nu} = \int \varphi_{\mu}^*(1) f(1) \varphi_{\nu}(1) \, d\mathbf{r}_1 = H_{\mu\nu}^{\text{core}} + G_{\mu\nu} \tag{1.39}$$

we insert the linear expansion for the MOs into the two-electron terms and get

$$F_{\mu\nu} = H_{\mu\nu}^{\text{core}} + G_{\mu\nu}$$

where we have introduced the matrix $\mathbf{G}$ as the two-electron part of the Fock matrix

### 1.4.2 Matrix Elements

We let $\psi_a$ denote

$$\psi_a(1) = \sum_{\lambda} C_{\lambda a} \phi_{\lambda}(1) \tag{1.40}$$

Allowing us to write $\mathbf{G}$ as

$$
\begin{aligned}
G_{\mu\nu} &= \sum_a^{N/2} \left[ 2 \int \phi_{\mu}^*(1) \psi_a^*(2) \frac{1}{r_{12}} \phi_{\nu}(1) \psi_a(2) \, d\mathbf{r}_1 d\mathbf{r}_2 - \int \phi_{\mu}^* \psi_a^*(2) \frac{1}{r_{12}} \psi_a(1) \phi_{\nu}(2) \, d\mathbf{r}_1 d\mathbf{r}_2 \right] \\
&= \sum_i^{N/2} \sum_{\lambda\sigma} C_{\lambda a} C_{\sigma a}^* \left[ 2 \int \phi_{\mu}^*(1) \phi_{\nu}^*(2) \frac{1}{r_{12}} \phi_{\nu}(1) \phi_{\lambda}(2) \, d\mathbf{r}_1 d\mathbf{r}_2 - \int \phi_{\mu}^*(1) \phi_{\sigma}^*(2) \frac{1}{r_{12}} \phi_{\lambda}(1) \phi_{\nu}(1) \, d\mathbf{r}_1 d\mathbf{r}_2 \right] \\
&= \sum_i^{N/2} \sum_{\lambda\sigma} C_{\lambda a} C_{\sigma a}^* \left[ 2 \langle \mu\nu | \nu\lambda \rangle - \langle \mu\sigma | \lambda\nu \rangle \right]
\end{aligned}
$$

The density matrix ($\mathbf{P}$), or *charge-density bond-order matrix*, is defined as

$$P_{\lambda\sigma} = 2 \sum_a^{N/2} C_{\lambda a} C_{\sigma a}^* \tag{1.41}$$

Here $a$ sums over occupied orbitals. We now have a form of $\mathbf{G}$ that is much simpler

$$G_{\mu\nu} = \sum_{\lambda\sigma} P_{\lambda\sigma} \left[ 2 \langle \mu\sigma | \nu\lambda \rangle - \langle \mu\sigma | \lambda\nu \rangle \right] \tag{1.42}$$

Because the Fock matrix depends on the density matrix

$$\mathbf{F} = \mathbf{F}(\mathbf{P}) \tag{1.43}$$

or, equivalently on the expansion coefficients

$$\mathbf{F} = \mathbf{F}(\mathbf{C}) \tag{1.44}$$

the Roothaan equations are nonlinear,

$$\mathbf{F}(\mathbf{C})\mathbf{C} = \mathbf{S}\mathbf{C}\boldsymbol{\varepsilon} \tag{1.45}$$

and they must be solved in an iterative fashion. Before discussing how to do this, let's discuss the solutions of the matrix equation

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon} \tag{1.46}$$

at each step in the iteration. If $\mathbf{S}$ where the unit matrix, we would have

$$\mathbf{FC} = \mathbf{C}\boldsymbol{\varepsilon} \tag{1.47}$$

and Roothaan's equations would just have the form of the usual matrix eigenvalue problem and we could find the eigenvectors $\mathbf{C}$ and eigenvalues $\boldsymbol{\varepsilon}$ by diagonalizing $\mathbf{F}$. Because the basis is nonorthogonal, we must reformulate the eigvenvalue problem $\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon}$.

