

Introduction to Hartree-Fock calculations using ORCA and Chemcraft

In this exercise, you will use the ORCA software to carry out calculations of the electron wavefunctions of molecules on the computer cluster Elja. *ChemCraft* is used (on your local computer) to generate the input atomic coordinates and to visualize the results of the ORCA calculation. Follow the instructions in this file and provide answers to the questions raised in a word processor of your choice (for example Word or L^AT_EX). At the end create a PDF document and e-mail to the instructor at yla1@hi.is.

ORCA can be used to calculate the wavefunction of electrons in molecules by solving the time-independent Schrödinger equation using various different approximations. At the basic level, these are so-called ‘self-consistent-field’ (SCF) calculations, or, in the limit of an infinite (complete) basis set, Hartree-Fock (HF) calculations. Other levels of approximation are also available in this software, both higher and lower. But, we will start with HF because it is the approach that underlies the way most chemists think about atoms and molecules, and it defines the terminology used when discussing and presenting other levels of approximation.

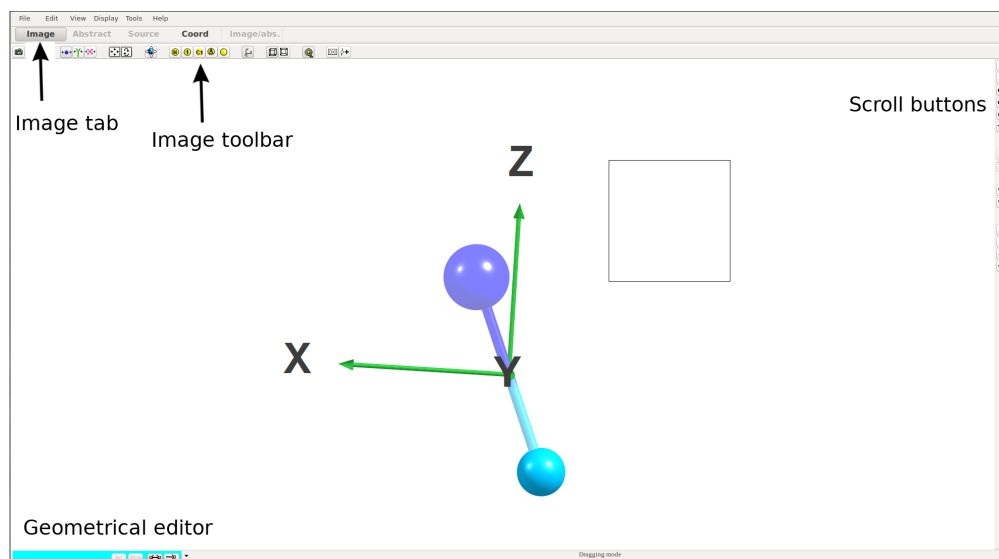
The instructions given here get you started using ORCA and Chemcraft. You will create the input file, run the calculation, retrieve the output and analyze it. For some of the later exercises, additional steps are necessary and they will then be described in detail.

In the following, three simple systems are studied, the H₂, HF and LiH molecules.

A. Molecular orbitals and electron density of H₂

In the first part of this exercise, you will calculate the electron wavefunction of the H₂ molecule using the Hartree-Fock (HF) approximation.

Chemcraft interface



In the uppermost part of the Chemcraft interface, there are ‘File’, ‘Edit’, ‘View’, ‘Display’, ‘Tools’ and ‘Help’ menus. Below those are the ‘Image’, ‘Abstract’, ‘Source’, ‘Image/abs’ and ‘Coords’ tabs. Further

below is the so-called image toolbar. The scroll buttons are to the right of the main window and the geometry editor is below the main window, see the figure on page 1.

Create the H₂ molecule

After starting Chemcraft, click on **Edit** and select **Add atom (Edit→Add atom)** to open the interactive periodic table. Select H and left-click on the main Chemcraft window. Now you should see a hydrogen atom represented by a light blue sphere. To add the second H atom, bonded to the former, left-click on the middle of the H atom. This will create a H₂ molecule with a reasonable initial guess for the bond length, r_{HH} . To construct larger molecules you can use the fragment database of Chemcraft, **Edit→Add fragment**. Once you have found your compound or fragment simply copy it to the clipboard (Ctrl+C) and then paste it to the screen by left-clicking on the main window.

