

An Iterative Technique for Solving the N -electron Hamiltonian: The Hartree-Fock method

Tony Hyun Kim

Abstract

The problem of electron motion in an arbitrary field of nuclei is an important quantum mechanical problem finding applications in many diverse fields. From the variational principle we derive a procedure, called the Hartree-Fock (HF) approximation, to obtain the many-particle wavefunction describing such a system. Here, the central physical concept is that of electron indistinguishability: while the antisymmetry requirement greatly complexifies our task, it also offers a symmetry that we can exploit. After obtaining the HF equations, we then formulate the procedure in a way suited for practical implementation on a computer by introducing a set of spatial basis functions. An example implementation is provided, allowing for calculations on the simplest heteronuclear structure: the helium hydride ion. We conclude with a discussion of how to derive useful chemical information from the HF solution.

1 Introduction

In this paper I will introduce the theory and the techniques to solve for the approximate ground state of the *electronic Hamiltonian*:

$$H = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{Ai}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} \quad (1)$$

which describes the motion of N electrons (indexed by i and j) in the field of M nuclei (indexed by A). The coordinate system, as well as the fully-expanded Hamiltonian, is explicitly illustrated for $N = M = 2$ in Figure 1. Although we perform the analysis

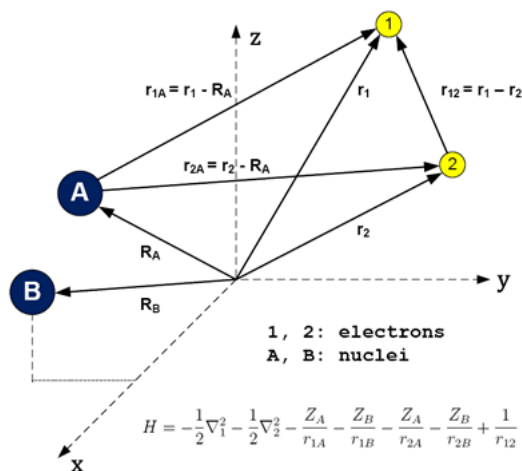


Figure 1: The coordinate system corresponding to Eq. 1 for $N = M = 2$.

in dimensionless form, the physical lengths and energies can be readily obtained by multiplying by the scale factors $a_0 = 5.3 \times 10^{-11} \text{m}$ and $\epsilon_a = 27.21 \text{eV}$ respectively.

The above Hamiltonian and the system it represents are of profound importance for many areas of study even beyond physics. In particular, it is the theoretical basis for quantum chemistry, a field in which one derives from first principles a wealth of chemical facts, including quantities of practical modern interest such as structures of molecules and mechanisms of reactions. All of this is possible without relying on empirical parameters beyond those that comprise a_0 and ϵ_a . Accordingly, this paper culminates with a computer implementation to solve for the ground state of the helium hydride ion. I then report the computed equilibrium HeH^+ bond length, the vibrational frequency, and the three-dimensional electrostatic potential map, all of which are useful parameters in characterizing a substance.

Our method, called the *Hartree-Fock (HF) approximation* or the *self-consistent field (SCF)*, iteratively treats each of the electrons of the N -particle wavefunction in sequence, on a one-by-one basis. We begin exploring such an iteration scheme by analyzing the simplest two-electron system, the He atom.

1.1 Hartree iteration on the helium atom

It is well-known that the difficulty in obtaining the two-electron wavefunction $\Psi(\vec{r}_1, \vec{r}_2)$ from the helium Hamiltonian originates from the Coulombic interaction term between

electrons 1 and 2. In fact, when one simply ignores this interaction, the two-particle Schrödinger equation is satisfied by a product wavefunction, $\Psi(\vec{r}_1, \vec{r}_2) = \psi_a(\vec{r}_1)\psi_b(\vec{r}_2)$, where ψ_i are the scaled single particle states obtained from the hydrogen atom.

With this background, we then follow the suggestion of the numerical analyst D.R. Hartree, who in the 1920's proposed that the many-electron wavefunction can still be expressed as a product of two single-particle states, even in the presence of mutual repulsion. In other words, we will now simply take for granted the functional form $\Psi(\vec{r}_1, \vec{r}_2) = \psi_a(\vec{r}_1)\psi_b(\vec{r}_2)$, but without yet specifying $\psi_{a,b}$.

Given this trial wavefunction, Hartree's iterative method describes how to determine the single-particle states $\psi_{a,b}$. Again, "iteration" implies that the algorithm analyzes each electron one at a time.

As a starting point, consider the doubly-charged helium nucleus stripped of both electrons. In assigning a wavefunction to the first electron around this nucleus, the relevant potential is clearly $V_{nuc} = 2/r_1$, for which we have available an analytic form of the ground-state eigenfunction; call this $\psi_a^{(0)}$. The potential and the initial wavefunction of electron 1 are shown in Figure 2(a).

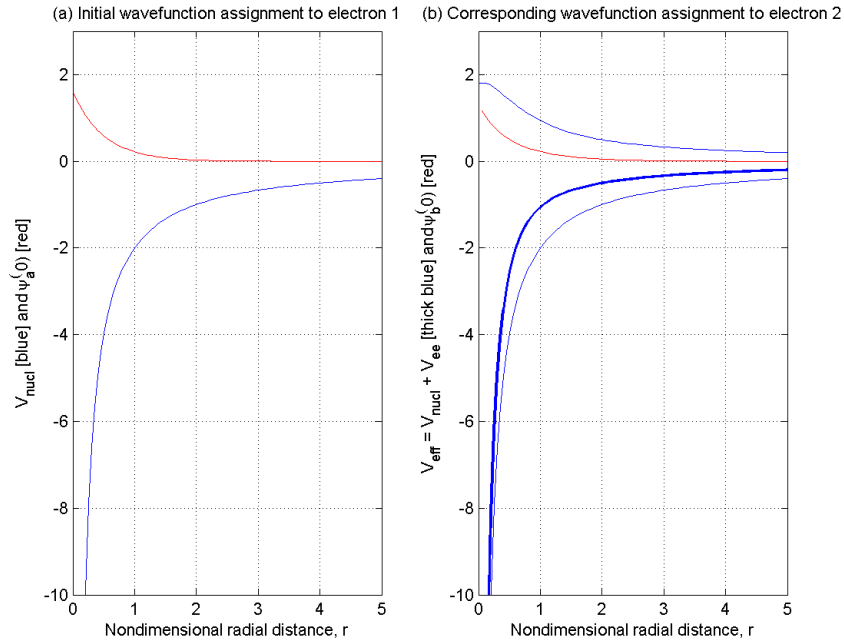


Figure 2: Individual wavefunctions and potentials of each electron.

