Lab 2

The chemical bond

The density matrix and Mulliken population analysis[1]

The electronic wave function, ψ, of an *n*-electron molecule is a function of the spatial and spin coordinates of the electrons of the molecule. We know that

$$\left|\psi\left(x_1,...,z_n,m_{s1},...,m_{sn}\right)\right|^2 dx_1dy_1dz_1....dx_ndy_ndz_n \tag{1}$$

is the probability of simultaneously finding electron 1 with spin m_{s1} in the volume $dx_1dy_1dz_1$ at $(x_{1,}y_1,z_1)$, electron 2 with spin m_{s2} in the volume $dx_2dy_2dz_2$ at $(x_{2,}y_2,z_2)$ and so on. The probability density ρ of finding an electron in the point (x,y,z) is

$$\rho(x,y,z) = n\sum_{all\ m_s}\int...\int\left|\psi(x,y,z,x_2,...,z_n,m_{s1},...,m_{sn})\right|^2 dx_2...dz_n \tag{2}$$

For the set of basis functions $\chi_1,\chi_2,...,\chi_b$ each molecular orbital ϕ_j has the form

$$\phi_j = \sum_{r=1}^{b} c_{rj}\chi_l = c_{1j}\chi_1 + c_{2j}\chi_2 + ... + c_{bj}\chi_b \tag{3}$$

In Hartree-Fock theory, the wave function is an antisymmetrized product of the one-electron molecular orbitals[2]. Thus, for the many-electron wave function, ρ is found by multiplying the probability-density function of each molecular orbital (MO) by the number of electrons occupying it and summing the results:

$$\rho\left(x,y,z\right) = \sum_j n_j\left|\phi_j\right|^2 \tag{4}$$

where the sum is over all the different orthogonal spatial MOs, and n_j is the number of electrons in the MO ϕ_j $(n_j = 0,1,2)$. So that

$$\rho = \sum_{j=1}^{m} n_j\phi_j^*\phi_j = \sum_{r=1}^{b}\sum_{s=1}^{b}\sum_{j=1}^{m} n_j c_{rj}^* c_{sj}\chi_r^*\chi_s = \sum_{r=1}^{b}\sum_{s=1}^{b} D_{rs}\chi_r^*\chi_s \tag{5}$$

where *m* is the number of MO.

The quantities D_{rs} are called ***density matrix elements*** and are defined as

---

[1] This theoretical background is based on the following reference:
I. N. Levine. Quantum Chemistry. 5th ed. 2000. Upper Saddle River, NJ: Prentice-Hall.

[2] As detailed in the theoretical background in the course book.

$$D_{rs} = \sum_{j=1}^{m} n_j c_{rj}^{*} c_{sj} \tag{6}$$

These matrix elements can be used for the population analysis of molecular systems. A widely used method to analyze the wave function is population analysis, introduced by ***Mulliken***. He proposed a method that apportions the electrons of an *n*-electron molecule into net populations n_r in the basis functions χ_r and overlap populations n_{r-s} for all possible pairs of basis functions. For simplicity, we shall assume that the c_{si}'s and χ_s's are real. The probability density associated with one electron in ϕ_j is $|\phi_j|^2$. Integrating $|\phi_j|^2$ over three-dimensional space and using the fact that ϕ_j and χ_s's are normalized we get

$$1 = c_{1j}^2 + c_{2j}^2 + \ldots + c_{bj}^2 + \sum_{r<s} 2c_{rj} c_{sj} S_{rs} \tag{7}$$

where the *S*'s are overlap integrals:

$$S_{rs} = \iint \chi_r \chi_s dv_r dv_s \tag{8}$$

Mulliken proposed that the terms in Eq (7) be apportioned as follows. One electron in the molecular orbital ϕ_j contributes c_{1j}^2 to the net population in χ_1, c_{2j}^2 to the net population in χ_2, etc., and contributes $2c_{1j}c_{2j}S_{12}$ to the overlap population of χ_1 and χ_2, $2c_{1j}c_{3j}S_{13}$ to the overlap population of χ_1 and χ_3, etc.

