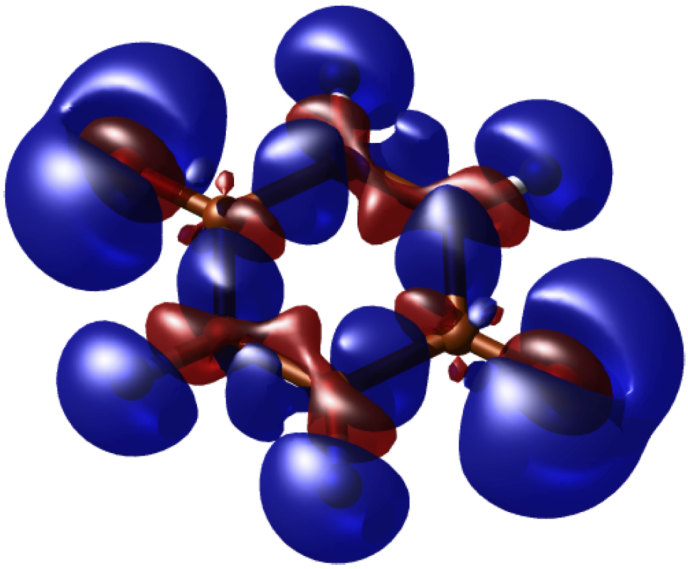
# PChem Lab 357: Spring 2019



## Excercise 5: Introduction to Computational Chemistry
## Excersise Problems for Gaussian16

**Adopted from Hands-On Workshop on Density-functional theory and beyond organized by FHI der MPG, Berlin.**
**Adopted for G16 by Mateusz Marianski, NYC, Feb 2019**
**Written fom FHI-aims code by Igor Y. Zhang et al.**

## *A quick summary of the exercises*

### *A guideline through the tutorial*

This tutorial aims to give a basic introduction to electronic structure calculations for very simple systems. As every quantum chemistry code has its own philosophy, this tutorial should familiarize you with general-purpose Gaussian16 (hereafter abbreviated as G16) software and GausView6 visualization tool. The experiments will also demonstrate predictive power of quantum-chemical calculations.

First, the basic structure of an input file to the G16 software will be explained. The second part will introduce scanning along the binding curve and computing observables. The third part introduces geometric optimization of a small molecule and how to assesse reliability of the result.

Prob. I: The hydrogen atom

Prob. II: Hydrofluoric acid: bond length and dipole moment

Prob. III: Hydronium cation: geometry relaxation, vibrations, and PES

### *First look at G16*

A good practice is to perform each calculations a separate directory. The calcuations are initialized by calling the g16 command on an input file:

```
g16 input &
```

By convention, always name the `input` file as `input.com`. The basic input file is composed as follow:

---

**A structure of a simple input file**:

%nproc=4
%mem=400MB
`# method basis-set`
# sp scf=tight
[ empty line ]
your-comment
[ empty line ]
`charge multiplicity`
`atom1 X Y Z`
`atom2 X Y Z`
...
`atomN X Y Z`
[ empty line ]

---

The above starts a calculation on two processors using 400 megabytes of memory. The output will be redirected to the `input.log` oputput file. This file contains the basic information and results of the calculation such as the total energy, atomic forces, and so forth. Additional output files might be generated according to the specified settings. Next paragraph explines the file step by step.

*`%nproc=4`* This keyword specifies the number of processors that will be employed for calculations. The laptops you use have 4 processors and you should use them all.

*`%mem=400MB`* This keyword manages the maximum memory usage during the calculations. The efficient memory specification is beyond this tutorial.

*`method`* Substitute the `method` with the method of choice for electron - electron interactions. In this tutorial, we will use several methods, namely Hartree-Fock (HF), Møller-Plesset second order perturbation theory (MP2) and few density-functionals. The details of each method will be covered in details during the lecture.

*`basis-set`* Substitute the `basis-set` word with the desired basis set which specifies a set of basis functions (for instance atomic orbitals, gaussian-type orbitals, plane waves) that will be used to express electronic configuration. The recommended basis sets are specified in each excercise.

