
Numerical Algorithms Applied to Computational Quantum Chemistry
Homework 3: Build your own extended Huckel program

Prof. Martin Head-Gordon, with Zhe Wang (C279) and Xiao Liu (C179)

Due Feb. 29, 2024

1 MOLECULE INPUT AND BASIS FUNCTION SETUP

The first issue you must address is reading input so that you can set up the calculation. To keep things simple, your code should be able to handle a user-supplied hydrocarbon molecule (i.e. the only elements are C and H), with an even number of electrons, using the valence AO basis described below. Your GSI will supply coordinates and debugging results for three test cases:

- H_2 : there will be 2 basis functions ($N = 2$), and 2 electrons ($n = 1$).
- C_2H_2 : $N = 10$, $n = 5$.
- C_2H_4 : $N = 12$, $n = 6$.

Note that you could also have built these yourself: these are so-called “standard” geometries, where bond-lengths, R , and angles, A are given specified values. $R(\text{C} - \text{H}) = 1.10 \text{ \AA}$, $R(\text{C} - \text{C}) = 1.54 \text{ \AA}$, $R(\text{C} = \text{C}) = 1.34 \text{ \AA}$, $R(\text{C} \equiv \text{C}) = 1.21 \text{ \AA}$, and finally $R(\text{H} - \text{H}) = 0.74 \text{ \AA}$. Finally, we change the units of length to atomic units (Bohr), where $1 \text{ Bohr} = 0.52917706 \text{ \AA}$, which is the Bohr radius of the H atom (in case you were wondering!).

Your program should perform the following tasks:

- Read in the coordinates, in the format: E X Y Z for each atom, where E is the element (handle at least H and C).
- Evaluate the number of basis functions, N from the molecular formula, C_aH_b , where the relation is $N = 4a + b$. Your matrices, such as $\mathbf{S}, \mathbf{H}, \mathbf{X}$, etc, will be $N \times N$, so this will enable you to define them.

- Evaluate the number of electrons $2n = 4a + b$. Throw an error if the number of electron pairs $n = 2a + b/2$ is not an integer. Knowing n is necessary to evaluate the energy later.
- Build a list of the basis functions, which are contracted gaussians, as described in detail below. For each basis function, $\omega_\mu(\mathbf{r})$ (for $\mu = 1 \cdots N$) there will be (i) a center, \mathbf{R} , (ii) 3 quantum numbers, (l, m, n) , and (iii) information about 3 primitive functions: 3 exponents, α_k (given below), 3 corresponding contraction coefficients, d_k (given below), and 3 normalization constants, N_k^{lmn} (to be worked out by your code in q. 2 below). See Eqs. 1.1 and 1.2 below for definitions.

As a reminder, the atomic orbital (AO) basis functions used to describe electrons the molecular orbitals (MO's) of electronic structure theory are often built from fixed linear combinations ("contractions") of primitive gaussians. These AO basis functions are called contracted gaussian functions. The main reason that contractions are necessary is that gaussian functions do not have the correct exponential form near each atomic nucleus. But the matrix elements over exponential functions are intractable in some cases, so contracted gaussians are a very good compromise. Like we already discussed for primitives, the contracted gaussians are organized into shells made up of shells of primitives.

The primitive shells that contribute to a contracted shell of basis functions share common angular momentum, and a common center: they differ only in their radial decay (i.e. the exponent of the gaussian). We can thus write a contracted function as a linear combination of primitive gaussian functions (taken to be normalized):

$$\omega_\mu(\mathbf{r}) = \sum_k^K d_{k\mu} \omega_k(\mathbf{r}) \quad (1.1)$$

The coefficients, $d_{k\mu}$, mix together the primitive gaussian functions, which are given by:

$$\omega_k(\mathbf{r}) = N_k^{lmn} (x - X)^l (y - Y)^m (z - Z)^n \exp[-\alpha_k (\mathbf{r} - \mathbf{R})^2] \quad (1.2)$$

The normalization constants, N_k^{lmn} , are defined such that the overlap of each primitive with itself is equal to one:

$$S = \int_x \int_y \int_z \omega_k(\mathbf{r}) \omega_k(\mathbf{r}) = 1 \quad (1.3)$$

The 3 exponents and contraction coefficients for the 1s AO orbital of the H atom are:

```
3.42525091 0.15432897
0.62391373 0.53532814
0.16885540 0.44463454
```

Those for C follow. Note that the same exponents (left column) are reused for both the 2s function (contraction coefficients in 2nd column) and the p shell (of 3 functions; contraction coefficients in 3rd column are identical for each of the 3 p functions).

```
2.94124940 -0.09996723 0.15591627
0.68348310 0.39951283 0.60768372
0.22228990 0.70011547 0.39195739
```

You can get the corresponding information for other elements from the basis set exchange: <https://www.basissetexchange.org>, as well as check the data we've provided here! The name of this basis set is STO-3G, and if you get additional data to add to your code, make sure you remove the 1s core orbital which is not needed for the extended Huckel program you are making in this problem set.

