VIRGINIA TECH
Department of Chemistry

Electronic Structure Theory (CHEM 6664)

Final Project 1

Name: _____

SID: _____

Prof. Edward Valeev, Spring 2023

1   Final Project Description

The goal of this final project is to carry out a complete electronic structure computation on a molecular system without using any data other than the coordinates of the nuclei, the number of the electrons, and the basis set. Specifically, you are going to perform a Hartree-Fock calculation on the H_2 molecule using an "uncontracted" 3-21G basis set. The end result will be the equilibrium geometry and the dissociation energy of the molecule. You will then verify the correctness of your results against an .

The main challenges of this project are:

1. derive an expression for the overlap integral between two Gaussian functions of arbitrary angular momenta;

2. derive an expression for the kinetic energy integral over s-type Gaussians;

3. use the derived and given expressions for the integrals to compute the Hartree-Fock energy and determine the equilibrium geometry of the molecule;

Guidelines:

1. The basis set that you will use is described in Gaussian format as:

```
!
BASIS="3-21G-UC"
 H   0
 S   1  1.00
       5.44717800           1.00000000
 S   1  1.00
      0.824547000           1.00000000
 S   1  1.00
      0.183192000           1.00000000
 ****
```

2   Gaussian Integrals

A *primitive* non-normalized *three-dimensional* Cartesian Gaussian function with orbital exponent α centered on $\mathbf{R} = (R_x, R_y, R_z)$ will be denoted as

$$|\mathbf{n}, \alpha, \mathbf{A}\rangle \equiv G(\mathbf{r}; \mathbf{n}, \alpha, \mathbf{R}) = (r_x - R_x)^{n_x}(r_y - R_y)^{n_y}(r_z - R_z)^{n_z} \exp(-\alpha|\mathbf{r} - \mathbf{R}|^2) \quad , \qquad (1)$$

where $\mathbf{r} = (r_x, r_y, r_z)$ and $\mathbf{n} = (n_x, n_y, n_z)$ (n_x, n_y, n_z are non-negative integers). The angular momentum of the Gaussian is $\lambda_{\mathbf{n}} = n_x + n_y + n_z$. Hence, $G(\mathbf{r}; \mathbf{0}, \alpha, \mathbf{R})$ is an s-type Gaussian, where $\mathbf{0} = (0, 0, 0)$.

1. Derive a closed expression for the following one-dimensional integral:

$$R_k = \int_{-\infty}^{\infty} dx\, x^k \exp(-\alpha x^2) \tag{2}$$

where k is a non-negative integer and α is a positive real number.

2. Using the Gaussian Product Theorem for s-type Gaussians (verify this identity),

$$G(\mathbf{r};\alpha,\mathbf{A})G(\mathbf{r};\beta,\mathbf{B}) = KG(\mathbf{r};\alpha+\beta,\mathbf{P}), \quad \text{where} \tag{3}$$

$$K = K_x K_y K_z, \tag{4}$$

$$K_x = \exp\left(-\frac{\alpha\beta}{\alpha+\beta}(A_x - B_x)^2\right), \quad \text{etc.,} \tag{5}$$

$$P_x = \frac{\alpha A_x + \beta B_x}{\alpha+\beta} \tag{6}$$

(i.e. product of 2 primitive s-Gaussians is another primitive s-Gaussian located at the "center-of-mass", with the exponents playing the role of the mass), show that the overlap integral between 2 primitive non-normalized Gaussians is

$$\langle \mathbf{n},\alpha,\mathbf{A}|\mathbf{m},\beta,\mathbf{B}\rangle = \int G(\mathbf{r};\mathbf{n},\alpha,\mathbf{A})G(\mathbf{r};\mathbf{m},\beta,\mathbf{B})\, d\mathbf{r} = \exp[-\alpha\beta|\mathbf{A}-\mathbf{B}|^2/(\alpha+\beta)]I_x I_y I_z \quad, \tag{7}$$

where

$$I_x = \int (r_x - A_x)^{n_x}(r_x - B_x)^{m_x} \exp(-(\alpha+\beta)(r_x - P_x)^2) \quad, \text{etc.} \tag{8}$$

and

$$P_x = \frac{\alpha A_x + \beta B_x}{\alpha+\beta} \quad, \text{etc.} \tag{9}$$

Then rewrite $r_x - A_x$ as $(r_x - P_x) + (P_x - A_x)$ and use the binomial expansion

$$(a+b)^c = \sum_{d=0}^{c} a^{c-d} b^d \binom{c}{d} \tag{10}$$

to show that

$$(r_x - A_x)^{n_x}(r_x - B_x)^{m_x} = \sum_{i=0}^{n_x}(P_x - A_x)^{n_x - i}\binom{n_x}{i}\sum_{j=0}^{m_x}(P_x - B_x)^{m_x - j}\binom{m_x}{j}(r_x - P_x)^{i+j}. \tag{11}$$