Let $n_{r,j}$ and $n_{r-s,j}$ symbolize the contributions of electrons in the MO ϕ_j to the net population in χ_r and to the overlap population between χ_r and χ_s, respectively. We have

$$n_{r,j} = n_j c_{rj}^2, \qquad n_{r-s,j} = n_j \left(2c_{rj} c_{sj} S_{rs}\right) \tag{9}$$

By summing over the occupied MOs, we obtain the Mulliken net population n_r in χ_r and the overlap population n_{r-s} for the pair χ_r and χ_s as

$$n_r = \sum_j n_{r,j} \text{ and } n_{r-s} = \sum_j n_{r-s,j} \tag{10}$$

Mulliken's matrix elements are defined as

$$M_{rr} = n_r \text{ and } M_{rs} = n_{r-s} \tag{11}$$

The sum of all the net and overlap populations equals the total number of electrons in the molecule:

$$\sum_r n_r + \sum_{r>s}\sum_s n_{r-s} = n \tag{12}$$

In this lab, we will use the orbitals analysis and the density matrix in order to identify single, double and triple bonds in molecular systems.

Example

We performed a population analysis calculation of ethyne molecule using the Restricted Hartree-Fock method and the 6-31G(d) basis set. The route section of the above calculation is:

```
# RHF STO-3G Pop=(Regular, bonding) SCF=Tight
```

The following section include some parts of the output file, remarks are in blue.

The population and bond analysis of ethyne (using the RHF method and the STO-3G minimal basis set)

```
**********************************************************************

            Population analysis using the SCF density.

 **********************************************************************
```

The following is the symmetries and the energies (Hartree) of the alpha occupied and virtual calculated molecular orbitals.

```
 Orbital Symmetries:
       Occupied  (SGG) (SGU) (SGG) (SGU) (SGG) (PIU) (PIU)
       Virtual   (PIG) (PIG) (SGU) (SGG) (SGU)
  The electronic state is 1-SGG.
 Alpha  occ. eigenvalues --  -11.00376 -10.99996  -0.96705  -0.70442  -
0.61205
 Alpha  occ. eigenvalues --   -0.36290  -0.36290
 Alpha virt. eigenvalues --    0.41049   0.41049   0.56507   0.72767
1.53352
```

In the following **Molecular Orbital Coefficients section**:

Columns 5-9 represents the molecular orbitals ϕ_j, where $j = 3,4,...,7$. The first row is the index number of the molecular orbital (*j*), the second row is the symmetry and the occupation of the molecular orbital and the third row is the eigenvalue (energy) of the molecular orbital (Hartree).

The syntax of next rows is as follows:

Column	Description
1	The index number (r) of the basis function χ_r.
2	The number of the atom that the basis function belongs to.
3	The symbol of the atom that the basis function belongs to.
4	The atomic orbital that the basis function approximates.
5-9	The coefficient of the basis function in the molecular orbital. That is, c_{rj}, where r is the index number of the basis function and j is the index number of the molecular orbital (Eq. (3)). For example, one can see that the coefficient of the seventh basis function $\chi_7 = 2S_{\text{second carbon atom}}$ in the fifth molecular orbital ϕ_5 is 0.06923. That is, $\phi_5 = \sum_{r=1}^{30} c_{r5}\chi_r$, where $c_{75} =$ 0.06923.

```
   Molecular Orbital Coefficients
                        3         4         5         6         7
                     (SGG)--O  (SGU)--O  (SGG)--O  (PIU)--O  (PIU)--O
   EIGENVALUES --    -0.96705  -0.70442  -0.61205  -0.36290  -0.36290
 1 1   C  1S         -0.19580  -0.12242   0.00196   0.00000   0.00000
 2        2S          0.48194   0.33522   0.06923   0.00000   0.00000
 3        2PX         0.00000   0.00000   0.00000   0.00000   0.61572
 4        2PY         0.00000   0.00000   0.00000   0.61572   0.00000
 5        2PZ        -0.17336   0.27415   0.44486   0.00000   0.00000
 6 2   C  1S         -0.19580   0.12242   0.00196   0.00000   0.00000
 7        2S          0.48194  -0.33522   0.06923   0.00000   0.00000
 8        2PX         0.00000   0.00000   0.00000   0.00000   0.61572
 9        2PY         0.00000   0.00000   0.00000   0.61572   0.00000
10        2PZ         0.17336   0.27415  -0.44486   0.00000   0.00000
11 3   H  1S          0.12147   0.33805   0.34830   0.00000   0.00000
12 4   H  1S          0.12147  -0.33805   0.34830   0.00000   0.00000
```