If you make a mistake and need to delete atoms, left-click on an atom to select it and press Del key to remove it. You can also delete the whole molecule and start over by selecting **Edit→Select all** and **Edit→Delete selected**, or **File→Clear**.

To drag atoms, fragments or molecules, click on the icon second from the left in the image toolbar ('finger-pointing-at-the-red-sphere') or simply Ctrl+D. This will activate three new buttons on the image toolbar. Place the cursor over the three buttons to see what they do. Now drag one of the H atoms to distort your initial H₂ configuration given by Chemcraft (but do not rupture the chemical bond). When doing the calculations later in this project use this distorted configuration of the H₂ molecule to better see how the Hartree-Fock calculation can optimize the molecular geometry (i.e. find the lowest energy structure of the molecule).

Rotate, Translate and Change Display

The right mouse button rotates the molecule, or translates it if Shift key is pressed, or rotates the "camera" if Ctrl key is pressed. Furthermore, you can click the icons on the scroll button to rotate, translate and zoom. Try it and get a feeling for how to change the orientation of the molecule on the screen. If the molecule rotates in an 'odd' way, try 'Center' (or press Space key) and 'Best View' (or press Enter key) buttons on the image toolbar.

You can change the way the molecule is displayed in the "Display" menu. Try the different options and finally, select any ball&stick representation, which shows both atoms and bonds.

Click on the magnifying glass on the image toolbar to zoom in on the picture.

For analysis of the results of a calculation it is good to have the molecule in some 'standard orientation', i.e., centered at the origin of the coordinate system and in some well defined orientation. You can visualize the axis by **View→Show axes**.

Now save the geometry in '.xyz' file format, copy the saved file using SFTP to Elja and use it as input for an ORCA calculation. An alternative way that is simpler for such a small molecule is to click the coords tab, select the option 'same with 12-digit coordinates' from the 'show in format' drop-down menu. This will display the geometry of your molecule in Cartesian coordinates with the appropriate chemical symbol (symbol, x-coordinate, y-coordinate z-coordinate). Then, copy-paste the geometry into the ORCA input file that you will create on Elja (see next section).

Set up the Hartree-Fock Calculation

Start by running the SSH client (for Windows users) or terminal (Mac/Linux users) and log on to Elja (ssh -X username@elja.hi.is). For Windows users MobaXterm is an excellent SSH client. **In the following text,**

commands that are to be typed into the terminal window are denoted by >>.

After you login to Elja, you should be located in your home directory (i.e. /users/home/username/). To verify this run the command 'print-working-directory' (>> pwd). Then enter your CompChem directory (>> cd CompChem). If you did not create a directory named CompChem in the last practical session, type >> mkdir CompChem. This will create a directory called CompChem in your home directory. Now create a sub-directory (within CompChem) and give it the name lab2 (>> mkdir lab2). Change directory to lab2 (>> cd lab2) and create a sub-directory (within lab2) called partA. To do this you use the mkdir command (>> mkdir partA). Enter the new directory, >> cd partA. Now if you run 'print-working-directory' you should be located in /users/home/username/CompChem/lab2/partA/. Now to create an ORCA input file in the partA directory, open an empty file with the nano program; >> nano orca.inp. In principle, the prefix before the '.inp' extension can be any name that you wish but then you need to modify the sub.sh as well. This file will give ORCA instructions on what type of calculation to do, i.e., in the input file the system, the electronic structure method, the basis set and the job-type are specified.

In the figure below you can find an example of an ORCA input file. Inspect the structure of the file well.

The line beginning with "!" is called the simple-input line.

2 jobs defined here by keyword "opt" and "freq". Order does not matter. A geometry optimization will always first be performed followed by a frequency calculation.

```
! HF-3c opt freq
*xyz 0 1
O 0.00000000 0.00000000 1.00000000 1.494112000
H 0.00000000 0.00000000 1.975896152 1.485595483
H 0.00000000 0.024103848 1.485595483
*
```

Electronic structure Method is HF-3c
Later, we will use other methods than HF-3c and we will also need basis set keywords.