## 1.5. Orthogonalization of the Basis

As it turns out, in molecular calculations, the basis sets we use are not actually orthonormal. They are normalized, but are not orthogonal to one another, which will give rise to an overlap matrix. In the Hartree-Fock method, in order to use Roothaan's equations they need to be put into an orthonormal form so the standard matrix eigenvalue problem can be solved. The procedure for doing this is, very unimaginatively, called 'orthogonalization' and proceeds in the following manner. For a set of functions $\{\phi_\mu\}$ that are nonorthogonal,

$$S_{\mu\nu} = \int \varphi_\mu^*(\mathbf{r})\varphi_\nu(\mathbf{r}) \, d\mathbf{r} \tag{1.48}$$

it is always be possible to get a transformation matrix $\mathbf{X}$ such that a transformed set of functions $\{\phi_\mu'\}$ given as

$$\varphi_\mu' = \sum_\nu X_{\nu\mu}\phi_\nu \qquad \mu = 1, 2, \ldots, K \tag{1.49}$$

will form an orthonormal set,

$$\int \varphi_\mu'^*(\mathbf{r})\varphi_\nu'(\mathbf{r}) \, d\mathbf{r} = \delta_{\mu\nu} \tag{1.50}$$

To derive the properties of $X$, we put the transformation (1.49) into (1.50) and get

$$\int \phi_\mu'^*(r)\phi_\nu'(r) \, dr = \int \left[ \sum_\lambda X_{\lambda\mu}^* \phi_\lambda^*(r) \right] \left[ \sum_\sigma X_{\sigma\nu}^* \phi_\sigma(r) \right] \, dr$$

$$= \sum_\lambda \sum_\sigma X_{\lambda\mu}^* \int \phi_\lambda^*(r)\phi_\sigma(r) X_{\sigma\nu} \, dr$$

$$= \sum_\lambda \sum_\sigma X_{\lambda\mu}^* S_{\lambda\sigma} X_{\sigma\nu} = \delta_{\mu\nu} \tag{1.51}$$

This last equation can actually be written as the infamous matrix equation

$$\mathbf{X}^\dagger \mathbf{S} \mathbf{X} = \mathbb{1} \tag{1.52}$$

which defines the relation that the matrix $\mathbf{X}$ must satisfy if the transformed orbitals are to form an orthonormal set. $\mathbf{X}$ must also be nonsingular, that is, it must possess an inverse $\mathbf{X}^{-1}$. We can obtain two different

transformation matrices $\mathbf{X}$, since $\mathbf{S}$ is Hermitian it can be diagonalized by a unitary matrix $\mathbf{U}$

$$\mathbf{U}^\dagger \mathbf{S} \mathbf{U} = \mathbf{s} \tag{1.53}$$

where $\mathbf{s}$ is a diagonal matrix of the eigenvalues of $\mathbf{S}$.

There are two ways of orthogonalizing this basis set $\{\phi_\mu\}$, the first procedure is called *symmetric orthogonalization* which uses the inverse square root of $\mathbf{S}$ for $\mathbf{X}$

$$\mathbf{X} \equiv \mathbf{S}^{-1/2} = \mathbf{U} \mathbf{s}^{-1/2} \mathbf{U}^\dagger \tag{1.54}$$

We can form $\mathbf{S}^{-1/2}$ by diagonalizing $\mathbf{S}$ to form $\mathbf{s}$, then taking the inverse square root fo each of the eigenvalues to form the diagonal matrix $\mathbf{s}^{-1/2}$ and then "undiagonalizing" by the transformation in (1.54). If $\mathbf{S}$ is Hermitian so to is $\mathbf{S}^{-1/2}$, putting (1.54) into (1.52) will give us

$$\mathbf{S}^{-1/2} \mathbf{S} \mathbf{S}^{-1/2} = \mathbf{S}^{-1/2} \mathbf{S}^{1/2} = \mathbf{S}^0 = \mathbb{1} \tag{1.55}$$

shows that indeed $\mathbf{X} = \mathbf{S}^{-1/2}$ is an orthogonalizing transformation matrix. When taking the square root of (1.54) there will be no difficulty as all eigenvalues of $\mathbf{S}$ are positive. However, if there is linear dependence or near linear dependence in your basis set, some of the eigenvalues will approach zero, so (1.54) will involve dividing by very small quantities. This will lead to problems in numerical precision for basis sets with these issues.