We then move on to electron 2. Recalling that $|\psi_a(\vec{r}_1)|^2$ represents the spatial

probability distribution of electron 1, it is plausible to associate a potential V_{ee} due to the corresponding charge density, $\rho_a = e|\psi_a(\vec{r}_1)|^2$. Therefore, the Schrödinger equation for electron 2 will involve the electron-electron potential V_{ee} in addition to V_{nucl} . In the dimensionless units we are working in, the fundamental charge e is not necessary to convert between spatial and charge distributions. The electrostatic potential energy of the electron-electron interaction can then be obtained as:

$$V_{ee}(\vec{r}_2; \psi_a) = \int d^3r_1 \frac{1}{r_{12}} \rho_a(\vec{r}_1) = \int d^3r_1 \frac{1}{r_{12}} |\psi_a(\vec{r}_1)|^2 \quad (2)$$

where $r_{12} = |\vec{r}_1 - \vec{r}_2|$. The notation for V_{ee} explicitly indicates its dependence on ψ_a . The numerical solution for electron 2's wavefunction ψ_b , in the effective potential $V_{ee} + V_{nucl}$, is shown in Figure 2(b).

Returning to electron 1, it is now possible to utilize ψ_b to similarly calculate $V_{ee}(\vec{r}_1; \psi_b)$. The new SE is solved, yielding $\psi_a^{(1)}(\vec{r}_1)$. The process is repeated, alternating between the two electrons, until the sequences of functions $\{\psi_a^{(i)}\}$ and $\{\psi_b^{(i)}\}$ converge within some desired precision. Formally, we are solving the equations:

$$\left[-\frac{1}{2}\nabla_1^2 + V_{nucl}(\vec{r}_1) + V_{ee}(\vec{r}_1; \psi_b) \right] \psi_a(\vec{r}_1) = E_a \psi_a(\vec{r}_1) \quad (3)$$

$$\left[-\frac{1}{2}\nabla_2^2 + V_{nucl}(\vec{r}_2) + V_{ee}(\vec{r}_2; \psi_a) \right] \psi_b(\vec{r}_2) = E_b \psi_b(\vec{r}_2) \quad (4)$$

in an iterative fashion. Note however, that it is not *a priori* necessary that Eqs. 3 and 4 are solved in this way. On the other hand, due to the coupling through V_{ee} , a direct solution to the above set of nonlinear Schrödinger equations is prohibitively difficult. Hence, the value of the Hartree approach is that it reduces the many-body Hamiltonian into several single-particle Hamiltonians, albeit with coupling difficulties. In virtually every implementation, the resulting equations are solved by fixed point iteration in the way we have already illustrated.

As with any iterative scheme, convergence and stability of solution are important concerns. However, in this paper I shy away from such issues, blissfully expecting all of my computations to converge without problem. Furthermore, recall that eigenvalue problems generally admit multiple solutions with different energy eigenvalues. Since we are interested in the ground state of the many-particle wavefunction, we seek (or assume we have obtained) the set of functions (ψ_a, ψ_b) that satisfies Eqs. 3 and 4 with the lowest associated energies. Having noted these complications, Hartree's method then gives an algorithm for producing multi-electron wavefunctions of the

form $\Psi(\vec{r}_1, \vec{r}_2) = \psi_a(\vec{r}_1)\psi_b(\vec{r}_2)$, valid for Hamiltonians involving electron-electron repulsion as in Eq. 1.

1.2 Shortcomings of the above iteration scheme

Thus far, I have given an uncritical exposition of the *Hartree method*. I now outline three of its inadequacies. The remainder of the paper is organized around each of their resolution.

1. The Hartree method ignores the antisymmetry requirement for the many electron wavefunction. In Section 2, we will rectify this by introducing the Hartree-Fock (HF) approximation, which takes as its trial wavefunction Ψ a Slater determinant.
2. Beyond heuristic arguments, it is not clear how the conversion from the many-body Hamiltonian to several single-particle Hamiltonians can be rigorously justified from the principles of QM. We show in Section 3 that the HF approximation is an advanced application of the variational principle.
3. In general, coordinate-space integration of the Schrödinger equation is difficult. Since we seek bound states for the electrons, there will be some R such that when $|\vec{r}| > R$, the electron is in a classically forbidden region, characterized by $E - V(\vec{r}) < 0$. This means that we are attempting to integrate an ODE over a region where one of the modes is (roughly) a growing exponential $Ae^{r/L}$. Since we seek bounded solutions, our conditions must be such that the coefficient A is precisely zero. Achieving this feat involves a careful choice of integration bounds and fine-tuning of other numerical knobs. For practical purposes, we seek a more robust approach that is applicable to a wider range of inputs. In Section 4, a mathematical technique is introduced that converts the integro-differential equations of HF into a matrix equation, which are more tractable, and are actually how Hamiltonian solvers are implemented in practice.

To summarize, our immediate task is then to derive from the N -particle Hamiltonian (Eq. 1) N nonlinear single-particle Schrödinger equations, analogous to Eqs. 3 and 4, that take into account the proper Fermi statistics.

2 Slater determinant wavefunctions

The defining property of the Hartree-Fock (HF) approximation, as an improvement on the Hartree method, is that the trial wavefunction Ψ is chosen to be a Slater determinant of mutually orthonormal single-particle states. Of course, the motivation arises from the fact that the mathematical properties of a determinant trivially satisfy the antisymmetry requirement. However, having increased the number of terms in Ψ from 1 to $N!$, in general we will have to pay an immense computational price for this modification!

Also, recall that an electron has a spin degree of freedom in addition to its spatial coordinates. In fact, the antisymmetry requirement applies to an exchange of both spatial and spin coordinates, whereas we dealt solely with space in our previous discussion of the Hartree iteration. Hence, we must now augment our previous notation to explicitly incorporate spin.