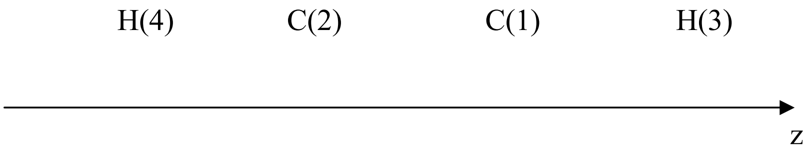
In the **density matrix** section the syntax is as follows:

The first row is the index of the basis function and the syntax of next rows as follows:

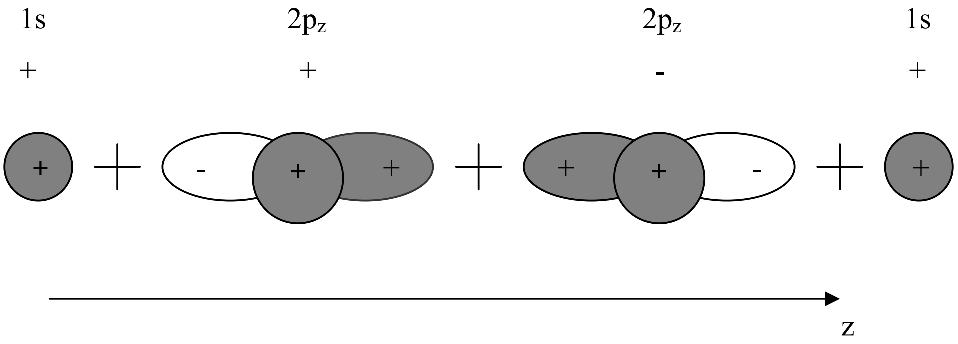
Column	Description
1	The index number (r) of the basis function χ_r .
2	The number of the atom that the basis function belongs to.
3	The symbol of the atom that the basis function belongs to.
4	The atomic orbital that the basis function approximates.
5-9	Is zero if the basis function of the row and the basis function of the column belong to the same atom, Otherwise, it is the density matrix elements defined in Eq. (6) $D_{rs} = \sum_{j=1}^{m} n_j c_{rj}^* c_{sj}$ where r is the index number of the basis function defined in the row and s is the index number defined in the column . For example, $D_{75} = \sum_{j=1}^{m} n_j c_{7j}^* c_{5j}$ = -0.28900, where χ_7 is approximation to 2S atomic orbital that belongs to the second carbon atom and χ_5 is approximation to 2PZ atomic orbital that belongs to the first carbon atom.

```
Bonding Mulliken population analysis:
      DENSITY MATRIX.
                           1         2         3         4         5
   1 1   C  1S          0.00000
   2        2S          0.00000   0.00000
   3        2PX         0.00000   0.00000   0.00000
   4        2PY         0.00000   0.00000   0.00000   0.00000
   5        2PZ         0.00000   0.00000   0.00000   0.00000   0.00000
   6 2   C  1S          0.04570  -0.12510   0.00000   0.00000   0.14378
   7        2S         -0.12510   0.24810   0.00000   0.00000  -0.28900
   8        2PX         0.00000   0.00000   0.75822   0.00000   0.00000
   9        2PY         0.00000   0.00000   0.00000   0.75822   0.00000
  10        2PZ        -0.14378   0.28900   0.00000   0.00000  -0.30554
  11 3   H  1S         -0.14065   0.39154   0.00000   0.00000   0.45317
  12 4   H  1S          0.03786  -0.06118   0.00000   0.00000   0.08238
```

