

The goal of this Project is to carry out a complete LCAO Hartree-Fock computation on a molecular system without using any external data other than the molecular geometry and a basis set. To simplify our work we will only consider primitive (non-contracted) spherical (s-type, or zero angular momentum) Gaussians; extension to higher angular momenta is straightforward but tedious and in practice it is recommended that an existing Gaussian integrals library, like Libint, is used.

As a reference example for checking your work we will use the hydrogen molecule at bond distance of 1.4 bohr in an *uncontracted* 3-21G basis, denoted as 3-21G(uc). Each hydrogen atom thus carries 3 primitive spherical Gaussians with exponents 5.447178, 0.82454724, and 0.18319158, for a total of 6 basis functions in the molecule.

A primitive Gaussian with orbital exponent  $\alpha_a$  centered on  $\mathbf{A} \equiv \{A_x, A_y, A_z\}$  and nonnegative integer quantum numbers  $\mathbf{a} \equiv \{a_x, a_y, a_z\}$  will be denoted as

$$G(\mathbf{r}; \alpha_a, \mathbf{A}, \mathbf{a}) \equiv (x - A_x)^{a_x} (y - A_y)^{a_y} (z - A_z)^{a_z} \exp(-\alpha_a |\mathbf{r} - \mathbf{A}|^2). \quad (1)$$

For convenience, we will use bracket notation and short-hand  $G(\mathbf{r}; \alpha_a, \mathbf{A}, \mathbf{a}) \equiv |a_{\alpha_a, \mathbf{A}}\rangle$ , or simply  $|\mathbf{a}\rangle$ . The quantum numbers of an s Gaussian are zero, hence it will be denoted as  $|\mathbf{0}_\mathbf{A}\rangle$ . We will repeatedly take advantage of the important property of primitive Gaussians that they factorize into Cartesian components:

$$|\mathbf{a}\rangle \equiv |a_x\rangle \times |a_y\rangle \times |a_z\rangle, \quad (2)$$

where  $|a_x\rangle \equiv (x - A_x)^{a_x} \exp(-\alpha_a (x - A_x)^2)$  is a 1-dimensional Gaussian with  $a_x$  quanta, exponent  $\alpha_a$ , and centered at  $A_x$ .

### 1.

Note that it is convenient to use Gaussians that are *unit-normalized*; we'll denote the unit-normalized analog of the Gaussian in Eq. (1) by  $|\tilde{\mathbf{a}}\rangle$  (Gaussians  $|\mathbf{a}\rangle$  without the normalization constant are sometimes referred to as *unnormalized*). In practice, formulas for Gaussian integrals are often given in terms of unnormalized Gaussians, with the result to be scaled by the normalization constants if the integrals over normalized Gaussians is desired.

The normalization constant for a primitive spherical Gaussian is

$$N(\alpha) = \left(\frac{2\alpha}{\pi}\right)^{3/4}, \quad (3)$$

Write a function that takes  $\alpha$  as input and returns the normalization constant.

EXTRA: Derive this formula by solving for  $N$  from

$$1 = \langle \tilde{\mathbf{0}}_\mathbf{A} | \tilde{\mathbf{0}}_\mathbf{A} \rangle \equiv |N(\alpha)|^2 \langle \mathbf{0}_\mathbf{A} | \mathbf{0}_\mathbf{A} \rangle = |N(\alpha)|^2 \int (G(\mathbf{r}; \alpha_a, \mathbf{A}, 0, 0, 0))^2 d\mathbf{r} \quad (4)$$

The three-dimensional integral can be evaluated in Cartesian coordinates because primitive Gaussians factorize into  $x$ ,  $y$ , and  $z$  components, e.g.