We use \vec{x}_i to denote the complete set of coordinates associated with the i -th electron, comprised of the spatial \vec{r}_i and spin $w_i = \pm\frac{1}{2}$ parts. Throughout this paper, the single particle state will be expressed in various forms as deemed convenient:

$$|m\rangle = \chi_m(\vec{x}_i) = \chi_m(i) = \psi_m(\vec{r}_i) \otimes |w_i\rangle = \psi_m(\vec{r}_i) |w_i\rangle \quad (5)$$

The first form is useful when we wish to emphasize the state, rather than the electron index. In the last two expressions, the complete state $\chi_m(\vec{x}_i)$ (“spin orbital” in chemistry parlance) is separated into its spatial $\psi_m(\vec{r}_i)$ and spin $|w_i\rangle$ parts.

With this convention, the trial determinantal wavefunction may be written:

$$\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\vec{x}_1) & \chi_2(\vec{x}_1) & \dots & \chi_N(\vec{x}_1) \\ \chi_1(\vec{x}_2) & \chi_2(\vec{x}_2) & \dots & \chi_N(\vec{x}_2) \\ \vdots & \vdots & & \vdots \\ \chi_1(\vec{x}_N) & \chi_2(\vec{x}_N) & \dots & \chi_N(\vec{x}_N) \end{vmatrix} \quad (6)$$

$$= \frac{1}{\sqrt{N!}} \sum_{n=1}^{N!} (-1)^{p_n} P_n \{ \chi_1(1) \chi_2(2) \dots \chi_N(N) \} \quad (7)$$

In Eq. 7, the index n runs over all $N!$ permutations of the N single particle states. The quantity p_n takes on 0 or 1 depending on whether the permutation P_n is even or odd, respectively. An even (odd) permutation is one that can be formed by an even (odd) number of exchanges of two elements. (Such minimal exchanges are called transpositions.) We may regard the action of P_n as permuting the electron indices,

so that if P_2 is the transposition of 1 and 2, then

$$P_2 \{\chi_1(1)\chi_2(2)\chi_3(3)\dots\chi_N(N)\} = \chi_1(2)\chi_2(1)\chi_3(3)\dots\chi_N(N)$$

and $(-1)^{p_2} = -1$ by definition.

2.1 Matrix elements involving determinantal wavefunctions

The Slater determinant state is fundamental to the HF theory. Therefore, in order to proceed, we will need ability in calculating matrix elements involving such wavefunctions. For the purposes of Hartree-Fock, it will suffice to focus on $\langle\Psi|H|\Psi\rangle$, where H is the electronic Hamiltonian given in Eq. 1.

Our task in computing $\langle\Psi|H|\Psi\rangle$ is made simpler by recognizing the “one- and two-electron” structure of the electronic Hamiltonian, and by using the *indistinguishability* of electrons to take advantage of that structure. Begin by writing Eq. 1 as

$$\begin{aligned} H &= \sum_{i=1}^N \left(-\frac{1}{2}\nabla_i^2 + \sum_{A=1}^M \frac{Z_A}{r_{iA}} \right) + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} \\ &= \sum_{i=1}^N h_1(i) + \sum_{i=1}^N \sum_{j>i}^N h_2(i, j) \end{aligned} \quad (8)$$

Here, we have identified the *one-electron operator* $h_1(i) = -\frac{1}{2}\nabla_i^2 + \sum_{A=1}^M Z_A/r_{iA}$ (of the i -th electron) and the *two-electron operator* $h_2(i, j) = r_{ij}^{-1}$ (involving electrons i and j). The terminology corresponds to the fact that only one and two sets of electron coordinates are involved in the matrix elements of $h_1(i)$ and $h_2(i, j)$ respectively. In particular, $h_1(i)$ is also termed the *core-Hamiltonian* of the i -th electron, describing its kinetic and potential energy in the field of the nuclei.

2.1.1 One-electron integrals

We now show explicitly that $\langle\Psi|h_1(i)|\Psi\rangle$ reduces to an integral over a single electron coordinate. Consider $i = 1$. Expanding Ψ as given in Eq. 7,

$$\begin{aligned} \langle\Psi|h_1(1)|\Psi\rangle &= \frac{1}{N!} \sum_i^{N!} \sum_j^{N!} (-1)^{p_i} (-1)^{p_j} \int d\vec{x}_1 d\vec{x}_2 \dots d\vec{x}_N \\ &\times P_i \{\chi_1^*(1)\chi_2^*(2)\dots\chi_N^*(N)\} h_1(1) P_j \{\chi_1(1)\chi_2(2)\dots\chi_N(N)\} \end{aligned}$$

Since the single particle states are chosen to be orthonormal, the above expression is zero unless electrons 2, 3, ..., N occupy the same spin orbitals in the i -th permutation

as in the j -th permutation. (Recall that $h_1(1)$ depends only on the coordinates of electron 1. The other electrons “go right through” $h_1(1)$.) This condition is equivalent to the two permutations being identical. In such a case, the sums over i and j can be accounted for by a single summation, and $(-1)^{p_i}(-1)^{p_j} = (-1)^{p_i}(-1)^{p_i} = 1$, yielding:

$$\begin{aligned} \langle \Psi | h_1(1) | \Psi \rangle &= \frac{1}{N!} \sum_i^{N!} \int d\vec{x}_1 d\vec{x}_2 \dots d\vec{x}_N \\ &\times P_i \{ \chi_1^*(1) \chi_2^*(2) \dots \chi_N^*(N) \} h_1(1) P_i \{ \chi_1(1) \chi_2(2) \dots \chi_N(N) \} \end{aligned}$$

Now note that in the sum over the $N!$ permutations, electron 1 occupies each spin orbital χ_m , $(N-1)!$ times. This follows since there are $(N-1)!$ ways to arrange electrons 2, 3, ..., N , after having fixed the orbital of electron 1. In conclusion, $\langle \Psi | h_1(1) | \Psi \rangle$ can be expressed as:

$$\begin{aligned} \langle \Psi | h_1(1) | \Psi \rangle &= \frac{(N-1)!}{N!} \sum_{m=1}^N \int d\vec{x}_1 \chi_m^*(\vec{x}_1) h_1(1) \chi_m(\vec{x}_1) \\ &= \frac{1}{N} \sum_{m=1}^N \langle m | h_1(1) | m \rangle \end{aligned} \quad (9)$$

where the sum is over the single-particle states, rather than the electron index.

So, note that the core-energy of electron 1 is an average of the expected core-energy of every single-particle state that comprise the determinant. This is a direct consequence of the indistinguishability of electrons: because we have applied the proper statistics to describe the many-particle wavefunction (i.e. a Slater determinant), it does not make sense to assign an electron into a distinguishable combination of the single-particle states. Instead, every electron must occupy every single-particle state in an exactly identical way! (And hence Eq. 9.)