The following is a schematic representation of ethyne nuclei according to the “standard orientation” of the Gaussian output:



1. We can see, for example, that the forth-molecular orbital is σ_u. σ denotes that the projection of the angular momentum on the molecular axis is zero for this molecular orbital. *u* tells us that inversion through the origin of coordinates will give the same orbital but with an opposite sign[3].
2. We can use the “population analysis using the SCF density” and the molecular orbital coefficients of the output in order to draw schematically the molecular orbitals. We only use atomic orbitals with *significant* coefficients (that is, larger than 0.1 or lower than –0.1). The following is a schematic representation[4] of some of the occupied molecular orbitals:
   a. The *third* molecular orbital (σ_g), would be



The 2p$_z$ orbitals of two C atoms have opposite directions and their wavefunctions have the same signs between the atoms.

This is a bonding orbital with respect to the C-C bond.

The wavefunctions of 2p$_z$ orbital C and the s one of H atoms have opposite signs between the atoms.

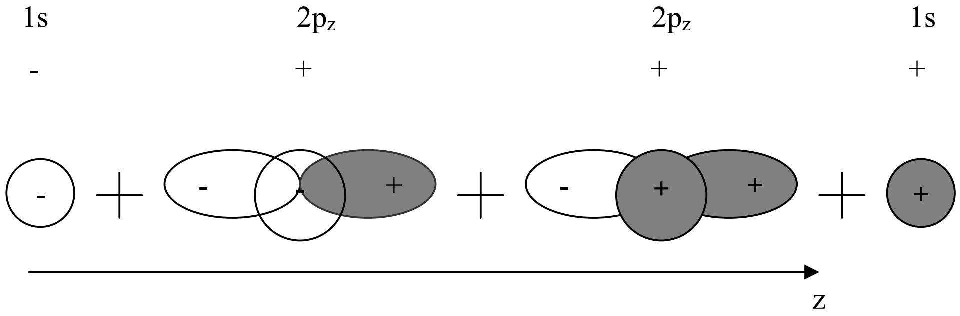
This is an anti-bonding orbital with respect to the C-H bonds.

---

[3] We use the letter “g” for molecular orbitals that are not changed under an inversion through the origin of coordinates.

[4] In this representation, gray represents plus and white represents minus.

b. The *fourth* molecular orbital (σ_u), would be



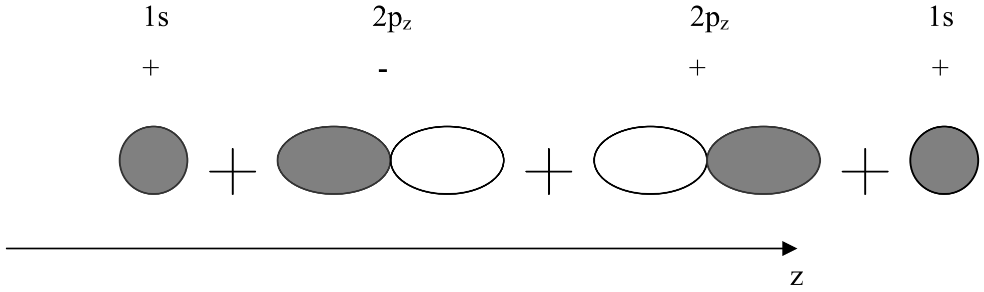
The $2p_z$ orbitals of two C atoms have the same directions and their wavefunctions have opposite signs between the atoms.

This is an anti-bonding orbital with respect to the C-C bond.

The wavefunctions of $2p_z$ orbital C and the s one of H atoms have the same signs between the atoms.

This is a bonding orbital with respect to the C-H bonds.