This is the coordinate block, it begins and end with a * character.
"xyz" keyword here tells ORCA that you are going to specify Cartesian coordinates (xyz) in Angstrom (default)

"0 1" refers to charge and spin multiplicity of the molecule.
(here the molecule is neutral (0) and calculated as a singlet).
Spin multiplicity is defined as 2S+1 where S is the spin of the molecule (S=0 for closed-shell nonradicals).

By choosing the molecule to be neutral (0) we are telling the program how many electrons to add. Because the nuclear charge is: 10 e+ as mentioned in the right block, choosing the total charge to be neutral means that we are telling ORCA to add 10 electrons to the system.
Total charge: 0 = 10e+ + 10e-

Cartesian coordinates in Angstrom, for each atom:
First column: element symbol (or atom number)
Columns 2 - 4: x-coordinate, y-coordinate, z-coordinate

The list of elements specified here determines the nuclear charge:
Oxygen: 8 e+
2 Hydrogens: 2e+
Total nuclear charge is here: 10 e+

The first line of the file is called the 'simple-input' line. This line is always started by an exclamation mark (!). The simplicity of the ORCA input is one of the main advantages that ORCA has over other quantum chemistry packages. In this line, the user should specify the level of approximation to use (i.e. the electronic structure method) and the job-type that ORCA should use for a given calculation.

Electronic structure method

There is a plethora of electronic structure methods implemented in ORCA. When choosing a method, we need to think of the system in question, the computational cost and the accuracy that is required. At this point in time, we will focus on the HF approximation.

SCF type

The SCF type can be specified in the simple-input line (not visible in the example above). It determines the way the electronic wavefunction is approximated. An unrestricted wavefunction (keyword in ORCA: UKS/UHF) is described by a Slater-determinant that can have different orbitals for the two spins of the electrons. A restricted wavefunction, on the other hand, uses the same orbitals for the electrons of both spins. This reduces the computational effort and usually gives a correct description of the ground state of 'normal' molecules with a spin multiplicity of zero. By default ORCA will chose restricted Slater determinant

for systems with all-paired electrons and unrestricted for systems with unpaired electrons.

SCF convergence

The convergence tolerance on the SCF affects to what degree the electronic structure of the system is optimized during the SCF iterations. The default values of ORCA should be fine for most calculations performed in this course.

Basis sets

Also, in the simple-input line, you can specify what basis set to use, i.e. how many Gaussian functions are used to represent the atomic orbitals. You can select from various basis set types, the most popular are the Pople, Dunning and Ahlrich basis sets. For this lab exercise, use Pople style basis. In particular, the 6-31G basis set (keyword in ORCA: 6-31G). For more information on basis sets in ORCA it is good to look at: <https://sites.google.com/site/orcainputlibrary/home>

Job specification.

You can define what type of calculation to do with ORCA using keywords (all in the simple-input line):

A single point energy (keyword in ORCA: ENGRAD) calculation evaluates the total energy and force acting on the atoms only for the single set of atom coordinates specified in the input file.

A minimum energy structure search (keyword in ORCA: OPT) will iteratively move the atoms so as to find the atomic coordinates that give lowest total energy for the molecule. The calculation will be repeated for different locations of the atomic nuclei until no force acts on the atoms. After each calculation, the nuclei are moved according to the forces acting on them. Such an optimization needs to be carried out to determine the optimal bond length (and angles in larger molecules).

Vibrational frequencies (keywords in ORCA: FREQ and NUMFREQ) can be calculated for a given geometry and used to compare to IR spectra of molecules and also to compute transition rates for reactions. The freq keyword computes the vibrational frequencies using the analytical Hessian (when available for a given electronic structure method). Otherwise, NUMFREQ can be used. Numfreq approximates the Hessian matrix using finite difference methods.

Transition state search (keyword in ORCA: OPTTS) instructs ORCA to search for a geometry that corresponds to a saddle point on the potential energy surface.

Nudged elastic band calculations (keyword in ORCA: CI-NEB and NEB-TS) instructs ORCA to search for minimum energy paths and saddle points on the potential energy surface, corresponding to a transition from a given reactant to a product state.

More experienced user can use advanced input 'blocks' to get more flexibility in the methods. The advanced input blocks are placed directly below the simple-input line (we will address these later).