$$\begin{aligned} \langle \mathbf{0}_\mathbf{A} | \mathbf{0}_\mathbf{A} \rangle &= \int_{\mathbb{R}^3} (G(\mathbf{r}; \alpha_a, \mathbf{A}, 0, 0, 0))^2 d\mathbf{r} = \int_{\mathbb{R}^3} \exp(-2\alpha_a |\mathbf{r} - \mathbf{A}|^2) d\mathbf{r} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(-2\alpha_a (x - A_x)^2) \exp(-2\alpha_a (y - A_y)^2) \exp(-2\alpha_a (z - A_z)^2) dx dy dz \\ &= \int_{-\infty}^{\infty} \exp(-2\alpha_a (x - A_x)^2) dx \times \int_{-\infty}^{\infty} \exp(-2\alpha_a (y - A_y)^2) dy \times \int_{-\infty}^{\infty} \exp(-2\alpha_a (z - A_z)^2) dz \\ &= \langle 0_{A_x} | 0_{A_x} \rangle \langle 0_{A_y} | 0_{A_y} \rangle \langle 0_{A_z} | 0_{A_z} \rangle \end{aligned} \quad (5)$$

where  $|0_{A_x}\rangle \equiv \exp(-\alpha_a (x - A_x)^2)$  is a 1-dimensional primitive Gaussian with 0 quantum number.

*Solution:*

□

2.

Implement a function to compute overlap integrals over unnormalized primitive Gaussians of arbitrary quantum numbers:

$$\langle \mathbf{a} | \mathbf{b} \rangle \equiv \int_{\mathbb{R}^3} G(\mathbf{r}; \alpha_a, \mathbf{A}, \mathbf{a}) G(\mathbf{r}; \alpha_b, \mathbf{B}, \mathbf{b}) d\mathbf{r} = S(\alpha_a, A_x, a_x, \alpha_b, B_x, b_x) S(\alpha_a, A_y, a_y, \alpha_b, B_y, b_y) S(\alpha_a, A_z, a_z, \alpha_b, B_z, b_z). \quad (6)$$

where

$$S(\alpha_a, A, a, \alpha_b, B, b) = \exp(-\alpha_a \alpha_b (A - B)^2 / \gamma) \sum_{i=0}^{(a+b)/2} f_{2i}(a, b, P - A, P - B) \frac{(2i-1)!!}{(2\gamma)^i} \left( \frac{\pi}{\gamma} \right)^{1/2} \quad (7)$$

is the overlap of 1-dimensional Gaussians. In Eq. (7)

$$\gamma \equiv \alpha_1 + \alpha_2 \quad (8)$$

$$P \equiv \frac{\alpha_a A + \alpha_b B}{\gamma} \quad (9)$$

$$f_k(l_1, l_2, L_1, L_2) = \sum_{i=\max(0, k-l_2)}^{\min(l_1, k)} L_1^{l_1-i} \binom{l_1}{i} L_2^{l_2-k+i} \binom{l_2}{k-i}. \quad (10)$$

Note that the upper limit of the summation in Eq. (7) should be interpreted as the ratio rounded down to the closest integer, i.e. if  $a + b = 3$  then the sum is over  $i \in \{0, 1\}$ .

Implement functions  $f_k(l_1, l_2, L_1, L_2)$  and  $S(\alpha_a, A, a, \alpha_b, B, b)$ . Use these functions to implement a function that takes two primitive Gaussians as input and returns their overlap.

*Solution:*

□

3. Write a function to evaluate the overlap of two *unit-normalized spherical* Gaussians using the general overlap function you just completed and the normalization constant function. Validate your function against the following analytic expression that can be obtained from Eq. (7) by setting  $\mathbf{a} = \mathbf{b} = \mathbf{0}$ :

$$\langle \tilde{\mathbf{0}}_{\mathbf{A}} | \tilde{\mathbf{0}}_{\mathbf{B}} \rangle = N(\alpha_a) N(\alpha_b) \exp(-\alpha_a \alpha_b |\mathbf{A} - \mathbf{B}|^2 / \gamma) \left( \frac{\pi}{\gamma} \right)^{3/2}. \quad (11)$$

Write a function that takes a *set* of primitive spherical Gaussians and evaluates the overlap matrix in their basis. Use this function to compute the matrix of overlap integrals for the reference example. Note that the matrix is symmetric, hence only  $21 = 6 \times 7/2$  integrals (diagonal + lower or upper triangle) are unique.