*`sp`* `sp` orders gaussian to perform single-point calculations, i.e. energy evaluation of a specified structure using `method` and `basis-set`

*`scf=tight`* The Schrödinger equation is solved in self-consistent manner. The `scf=tight` specifies tight convergence criteria for the self-consistent cycle.

*`your-comment`* This line, surrounded by two empty lines, holds your comment.

*`charge`* The `charge` keyword should be substituted with a total charge of your system.

*`multiplicity`* The `multiplicity` keyword should be substituted with multiplicity of your system (2S+1, where S is total spin)

*`atomX X Y Z`* This block specifies the geometry of the system. You can use either symbol or atomic number for specifing the atom type, followed by it's cartesian coordinates in units of Angstoms [Å]. This block must be followed by an empty line.

Remember, there is no bonds (sticks) in quantum chemistry. The bonding is the result of the respective positions of atoms in space. The 'stick' visible in visualization program is simply a rendering for more intuitive display.

### *Additional tools and programs*

**Bash shell:**

A short list of the basic bash (command line) commands is given in Appendix I.

**vi:**

A number of introductions to the `vi` editor are available online. Two such examples are `https://www.openvim.com` and `https://vim-adventures.com/`.

**Scripts:**

For some exercises, scripts are required for dedicated tasks. All scripts you will need for this tutorial can be found in respective directories.

*Appendix I: Bash and vi*

Bash is a Unix shell and command language for the GNU Project and the default shell on Linux and OS X systems. We will use it to execute most programs and exercises. Below you find a list of the most import commands. ''<>' are used here to visually differentiate directories and files in the cheat-sheeet, do not use them (they redirect the flow, see below). It furthermore offers a full programming language (shell script) to automatize tasks e.g. via loops.

**Basic navigation:**

```
ls - list all files and folders
ls <dir-name> - list files in the directory.
ls -lh - Detailed list, Human readable
ls -l *.jpg - list jpeg files only

cd <folderName> - change directory
cd ..- go up one folder, tip: ../../../
```

**Basic file operations:**

```
cat file - show content of a file
head - from the top
tail - from the bottom

mkdir <directory-name>  - creates a new directory dir-
name (called folder in windows)
cp file1 file2 - copy file1 to file2
cp image.jpg <dir-name>/ - cp file1 to the dir-name directory
cp *.txt stuff/ - copy all of txt-type files to the directory stuff

mv file1 file2 - move (rename) file1 to file2
mv file1 <dir-name>/ - move file1 to directory dir-name
mv <folderName>/ .. - move directory up

rm file1 - delete file1
rm -r <dir-name> - delete directory dir-name
```

**Extract, sort and filter data:**

```
grep <someText> <fileName> - search for text in file
   -i - Doesn't consider uppercase words
grep -r <text> <folderName>/ - search for file names
with occurrence of the text
```

**Flow redirection and chain commands -redirecting results of commands:**

```
'>' at the end of a command to redirect the result to a file
'>>' to redirect the result to the end of a file
'|' at the end of a command to enter another one
'&' run the command in the background
```

**Basic control:**

```
TAB - auto completion of file or command
Up/Down - See previous commands
CTRL R - reverse search history
CTRL L - Clear the terminal
!! - repeat last command
```

**vi** is a terminal-based file edit program. By typing `vi` you open the program and create a new file that can be save later. By typing `vi file` , you open `file` to edit it. You can create new file by that you can edit immediately by this program.

The editor, despite its simplicity in appearance, is a very powerful terminal-based tool with numerous key-bindings. Therefore, be careful what you press. In order to start editing the file, you first need to press `i` ('insert') and then you can start typing. In order to save the file, press escape to exit the editing mode, `:` (shift + `;`) to enter the command mode in the bottom of the editor and type wq (for 'write quit'). Confirm with enter. If you want to quit without saving the file, type `q!`.

