Functional derivative in Hartree-Fock theory to derive the Roothaan Hall equations

Corey Curran

February 25, 2025

1 Introduction

We aim to demonstrate the functional derivative in Hartree-Fock theory to derive the Roothaan Hall equations. In order to accomplish this, we need to set a foundation of linear algebra and matrix calculus.

2 Mathematical Background

We will start with some useful identities from linear algebra.

2.1 Linear Algebra

Let $A, B \in \mathbb{C}^{m\times n}$

- $\mathrm{Tr}\left(AB^\dagger\right) = \mathrm{Tr}\left(B^\dagger A\right)$

$$\mathrm{Tr}\left(AB^\dagger\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} A_{i,j}B_{i,j} = \sum_{i=1}^{m}\sum_{j=1}^{n} B_{i,j}A_{i,j} = \mathrm{Tr}\left(B^\dagger A\right)$$

- $\mathrm{Tr}\left(AB^\dagger\right) = \mathrm{Tr}\left(\left(AB^\dagger\right)^\dagger\right)^*$

$$\mathrm{Tr}\left(AB^\dagger\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} A_{i,j}B_{i,j} = \sum_{i=1}^{m}\sum_{j=1}^{n} \left(A^\dagger\right)^*_{j,i}\left(B^\dagger\right)^*_{j,i} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} \left(A^\dagger\right)_{j,i}\left(B^\dagger\right)_{j,i}\right)^* = \mathrm{Tr}\left(\left(AB^\dagger\right)^\dagger\right)^*$$

Now let us consider some matrix calculus results.

2.2 Matrix Calculus

Note that we will be using the denominator layout convention, meaning that the dimension of our derivatives will be the same as the dimension of the variable over which we are differentiating (for instance, if we take have some function f where $f(X) = a$ maps an $m \times n$ matrix X to a scalar a and we differentiate over X, our result will be an $m \times n$ matrix). First we will define some basic identities, then we will use them to prove more relevant results.

Once again, let $A, B \in \mathbb{C}^{m\times n}$

- $\nabla_A \mathrm{Tr}\left(AB^\dagger\right) = \nabla_A \mathrm{Tr}\left(B^\dagger A\right) = B$
- $\nabla_{A^\dagger} \mathrm{Tr}\left(AB^\dagger\right) = \nabla_{A^\dagger} \mathrm{Tr}\left(B^\dagger A\right) = B^\dagger$
- Matrix product rule: $\nabla_A[f(A)g(A)] = \nabla_A[f(A)] \cdot g(A) + \nabla_A[g(A)]f(A)$

Combining these rules, we immediately arrive at the following, very useful identity: let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{m\times m}$. Then, we have

$$\begin{aligned}
\nabla_A \mathrm{Tr}\left(A^\dagger BA\right) =& \nabla_{A_1} \mathrm{Tr}\left(A_1^\dagger BA_2\right) + \nabla_{A_2} \mathrm{Tr}\left(A_1^\dagger BA_2\right) && (1)\\
=& BA + \left[A^\dagger B\right]^\dagger && (2)\\
=& BA + B^\dagger A && (3)
\end{aligned}$$

Further generalizing this, if we have $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{m\times m}$, and $C \in \mathbb{C}^{n\times n}$, then we have

$$\begin{aligned}
\nabla_A \mathrm{Tr}\left(CA^\dagger BA\right) =& \nabla_{A_1} \mathrm{Tr}\left(CA_1^\dagger BA_2\right) + \nabla_{A_2} \mathrm{Tr}\left(CA_1^\dagger BA_2\right) && (4)\\
=& \nabla_{A_1} \mathrm{Tr}\left(A_1^\dagger BA_2C\right) + \nabla_{A_2} \mathrm{Tr}\left(CA_1^\dagger BA_2\right) && (5)\\
=& BAC + \left[CA^\dagger B\right]^\dagger && (6)\\
=& BAC + B^\dagger AC^\dagger && (7)
\end{aligned}$$