Given the indistinguishability of the electrons, it is then clear that $h_1(i) = h_1(j)$ for every i, j . We can thus conclude:

$$\langle \Psi | \sum_{i=1}^N h_1(i) | \Psi \rangle = \sum_{m=1}^N \langle m | h_1(1) | m \rangle \quad (10)$$

Conventionally, the integration variable of the one-electron integral is taken to be \vec{x}_1 .

2.1.2 Two-electron integrals

This time, we exploit indistinguishability from the beginning, and write

$$\langle \Psi | \sum_{i=1}^N \sum_{j>i}^N h_2(i, j) | \Psi \rangle = \binom{N}{2} \langle \Psi | h_2(1, 2) | \Psi \rangle = \frac{N(N-1)}{2} \langle \Psi | h_2(1, 2) | \Psi \rangle \quad (11)$$

This is valid since any pair of electrons will have identical $\langle \Psi | h_2(i, j) | \Psi \rangle$ according to indistinguishability. Furthermore, the double sum accounts for all of the unique pairs among N electrons, of which there are $N(N-1)/2$.

Proceeding as before, we obtain:

$$\begin{aligned} \frac{N(N-1)}{2} \langle \Psi | h_2(1, 2) | \Psi \rangle &= \frac{1}{2(N-2)!} \sum_i^{N!} \sum_j^{N!} (-1)^{p_i} (-1)^{p_j} \int d\vec{x}_1 d\vec{x}_2 \dots d\vec{x}_N \\ &\times P_i \{ \chi_1^*(1) \chi_2^*(2) \dots \chi_N^*(N) \} h_2(1, 2) P_j \{ \chi_1(1) \chi_2(2) \dots \chi_N(N) \} \end{aligned}$$

However, unlike the one-electron case, the orthogonality of single-particle states only stipulates that electrons 3, 4, \dots , N be assigned the same spin orbital by permutations P_i and P_j . With fixed P_i , there are actually two possible choices for P_j that satisfy this constraint: P_j can either be identical to P_i , or be the composition of P_i and the transposition of electrons 1 and 2 (which we have previously called P_2). In the latter case, note that $(-1)^{p_i} (-1)^{p_j} = (-1)^{p_i} (-1)^{p_i+p_2} = -1$.

In the sums over $N!$ permutations, electrons 1 and 2 will occupy any two different spin orbitals χ_m and χ_n , $(N-2)!$ times. (For each pair, there are $(N-2)!$ ways to permute the other $N-2$ electrons among the $N-2$ remaining states.) Hence,

$$\begin{aligned} \frac{N(N-1)}{2} \langle \Psi | h_2(1, 2) | \Psi \rangle &= \frac{1}{2(N-2)!} (N-2)! \sum_m^N \sum_{n \neq m}^N \int d\vec{x}_1 d\vec{x}_2 \\ &\times \chi_m^*(1) \chi_n^*(2) h_2(1, 2) [\chi_m(1) \chi_n(2) - \chi_m(2) \chi_n(1)] \\ &= \frac{1}{2} \sum_{m=1}^N \sum_{n \neq m}^N \langle mn | h_2(1, 2) | mn \rangle - \langle mn | h_2(1, 2) | nm \rangle \end{aligned}$$

As a final modification, note that $\langle mn | h_2(1, 2) | mn \rangle - \langle mn | h_2(1, 2) | nm \rangle$ vanishes when $n = m$, so we can eliminate the restriction on the inner sum to conclude:

$$\frac{N(N-1)}{2} \langle \Psi | h_2(1, 2) | \Psi \rangle = \frac{1}{2} \sum_{m=1}^N \sum_{n=1}^N \langle mn | h_2(1, 2) | mn \rangle - \langle mn | h_2(1, 2) | nm \rangle \quad (12)$$

Referring back to Eq. 8, we have then finished the task of computing $\langle \Psi | H | \Psi \rangle$:

$$\begin{aligned} \langle \Psi | H | \Psi \rangle &= \langle \Psi | \sum_{i=1}^N h_1 | \Psi \rangle + \langle \Psi | \sum_{i=1}^N \sum_{j>i}^N h_2 | \Psi \rangle \\ &= \sum_{m=1}^N \langle m | h_1 | m \rangle + \frac{1}{2} \sum_{m=1}^N \sum_{n=1}^N \langle mn | h_2 | mn \rangle - \langle mn | h_2 | nm \rangle \quad (13) \end{aligned}$$

Due to indistinguishability, we are able to suppress the electron coordinate indices in the operators h_1 and h_2 without ambiguity.

3 Derivation of HF from the Variational Principle

In this section we derive the Hartree-Fock equations, analogous to Eqs. 3 and 4, that are applicable to the Slater-determinant trial function. The HF equations will be a set of coupled, nonlinear Schrödinger equations for the each of the single-particle states. The arguments presented here originate from John Slater (he of the determinant!) who first performed the following analysis on the Hartree iteration.

We begin by reminding the reader of the *variational principle*. For any choice of Ψ , the energy expectation $\langle \Psi | H | \Psi \rangle$ represents an upper bound to the true ground state energy of H . Having constrained Ψ to be an element of some set of functions, the best approximation to the ground state is obtained by minimization of the energy expectation over that set.

In the Hartree-Fock theory, the many-particle wavefunction Ψ is constrained to remain a Slater determinant formed by mutually orthonormal single-particle states $\{|m\rangle \mid m = 1, 2, \dots, N\}$. However, as in the original Hartree procedure, the single particle states are not yet identified, and therein lie our variational degrees of freedom. More precisely, we view the energy expectation $\langle \Psi | H | \Psi \rangle$ as a *functional* on $\{|m\rangle\}$. We can then apply the standard techniques of the calculus of variations, seeking an optimal set of single-particle states that makes $\langle \Psi | H | \Psi \rangle$ stationary under arbitrary infinitesimal changes, $|m\rangle \rightarrow |m\rangle + |\delta m\rangle$. The variational principle then shows that the resulting set produces the best single-determinant approximation to the ground state.

An useful analogy is available. In classical mechanics, variational minimization of the Lagrangian functional yields a set of differential equations of motion in the individual coordinates called the Euler-Lagrange equations. Here, we are similarly expecting variational minimization of $\langle \Psi | H | \Psi \rangle$ to yield the “equations of motion” i.e. the HF equations in each of the “coordinates,” namely $\{|m\rangle\}$.