Finally, use Eq. (2) to derive an expression for I_x. Program a function to evaluate I_x for given parameters $A_x, B_x, n_x, n_y, \alpha, \beta$. Realize how this function can be used to evaluate I_y and I_z and implement a function to evaluate the overlap integral in Eq. (7) for given $\mathbf{n}$, α, $\mathbf{A}$, $\mathbf{m}$, β, and $\mathbf{B}$.

3. Now let's turn our attention to the kinetic energy integrals over s-type Gaussians. First, show that

$$\frac{\partial^2}{\partial r_x^2}|\mathbf{0}, \alpha, \mathbf{A}\rangle = -2\alpha|\mathbf{0}, \alpha, \mathbf{A}\rangle + 4\alpha^2|\mathbf{2}_x, \alpha, \mathbf{A}\rangle \quad , \tag{12}$$

where $\mathbf{2}_x = (2, 0, 0)$. Then express the matrix element of the kinetic energy operator over s-type Gaussians, $\langle \mathbf{0}, \alpha, \mathbf{A}|\hat{T}|\mathbf{0}, \beta, \mathbf{B}\rangle$, where $\hat{T} = -\frac{1}{2}\left(\frac{\partial^2}{\partial r_x^2} + \frac{\partial^2}{\partial r_y^2} + \frac{\partial^2}{\partial r_z^2}\right)$, as a linear combination of the overlap integrals (Eq. (7)). Write a function to compute the kinetic energy integral for given α, $\mathbf{A}$, β, and $\mathbf{B}$.

4. The nuclear attraction energy integral over s-type Gaussians can be expressed in closed form as

$$\langle \mathbf{0}, \alpha, \mathbf{A}|\hat{V}_{\text{ne}}|\mathbf{0}, \beta, \mathbf{B}\rangle = -\sum_C \int e^{-\alpha r_A^2}\frac{Z_C}{|\mathbf{r} - \mathbf{C}|}e^{-\beta r_B^2}d\mathbf{r} \tag{13}$$

$$= -\frac{2\pi \exp[-\alpha\beta(|\mathbf{A} - \mathbf{B}|^2)/\gamma]}{\gamma}$$

$$\times \left(\sum_{C:\mathbf{P}\neq\mathbf{C}} Z_C\frac{\pi^{1/2}\text{erf}(\gamma^{1/2}|\mathbf{P} - \mathbf{C}|)}{2\gamma^{1/2}|\mathbf{P} - \mathbf{C}|} + \sum_{C:\mathbf{P}=\mathbf{C}} Z_C\right) \quad , \tag{14}$$

where $\gamma = \alpha + \beta$, erf is the error function, and the sum in Eq. (13) is over all nuclei in the molecule (Z_C is the charge of nucleus C and $\mathbf{C}$ is its coordinate vector). The first sum in Eq. (14) is over all nuclei whose positions do not coincide with $\mathbf{P}$, etc. Write a function to compute the nuclear attraction integral for given α, $\mathbf{A}$, β, and $\mathbf{B}$.

5. The electron repulsion integral over s-type Gaussians can be expressed in closed form as

$$\langle \mathbf{0}, \alpha_1, \mathbf{A}_1; \mathbf{0}, \alpha_2, \mathbf{A}_2|\frac{1}{r_{12}}|\mathbf{0}, \beta_1, \mathbf{B}_1; \mathbf{0}, \beta_2, \mathbf{B}_2\rangle$$

$$= \int e^{-\alpha_1|\mathbf{r}_1-\mathbf{A}_1|^2}e^{-\alpha_2|\mathbf{r}_2-\mathbf{A}_2|^2}\frac{1}{r_{12}}e^{-\alpha_1|\mathbf{r}_1-\mathbf{B}_1|^2}e^{-\alpha_2|\mathbf{r}_2-\mathbf{B}_2|^2}d\mathbf{r}_1 d\mathbf{r}_2$$

$$= \frac{2\pi^{5/2}K_1 K_2}{\gamma_1\gamma_2(\gamma_1 + \gamma_2)^{1/2}}F_0(|\mathbf{P}_1 - \mathbf{P}_2|^2\gamma_1\gamma_2/(\gamma_1 + \gamma_2)) \tag{15}$$