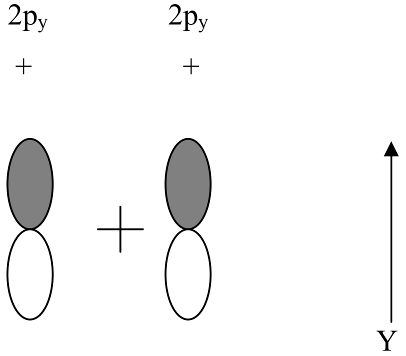
c. A schematic representation of the *fifth* molecular orbital (σ_g), would be



This is a bonding orbital with respect to the C-C bond.

This is also a bonding orbital with respect to the C-H bonds .

d. A schematic representation of the *sixth* molecular orbital (π_u), would be



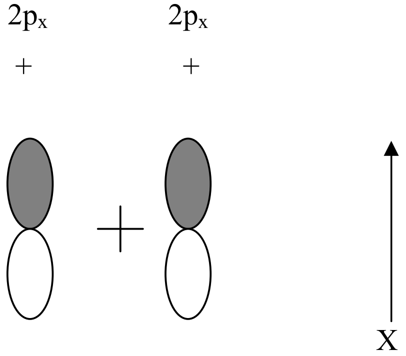
The $2p_y$ orbitals of two C atoms have the same directions and their wavefunctions have the same signs between the atoms.

This is a bonding orbital with respect to the C-C bond.

Orbitals of H atoms do not contribute to this molecular orbital.

This is a non-bonding orbital with respect to the C-H bond.

e. A schematic representation of the *seventh* molecular orbital (π_u), would be



This is a bonding orbital with respect to the C-C bond and a non-bonding orbital with respect to the C-H .

From the molecular orbital analysis, we see four bonding orbitals and one anti-bonding (with respect to the C-C bond). Therefore, the bond order of the C-C bond is three. To determine the bond order of the C-C bond, we take into account only molecular orbitals that are occupied by valence electrons, explain why.

We can also look at the "bonding Mulliken population analysis" in order to find the bond order of the C-C bond:

There are two π orbitals composed of two *2px* and two *2py* atomic orbitals of the two atoms and a σ bond composed of the *1s*, *2s* and *2pz* orbitals.

One can replace columns and replace rows in order to get separated blocks, each block represent a chemical bond. In this case, one can replace the third and the fifth columns and the tenth and the eighth rows in the original "bonding density matrix" above and get the following three separated blocks[5]:

C2\C1	1S	2S	2PZ	2PX	2PY
1S	0.04570	-0.12510	0.14378	0.00000	0.00000
2S	-0.12510	0.24810	-0.28900	0.00000	0.00000
2PZ	-0.14378	0.28900	-0.30554	0.00000	0.00000
2PX	0.00000	0.00000	0.00000	0.75822	0.00000
2PY	0.00000	0.00000	0.00000	0.00000	0.75822

The first block represents a σ bond, the second and third blocks represent two π bonds. According to this analysis, the order of the C-C bond in the ethyne molecule is three, as concluded by the MO analysis.

Part 1

Use Hartree-Fock method (RHF, ROHF or UHF[6]) and 6-31G(d)[7] basis set, in order to perform an energy optimization calculation and population analysis of oxygen, nitrogen and fluoride molecules. Use SCF=Tight and Pop=Full additional keywords.

Steps of calculation in GaussView:

- Use the GaussView guide to build a model for each molecule (document the distance between the atoms of the molecule before the geometry optimization).
- L-click the Calculate button in the GaussView toolbar and select Gaussian.
- Select 'Optimization to minimum' and 'calculate force constant – never' in the Job Type dialog box.
- The occupation states of the above molecules are the following:
  N_2: $(1\sigma_g)^2\ (1\sigma_u)^2(2\sigma_g)^2\ (2\sigma_u)^2(1\pi_u)^4(3\sigma_g)^2$
  O_2: $(1\sigma_g)^2\ (1\sigma_u)^2(2\sigma_g)^2\ (2\sigma_u)^2(1\pi_u)^4(3\sigma_g)^2(1\pi_g)^2$

---

[5] Of course, one need not do the actual replacements in order to identify the separated blocks. If two matrix elements are in the same row or the same column, they belong to the same block, and correspond to the same chemical bond.