# *Part I: Basic electronic structure with FHI-aims*

## *Problem I: The hydrogen atom*

In this exercise, we will look at different basis sets uning the hydrogen atom. The hydrogen atom is the only non-trivial system for which the exact analytic solution is known. By the end of the first exercise, we will see how various computational methods compare to each other and to the exact solution. From a technical perspective, we will learn how to compose input files, run basic gaussian calculations, search for energy in the gaussian output, and perform basis set convergence tests.

### *Getting started - the hydrogen atom*

**Tasks**

1. First, go to the Problem_I directory by typing in the terminal `cd Desktop/Pchem_Lab/Problem_I`. There, create a test directory ( `mkdir dir-name`) and generate inside a simple `input.com` file by hand, which contains only a single hydrogen atom, using the example shown in the introduction. This corresponds to a single hydrogen atom in a hypothetical ideal gas phase. It is located at the origin of the coordinate system, although its position does not matter here.

2. For the method, use HF (Hartree-Fock method) and minimal STO-3G basis set which represents each available atomic orbital with 3 cotracted gaussian functions.

3. Now, inside the directory, run G16 using the command:

   ```
   g16 input.com &
   ```

   One the calculation has finished, open the `input.log` file with a text editor (type, for instance, `kwrite input.log`). If you find the line "`Normal termination of Gaussian`" near the end, then your calculation is converged. We are now interested in the total energy. Search for "SCF Done:" keywords inside the output file. You should find a following line:

   ```
   SCF Done: E(UHF) = XXXXX A.U. after X cycles
   ```

   This is the computed electronic energy of H atom using Hartree-Fock theory in STO-3G basis set. Compare it with the exact result for the hydrogen atom (0.5 Hartree $\approx$ 13.6057eV $\approx$ 313.7545 kcal mol$^{-1}$).

---

**TIP**: In later exercises, to find this value fast and efficiently, use the command
`grep `Done' input.log`
The `grep` command searches the input.log file looking for 'Done' phrase and outputs each line with such a word. Since the file contains only one such a phrase (it solved the electron-theory problem only once), there is only one such line. Please note that the capitalization matters.

---

4. Redo the calculation with different basis sets ( `cc-pVDZ` , `cc-pVTZ`, `cc-pVQZ`) by copying the input file into new directory and changing respective keyword in the input file. Search the output file to find out how many basis functions are actually used in the calculations. Then, plot the total energy as function of the basis set size. At which basis set the energy converges to exact solution?

### *Method performance*

Repeat the calculations different methods using prepared bash-script `performance.sh`. In the scirpt, replace the `METHODS` with follwoing list of density functionals:

- `SWVN`
- `PBEPBE`
- `PBE1PBE`

Next, execute the script by typing:

```
bash performance.sh
```

The scirpt will iterate over specified methods and tested basis sets (STO-3G, cc-pVXZ, where X=D,T,Q) and create nested directories for each method/basis set pair. Next, it will execute the calculations. Finally, it creates a `performance.dat` file which contains a list of basis sets, number of basis functions and computed energy for different methods. **Use this data to prepare a plot showing convergence of different methods to the exact value of 0.5 Ha.** Do all of them converge correctly to the same solution? The details of listed theoretical methods to evaluate electron-electron interactions and why they converge to different values for apparently trivial 1-electron system are beyond this tutorial and will be covered in lecture.

## *Problem II: Hydrofluoric acid (HF): bond length and dipole moment*

### *Hydrofluoric acid (HF)*

In the exercise, we will calculate the binding curve, atomization energy ($\Delta H_{at}$), and the dipole moment for hydrogen fluoride (HF) molecule with two methods. From a technical perspective, this exercise teaches how simple shell scripting can be used to make your (computational) life easier.

**Tasks:**

1. The first task of this exercise will be to find the equilibrium bond distance of HF from a series of calculations. Start by creating an input file (name it `input.temp`) which contains a F atom in the center of the coordinate system and a H atom at distance `DIST` along z-axis. `DIST` is a variable name for a H-F distance in different computational steps. Please note that HF is a neutral closed-shell system (change the multiplicity).