In the current problem, the variations in $\{|m\rangle\}$ are constrained by the orthonormality requirement, $\langle m | n \rangle - \delta_{mn} = 0$. The standard procedure in accomodating such constraints in an optimization problem is to introduce Lagrange multipliers ϵ_{mn} ; a technique in which one minimizes the Lagrange function L'

$$L'(\{|m\rangle\}) = \langle \Psi | H | \Psi \rangle - \sum_{m=1}^N \sum_{n=1}^N \epsilon_{mn} (\langle m | n \rangle - \delta_{mn})$$

rather than the energy expectation $\langle \Psi | H | \Psi \rangle$ directly.

However, for the sake of arriving at the HF equations more quickly, I will only enforce the normalization constraint $\langle m|m \rangle - 1 = 0$ in the Lagrange function. Through the resulting equations, we can check *a posteriori* that the single-particle states are also orthogonal as required. (We don't have the space to actually perform the verification, I'm afraid) The possibility of this particular approach was shown in the reference by Bethe and Jackiw. The relevant Lagrange function L is then,

$$\begin{aligned} L &= \langle \Psi | H | \Psi \rangle - \sum_{m=1}^N \epsilon_m (\langle m|m \rangle - 1) \\ &= \sum_{m=1}^N \langle m | h_1 | m \rangle + \frac{1}{2} \sum_{m=1}^N \sum_{n=1}^N \{ \langle mn | h_2 | mn \rangle - \langle mn | h_2 | nm \rangle \} - \sum_{m=1}^N \epsilon_m (\langle m|m \rangle - 1) \end{aligned}$$

Applying the arbitrary changes $|m\rangle \rightarrow |m\rangle + |\delta m\rangle$ in the single-particle states, we find the first variation of L to be

$$\begin{aligned} \delta L &= \sum_{m=1}^N \langle \delta m | h_1 | m \rangle + \langle m | h_1 | \delta m \rangle \\ &+ \frac{1}{2} \sum_{m=1}^N \sum_{n=1}^N \{ \langle (\delta m)n | h_2 | mn \rangle + \langle m(\delta n) | h_2 | mn \rangle + \langle mn | h_2 | (\delta m)n \rangle + \langle mn | h_2 | m(\delta n) \rangle \\ &- \frac{1}{2} \sum_{m=1}^N \sum_{n=1}^N \{ \langle (\delta m)n | h_2 | nm \rangle + \langle m(\delta n) | h_2 | nm \rangle + \langle mn | h_2 | (\delta n)m \rangle + \langle mn | h_2 | n(\delta m) \rangle \\ &- \sum_{m=1}^N \epsilon_m (\langle \delta m | m \rangle + \langle m | \delta m \rangle) \end{aligned} \quad (14)$$

The key to making sense of the above expression is to recognize that the two-electron operator is invariant to an exchange in the order of electrons: $h_2(i, j) = h_2(j, i)$. This of course follows from the fact that h_2 represents Coulombic interaction, which depends only on the relative distance of the two interacting particles. In particular, the second term in the first double sum of Eq. 14 may be rewritten:

$$\begin{aligned} \langle m(\delta n) | h_2 | mn \rangle &= \int d\vec{x}_1 d\vec{x}_2 \chi_m^*(\vec{x}_1) \delta \chi_n^*(\vec{x}_2) h_2 \chi_m(\vec{x}_1) \chi_n(\vec{x}_2) \\ &= \int d\vec{x}_1 d\vec{x}_2 \chi_m^*(\vec{x}_2) \delta \chi_n^*(\vec{x}_1) h_2 \chi_m(\vec{x}_2) \chi_n(\vec{x}_1) = \langle (\delta n)m | h_2 | nm \rangle \end{aligned}$$

where, in the second line, we have essentially swapped the names of the ‘‘dummy’’ integration variables. Similarly, the second term in the second double sum can be shown to be $\langle m(\delta n) | h_2 | nm \rangle = \langle (\delta n)m | h_2 | mn \rangle$.

With the above changes, it can be seen that every term in Eq. 14 occurs in complex conjugate pairs. For notational simplicity, we will show only the terms involving variations that are conjugated (i.e. a δ in the bra) and suppress their complex conjugate partners. With this convention, δL can be expressed as

$$\begin{aligned}\delta L &= \sum_{m=1}^N \langle \delta m | h_1 | m \rangle \\ &+ \frac{1}{2} \sum_m^N \sum_n^N \langle (\delta m) n | h_2 | mn \rangle + \langle (\delta n) m | h_2 | nm \rangle \\ &- \frac{1}{2} \sum_m^N \sum_n^N \langle (\delta m) n | h_2 | nm \rangle + \langle (\delta n) m | h_2 | mn \rangle \\ &- \sum_{m=1}^N \epsilon_m \langle \delta m | m \rangle\end{aligned}$$

Now note that the two-electron integral $\langle (\delta a) b | h_2 | ab \rangle$ involving any two single particle states $|a\rangle$ and $|b\rangle$ will be attained twice in the first double summation, namely, when $(m = a, n = b)$, and also when $(m = b, n = a)$. We can say the same for $\langle (\delta a) b | h_2 | ba \rangle$ in the second double sum. Therefore, we finally formulate δL in a form suitable for use:

$$\delta L = \sum_{m=1}^N \langle \delta m | h_1 | m \rangle + \sum_m^N \sum_n^N \{ \langle (\delta m) n | h_2 | mn \rangle - \langle (\delta m) n | h_2 | nm \rangle \} - \sum_{m=1}^N \epsilon_m \langle \delta m | m \rangle$$

Expressing the inner products in coordinate-space, and factoring out $\delta \chi_m^*(\vec{x}_1)$, which is common to all terms, yields:

$$\begin{aligned}\delta L &= \sum_{m=1}^N \int d\vec{x}_1 \delta \chi_m^*(\vec{x}_1) \\ &\times \left[h_1 \chi_m(\vec{x}_1) + \sum_{n=1}^N \int d\vec{x}_2 \chi_n^*(\vec{x}_2) h_2 \chi_m(\vec{x}_1) \chi_n(\vec{x}_2) - \sum_{n=1}^N \int d\vec{x}_2 \chi_n^*(\vec{x}_2) h_2 \chi_n(\vec{x}_1) \chi_m(\vec{x}_2) - \epsilon_m \chi_m(\vec{x}_1) \right] \\ \delta L &= \sum_{m=1}^N \int d\vec{x}_1 \delta \chi_m^*(\vec{x}_1) \left[h_1 + \sum_{n=1}^N \int d\vec{x}_2 \frac{1}{r_{12}} |\chi_n(\vec{x}_2)|^2 - \sum_{n=1}^N \int d\vec{x}_2 \chi_n^*(\vec{x}_2) \frac{1}{r_{12}} P_2 \chi_n(\vec{x}_2) - \epsilon_m \right] \chi_m(\vec{x}_1)\end{aligned}$$