Lastly, consider once again $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{m\times m}$. Then, we have

$$
\begin{aligned}
\nabla_A \mathrm{Tr}\left(AA^\dagger BAA^\dagger\right) =& \nabla_{A_1}\mathrm{Tr}\left(A_1A_2^\dagger BA_3A_4^\dagger\right) + \nabla_{A_2}\mathrm{Tr}\left(A_1A_2^\dagger BA_3A_4^\dagger\right) && (8)\\
&+ \nabla_{A_3}\mathrm{Tr}\left(A_1A_2^\dagger BA_3A_4^\dagger\right) + \nabla_{A_4}\mathrm{Tr}\left(A_1A_2^\dagger BA_3A_4^\dagger\right) && (9)\\
=& \left[A_2^\dagger BA_3A_4^\dagger\right]^\dagger + \nabla_{A_2}\mathrm{Tr}\left(A_2^\dagger BA_3A_4^\dagger A_1\right) + \nabla_{A_3}\mathrm{Tr}\left(A_4^\dagger A_1A_2^\dagger BA_3\right) + A_1A_2^\dagger BA_3 && (10)\\
=& AA^\dagger B^\dagger A + BAA^\dagger A + B^\dagger AA^\dagger A + AA^\dagger BA && (11)\\
=& AA^\dagger B^\dagger A + \left(AA^\dagger B^\dagger\right)^\dagger A + \left(AA^\dagger B\right)^\dagger A + AA^\dagger BA && (12)
\end{aligned}
$$

3 Hartree-Fock Relevance

Now that we have derived these identities, we can consider how we may be able to apply them within the framework of Hartree-Fock theory. In particular, consider the Lagrangian associated with the Hartree-Fock method viewed as a constrained minimization problem $\mathcal{L}[\mathbf{C}, \Lambda] = \mathcal{E}_{HF}(\mathbf{C}) - \mathrm{Tr}\left[\Lambda\left(\mathbf{I} - \mathbf{C}^\dagger\mathbf{S}\mathbf{C}\right)\right]$ (equation 154 in the notes). In order to solve this minimization, we need to differentiate this Lagrangian over C, the result of which can be seen in equation 155 in the notes. Our task now, then, is to verify the provided formula by applying the methods established above. Specifically, we aim to verify that

$$
\nabla_{\mathbf{C}}\mathcal{L}[\mathbf{C}, \Lambda] = 2\mathbf{h}\mathbf{C} + \frac{1}{2}4\mathbf{J}(\mathbf{C}\mathbf{C}^\dagger)\mathbf{C} - \frac{1}{2}4\mathbf{K}(\mathbf{C}\mathbf{C}^\dagger)\mathbf{C} + 2\mathbf{S}\mathbf{C}\Lambda
$$

To complete this, let's recall some relevant definitions:

- $\mathbf{S}_{i,j} = \langle\chi_i|\chi_j\rangle$
- $\mathcal{E}_{HF}(\mathbf{C}) = \mathrm{Tr}\left(\mathbf{C}^\dagger\mathbf{h}\mathbf{C}\right) + \frac{1}{2}\mathrm{Tr}\left(\mathbf{J}\left(\mathbf{C}\mathbf{C}^\dagger\right)\mathbf{C}\mathbf{C}^\dagger\right) - \frac{1}{2}\mathrm{Tr}\left(\mathbf{K}\left(\mathbf{C}\mathbf{C}^\dagger\right)\mathbf{C}\mathbf{C}^\dagger\right)$
- $\mathbf{h} = \frac{1}{2}\mathbf{T} - \mathbf{V}$
- $\mathbf{T}_{j,k} = \frac{1}{2}\langle\nabla\phi_j|\nabla\phi_k\rangle$
- $\mathbf{V}_{j,k} = \langle\phi_j|V|\phi_k\rangle$
- $\phi_i = \sum_{j=1}^{K}\mathbf{C}_{j,i}\chi_j$
- $\mathbf{J}\left(\mathbf{C}\mathbf{C}^\dagger\right)_{k,m} = \sum_{l,n=1}^{K}\left(\mathbf{C}\mathbf{C}^\dagger\right)_{n,l}\mathbf{V}_{k,l,m,n}$

- $\mathbf{K}\left(\mathbf{C}\mathbf{C}^{\dagger}\right)_{k,n} = \sum_{l,m=1}^{K} \left(\mathbf{C}\mathbf{C}^{\dagger}\right)_{m,l} \mathbf{V}_{k,l,m,n}$
- $\mathbf{V}_{i,j,k,l} = \langle \chi_i \chi_j \| \chi_k \chi_l \rangle$

Ok, now we are ready to consider the Lagrangian.