EXTRA: derive Eq. (12) by factorizing it into 1-dimensional overlaps,

$$\langle \tilde{\mathbf{0}}_{\mathbf{A}} | \tilde{\mathbf{0}}_{\mathbf{B}} \rangle \equiv N(\alpha_a) N(\alpha_b) \langle 0_{A_x} | 0_{B_x} \rangle \langle 0_{A_y} | 0_{B_y} \rangle \langle 0_{A_z} | 0_{B_z} \rangle, \quad (12)$$

and evaluating the 1-dimensional overlaps analytically (Mathematica can do it trivially).

*Solution:*

The reference overlap matrix:

$$\begin{pmatrix} 1. & 0.55559 & 0.211372 & 0.00480435 & 0.136509 & 0.149343 \\ 0.55559 & 1. & 0.677429 & 0.136509 & 0.445724 & 0.504981 \\ 0.211372 & 0.677429 & 1. & 0.149343 & 0.504981 & 0.835665 \\ 0.00480435 & 0.136509 & 0.149343 & 1. & 0.55559 & 0.211372 \\ 0.136509 & 0.445724 & 0.504981 & 0.55559 & 1. & 0.677429 \\ 0.149343 & 0.504981 & 0.835665 & 0.211372 & 0.677429 & 1. \end{pmatrix} \quad (13)$$

□

4. Show that the matrix element of the kinetic energy operator can be expressed in terms of the

$$\langle \tilde{\mathbf{0}}_{\mathbf{A}} | \hat{T} | \tilde{\mathbf{0}}_{\mathbf{B}} \rangle \equiv -\frac{1}{2} N(\alpha_a) N(\alpha_b) \langle \mathbf{0}_{\mathbf{A}} | \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) | \mathbf{0}_{\mathbf{B}} \rangle \quad (14)$$

$$= N(\alpha_a) N(\alpha_b) (T_x + T_y + T_z) \quad (15)$$

where

$$T_x \equiv \alpha_b \langle \mathbf{0}_{\mathbf{A}} | \mathbf{0}_{\mathbf{B}} \rangle - 2\alpha_b^2 \langle \mathbf{0}_{\mathbf{A}} | \{2, 0, 0\}_{\mathbf{B}} \rangle \quad (16)$$

and  $|\{2, 0, 0\}_{\mathbf{B}}\rangle \equiv (x - B_x)^2 \exp(-\alpha_b |\mathbf{r} - \mathbf{B}|^2)$  is a primitive  $d_{x^2}$  Gaussian centered at  $\mathbf{B}$  with exponent  $\alpha_b$ .

Write a function that takes two primitive spherical Gaussians and returns the kinetic energy integral. Write a function that takes a *set* of primitive spherical Gaussians and evaluates the kinetic energy matrix in their basis. Use this function to compute the matrix of kinetic energy integrals for the reference example.

*Solution:*

The reference kinetic energy matrix:

$$\begin{pmatrix} 8.17077 & 1.19365 & 0.112385 & -0.100447 & 0.01884 & 0.0610161 \\ 1.19365 & 1.23682 & 0.30462 & 0.01884 & 0.254303 & 0.182601 \\ 0.112385 & 0.30462 & 0.274787 & 0.0610161 & 0.182601 & 0.202147 \\ -0.100447 & 0.01884 & 0.0610161 & 8.17077 & 1.19365 & 0.112385 \\ 0.01884 & 0.254303 & 0.182601 & 1.19365 & 1.23682 & 0.30462 \\ 0.0610161 & 0.182601 & 0.202147 & 0.112385 & 0.30462 & 0.274787 \end{pmatrix} \quad (17)$$