## 1.6. Canonical Orthogonalization

So, how does deal with these linear dependences? Simple, orthogonalize things in a different way, namely canonical orthogonalization. To begin we use the transformation matrix

$$\mathbf{X} = \mathbf{U} \mathbf{S}^{-1/2} \tag{1.56}$$

where we divide the columns of the unitary matrix $\mathbf{U}$ by the square root of the corresponding eigenvalue

$$X_{ij} = U_{ij} / s_j^{1/2} \tag{1.57}$$

If we throw this into the matrix equation

$$\mathbf{X}^\dagger \mathbf{S} \mathbf{X} = \mathbb{1} \tag{1.58}$$

we will get

$$\begin{aligned} \mathbf{X}^\dagger \mathbf{S} \mathbf{X} &= (\mathbf{U} \mathbf{s}^{-1/2})^\dagger \mathbf{S} \mathbf{U} \mathbf{s}^{-1/2} \\ &= \mathbf{s}^{-1/2} \mathbf{U}^\dagger \mathbf{S} \mathbf{U} \mathbf{s}^{-1/2} \\ &= \mathbf{s}^{-1/2} \mathbf{s} \mathbf{s}^{-1/2} = \mathbb{1} \end{aligned}$$

and also showing that $\mathbf{X} = \mathbf{U} \mathbf{s}^{-1/2}$ is an orthogonalizing transformation matrix. But if you look at (1.57), you might see that if any of the $s_i$ eigenvalues approach zero, which happens when you have linear dependence in your basis set, it's not going to be good. Your numerical calculations will blow up (figuratively, not literally). But we can bypass this problem with canonical orthogonalization. In the matrix eigenvalue problem provided we order the columns of $\mathbf{U}$ in the same way. Suppose we order the positive eigenvalues $s_i$ in the order $s_1 > s_2 > s_3 \cdots$. We can then screen these eigenvalues and determine which ones are 'undesirable' that is are

too small and cause numerical instability issues. We can then construct a new transformation matrix

$$
\tilde{\mathbf{X}} = \begin{pmatrix} U_{1,1}/s_1^{1/2} & U_{1,2}/s_2^{1/2} & \dots & U_{1,K-m}/s_{K-m}^{1/2} \\ U_{2,1}/s_1^{1/2} & U_{2,2}/s_2^{1/2} & \dots & U_{2,K-m}/s_{K-m}^{1/2} \\ \vdots & \vdots & & \vdots \\ U_{K,1}/s_1^{1/2} & U_{K,2}/s_2^{1/2} & \dots & U_{K,K-m}/s_{K-m}^{1/2} \end{pmatrix} \tag{1.59}
$$

where the undesirable, last $m$ columns of $\mathbf{X}$ are eliminated to give us the nondegenerate $K \times (K - m)$ matrix $\tilde{\mathbf{X}}$. With the new transformation matrix, we get only $K - m$ transformed orthonormal basis functions

$$
\phi'_\mu = \sum_{\nu=1}^{K} \tilde{\mathbf{X}}_{\nu\mu} \phi_\nu \qquad \mu = 1, 2, \dots, K - m \tag{1.60}
$$

These would span exactly the same region of space as the original set, provided the eliminated eigenvalues were exactly zero. However, in practice, one will often find linear dependence issues with eigenvalues around the order of $s_i \leq 10^{-4}$, depending on the machine precision. When we eliminate the columns with these small eigenvalues we are discarding part of the basis set, but only a very small part.