Finally, the last part of the ORCA input file should contain information about the system: the charge, spin, chemical symbols and molecular geometry in Cartesian coordinates (ORCA can also use internal coordinates). You can copy the coordinates generated by Chemcraft to your input file.

Geometry block

In the first line of the geometry block, the user needs to specify the format for the input coordinates, charge and spin-multiplicity of the system. For Cartesian coordinates it would be '*xyz charge multiplicity', where you replace the words 'charge' and 'multiplicity' with the appropriate integers. The H₂ molecule has

a charge of 0 and multiplicity of 1. The multiplicity is the number $2S + 1$ where S is the net spin of all the electrons in the molecule. A singlet corresponds to $S = 0$ and is thus designated by multiplicity of 1.

The molecular geometry should be contained in the lines following the '*xyz charge multiplicity' line. The most simple approach is to paste the molecular geometry of H₂ that was generated with Chemcraft into the ORCA input file (on many SSH clients you can paste by right-clicking on the terminal window). The pasted text should then contain the type of the atom (chemical symbol) followed by the x-, y- and z-coordinates of the atom. The geometry block and the input file then end with an asterisk (*) in a new line.

You can add comments in an ORCA input file by using #. Such comments can help you remember why you did the calculation but will not have any effect on the calculation.

The ORCA input file for the H₂ molecule could look like this:

```
!HF 6-31G opt

*xyz 0 1
  H   -1.614714750785      0.000000000000      -5.520392310377
  H   -1.614714750785      0.000000000000      -6.260392310377
*
```

But, the initial molecular geometry that you have generated will be different from the one shown above.

Run the calculation

Now that you have created the input file, you are almost ready to submit the calculation to Elja. In order to do that you need a jobscrip. You should already have a 'sub.sh' in your home directory on Elja, in the lab1 directory from the previous practical session. Copy that to your working directory (>> cp /home/username/previous_location/sub.sh /home/username/new_location/)

Elja has several queues for calculations. You should use the short queue where the maximum computer time allowed per calculation is two hours.

To submit the calculation to Elja write; >> sbatch sub.sh in the terminal window.

You can monitor the status of your job with the command; >> squeue -u username, where 'username' is your personal username on Elja. The following screenshot is an output from the execution of >> squeue -u via9 (by a user called via9).

JOBID	PARTITION	NAME	USER	ST	TIME	NODES	NODELIST(REASON)
472066	short	orca	yla1	R	0:00	1	compute-1

The first number: 472066 is the job-id and is unique to this particular job. To delete a job, you can write; >> scancel job-id. 'NODES' is the number of nodes (node is simply a cluster of central processing units (CPUs)) that you wish to use to perform the calculation. The 'R' tells you that your calculation is already underway. However, if you find 'Q' instead of 'R' the calculation is in queue and will start when the requested number of nodes/CPUs becomes available. Moreover, 'C' tells you that your job has finished. This could either mean that your calculation was successful and the job finished or the job returned an error.

The output of the calculation is written to a file with the name 'orca.out'. To read this file on Elja, you can experiment with any of the following commands; >> nano orca.out, >> less orca.out (to exit 'less' press Ctrl+Q), >> tail -100 orca.out and >> cat orca.out. If the calculation has successfully finished, there will

be a line printed at the end of the output file with the following message ******ORCA TERMINATED NORMALLY******. If the calculation is still running the output file will be incomplete. You can copy the output file, orca.out, back to your local machine and open it in Chemcraft (**File**→**Open**). Then, under the 'Source' tab you can view and scroll through the output file. The 'Abstract' tab will give you a short summary of the calculation settings.

If the calculation finished successfully, few additional files will be created in your working directory, namely 'orca.xyz', 'orca_trj.xyz' and 'orca.gbw'. The file 'orca.xyz' is the optimized equilibrium geometry of the H₂ molecule. The 'orca.gbw' is the wavefunction file for this particular geometry. Wavefunction files can occupy large amount of space on the hard drive and should not be stored for too long. However, they do provide an excellent initial guess for the electronic structure in calculations where the same molecular geometry is used. The file 'orca_trj.xyz' will show you the trajectory of how the geometry of the molecule changes as the potential energy of the molecule is minimized with respect to the atomic positions.