In the last line, we have substituted in $h_2 = r_{12}^{-1}$ and again used the operator P_2 that performs transposition of electrons 1 and 2, i.e. $P_2 \chi_n(\vec{x}_2) \chi_m(\vec{x}_1) = \chi_n(\vec{x}_1) \chi_m(\vec{x}_2)$. As is the usual in the calculus of variations, we then argue that since $\delta \chi_m^*$ is arbitrary,

the stationary condition $\delta L = 0$ is obtained when each of the multiplicative factors in Eq. 15 are zero. In other words, the minimization condition is equivalent to:

$$\begin{aligned} f(\vec{x}_1)\chi_m(\vec{x}_1) &= \epsilon_m\chi_m(\vec{x}_1) \\ f(\vec{x}_1) &= h_1 + \sum_{n=1}^N \int d\vec{x}_2 \frac{1}{r_{12}} |\chi_n(\vec{x}_2)|^2 - \sum_{n=1}^N \int d\vec{x}_2 \chi_n^*(\vec{x}_2) \frac{1}{r_{12}} P_2 \chi_n(\vec{x}_2) \end{aligned} \quad (16)$$

which holds for all $m = 1, 2, \dots, N$. These equations, having the form of a single-particle Schrödinger equation, are the *Hartree-Fock equations* that characterize the optimal single-particle states to be used in the Slater determinant. The operator $f(\vec{x}_1)$ is called the *Fock operator*, and the orthogonality of its eigenfunctions is proved in Ref 2. The coupling between the states is contained within the Fock operator, and makes Eq. 16 a nonlinear SE; hence the need for iterative methods. We now note that the conjugate terms we have suppressed in δL do not give independent constraints. Instead, they merely produce conjugated forms of Eq. 16.

Thus ends our derivation of the Hartree-Fock equations from the variational principle. We have reduced the N -electron Hamiltonian into N coupled single-particle problems that are then typically solved by fixed point iteration, as shown in the introduction.

It is reassuring to find the electronic repulsion term $\sum_{n=1}^N \int d\vec{x}_2 \frac{1}{r_{12}} |\chi_n(\vec{x}_2)|^2$ in the HF equations, which agrees with the intuition embodied in the original Hartree method. However, the *exchange term* $\sum_{n=1}^N \int d\vec{x}_2 \chi_n^*(\vec{x}_2) \frac{1}{r_{12}} P_2 \chi_n(\vec{x}_2)$ originates from the antisymmetry requirement, and defies simple interpretation. (To reviewers: I have some ideas, but would like to first discuss them with you before committing. Also, any suggestions?)

4 Computer implementation of HF

It is apparent that the Hartree-Fock equations are more complicated than their counterparts in the Hartree procedure. Therefore, this section is devoted to simplifying and converting Eq. 16 into a form suitable for practical implementation. Our mathematical method is to introduce a set of *basis functions* that span the space of physically relevant wavefunctions. It is an explicit, numerical embodiment of the type of linear algebra we have encountered so far in QM.

4.1 Integration over the spin coordinate

The first concession we make in exchange for simplification is to specialize in systems for which the number of electrons N is even.

Recall that to form the many-particle ground state, we seek the N lowest-energy single-particle eigenfunctions of the Hartree-Fock equation (Eq. 16). When N is even, the Fock operator does not depend on electron spin, as will be shown below. For the lowest overall energy, it then follows that we can focus on the $\frac{N}{2}$ lowest-energy spatial states $\{\psi_n \mid n = 1, 2, \dots, N/2\}$, and then doubly occupy each with electrons of opposite spin. In other words, we let:

$$\chi_m = \psi_{n(m)} \otimes |w(m)\rangle \quad m = 1, 2, \dots, N \quad (17)$$

where $n(m)$ is the quotient of $m \div 2$. The spin variable is $w(m) = \pm\frac{1}{2}$ depending on whether m is even or odd.

Applying the Fock-operator (Eq. 16) on the single-particle state of Eq. 17 yields:

$$\begin{aligned} f(\vec{x}_1)\chi_m(\vec{x}_1) &= f(\vec{x}_1) [\psi_{n(m)}(\vec{r}_1) |w_1(m)\rangle] \\ &= \{h_1\psi_{n(m)}(\vec{r}_1)\} |w_1(m)\rangle \\ &+ \left\{ \sum_{l=1}^N \int d\vec{r}_2 dw_2 \frac{1}{r_{12}} |\psi_{n(l)}^*(\vec{r}_2) |w_2(l)\rangle|^2 \cdot \psi_{n(m)}(\vec{r}_1) \right\} |w_1(m)\rangle \\ &- \sum_{l=1}^N \int d\vec{r}_2 dw_2 \frac{1}{r_{12}} \{ \psi_{n(l)}^*(\vec{r}_2) \langle w_2(l) | \} P_2 \{ \psi_{n(l)}(\vec{r}_2) |w_2(l)\rangle \} \{ \psi_{n(m)}(\vec{r}_1) |w_1(m)\rangle \} \end{aligned}$$

In the second term, the spin variable w_2 can simply be “integrated out.” On the other hand, we analyze the last term more carefully, where the transposition operator P_2 swaps electron indices 1 and 2 to yield:

$$P_2 \{ \psi_{n(l)}(\vec{r}_2) |w_2(l)\rangle \} \{ \psi_{n(m)}(\vec{r}_1) |w_1(m)\rangle \} = \{ \psi_{n(l)}(\vec{r}_1) |w_1(l)\rangle \} \{ \psi_{n(m)}(\vec{r}_2) |w_2(m)\rangle \}$$