□

5. Write a function that takes two primitive spherical Gaussians and the list of nuclei (each nucleus has a charge  $Z_C$  and its position vector  $\mathbf{C}$ ) and returns the nuclear attraction energy integral:

$$\langle \tilde{\mathbf{0}}_{\mathbf{A}} | \hat{V}_{ne} | \tilde{\mathbf{0}}_{\mathbf{B}} \rangle = N(\alpha_a) N(\alpha_b) \sum_{\mathbf{C}}^{\text{all nuclei}} \langle \mathbf{0}_{\mathbf{A}} | \left( -\frac{Z_C}{r_C} \right) | \mathbf{0}_{\mathbf{B}} \rangle \quad (18)$$

$$= -N(\alpha_a) N(\alpha_b) \frac{2\pi \exp(-\alpha_a \alpha_b |\mathbf{A} - \mathbf{B}|^2) / \gamma}{\gamma} \sum_{\mathbf{C}} Z_C F(\gamma |\mathbf{P} - \mathbf{C}|^2) \quad (19)$$

where  $\mathbf{P} = \frac{\alpha_a \mathbf{A} + \alpha_b \mathbf{B}}{\gamma}$  and  $F_0(T)$  is the so-called Boys function (also sometimes referred to as an incomplete gamma function):

$$F_0(T) \equiv \int_0^1 e^{-Tt^2} dt = \begin{cases} \frac{\pi^{1/2}}{2\sqrt{T}} \text{erf}(\sqrt{T}) , & T > 0 \\ 1, & T = 0 \end{cases} \quad (20)$$

$\text{erf}(x)$  in Eq. (20) is the standard error function; it is available as part of the math library in C++ as well as in other computer languages. Note that  $F_0(T)$  is the  $m = 0$  case of  $F_m(T) \equiv \int_0^1 t^{2m} e^{-Tt^2} dt$ , which is necessary to evaluate Coulomb integrals over non-spherical Gaussians.

Write a function that takes a *set* of primitive spherical Gaussians and the list of nuclei, and returns the nuclear attraction energy matrix in their basis. Use this function to compute the matrix of nuclear attraction energy integrals for the reference example.

*Solution:*

The reference nuclear potential matrix:

$$\begin{pmatrix} -4.43868 & -1.96686 & -0.71692 & -0.0137118 & -0.472354 & -0.50857 \\ -1.96686 & -2.15546 & -1.22855 & -0.472354 & -1.01417 & -0.954841 \\ -0.71692 & -1.22855 & -1.23247 & -0.50857 & -0.954841 & -1.07674 \\ -0.0137118 & -0.472354 & -0.50857 & -4.43868 & -1.96686 & -0.71692 \\ -0.472354 & -1.01417 & -0.954841 & -1.96686 & -2.15546 & -1.22855 \\ -0.50857 & -0.954841 & -1.07674 & -0.71692 & -1.22855 & -1.23247 \end{pmatrix} \quad (21)$$

□

6. DETOUR: solve the Schrödinger equation for the 1-e atom using the functionality we developed. For the wave function expressed as a linear combination of primitive spherical Gaussians centered at the nucleus the optimal values of their coefficients are obtained via the variational principle, by solving the generalized eigenvalue problem:

$$\mathbf{H}\mathbf{c} = E\mathbf{S}\mathbf{c} \quad (22)$$

where  $\mathbf{S}$  is the overlap matrix in the basis of Gaussians,  $\mathbf{H}$  is the Hamiltonian matrix in the same basis (= the kinetic energy matrix + the nuclear attraction energy matrix),  $\mathbf{c}$  is the vector of coefficients of the Gaussians, and the lowest eigenvalue  $E$  is the upper bound to the ground-state energy.

Use an even-tempered set of Gaussians with exponents  $\{...1/4, 1/2, 1, 2, 4, 8, 16, ...\}$  (start with a small subset and use as large of a subset as possible) to determine the energy of the hydrogen atom. Plot the wave function that you calculated and compare to the exact solution. Where is the error in the wave function localized? How do the error in the energy and the norm of the wave function error behave as a function of the basis set size?