If you see an error message in the output file (orca.out), check the input file. If you can't find a mistake, please ask the instructor to have a look at the problem.

Visualize the optimized geometry

Now copy the 'orca.xyz' file back to your local machine and visualize it with Chemcraft. Open the Chemcraft program, go to **File**→**Open** and select the new 'orca.xyz' file. Alternatively, you can also copy back the output file ('orca.out') and open that file with Chemcraft.

You should now see the H₂ molecule in the main window of Chemcraft. To find the bond distance, select both of the H atoms by left-clicking on the two atoms. The bond length (in Å) is reported in the geometrical editor. Write the bond length down as you will need it later.

Visualize 3-dimensional isosurfaces

To visualize the 3-dimensional isosurfaces illustrating the electron density or molecular orbitals, we need to make a new calculation. Create a sub-directory for the calculation. Copy the jobscript and ORCA input file to the new directory. Replace the molecular geometry in the new copy of the ORCA input file with the optimized H₂ molecular geometry, i.e. from the 'orca.xyz' file of the previous calculation. Remove the 'opt' keyword from the input file. Now we need to extend the input file (to increase the level of print from ORCA using an 'output print block').

Add the following output-block after the simple-input line:

```
%output
Print[ P_Basis ] 2
Print[ P_MOs ] 1
end
```

This will tell ORCA to write an orbital information to the output file so that Chemcraft can then reconstruct and visualize the orbitals. Now submit the job from with the new edited input file.

Once the calculation is done, copy the output file ('orca.out') to your local machine and open it in Chemcraft.

From the menu select **Tools**→**Orbitals**→**Render molecular orbitals**. Now you will get a list of occupied and unoccupied orbitals sorted by potential energy of the orbitals. Select the HOMO (highest occupied molecular orbital) and click 'ok'. You can now click 'Molecular Orbital 1 ...' in the drop-down menu to the left of the Chemcraft main screen and then click the 'show isosurface' toggle option. Typically,

a good isocontour value is 0.1–0.05 Hartree/Bohr³. To make the ball&stick molecule visible, click 'style' and check 'add transparency'. Orbitals are usually displayed by *isosurfaces*. This means, a surface is drawn in 3-dimensional space, at which the orbital has a constant value, the *isovalue* (similar to a contour line as a 2-dimensional projection of 3-dimensional data).

To export an image of the displayed isosurface, go to menu **File**→**Save as**→**Save image** and select a suitable image format. In general '**JPG format**' should be fine for use in a word processor such as Word. Choose a filename for the image and save it.

To display the LUMO (lowest unoccupied molecular orbital) and total electron density, go back to **Tools**→**Orbitals**→**Render molecular orbitals**. Select 'Molecular Orbital 2 ...' and 'Total SCF density' and click 'ok'. Choose the same parameters for the grid and the isocontour value as before. Now to see the sign of the wavefunction (phase) check the option 'both-signed'. A part of the orbital is then colored red (minus sign) and another part blue (plus sign). For the electron density a 'bubble' should appear around the molecule and hide it, like in the case of the HOMO isosurface. Notice that the density is almost spherical, just slightly elongated along the molecular axis. Export this image as well. Save an image of the LUMO and total electron density.

Q1: Present the images of the HOMO and the LUMO in your report. Take care that the windows were all maximized and the camera settings were identical before you saved the images. Arrange the images of the HOMO and LUMO side by side. Write text underneath to label each molecular orbital as bonding or antibonding. Also, specify whether the molecular orbitals are σ or π orbitals.

Q2: What is the optimal bond length in a H_2 molecule predicted by the Hartree-Fock calculations in the 6-31G basis set? How many iterations did ORCA need to find the optimal bond length? Compare with the experimental value for the bond length, 0.74.

To get the total energy of the molecule in this configuration, you can search for the last occurrence of the words "**FINAL SINGLE POINT ENERGY**" in the log file. The energy of the molecule is displayed in units of Hartree (1 Hartree = 27.2114 eV, twice the ground state energy of a H-atom).