Since this last result is combined with $\psi_{n(l)}^*(\vec{r}_2) \otimes \langle w_2(l) |$, the terms in the last summation are nonzero only when the spins $|w_2(l)\rangle$ and $|w_1(m)\rangle$ align. With N even and with the electron assignment as in Eq. 17, for either $w_1(m) = \pm\frac{1}{2}$, there are $\frac{N}{2}$ terms that meet the spin constraint. Evidently, the Fock operator $f(\vec{x}_1)$ does not affect nor depend on the spin of $\chi_m(\vec{x}_1)$, so that we can write:

$$f(\vec{x}_1) = f(\vec{r}_1) \otimes \mathbf{1} \quad (18)$$

$$f(\vec{r}_1) = h_1 + \sum_{n=1}^{N/2} \left\{ 2 \int d\vec{r}_2 \frac{1}{r_{12}} |\psi_n(\vec{r}_2)|^2 - \int d\vec{r}_2 \frac{1}{r_{12}} \psi_n^*(\vec{r}_2) P_2 \psi_n(\vec{r}_2) \right\} \quad (19)$$

where $f(\vec{r}_1)$ is the spatial Fock operator, and $\mathbf{1}$ is the identity operator in the spin space. The transposition operator P_2 is understood in this context to swap the spatial coordinates. As indicated earlier, we have now proven that the full Fock operator is independent of spin in the case when N is even. Our task is then to obtain the $\frac{N}{2}$ lowest-energy single-particle eigenstates of the spatial Hartree-Fock equation:

$$f(\vec{r}_1)\psi_n(\vec{r}_1) = \epsilon_n\psi_n(\vec{r}_1) \quad (20)$$

4.2 Introduction of a basis

In the introduction we remarked on the difficulty of solving the iteration equations in coordinate space. The contribution of C.J. Roothaan in 1951 was to demonstrate that, by introducing a set of known spatial *basis functions*, the differential Hartree-Fock equations could be converted into a set of algebraic equations to be solved by standard matrix techniques.

Suppose that $\{\phi_\mu\}$ represents a set of basis functions for the space of square integrable functions. In practice, we must choose some K -element subset of this basis for actual use. We can then approximate the i -th spatial wavefunction by a linear combination

$$\psi_i = \sum_{\mu=1}^K C_{\mu i} \phi_\mu \quad i = 1, 2, \dots, (K \geq \frac{N}{2})$$

The above expression would be exact if the limited basis set $\{\phi_\mu \mid \mu = 1, 2, \dots, K\}$ were in fact complete. However, in practice, the basis functions typically have no claim on rigorous completeness. In fact, in the calculations on HeH^+ , I use a set of just two functions! The justification for such an audacious move is that typical applications do not demand mathematical completeness. Rather, the more relevant requirement is that our wavefunctions be expressible in the chosen basis. Therefore, by choosing functions that are appropriate for the particular physics, one gets away with even dramatically incomplete basis sets. In the current problem, the physically motivated basis set involves the atomic wavefunctions centered at each of the nuclei.

Finally, we also note that the choice of a particular basis set does not affect the mathematical theory that follows. In fact, in most implementations, the choice of basis set is reconfigurable at runtime. (Not my application, however!)

4.3 Roothaan Equation

To obtain the HF equations in matrix form, consider again the spatial Hartree-Fock equation (Eq. 20). Begin by expanding $\psi_n(\vec{r}_1)$ in the chosen basis,

$$f(\vec{r}_1) \sum_{\nu} C_{\nu n} \phi_{\nu}(\vec{r}_1) = \epsilon_n \sum_{\nu} C_{\nu n} \phi_{\nu}(\vec{r}_1) \quad (21)$$

We then multiply by $\phi_{\mu}^*(\vec{r}_1)$ on the left and integrate, to obtain:

$$\sum_{\nu} \left\{ \int d\vec{r}_1 \phi_{\mu}^*(\vec{r}_1) f(\vec{r}_1) \phi_{\nu}(\vec{r}_1) \right\} C_{\nu n} = \epsilon_n \sum_{\nu} \left\{ \int d\vec{r}_1 \phi_{\mu}^*(\vec{r}_1) \phi_{\nu}(\vec{r}_1) \right\} C_{\nu n} \quad (22)$$

This motivates the definition of two matrices. The first is the *Fock matrix* $F_{\mu\nu} = \int d\vec{r}_1 \phi_{\mu}^*(\vec{r}_1) f(\vec{r}_1) \phi_{\nu}(\vec{r}_1)$. The second is the *overlap matrix* $S_{\mu\nu} = \int d\vec{r}_1 \phi_{\mu}^*(\vec{r}_1) \phi_{\nu}(\vec{r}_1)$, which can be shown (in the appendix) to be positive definite. With these definitions Eq. 22 becomes,

$$\sum_{\nu} F_{\mu\nu} C_{\nu n} = \epsilon_n \sum_{\nu} S_{\mu\nu} C_{\nu n} \quad (23)$$

This result may more succinctly written as a single matrix equation, known as the *Roothaan equation*:

$$\mathbf{FC} = \mathbf{SC}\epsilon \quad (24)$$

Here, the matrix ϵ is diagonal and contains the single-particle energy eigenvalues ϵ_n as the n -th element. Furthermore, \mathbf{C} is the $K \times K$ *coefficient matrix* whose n -th column denotes the expansion coefficients of ψ_n in the basis set $\{\phi_{\mu}\}$. Hence, solving for the optimal single-particle states in the Hartree-Fock approximation is equivalent solving for the coefficient matrix \mathbf{C} that solves the Roothaan equation!

4.3.1 Structure of the Fock matrix

In Eq. 24 \mathbf{S} is known; and \mathbf{C} and ϵ are to be determined through a pseudo-eigenvalue problem. The only remaining component in the Roothaan equation is the Fock matrix, which we now explicitly calculate terms of the chosen basis. From the definition of the Fock matrix and $f(\vec{r}_1)$, we have,

$$\begin{aligned} F_{\mu\nu} &= \int d\vec{r}_1 \phi_{\mu}^*(\vec{r}_1) h_1 \phi_{\nu}(\vec{r}_1) \\ &+ \sum_{n=1}^{N/2} \int d\vec{r}_1 \phi_{\mu}^*(\vec{r}_1) \left[2 \int d\vec{r}_2 \frac{1}{r_{12}} |\psi_n(\vec{r}_2)|^2 - \int d\vec{r}_2 \psi_n^*(\vec{r}_2) \frac{1}{r_{12}} P_2 \psi_n(\vec{r}_2) \right] \phi_{\nu}(\vec{r}_1) \end{aligned}$$