   The template file should specify the method and the basis set used for computation. In this excercise, use `HF` (Hartree-Fock) and `6-31G(d,p)` basis set. Use same keywords ( `scf=tight` and `sp`) from previous excercise.

2. Next, take a look at the bash script `run_scan.sh` which runs G16 calculations for a series of bond distances between 0.7Å and 1.3Å with 0.1Å steps, and a denser step width of 0.02 Å between 0.85 Å and 0.95 Å).

   ---

   In details, the script performs following tasks:

   - create a unique directory for each computations
   - copies your template input file
   - replace the bond distance place holder `DIST` with the bond distance and
   - start G16 calculations
   - greps for distance/energy and writes it to a respective file.

   ---

   Run the bash script `bash run.sh` which will run the calculations and return you a file with a bond length vs energy. Which bond length corresponds to the lowest energy? How does the bond length compare to the experimental bond length of 0.917Å?

3. To compare with experimental values, we compute the atomization energy ($\Delta H_{at}$). In order to calculate $\Delta H_{at}$, we will also need the total energy of the isolated H and F atoms. Compute the total energies for the single atoms using the methods `HF` and `6-31G(d,p)` basis set.

   Next, calculate the atomization energy ($\Delta H_{at}$) of HF by subtracting the free-atom energies from the predicted total energy of HF (i.e. the minimum total energy found when varying bond distances).

   $$\Delta H_{at} = E_{tot}^{HF} - E_{atom}^{H} - E_{atom}^{F} \quad (1)$$

   How does this compare to the experimental value of $\Delta H_{at} = 135.2\text{kcal mol}^{-1}$ (5.86 eV)?

4. Now, let us look at the dipole moment. Search for the corresponding line in the output file. You can use a grep function for this task:

   ```
   grep Dipole file-name -A1 | grep Tot | tail -n1 |
   awk '{print $8}'
   ```

   The above command is a great example of an ugly bash one-liner that does the job and you don't question it. You can test respective parts of the one-liner if you want to understand it better. How does the dipole at the equilibrium distance compare with the experimental value of 1.82 Debye? Plot the dipole moment vs. the bond distance. You will find a (mostly) linear correspondence. Do you expect this trend to continue at large distances? Why or why not?

5. Next, repeat the bond length determination using `PBE1PBE` method and same basis set. In order to do so, modify the input-template, mv all the results into `HF` directory ( `mkdir HF ; mv dist_ HF/`) and rerun the bash scirpt. In addition, you need to compute energies for H and F again using new method. How does the optimal bond length, atomization energy and dipole moment change? In the lab report, prepare a plot with both dissociation curves, dipole moments and computed atomization energies.

## *Problem III: Hydronium cation*

hydronium cation

### *Planar hydronium cation*

This exercise covers how to perform geometry optimizations. Specifically, we will relax the $H_3O^+$ molecule starting from an initial planar guess for the geometry.

1. Create the planar $H_3O^+$ geometry in geom_planar.xyz file. Use the following coordinates:

| | | | |
|---|---|---|---|
| O | 0.00 | 0.00 | 0.00 |
| H | 0.92 | −0.53 | 0.00 |
| H | −0.92 | −0.53 | 0.00 |
| H | 0.00 | 1.06 | 0.00 |

2. Create a template `input.com` file, using the template provided in the first problem. Use HF and 6-31G(d,p) basis set. Moreover, we want to relax the geometry and perform the vibrational analysis of the ion. Therefore, replace the `sp` keyword ('single-point') with `opt` ('optimization') and `freq` ('frequency'). Copy the geometry of the cation at the end of the input file. (`cat geom_planar.xyz >> input.com` and add and empty line at the end).

3. Run gaussian.

```
g16 input.com &
```

4. To visualize the results, open gaussview by typing `gaussview &` command. Select `file/open` and open the `output` file. What does the fully relaxed structure look like? Do you think that this is the structure of $H_3O^+$ in the gas phase? Save a picture of the ion.