4 Procedure & Results

We will start by simplifying the formulation, then we will apply the differential operator. We have

$$\begin{aligned}\mathcal{L}[\mathbf{C}, \Lambda] =& \mathcal{E}_{HF}(\mathbf{C}) - \mathrm{Tr}\left[\Lambda\left(\mathbf{I} - \mathbf{C}^{\dagger}\mathbf{S}\mathbf{C}\right)\right] \\ =& \mathcal{E}_{HF}(\mathbf{C}) - \mathrm{Tr}\left(\Lambda\mathbf{I} - \Lambda\mathbf{C}^{\dagger}\mathbf{S}\mathbf{C}\right) \\ =& \mathcal{E}_{HF}(\mathbf{C}) - \mathrm{Tr}(\Lambda\mathbf{I}) + \mathrm{Tr}\left(\Lambda\mathbf{C}^{\dagger}\mathbf{S}\mathbf{C}\right)\end{aligned}$$

Now we apply the differential operator. Note that $\mathrm{Tr}(\Lambda\mathbf{I})$ does not depend on $\mathbf{C}$, so it's derivative will be $\mathbf{0}$. Thus, we have

$$\begin{aligned}\nabla_{\mathbf{C}}\mathcal{L}[\mathbf{C}, \Lambda] =& \nabla_{\mathbf{C}}\mathcal{E}_{HF}(\mathbf{C}) + \nabla_{\mathbf{C}}\mathrm{Tr}\left(\Lambda\mathbf{C}^{\dagger}\mathbf{S}\mathbf{C}\right) \\ =& \nabla_{\mathbf{C}}\mathrm{Tr}\left(\mathbf{C}^{\dagger}\mathbf{h}\mathbf{C}\right) + \frac{1}{2}\nabla_{\mathbf{C}}\mathrm{Tr}\left(\mathbf{J}\left(\mathbf{C}\mathbf{C}^{\dagger}\right)\mathbf{C}\mathbf{C}^{\dagger}\right) - \frac{1}{2}\nabla_{\mathbf{C}}\mathrm{Tr}\left(\mathbf{K}\left(\mathbf{C}\mathbf{C}^{\dagger}\right)\mathbf{C}\mathbf{C}^{\dagger}\right) + \nabla_{\mathbf{C}}\mathrm{Tr}\left(\Lambda\mathbf{C}^{\dagger}\mathbf{S}\mathbf{C}\right)\end{aligned}$$

Let's consider this term by term.

First, we have $\nabla_{\mathbf{C}}\mathrm{Tr}\left(\mathbf{C}^{\dagger}\mathbf{h}\mathbf{C}\right)$. This is of the same form as (1). Thus, we know that $\nabla_{\mathbf{C}}\mathrm{Tr}\left(\mathbf{C}^{\dagger}\mathbf{h}\mathbf{C}\right) = \mathbf{h}\mathbf{C} + \mathbf{h}^{\dagger}\mathbf{C}$. Due to the hermiticity of $\mathbf{h}$, we then have that $\nabla_{\mathbf{C}}\mathrm{Tr}\left(\mathbf{C}^{\dagger}\mathbf{h}\mathbf{C}\right) = 2\mathbf{h}\mathbf{C}$.

Next, let's consider $\nabla_{\mathbf{C}}\mathrm{Tr}\left(\Lambda\mathbf{C}^{\dagger}\mathbf{S}\mathbf{C}\right)$. This is of the same form as (4). Thus, we know that $\nabla_{\mathbf{C}}\mathrm{Tr}\left(\Lambda\mathbf{C}^{\dagger}\mathbf{S}\mathbf{C}\right) = \mathbf{S}\mathbf{C}\Lambda + \mathbf{S}^{\dagger}\mathbf{C}\Lambda^{\dagger}$. Due to the hermiticity of $\mathbf{S}$ and Λ, we then have that $\nabla_{\mathbf{C}}\mathrm{Tr}\left(\Lambda\mathbf{C}^{\dagger}\mathbf{S}\mathbf{C}\right) = 2\mathbf{S}\mathbf{C}\Lambda$