Consider first a new coefficient matrix $\mathbf{C}'$ related to the old coefficient matrix $\mathbf{C}$ by

$$
\mathbf{C}' = \mathbf{X}'\mathbf{C} \qquad \mathbf{C} = \mathbf{X}\mathbf{C}' \tag{1.61}
$$

we assume that $\mathbf{X}$ possess an inverse, which will be the case if we have eliminated linear dependencies. Substituting in $\mathbf{C} = \mathbf{X}\mathbf{C}'$ into the Roothaan equations gives

$$
\mathbf{F}\mathbf{X}\mathbf{C}' = \mathbf{S}\mathbf{X}\mathbf{C}'\boldsymbol{\varepsilon} \tag{1.62}
$$

Multiplying the left by $\mathbf{X}^\dagger$ will give

$$
(\mathbf{X}^\dagger\mathbf{F}\mathbf{X})\mathbf{C}' = (\mathbf{X}^\dagger\mathbf{S}\mathbf{X})\mathbf{C}'\boldsymbol{\varepsilon} \tag{1.63}
$$

Defining a new matrix

$$
\mathbf{F}' = \mathbf{X}^\dagger\mathbf{F}\mathbf{X} \tag{1.64}
$$

then using (1.58), we get

$$
\mathbf{F}'\mathbf{C}' = \mathbf{C}'\boldsymbol{\varepsilon} \tag{1.65}
$$

These happen to be the transformed Roothaan equations, which we can solve for $\mathbf{C}'$ by diagonalizing $\mathbf{F}'$, using equations (1.64), (1.65), and (1.61) to solve the Roothaan equations $\mathbf{F}\mathbf{C} = \mathbf{S}\mathbf{C}\boldsymbol{\varepsilon}$ for $\mathbf{C}$ and $\boldsymbol{\varepsilon}$. The intermediate primed matrices are just the Fock matrix and expansion coefficients in the orthogonalized basis, i.e.

$$
\psi_i = \sum_{\mu=1}^{K} C'_{\mu i}\phi'_\mu \qquad i = 1, 2, \dots, K \tag{1.66}
$$

$$
F'_{\mu\nu} = \int \phi'^*_\mu(1) f(1) \phi'_\nu(1) \, d\mathbf{r}_1 \tag{1.67}
$$

## 1.7. Self-Consistent Field

Let us review what we need to do in order to solve the Roothaan equations using the self-consistent field procedure. In the SCF procedure we need to solve the pseudoeigenvalue matrix equations

$$\sum_\nu F_{\mu\nu} C_{\nu i} = \varepsilon_i \sum_\nu S_{\mu\nu} C_{\nu i}$$

$$\mathbf{FC} = \mathbf{SC\varepsilon}$$

known as the Roothaan equations. These can be solved self-consistently for the orbital coefficient matrix $\mathbf{C}$ and the orbital energy eigenvalues $\varepsilon_i$. The Fock matrix, $\mathbf{F}$, has elements $F_{\mu\nu}$ given (in the atomic orbital basis) as

$$F_{\mu\nu} = H_{\mu\nu} + 2(\mu\nu|\lambda\sigma)\, P_{\lambda\sigma} - (\mu\lambda|\nu\sigma)\, P_{\lambda\sigma}$$

where $P_{\lambda\sigma}$ is an element of the one-particle density matrix $\mathbf{P}$, constructed from the orbital coefficient matrix $\mathbf{C}$

$$P_{\lambda\sigma} = C_{\sigma i} C_{\lambda i}$$

Formally, the orbital coefficient matrix $\mathbf{C}$ is a $N \times M$ matrix, where $N$ is the number of atomic basis functions, and $M$ is the total number of molecular orbitals. Physically, this matrix describes the contribution of every atomic basis function (the columns) to a particular molecular orbital (e.g., the $i^{\text{th}}$ row). The density matrix $\mathbf{P}$ is a square matrix describing the electron density contained in each orbital. In the molecular orbital basis, the density matrix has elements

$$P_{pq} = \begin{cases} 2\delta_{pq} & p \text{ occupied} \\ 0 & p \text{ virtual} \end{cases}$$