[6] RHF- Restricted Hartree-Fock.
UHF – Unrestricted Hartree-Fock
ROHF-Restricted Open (Shell) Hartree-Fock.

[7] The 6-31G(d) basis set is sometimes indicated 6-31G*.

F_2: $(1\sigma_g)^2\ (1\sigma_u)^2(2\sigma_g)^2\ (2\sigma_u)^2(1\pi_u)^4(3\sigma_g)^2(1\pi_g)^4$

Thus, the total spin of the nitrogen and the fluoride molecules are zero and the total spin of the oxygen molecule is 1. Therefore, the spin multiplicity of the nitrogen and the fluoride molecules are singlet and the spin multiplicity of the oxygen is triplet.

- In the Method dialog box:

  Select ground state. For the nitrogen and fluoride calculations select the Restricted Hartree-Fock method. For the oxygen calculations use both Unrestricted Hartree-Fock and Restricted Open (Shell) Hartree-Fock methods. Select Charge - 0 for all molecules, Spin - Singlet for the nitrogen and the fluoride molecules and triplet for the oxygen molecule, Basis set – 6-31G(d).
- Insert some convenient title in the Title section (For example, geometry optimization and population analysis of the oxygen molecule with 6-31G(d) basis set).
- Press the Link 0 button to open a dialog box, save a checkpoint file (.chk file) in the sane directory of the input and output files by pressing Checkpoint file. This checkpoint file will enable you to plot the calculated molecular orbitals.
- In the General dialog box deselect any option.
- Insert the following line in the Additional Keywords section:

  SCF=Tight Pop=Full

  The last option will give a full Mulliken population analysis of all molecular orbitals.
- Save an input file with a meaningful name.
- Start the calculation by L-clicking the Submit button.
- You can open the Gaussian dialog box to check the calculation state.
- Once the calculation has terminated, you can open the output file (~.out file) in Notepad. Verify that the calculation has terminated successfully and that convergence was accomplished.
- Open the output file (~.log file) in GaussView, this file contains the optimized structure of the molecule. Document the optimized distance between the atoms of the molecule.
- L-click the Results button in GaussView toolbar, select Summary to display a summary of the calculation results.

In order to plot the calculated molecular orbitals:

- Open the checkpoint file (~.chk file) in GaussView.
- L-click on the Molecular Orbital Editor button on GaussView toolbar.
- L-click the Visualize button.
- Select Isovalue=0.1 in the Visualize dialog box.
- Highlight a molecular orbital in the molecular orbitals diagram.
- Press the Update button to plot the molecular orbitals.

Exercises:

1. Perform the calculations for the oxygen, nitrogen and fluoride molecules and compare the optimized distance between the atoms with distance in the original model (starting point of optimization).
2. What is the axis of each molecule according to the "standard orientation" of Gaussian output files.
3. Use the ROHF and the UHF methods in order to calculate the electronic structure of the oxygen molecule. Which of these methods gives a more accurate energy and which is better for orbital analysis? Why can't we use the RHF method in order to perform this calculation?
4. What is the symmetry, energy, and occupation number of the HOMO and LUMO molecular orbitals of each molecule?
5. What is the symmetry of occupied molecular orbitals? From which atomic orbitals are they consisted? (Write only atomic orbitals with substantial weight, larger than 0.1 or lower than –0.1).
6. Identify the bonded and anti-bonded occupied orbitals.
7. According to the occupied molecular orbitals, identify the bond-order (single, double or triple bonds).
8. Perform calculations of the fluoride, oxygen and nitrogen atoms. The spin multiplicity of the nitrogen atom is quartet, of the oxygen is triplet and of the fluoride is doublet. Use ROHF for fluoride and nitrogen, and both ROHF and UHF for oxygen. Calculate the dissociation energy $E_D=2E_{atom}-E_{mol}$ of each molecule. The energy of atom has to be calculated with the same version of HF (restricted of unrestricted) as the corresponding molecular energy. Compare your results with the literature.