5. Select Results/Summary and note down the total energy of the ion. Next, open Results/Vibrations and you will see a list of normal modes/vibrations sorted by the wavenumber [$cm^{-1}$]. Animate some of the vibration. You should see that one of the frequencies is negative - check to what normal modes it corresponds. Write down all vibrations for the lab report and indicate to what kind of molecular motion they correspond to.

### *Pyramidal hydronium cation*

hydronium cation

Next, repeat the calculations for a pyramidal hydronium cation:

| | | | |
|---|---|---|---|
| O | 0.00 | 0.00 | 0.00 |
| H | 0.92 | −0.53 | −0.66 |
| H | −0.92 | −0.53 | −0.66 |
| H | 0.00 | 1.06 | −0.66 |

Visualize the results with Gaussview. You should see the $H_3O^+$ in pyramidal conformation now. Select Results/Summary and note down the total energy of the ion and compare it with the planar structure. Which conformation has lower energy? Next, open Results/Vibrations to inspect normal modes/vibrations. If caclulations were done properly, all vibrations should have positive wavenumbers. Save the vibrations and descibe the motion they correspond to.

### *Potential-Energy Surface Scan*

In the final problem, we are going to inspect the potential-energy surface of the hydronium ion along its umbrella mode. In Problem planar cation you have seen that the negative (in fact, it's imaginary, `i` is dropped by convention) frequency corresponds to such 'umbrella' mode.

- Change to the `PES` directory. You will an template input file already prepared. If you open it, you should see that the xyz-cartesian coordinates has been replaced with a z-matrix. The z-matrix allows precise control of the geometry within single calculations.

```
O
X                   1 a
H                   1 a   2 HOX
H                   1 a   2 HOX   3 120.0
H                   1 a   2 HOX   3 -120.0
a=OH
HOX= 135. -1. 45
[ empty line ]
```

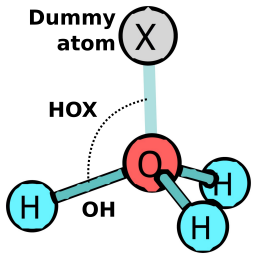



Figure 1: Definition of hydronium ion internal coordinates. The calculations perform a scan along HOX coordinates for all 3 hydrogens between 135 and 90 degrees. 120 degrees dihedral angle indicates the relative position of hydrogen atoms.

- The first column shows the bonding, second shows the angles between atoms and the third column specifies the dihedral angle. X is a dummy

(non-existent) atom that enables to control the umbrella motion. Figure below explains the z-matrix graphically. First, open the optimized pyramidal ion and measure the O-H distance and replace `OH` in the input file. Next, run the calculations. The calculations will perform a scan along the X-O-H angle, performing a single point calculations every 1 degree from 135 to 90 degrees. When the calcualtions are finished, visualize the results with gaussview (results/scan). Save the results, the Potenstial-Energy Surface (PES) along the specified coordinate (X-O-H angle). The PES for angles below 90 degrees is the mirror image. Use kcal/mol instead of Ha in the report.

- Localize the lowest-energy and transition structure along the PES and calculate the reaction barrier of the internal flip of the hydronium ion. Whereas the energy of the pyramidal ion is comparable with the minimum on the PES, the energy of the optimized planar cation and the maximum on PES is different. Why? Compare the geometries.
- In the final step, rerun the calculations using `MP2` method.

### *Lab report*

For the lab report, please prepare following data. This is bare minimum, there are couple of open questions in the text that you should try to asses.

1. Plot the Energy vs number of basis function for a hydrogen atom for all methods used in problem 1.
2. Plot the binding curves for HF using Hartree-Fock and PBE1PBE methods. Find the minimum distance and compute the atomization energy. Plot the dipole moment as a function of the bond distance.
3. Prepare tables listing molecular vibrations in hydronium ion in planar and pyramidal geometries. Assign the wavenumbers to molecular vibrations.
4. Plot the PES along the HOX coordinate using HF and MP2 methods. Use kcal/mol for the y-axis. Compute the heigh of the barrier separating two pyramidal structures.