where

$$\gamma_1 = \alpha_1 + \beta_1 \tag{16}$$

$$\gamma_2 = \alpha_2 + \beta_2 \tag{17}$$

$$\mathbf{P}_1 = \frac{\alpha_1 \mathbf{A}_1 + \beta_1 \mathbf{B}_1}{\gamma_1} \tag{18}$$

$$\mathbf{P}_2 = \frac{\alpha_2 \mathbf{A}_2 + \beta_2 \mathbf{B}_2}{\gamma_2} \tag{19}$$

$$K_1 = e^{-[\alpha_1 \beta_1 (|\mathbf{A}_1 - \mathbf{B}_1|^2)/\gamma_1]} \tag{20}$$

$$K_2 = e^{-[\alpha_2 \beta_2 (|\mathbf{A}_2 - \mathbf{B}_2|^2)/\gamma_2]} \tag{21}$$

$$F_0(T) = \frac{\pi^{1/2}}{2\sqrt{T}} \mathrm{erf}(\sqrt{T}) \ , \ T > 0 \tag{22}$$

$$F_0(0) = 1 \tag{23}$$

Write a function to evaluate the electron repulsion integral for given α_1, $\mathbf{A}_1$, β_1, $\mathbf{B}_1$, α_2, $\mathbf{A}_2$, β_2, and $\mathbf{B}_2$.

3 Hartree-Fock Method

In this section you will compute the Hartree-Fock energy E for a given internuclear distance R. Although we did not cover Hartree-Fock theory in complete technical detail, only a short introduction should be sufficient. Spin-free molecular orbitals $u_i(r)$ will expanded in terms of the atomic basis functions as

$$u_i(\mathbf{r}) = \sum_{j=1}^{M} c_{ij} \phi_j(\mathbf{r}) \tag{24}$$

where M is the number of atomic basis functions (in our case we have 2 atoms with 3 primitive s Gaussian basis functions each, hence $M = 6$) and c_{ij} is the coefficient of the jth atomic basis function to the ith molecular orbital (MO). A vector of coefficients for the ith MO will be denoted as $\mathbf{c}_i$. The coefficients are determined by solving the Hartree-Fock equations:

$$\mathbf{F}\mathbf{c}_i = e_i \mathbf{S}\mathbf{c}_i \quad , \tag{25}$$

where $\mathbf{F}$ and $\mathbf{S}$ are the Fock and overlap matrices, respectively, expressed in the basis of AO basis functions. These are M by M matrices, hence M solutions to Eq. (25) can be found that correspond to $N/2$ occupied MOs and $M - N/2$ virtual MOs (we are assuming here that the number of electrons, N, is even and all MOs are doubly occupied).[1] Note that Mathematica produces solutions that are not sorted by the

---

[1]If the overlap matrix is (near-)singular, as happens when the AO basis set approaches linear dependence, then the number of solutions may be fewer than M.

eigenvalues, hence to determine which eigenvectors correspond to the occupied orbitals you must sort the eigenvalues.

Caution: solving the Fock equations may not produce normalized eigenvectors! Therefore you must explicitly normalize the solutions to Eq. (25) using the normalization condition $\mathbf{c}_i^\dagger \mathbf{S} \mathbf{c}_i = 1$.

At this point you have all the necessary machinery to evaluate the matrices $\mathbf{F}$ and $\mathbf{S}$ and solve the Hartree-Fock equations. You may implement the following steps in any manner that you choose, but I recommend that you have a top-level function that will evaluate the Hartree-Fock energy as a function of the internuclear distance R. This function will make it much easier to implement the geometry optimization algorithm.

6. Compute the overlap matrix $\mathbf{S}$. The overlap matrix element $S_{ij} = \langle \phi_i | \phi_j \rangle$ between atomic basis functions $\phi_i(\mathbf{r})$ and $\phi_j(\mathbf{r})$ was derived in Item 2. HINT: Compute the eigenvalues of $\mathbf{S}$ – they must be all nonnegative (if we used normalized Gaussian functions they would be between 0 and 1). The presence of eigenvalues outside of this range, or too close to 0 ($< 10^{-6}$), suggests something is wrong!