Now let's consider $\nabla_{\mathbf{C}}\mathrm{Tr}\left(\mathbf{J}\left(\mathbf{C}\mathbf{C}^{\dagger}\right)\mathbf{C}\mathbf{C}^{\dagger}\right)$. Based on the definition of $\mathbf{J}$, if sufficient care is taken of the relevant indices, this is functionally equivalent to $\nabla_{\mathbf{C}}\mathrm{Tr}\left(\mathbf{C}\mathbf{C}^{\dagger}\mathbf{V}\mathbf{C}\mathbf{C}^{\dagger}\right)$, which has the same form as (8). Thus, we know that, again subject to careful indexing, $\nabla_{\mathbf{C}}\mathrm{Tr}\left(\mathbf{J}\left(\mathbf{C}\mathbf{C}^{\dagger}\right)\mathbf{C}\mathbf{C}^{\dagger}\right) = \mathbf{C}\mathbf{C}^{\dagger}\mathbf{V}^{\dagger}\mathbf{C} + \left(\mathbf{C}\mathbf{C}^{\dagger}\mathbf{V}^{\dagger}\right)^{\dagger}\mathbf{C} + \left(\mathbf{C}\mathbf{C}^{\dagger}\mathbf{V}\right)^{\dagger}\mathbf{C} + \mathbf{C}\mathbf{C}^{\dagger}\mathbf{V}\mathbf{C}$. Due to the hermiticity of $\mathbf{V}$, we then have that $\mathbf{C}\mathbf{C}^{\dagger}\mathbf{V} = \mathbf{C}\mathbf{C}^{\dagger}\mathbf{V}^{\dagger} = \mathbf{J}\left(\mathbf{C}\mathbf{C}^{\dagger}\right)$. Additionally, due to the hermiticity of $\mathbf{J}$, we have that $\nabla_{\mathbf{C}}\mathrm{Tr}\left(\mathbf{J}\left(\mathbf{CC}^\dagger\right)\mathbf{CC}^\dagger\right) = 4\mathbf{J}\left(\mathbf{CC}^\dagger\right)\mathbf{C}$.

Finally, let's consider $\nabla_{\mathbf{C}}\mathrm{Tr}\left(\mathbf{K}\left(\mathbf{CC}^\dagger\right)\mathbf{CC}^\dagger\right)$. Similarly to $\mathbf{J}$, based on the definition of $\mathbf{K}$, if sufficient care is taken of the relevant indices, this is functionally equivalent to $\nabla_{\mathbf{C}}\mathrm{Tr}\left(\mathbf{CC}^\dagger\mathbf{VCC}^\dagger\right)$, which has the same form as (8). Thus, we know that, again subject to careful indexing, $\nabla_{\mathbf{C}}\mathrm{Tr}\left(\mathbf{K}\left(\mathbf{CC}^\dagger\right)\mathbf{CC}^\dagger\right) = \mathbf{CC}^\dagger\mathbf{V}^\dagger\mathbf{C} + \left(\mathbf{CC}^\dagger\mathbf{V}^\dagger\right)^\dagger\mathbf{C} + \left(\mathbf{CC}^\dagger\mathbf{V}\right)^\dagger\mathbf{C} + \mathbf{CC}^\dagger\mathbf{VC}$. Due to the hermiticity of $\mathbf{V}$, we then have that $\mathbf{CC}^\dagger\mathbf{V} = \mathbf{CC}^\dagger\mathbf{V}^\dagger = \mathbf{K}\left(\mathbf{CC}^\dagger\right)$. Additionally, due to the hermiticity of $\mathbf{K}$, we have that $\nabla_{\mathbf{C}}\mathrm{Tr}\left(\mathbf{K}\left(\mathbf{CC}^\dagger\right)\mathbf{CC}^\dagger\right) = 4\mathbf{K}\left(\mathbf{CC}^\dagger\right)\mathbf{C}$.

Combining all of these pieces, we finally have that

$$\nabla_{\mathbf{C}}\mathcal{L}[\mathbf{C}, \Lambda] = 2\mathbf{hC} + \frac{1}{2}4\mathbf{J}\left(\mathbf{CC}^\dagger\right)\mathbf{C} - \frac{1}{2}4\mathbf{K}\left(\mathbf{CC}^\dagger\right)\mathbf{C} + 2\mathbf{SC}\Lambda$$

as desired. Setting this expression equal to 0, as one must do to perform the desired optimization, yields the Roothaan Hall generalized eigenvalue equations.