2 OVERLAP MATRIX OF CONTRACTED GAUSSIAN ATOMIC ORBITALS

It is now time for you to leverage your code that does overlap integrals over *unnormalized* primitive gaussians from homework 2:

$$S^{AB} = \int_x \int_y \int_z \omega_A(\mathbf{r}) \omega_B(\mathbf{r}) \quad (2.1)$$

$$= S_x^{AB} S_y^{AB} S_z^{AB} \quad (2.2)$$

Depending on how you set things up for homework 2, you may want to do a bit of restructuring to prepare for the first task: *evaluating the normalization constants for each primitive gaussian in each basis function*. For a given *unnormalized* primitive gaussian, $\omega_A(\mathbf{r})$ (i.e. like Eq. 1.2, but omitting N_k^{lmn}):

$$\left(N_k^{lmn}\right)^2 S^{AA} = 1 \Rightarrow N_k^{lmn} = \frac{1}{\sqrt{S^{AA}}} \quad (2.3)$$

Run a loop over all your basis functions and get the normalization constants for the 3 primitives that make up each basis function and save them in an array.

It is now possible to evaluate the overlap integrals over *contracted functions*, which are the quantities that enter our overlap integral, S :

$$S_{\mu\nu} = \int_x \int_y \int_z \omega_\mu(\mathbf{r}) \omega_\nu(\mathbf{r}) \quad (2.4)$$

These contracted overlap integrals, using the definition, Eq. 1.1 are sums of primitive unnormalized overlap integrals, S^{kl} :

$$S_{\mu\nu} = \sum_k^3 \sum_l^3 d_{k\mu} d_{l\nu} N_{k\mu} N_{l\nu} S^{kl} \quad (2.5)$$

The primitive unnormalized integrals, S^{kl} , are exactly what your code from the last homework can evaluate! So (perhaps after a little restructuring), it has provided you with the normalization constants and the S^{kl} needed to evaluate entries of the overlap matrix via Eq. 2.5.

Some debugging may well be called for here after the satisfaction of finishing the code for the overlap matrix wears off... Start with the H_2 case, which is a 2×2 matrix, and you should have numbers virtually equal to 1 on the diagonal, and a non-zero positive result of about **0.6** on the off-diagonal. Then progress to the case of C_2H_2 which is a grander 10×10 matrix. Your GSI will provide reference values for the overlap matrix.

3 HAMILTONIAN ASSEMBLY AND DIAGONALIZATION

Given the ability to read in a user-supplied molecule and set up the atomic orbital basis (first problem), and make the corresponding overlap matrix (second problem), the next step is to complete your extended Huckel program that will evaluate the molecular orbitals, and the energy of the input hydrocarbon.

The first step is to make the matrix of the extended Huckel hamiltonian, **H**. The diagonal elements are assigned values corresponding to the energy needed to remove an electron from that orbital of the corresponding atom. The accepted parameters (in units of eV, where 1 eV = 96.487 kJ/mol = 23.061 kcal/mol) are:

$$h[H] = -13.6 \quad (3.1)$$

$$h[C(2s)] = -21.4 \quad (3.2)$$

$$h[C(2p)] = -11.4 \quad (3.3)$$

Parameters are available for other elements too! For instance:

$$h[N(2s)] = -26.0 \quad (3.4)$$

$$h[N(2p)] = -13.4 \quad (3.5)$$

$$h[O(2s)] = -32.3 \quad (3.6)$$

$$h[O(2p)] = -14.8 \quad (3.7)$$

Make sure your code can do hydrocarbons at least, but feel free to add the capability to do other elements (i.e. larger classes of molecules) as well!