7. Compute the core Hamiltonian matrix $\mathbf{H}$. The core Hamiltonian matrix element $H_{ij} = \langle \phi_i | \hat{h} | \phi_j \rangle = \langle \phi_i | \hat{T} | \phi_j \rangle + \langle \phi_i | \hat{V}_{\text{ne}} | \phi_j \rangle$ is a sum of the kinetic energy and nuclear attraction integrals: $H_{ij} = T_{ij} + V_{ij}$.

8. The Hartree-Fock equations are solved iteratively. Let's denote as $\mathbf{c}_i^{(k)}$ a vector of coefficients for MO i at iteration k. One usually obtains an initial set of MO coefficients, $\mathbf{c}_i^{(0)}$, by diagonalizing the core Hamiltonian matrix (this is known as the core Hamiltonian guess):

$$\mathbf{H} \mathbf{c}_i^{(0)} = e_i \mathbf{S} \mathbf{c}_i^{(0)} \tag{26}$$

Compute the initial set of MO coefficients.

9. Compute the one-electron density matrix in the AO basis, $\mathbf{P}$, which is defined as

$$P_{ij} = \sum_{m=1}^{N/2} c_{mi} c_{mj} \quad , \tag{27}$$

using the initial set of coefficients $\mathbf{c}_i^{(0)}$. Note that the sum in Eq. (27) is only over the occupied MOs. The coefficients are assumed to be normalized properly. As a test of normalization use the following relation:

$$\frac{N}{2} = \sum_{i=1}^{M} \sum_{j=1}^{M} S_{ij} P_{ij} \tag{28}$$

10. Compute the closed-shell Fock matrix:

$$F_{ij} = H_{ij} + G_{ij} \quad , \tag{29}$$

where the two-electron part of the Fock matrix, $\mathbf{G}$, is

$$G_{ij} = \sum_{m=1}^{M} \sum_{n=1}^{M} P_{mn}(2\langle\phi_i\phi_m|\frac{1}{r_{12}}|\phi_j\phi_n\rangle - \langle\phi_i\phi_m|\frac{1}{r_{12}}|\phi_n\phi_j\rangle) \quad , \tag{30}$$

which includes the electron repulsion integrals that you evaluated in Item 5. Notice that $\mathbf{G}$ depends on the density matrix, which, in turn, depends on the coefficients. The coefficients, and thus $\mathbf{P}$ and $\mathbf{G}$, will change every iteration until convergence.

11. Evaluate the electronic Hartree-Fock energy

$$E_{\mathrm{e}} = \sum_{i=1}^{M} \sum_{j=1}^{M} (2H_{ij} + G_{ij})P_{ij} \tag{31}$$

and the total energy, which is a sum of the electronic energy and the nuclear repulsion energy:

$$E_{\mathrm{HF}} = E_{\mathrm{e}} + \frac{1}{R} \tag{32}$$

12. Solve the Hartree-Fock equations (Eq. (25)). If this is the first iteration, use the new set of coefficients to repeat the procedure starting with Item 9. If this is the second or later iteration, compare the Hartree-Fock energy obtained in the previous Item to the value obtained in the previous iteration. If the values differ by more than 10^{-12} (in practice, convergence criteria depend on the type of calculation task), repeat starting with Item 9, else stop. Congratulations, the Hartree-Fock solver has converged!

4 Geometry Optimization and Vibrational Analysis

13. Starting at $R = 1$a.u., approximate the first and second derivative of the Hartree-Fock energy by

$$E'(R) \approx \frac{E(R+\delta) - E(R-\delta)}{2\,\delta} \tag{33}$$

$$E''(R) \approx \frac{E(R+\delta) + E(R-\delta) - 2E(R)}{\delta^2} \tag{34}$$

where $\delta = 0.001$ a.u. If the absolute magnitude of the first derivative is greater than 10^{-6} a.u., take a Newton-Raphson step:

$$R^{new} = R - \frac{E'(R)}{E''(R)} \tag{35}$$

and repeat, otherwise report the bond distance and the Hartree-Fock energy.

14. Evaluate the harmonic force constant at the equilibrium geometry and use it to compute harmonic vibrational frequency.

15. (Optional) Use an existing quantum chemistry program (e.g. ORCA) to verify your results for R_e and $\tilde{\nu}$.

16. Using your Hartree-Fock energy code, evaluate the bond dissociation energy with the help of the analytic energy of the hydrogen atom.

17. Compare the computed bond distance, vibrational frequency, and dissociation energy to the literature values. Comment on the discrepancy.

5   Wave function analysis (optional)

18. Plot the highest occupied and lowest unoccupied orbitals along the line connecting the nuclei.