Using the following expansions: $\psi_n(\vec{r}_2) = \sum_\lambda C_{\lambda n} \phi_\lambda(\vec{r}_2)$, and $\psi_n^*(\vec{r}_2) = \sum_\sigma C_{\sigma n}^* \phi_\sigma^*(\vec{r}_2)$, the last term may be written as (after also exchanging the order of sums):

$$\sum_{\lambda\sigma} \sum_{n=1}^{N/2} C_{\lambda n} C_{\sigma n}^* \left\{ 2 \int d\vec{r}_1 d\vec{r}_2 \phi_\mu^*(\vec{r}_1) \phi_\sigma^*(\vec{r}_2) \frac{1}{r_{12}} \phi_\nu(\vec{r}_1) \phi_\lambda(\vec{r}_2) - \int d\vec{r}_1 d\vec{r}_2 \phi_\mu^*(\vec{r}_1) \phi_\sigma^*(\vec{r}_2) \frac{1}{r_{12}} \phi_\lambda(\vec{r}_1) \phi_\nu(\vec{r}_2) \right\} \quad (25)$$

For sanity, we shall define $P_{\lambda\sigma} = 2 \sum_{n=1}^{N/2} C_{\lambda n} C_{\sigma n}^*$ (sometimes called the *density matrix*) and adopt a new “bracket-like” notation for the basis function integrals:

$$(\mu\sigma|r_{12}^{-1}|\nu\lambda) = \int d\vec{r}_1 d\vec{r}_2 \phi_\mu^*(\vec{r}_1) \phi_\sigma^*(\vec{r}_2) \frac{1}{r_{12}} \phi_\nu(\vec{r}_1) \phi_\lambda(\vec{r}_2)$$

then $F_{\mu\nu}$ takes the tolerable form:

$$F_{\mu\nu} = H_{\mu\nu}^{core} + \sum_{\lambda\sigma} P_{\lambda\sigma} \left\{ (\mu\sigma|r_{12}^{-1}|\nu\lambda) - \frac{1}{2}(\mu\sigma|r_{12}^{-1}|\lambda\nu) \right\} \quad (26)$$

where the sums on λ and σ are over all of the basis functions. In the above, we have also defined the core-Hamiltonian matrix $H_{\mu\nu}^{core}$ which has its own internal structure:

$$\begin{aligned} H_{\mu\nu}^{core} &= \int d\vec{r}_1 \phi_\mu^*(\vec{r}_1) h_1 \phi_\nu(\vec{r}_1) = \int d\vec{r}_1 \phi_\mu^*(\vec{r}_1) \left\{ -\frac{1}{2} \nabla_1^2 + \sum_{A=1}^M \frac{Z_A}{r_{1A}} \right\} \phi_\nu(\vec{r}_1) \\ &= \int d\vec{r}_1 \phi_\mu^*(\vec{r}_1) \left\{ -\frac{1}{2} \nabla_1^2 \right\} \phi_\nu(\vec{r}_1) + \int d\vec{r}_1 \phi_\mu^*(\vec{r}_1) \left\{ \sum_{A=1}^M \frac{Z_A}{r_{1A}} \right\} \phi_\nu(\vec{r}_1) \\ &= T_{\mu\nu} + V_{\mu\nu}^{nucl} \end{aligned} \quad (27)$$

In closing, we have defined the *kinetic energy* \mathbf{T} and the *nuclear potential* \mathbf{V}^{nucl} matrices that comprise \mathbf{H}^{core} . Given the bewildering number of new quantities we have named in this section, we will revisit the procedure for constructing the Fock matrix in Section 4.4 with the help of a flow chart.

4.3.2 Solving the Roothaan equation

In exploring the structure of the Fock Matrix, it was shown that \mathbf{F} depends on the coefficient matrix (through \mathbf{P}). It then follows that the Roothaan equation is nonlinear, and cannot be directly solved by standard linear techniques.

Instead, we use an iterative approach in which we first compute $\mathbf{F}^{(i-1)}$ based on the previous set of coefficients $\mathbf{C}^{(i-1)}$ (or by an initial guess). The Fock matrix thus

generated is then considered to be fixed, which allows us to solve for the next set of coefficients $\mathbf{C}^{(i)}$ via the Roothaan equation, which may now be notated as:

$$\mathbf{F}^{(i-1)}\mathbf{C}^{(i)} = \mathbf{S}\mathbf{C}^{(i)}\epsilon \quad (28)$$

As in the introduction, such iteration is tantamount to holding the wavefunctions of the other electrons ($j \neq i$) fixed as we generate the new wavefunction for the i -th electron. The limit of the sequence of matrices $\mathbf{C}^{(i)} \rightarrow \mathbf{C}$ is then taken to be the solution of the Roothaan equations. The entries of the coefficient matrix can then be used to express (in the chosen basis) the single-particle wavefunctions of the Slater-determinant, thus completing our implementation.

There are several matrix techniques for solving Eq. 28, which differs from standard eigenvalue equations by the presence of the overlap matrix. In the appendix, we discuss one possibility called symmetric orthogonalization.

4.4 Summary of the HF algorithm

With the help of a flowchart, we now summarize the steps involved in a HF program.

5 Calculations on HeH+

This section is not completed for the first draft. I spent the time instead “polishing” the previous sections.

6 Conclusion

In this paper, we have given a thorough account of the basic Hartree-Fock theory. From the variational principle, we derived the Hartree-Fock equations which identify the optimal single-particle states to compose the Slater-determinant many-body wavefunction.

By introducing a set of basis functions, we then converted the resulting differential equations into matrix form, which can be solved efficiently on a computer. (To reviewers: My hope is that the all the necessary details were covered so that the reader could implement his or her own HF solver. Probably failed in this draft...)

Finally, the example implementation was utilized to calculate various of quantities of interest for helium hydride. (To reviewers: Another hope – maybe you now want to try building an HF solver to perform calculations on your favorite molecule?)

(Final note to reviewers: Please tear Section 4 to shreds. There is a lot of information and I want to present it in the most palatable way possible. I do think that it is absolutely necessary for the paper, because it is the most common way Hamiltonian solvers are implemented in practice. I also very much like that it is down-to-earth, honest-man's linear algebra.)

References

- [1] Szabo and Ostlund. *Modern Quantum Chemistry*. (New York: McGraw-Hill, Inc, 1989)
- [2] Bethe and Jackiw. *Intermediate Quantum Mechanics*. (California: Benjamin/Cummings Publishing Company, 1986)
- [3] Griffiths, David. *Introduction to Quantum Mechanics*. (New York: Prentice Hall, 2005)