Q3: Record the energy of the H_2 molecule in your report. What is the zero of energy in this case? Using the exact value of the energy of an H-atom, what is the bond energy? (This is, actually, not a good procedure since the two numbers subtracted are not calculated at the same level of theory. Generally, one relies on cancellation of errors in calculations of bond energies and should, therefore, consistently use the same level of approximation for all numbers subtracted or added.)

Repeat the H_2 calculation using the larger 6-31G** basis set. For this create a new directory, copy your input (.inp) file and the job-script (.sh) to the new directory.

Now you are ready to change the basis set keyword from 6-31G to 6-31G** in the fresh copy of the ORCA input file and submit the new job. In this basis set, the basis functions of hydrogen are extended by a set of the three p-type orbitals.

Q4: What is the predicted bond length and what is the energy of the molecule in this basis set? Is the energy higher or lower than for the 6-31G basis set? Is the bond length longer or shorter?

Q5: Explain why the energy of the molecule is different and why it must necessarily change in the direction it does (become lower or higher). Recall, SCF calculations are variational calculations. Is there any such rule for how the bond length changes with the basis set?

Q6: Does the bond energy necessarily also change in the same way as the energy of the molecule when the basis set is changed and both the molecule and the atoms are calculated at the same level of theory (why/why not)?

B. Electron density of HF and LiH molecules

HF molecule

Perform the same calculations on the HF molecule using Hartree-Fock/6-31G basis set. You can create the molecule, either by using the **Edit**→**Add atom** or search through the fragment/molecule database of Chemcraft. But first delete the H₂ molecule by **File**→**Clear**. Create a new input file, as described above in section A. Alternatively, you can copy an old input file to a new directory, open it with nano and replace one of the H atoms by a F atom.

As for the H₂ molecule above, load the output file in Chemcraft and display the electron density isosurface with a reasonable isovalue (e.g, 0.01 Hartree/Bohr³). Compare the electron density with that of the H₂ molecule. Save an image of the electron density.

Q7: *Does the electron density near the H atom increase or decrease when it binds to an F atom as compared to H₂? How does the size of the isosurface of the electron density near the H-atoms in the two molecules correlate with the electronegativity of the atoms?*

LiH molecule

Repeat the calculation for a LiH molecule. Save an image of the electron density.

Q8: *Compare the results for LiH with those for the H₂ and HF molecules. Comment on how the electron density at the H-atom in LiH compares with those in H₂ and HF. Explain the trend.*

Q9: *Include the images of the electron densities in your report. From the electron distribution, assign partial charges (δ^- , δ^+) to the atoms in the molecules.*

Basis set errors

A general rule in quantum chemistry is that one should only add or subtract values calculated at the same level of theory (unlike above where we made use of the fact that the exact value of the energy of the H atom is known). Calculate the Hartree-Fock energy of a hydrogen atom using the 6-31G basis set and then the larger 6-31G** basis set.

Create a new input file with only one H atom, select the right charge and spin multiplicity. Select same electronic structure method and basis set as before. Alternatively, you can also edit one of the previous input files. However, to run a calculation of a single atom, you need to remove the `OPT` keyword and simply perform a single point energy (and gradient) evaluation with the keyword `ENGRAD`.

Q10: *Which spin multiplicity do you have to select for the H-atom? Why? Why not select 'Equilibrium structure search' or 'OPT' in this case?*

Run the calculation and repeat for the larger basis set.

Q11: *What is the absolute and relative deviation (in %) of the calculated energy of a hydrogen atom using the two basis sets compared with the exact value?*

Calculate the bond energy of H₂ using the two basis sets and compare. Make sure you use the same level of approximation for both molecule and atom in each calculation of the bond energy.

Q12: *Does the larger basis set yield a stronger or weaker bond?*

Q13: *How does the bond energy obtained here deviate from the values you obtained earlier (see questions Q3 and Q4)? (give absolute and relative deviation in %).*

Here are more subtle questions about basis sets:

Q14: *Do you agree with the following statement (if so, explain; if not, try to present a counter example)? 'Adding basis functions to a basis set will always yield a lower or equal total energy in a variational calculation.'*

Q15: *Do you agree with the following statement (if so, explain; if not, try to present a counter example)? 'A variational calculation using a basis set B which has a larger number of functions than basis set A, will always yield a lower or equal total energy as a calculation using A.'*