The total RHF energy is given by

$$E_{\text{total}}^{\text{RHF}} = E_e^{\text{RHF}} + E_n^{\text{BO}}$$

where $E_e^{\text{RHF}}$ is the final electronic RHF energy, and $E_n^{\text{BO}}$ is the total nuclear repulsion energy within the Born-Oppenheimer approximation. To compute the electronic energy, we may use the density matrix in the AO basis:

$$E_e^{\text{RHF}} = (F_{\mu\nu} + H_{\mu\nu}) P_{\mu\nu},$$

and the nuclear repulsion energy is simply

$$E_n^{\text{BO}} = \sum_{A>B} \frac{Z_A Z_B}{r_{AB}}$$

where $Z_A$ is the nuclear charge of atom $A$, and the sum runs over all unique nuclear pairs.

### 1.7.1 Procedure

The SCF procedure involves solving for the set of orbitals that is a minimum in the entire set of orbitals, so that they are all *consistent* with themselves. To begin we first calculate the overlap matrix $S_{\mu\nu}$, the core Hamiltonian $H_{\mu\nu}^{\text{core}}$, and the two-electron integral $\langle \mu\nu|\lambda\sigma\rangle$. This two-electron integral is notorious in quantum chemistry as it is a four-index transformation (there are 4 'for/do' loops required) and is the most computationally expensive, often requiring a lot of memory. Then we must diagonalize $\mathbf{S} \rightarrow \mathbf{X}$ and take a guess for $\mathbf{P}$. From there we calculate $\mathbf{G}$ from $\mathbf{P}$ and then calculate $\mathbf{F} = \mathbf{H}^{\text{core}} + \mathbf{P}$, but remember we must orthogonalize the basis, to get $\mathbf{F}' = \mathbf{X}^\dagger \mathbf{F} \mathbf{X}$. We can then diagonalize this new $\mathbf{F}'$ to get $\mathbf{C}'$ and $\mathbf{E}$. Then we transform our

basis set back $\mathbf{C} = \mathbf{XC}'$. We form a new $\mathbf{P}$ and then check for *convergence*, that is, we check to see how $\mathbf{P}$ has changed from our initial guess. If the change is significant enough we repeat the process until the new change is below some set threshold. Hence we repeat until the changes in orbitals are consistent with each other.

$$\mathbf{FC} = \mathbf{SC\varepsilon} \qquad (1.68)$$

Here are the steps in a list format

1. Define your molecule so that it has a set of nuclear coordinates $\{\mathbf{R}_A\}$, atomic numbers $\{Z_A\}$, number of electrons $N$ and a basis set $\{\phi_\mu\}$.
2. Calculate all molecular one and two electron integrals $S_{\mu\nu}$, $H_{\mu\nu}^{\text{core}}$, and $\langle\mu\nu|\lambda\sigma\rangle$.
3. Diagonalize the overlap matrix $\mathbf{S}$ and obtain a transformation matrix $\mathbf{X}$ using either symmetric or canoncial orthogonalization.
4. Get a guess for the density matrix $\mathbf{P}$
5. Calculate the matrix $\mathbf{G}$ of equation from the density matrix $\mathbf{P}$ and the two-electron integrals $\langle\mu\nu|\lambda\sigma\rangle$.
6. Add $\mathbf{G}$ to the core-Hamiltonin to get the Fock matrix $\mathbf{F} = \mathbf{H}^{\text{core}} + \mathbf{G}$
7. Calculate the transformed Fock matrix $\mathbf{F}' = \mathbf{X}^\dagger \mathbf{F} \mathbf{X}$
8. Diagonalize $\mathbf{F}'$ to get $\mathbf{C}'$ and $\varepsilon$.
9. Calculate $\mathbf{C} = \mathbf{XC}'$.
10. Form a new density matrix $\mathbf{P}$ from $\mathbf{C}$ using Eq. (1.41) .
11. See if the procedure converged, by determining if the new density matrix of step (10) has changed very little, (within a specified criterion). If no convergence, go back to step (5) with the new density matrix
12. If the procedure converges, then we can use the resulting solution, represented by $\mathbf{C}, \mathbf{P}, \mathbf{F}$, etc, to calculate expectation values and other such quantities.