With the diagonal elements assigned, the off-diagonal elements of the Huckel hamiltonian matrix (for $\mu \neq \nu$ are given in terms of the average of the corresponding diagonals multiplied by the same element of the overlap matrix, $S_{\mu\nu}$:

$$H_{\mu\nu} = \frac{1}{2} K \{h[A(\mu), L(\mu)] + h[A(\nu), L(\nu)]\} S_{\mu\nu} \quad (3.8)$$

The constant is conventionally chosen as $K = 1.75$, and $A(\mu)$ is the atom associated with basis function $\omega_\mu(\mathbf{r})$, while $L(\nu)$ is its angular momentum. Verify that this is correctly coded and debugged (Your GSI will provide the matrix for the 3 test cases – make sure you debug H_2 first, because it is easiest!)

The second step is to solve the generalized eigenvalue problem to obtain the molecular orbital coefficients, **C** and the eigenvalues ϵ . You have seen this many times in lecture:

$$\mathbf{HC} = \mathbf{SC}\epsilon \quad (3.9)$$

The steps to do this are as follows:

1. make the orthogonalization transformation: $\mathbf{X} = \mathbf{S}^{-1/2}$

2. form the hamiltonian in the orthogonalized basis: $\mathcal{H} = \mathbf{X}^T \mathbf{H} \mathbf{X}$
3. diagonalize: $\mathcal{H} \mathbf{V} = \mathbf{V} \epsilon$
4. form the MO coefficients: $\mathbf{C} = \mathbf{X} \mathbf{V}$

Thanks to the wonders of linear algebra libraries like Armadillo, your code for these steps can be remarkably compact, as your GSI will discuss and demonstrate to you in lab for a related example. Debug your code using the H_2 case where your matrices are only 2×2 in size: Your GSI will provide his reference results for you to compare against.

The third step is to assemble the total energy, based on occupying the lowest n MOs for a system of $2n$ valence electrons:

$$E = \sum_i^n 2\epsilon_i \quad (3.10)$$

For H_2 , $n = 1$, and it will be your first test case: your result should be around -35 eV.

To celebrate completing your program, try evaluating some chemical energy differences. Here are the first two that we'd like you to test:

1. The bond energy of H_2 . While this is a trivial case, it is still interesting because your code should be able to tell us why molecules form from atoms. You evaluate this by subtracting twice the energy of the H atom (-27.2 eV) from your computed H_2 energy. This result is also your next debugging case. The result should be around -4 eV (the minus sign indicates binding).
2. The energy difference for the chemical reaction: $\text{C}_2\text{H}_2 + \text{H}_2 \rightleftharpoons \text{C}_2\text{H}_4$. Your prediction (evaluated as $E(\text{C}_2\text{H}_4) - E(\text{C}_2\text{H}_2) - E(\text{H}_2)$) can be compared against the experimental enthalpy change for this reaction of about -174 kJ/mol. Note that your total energy for C_2H_2 should be around -178 eV, while that for C_2H_4 should be around -210 eV (pay attention to units and unit conversions!).

4 GOING FURTHER: BUT ONLY IF YOU WANT TO!

- Other applications of your code: one that would be interesting is the energy difference between the benzene molecule, C_6H_6 , and the corresponding straight-chain hydrocarbon, hexatriene, C_6H_8 . To make things balance you also need a hydrogen molecule. In other words, evaluate the energy difference: $E(\text{C}_6\text{H}_6) + E(\text{H}_2) - E(\text{C}_6\text{H}_8)$. The difference is an approximation to the so-called resonance energy of benzene.
- Extend your code to molecules beyond hydrocarbons. Parameters are available for almost all elements in the periodic table, as documented up to Cl in the figure below. Note that d-type orbitals are recommended for elements beyond Si!

Atom	# Elec.	$e_s=e_p$	e_d	$\alpha_s(\text{eV})$	$\alpha_p(\text{eV})$	$\alpha_d(\text{eV})$
H	1	1.3		-13.6		
Li	1	0.650		-5.4	-3.5	
Be	2	0.975		-10.0	-6.0	
B	3	1.300		-15.2	-8.5	
C	4	1.625		-21.4	-11.4	
N	5	1.950		-26.0	-13.4	
O	6	2.275		-32.3	-14.8	
F	7	2.425		-40.0	-18.1	
Na	1	0.733		-5.1	-3.0	
Mg	2	0.950		-9.0	-4.5	
Al	3	1.167		-12.3	-6.5	
Si	4	1.383	1.383	-17.3	-9.2	-6.0
P	5	1.600	1.400	-18.6	-14.0	-7.0
S	6	1.817	1.500	-20.0	-13.3	-8.0
Cl	7	2.033	2.033	-30.0	-15.0	-9.0

Figure 4.1: Extended Huckel parameters for the elements H-Cl. From Simons:
<https://chem.libretexts.org/@go/page/60564>