*Solution:*

□

7. The electron-repulsion integral between two pairs of spherical primitive Gaussians reduces to:

$$\begin{aligned} (\tilde{\mathbf{0}}_A \tilde{\mathbf{0}}_B | \tilde{\mathbf{0}}_C \tilde{\mathbf{0}}_D) &\equiv \langle \tilde{\mathbf{0}}_A \tilde{\mathbf{0}}_C | \tilde{\mathbf{0}}_B \tilde{\mathbf{0}}_D \rangle = \\ &= N(\alpha_a)N(\alpha_b)N(\alpha_c)N(\alpha_d) \int e^{-\alpha_a|\mathbf{r}_1-\mathbf{A}|^2} e^{-\alpha_b|\mathbf{r}_1-\mathbf{B}|^2} \frac{1}{r_{12}} e^{-\alpha_c|\mathbf{r}_2-\mathbf{C}|^2} e^{-\alpha_d|\mathbf{r}_2-\mathbf{D}|^2} d\mathbf{r}_1 d\mathbf{r}_2 \\ &= N(\alpha_a)N(\alpha_b)N(\alpha_c)N(\alpha_d) \frac{2\pi^{5/2} K_{ab} K_{cd}}{\gamma_p \gamma_q (\gamma_p + \gamma_q)^{1/2}} F_0(|\mathbf{P} - \mathbf{Q}|^2 \gamma_p \gamma_q / (\gamma_p + \gamma_q)) \end{aligned} \quad (23)$$

where

$$\gamma_p = \alpha_a + \alpha_b \quad (24)$$

$$\gamma_q = \alpha_c + \alpha_d \quad (25)$$

$$\mathbf{P} = \frac{\alpha_a \mathbf{A} + \alpha_b \mathbf{B}}{\gamma_p} \quad (26)$$

$$\mathbf{Q} = \frac{\alpha_c \mathbf{C} + \alpha_d \mathbf{D}}{\gamma_q} \quad (27)$$

$$K_{ab} = e^{-\alpha_a \alpha_b |\mathbf{A} - \mathbf{B}|^2 / \gamma_p} \quad (28)$$

$$K_{cd} = e^{-\alpha_c \alpha_d |\mathbf{C} - \mathbf{D}|^2 / \gamma_q}, \quad (29)$$

and  $F_0(T)$  is the Boys function introduced in Eq. (20). Write a function that takes four spherical Gaussians as input and returns the electron repulsion integral.

All possible electron repulsion integrals is a set of  $6^4$  integrals (not all of them unique) which can be represented as a order-4 tensor. To avoid checking so many integrals, only evaluate all electron repulsion integrals for which  $\tilde{\mathbf{0}}_A = \tilde{\mathbf{0}}_C$  and  $\tilde{\mathbf{0}}_B = \tilde{\mathbf{0}}_D$ ; these integrals form a 6 by 6 matrix.

*Solution:*

The  $(\tilde{\mathbf{0}}_A \tilde{\mathbf{0}}_B | \tilde{\mathbf{0}}_A \tilde{\mathbf{0}}_B)$  matrix:

$$\begin{pmatrix} 2.63355 & 1.35042 & 0.6718 & 0.714283 & 0.701563 & 0.543935 \\ 1.35042 & 1.02462 & 0.617812 & 0.701563 & 0.662712 & 0.515466 \\ 0.6718 & 0.617812 & 0.482956 & 0.543935 & 0.515466 & 0.430883 \\ 0.714283 & 0.701563 & 0.543935 & 2.63355 & 1.35042 & 0.6718 \\ 0.701563 & 0.662712 & 0.515466 & 1.35042 & 1.02462 & 0.617812 \\ 0.543935 & 0.515466 & 0.430883 & 0.6718 & 0.617812 & 0.482956 \end{pmatrix} \quad (30)$$

□

8. Now we know how to compute all integrals we need to solve the Hartree-Fock equations. The canonical Hartree-Fock molecular orbitals  $|\psi_i\rangle$  will be approximated as a linear combination of a finite basis set of AOs  $|\phi_\mu\rangle$  ( $\mu \in 0..n_{\text{AO}} - 1$ ) as:

$$|\psi_i\rangle \approx \sum_{\mu=0}^{n_{\text{AO}}-1} |\phi_\mu\rangle c_{\mu i}. \quad (31)$$