9. According to your calculations are the above molecules bonded? If not, use the 6-31G(d) and the B3LYP[8] energy functional of the Density Functional Theory (DFT) method (to include electron-electron correlation) in order to calculate the dissociation energy of the molecules[9]. According to the DFT calculations, are these molecules bonded? Explain the obtained results.

---

[8] Use the B3LYP energy functional for closed shell systems and the UB3LYP (unrestricted B3LYP) for open shell systems.

Part 2:

Use the Hartree-Fock method and the 6-31G(d) basis set, in order to perform an energy optimization calculation and population analysis of ethane, ethene and ethyne molecules. Use SCF=Tight and Pop=(Full, Bonding) additional keywords.

Steps of calculation in GaussView:

- Build a model for each molecule using GaussView (document the distance between the atoms of the molecule before the geometry optimization).
- L-click the Calculate button in the GaussView toolbar and select Gaussian.
- Select 'Optimization to minimum' and 'calculate force constant – never' in the Job Type dialog box.
- In the Method dialog box:
  Select ground state, Restricted Hartree-Fock method for, Charge - 0, Spin – Singlet and Basis set – 6-31G(d).
- Insert some convenient title in the Title section
- In the General dialog box deselect any option.
- Insert the following line in the Additional Keywords section:
  SCF=Tight Pop=(Full, Bonding).
  Pop=Full – full Mulliken population analysis of all molecular orbitals.
  Pop=Bonding – perform a bonding analysis in addition to the standard analysis, this is Mulliken analysis in which only density terms involving pairs of basis functions on different canters are retained.
- Press the Link 0 button to open a dialog box, save a checkpoint file (.chk file) in the sane directory of the input and output files by pressing Checkpoint file. This checkpoint file will enable you to plot the calculated molecular orbitals.
- Save an input file with a meaningful name.
- Start the calculation by L-clicking the Submit button.
- You can open the Gaussian dialog box to check the calculation state.
- Once the calculation has terminated, you can open the output file (~.out file) in Notepad. Verify that the calculation has terminated successfully and that convergence was accomplished.

- Open the output file (~.LOG file) in GaussView, this file contains the optimized structure of the molecule. Document the optimized distance between the atoms of the molecule.
- L-click the Results button in GaussView toolbar, select Summary to display a summary of the calculation results.

In order to plot the calculated molecular orbitals:

- Open the checkpoint file (~.chk file) in GaussView.
- L-click on the Molecular Orbital Editor button on GaussView toolbar.
- L-click the Visualize button.
- Select Isovalue=0.1 in the Visualize dialog box.
- Highlight a molecular orbital in the molecular orbitals diagram.
- Press the Update button to plot the molecular orbitals.

Exercises:

1. Perform the calculations for the ethane, ethene and ethyne molecules and compare the optimized distance between the atoms with distance in the original model (starting point of optimization).
2. What is the total dipole moment of the molecules? Explain the obtained results.
3. What are the optimized angles and bond length of each molecule? Compare the C-C bond length of each molecule and explain the obtained results.
4. What are the energy and symmetry of the HOMO and LUMO molecular orbitals?
5. Tabulate the symmetry, energy, and occupation number for each occupied MO, atomic orbitals of C atoms yielding dominant contributions to it, and is the MO bonding/anti-bonding with respect to the CC bond. What can be said about the relation between energies of bonding and anti-bonding MOs created of the same AOs? Identify the single, double and triple C-C bonds of each molecule using the “Molecular Orbitals Coefficients”.

6. Use the “Bonding Mulliken Population - Density Matrix” in order to identify the single, double and triple C-C bonds in each case. (Neglect all the Density Matrix elements less than 0.1 by absolute value.)