The expansion coefficient matrix  $c_{\mu i} \equiv (\mathbf{c})_{\mu i}$  is obtained by solving the canonical Hartree-Fock equations:

$$\mathbf{F}\mathbf{c} = \mathbf{S}\mathbf{c}\epsilon, \quad (32)$$

where  $\mathbf{S}$  is the AO basis overlap matrix,  $\epsilon$  is the diagonal matrix of orbital energies, and  $\mathbf{F}$  is the Fock operator matrix in AO basis. If we assume that for a closed-shell molecule the spatial parts of spin-up and spin-down orbitals are identical (this is known as the (spin-)restricted Hartree-Fock, or RHF), the Fock matrix is obtained as

$$F_{\mu\nu} = h_{\mu\nu} + 2J_{\mu\nu} - K_{\mu\nu}, \quad (33)$$

where core Hamiltonian ( $\mathbf{h}$ ), Coulomb ( $\mathbf{J}$ ), and exchange ( $\mathbf{K}$ ) matrices are expressed in terms of the AO basis integrals as

$$h_{\mu\nu} \equiv \langle \phi_\mu | \hat{T} | \phi_\nu \rangle + \langle \phi_\mu | \hat{V}_{\text{ne}} | \phi_\nu \rangle \quad (34)$$

$$J_{\mu\nu} \equiv \sum_{\rho,\sigma=0}^{n_{\text{AO}}-1} (\phi_\mu \phi_\nu | \phi_\rho \phi_\sigma) P_{\rho\sigma} \quad (35)$$

$$K_{\mu\nu} \equiv \sum_{\rho,\sigma=0}^{n_{\text{AO}}-1} (\phi_\mu \phi_\sigma | \phi_\rho \phi_\nu) P_{\rho\sigma}, \quad (36)$$

and the 1-electron reduced *density* matrix  $\mathbf{P}$  is expressed in terms of the coefficients of *occupied* Hartree-Fock orbitals as

$$P_{\rho\sigma} \equiv \sum_{i=0}^{n_{\text{occ}}-1} c_{\rho i} c_{\sigma i}. \quad (37)$$

Note that the summation in Eq. (37) is performed over the occupied orbitals only, which correspond to those eigenvectors  $c_{\mu i}$  that correspond to the  $n_{\text{occ}}$  lowest eigenvalues  $\epsilon_{ii}$ .

To solve the Hartree-Fock equations, let's start out by assuming that  $\mathbf{P} = 0$ ; this is known as the core Hamiltonian guess, since the initial value of the Fock matrix is just  $\mathbf{h}$ . Then we solve the Hartree-Fock equations, Eq. (32), use the occupied eigenvectors to construct the density matrix via Eq. (37), compute the Coulomb and exchange matrices via Eqs. (35) and (36), compute the Fock matrix via Eq. (33), solve the Hartree-Fock equations again, and so on until convergence. The iterations are stopped when the number of iterations exceed the predetermined maximum number, or the *Frobenius norm* of the difference between the density matrices in consecutive iterations is below some threshold (say,  $10^{-6}$ ). The electronic Hartree-Fock energy is obtained

$$E_{\text{HF}} \equiv \frac{1}{2} \sum_{\mu\nu} (h_{\mu\nu} + F_{\mu\nu}) P_{\mu\nu}. \quad (38)$$

The total molecular energy is obtained as the sum of the electronic Hartree-Fock energy plus the nuclear repulsion energy:

$$E_{\text{nuc}} \equiv \sum_{I < J}^{n_{\text{atom}}-1} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \quad (39)$$

Write a function to solve the Hartree-Fock equations that takes as input the AO basis set, the nuclear positions (for the nuclear-electron attraction and the nuclear repulsion), the density matrix convergence criterion, and maximum number of iterations. Determine the Hartree-Fock energy of the reference system.

*Solution:*

