

GAUSSIAN 09W TUTORIAL

AN INTRODUCTION TO COMPUTATIONAL CHEMISTRY USING G09W AND

AVOGADRO SOFTWARE

Anna Tomberg
anna.tomberg@mail.mcgill.com

This is a quick tutorial that will help you to make your way through the first steps of computational chemistry using Gaussian 09W software (G09). The tutorial is oriented to beginners and describes in detail the most used calculations done using G09. However, the theoretical basis of these calculations will not be covered here. If you are interested in understanding the details, please refer to textbooks targeting this subject. I found [1] and [2] very helpful, and strongly recommend to take a look at these wonderful books.

CONTENTS

1	Program specifications	3
1.1	Input	3
1.2	Output	5
1.3	Visualization	5
1.4	Computation model	5
2	First look at a computation	5
2.1	EXAMPLE 1: Single Point Energy of Water	5
3	Gaussian Input details	7
3.1	Link 0	7
3.2	Route Section	7
3.3	Molecular Structure	8
4	Theoretical Models	8
4.1	What is a theoretical model?	8
4.2	<i>Ab initio</i> methods	9
4.2.1	Examples of <i>Ab initio</i> methods	9
4.3	Semiempirical methods	10
4.3.1	Examples of semiempirical methods	10
4.4	Density Functional Theory	10
4.4.1	Examples of DFT methods	10
4.5	Molecular Mechanics	10
4.5.1	Examples of MM methods	10
5	Basis Sets	11
5.1	What is a basis set and why is its selection important?	11
5.2	A bit of theory: Slater VS Gaussian	11
5.3	Types of Basis Sets and Notation	11
5.3.1	Minimal	11
5.3.2	Split Valence	12
5.3.3	Correlation-consistent	13
5.3.4	Useful Tips from David Sherrill [5]	13
5.3.5	Comparison between Pople and CC basis sets	13
6	Types of calculation	13
6.1	Geometry Optimization	14
6.1.1	What is it?	14
6.1.2	For your Job	14
6.1.3	What information do you get out of this calculation?	15
6.2	Single point energy	15
6.2.1	What is it?	15
6.2.2	For your Job	15
6.2.3	What information do you get out of this calculation?	15
6.3	Frequencies and Thermochemistry	16
6.3.1	What is it?	16
6.3.2	For your Job	16
6.3.3	What information do you get out of this calculation?	17

6.4	Stability check	17
6.4.1	What is it?	18
6.4.2	In G09	18
6.4.3	For your Job	18
6.5	Molecular Orbitals and Population Analysis	18
6.5.1	What is it?	18
6.5.2	For your Job	19
6.5.3	What information do you get out of this calculation?	20
6.6	UV-Vis and Electronic transitions	20
6.6.1	What is it?	20
6.6.2	For your Job	20
6.6.3	What information do you get out of this calculation?	21
6.7	Potential Energy Surface	21
6.7.1	What is it?	21
6.7.2	For your Job	21
6.8	Solvation effect	21
6.8.1	What is it?	22
6.8.2	For your Job	22
6.9	Other Molecular Properties	22
6.9.1	Polarizability	22
6.9.2	Forces on Nuclei	22
6.9.3	Molecular Volume	22
6.9.4	NMR analysis	22
6.9.5	Electrostatic potential and Electron density	22
7	For a successfull analysis	22
7.1	Main steps of a successfull computational study	23
7.1.1	Step 1: Define the boundaries	23
7.1.2	Step 2: Set up you computer	23
7.1.3	Step 3: Define a good way to name the jobs	23
7.1.4	Step 4: Define a parent molecule	23
7.1.5	Step 5: Succesfull sequence of calculations	23
7.1.6	Step 6: Read the output	24
8	Batch files: avoid waiting around	24
8.1	What are Batch Files?	24
8.2	How to create a .bcf file?	24
8.3	Batch File Generator	25
9	Examples	25
9.1	EXAMPLE 1: SP of H2O	25
9.2	EXAMPLE 2: Opt of Ethanol	25
9.3	EXAMPLE 3: Molecular Orbitals Calculation and Visualization of HF	26
9.4	EXAMPLE 4: The Energy of Stereisomers and Scaling Factors (butene)	29
10	TroubleShooting	30
10.1	Convergence cannot be achieved?	30
10.2	My job froze, what can I do to avoid restarting from scratch?	30
10.3	A double bond disappeared from the structure after computation?	30
10.4	My batch file does not run properly, why?	30

11 Appendix	31
12 References	33

1 PROGRAM SPECIFICATIONS

Gaussian 09W (G09) is a computational chemistry program that runs on any modern Windows 32-bit PC. If you want to install G09 on a 64bit PC, there is a special procedure you must follow:

1. Insert the CD with G09 and copy its content onto your computer. Any folder will do; I copied directly into the :C\ directory.
2. Open directory containing G09
3. Find the g09w.exe file
4. Right click on exe file, select Properties, a new window should appear
5. Go into the Compatibility menu
6. Put a checkmark next to : Run as administrator (this should enable other checkboxes)
7. Put a checkmark next to : Run this program in compatibility with (select Windows version that you are using)

The installation requires the Gaussian CD and a registration key.

VISUALIZATION SOFTWARE ChemDraw (ChemBio 3D Ultra) and Avogadro (v.1.0.3) softwares can be used for visualization.

VIDEO TUTORIALS If you prefer watching video tutorials for a better understanding, please see below:

- Part 1: <http://www.youtube.com/watch?v=LOZswm07j0U&feature=plcp>
- Part 2: <http://www.youtube.com/watch?v=604DEoETAnU&feature=plcp>
- Part 3: http://www.youtube.com/watch?v=_REn2MsqUc4&feature=plcp
- Part 4: <http://www.youtube.com/watch?v=WpyqqDwspoo&feature=plcp>
- Part 5: <http://www.youtube.com/watch?v=AScn10uyBoY&feature=plcp>

Or search for "Avogadro with Gaussian Tutorial" in Youtube.

1.1 INPUT

The input for G09 can have the following extensions:

- Gaussian Input File: .gjf
- Batch Control File: .bcf
- Avogadro Input File: .com
- Text File: .txt

The input can be done manually or come from another software, such as Chem-Draw (3D) or Avogadro. G09's input consists of the following parts, shown on Figure 1:

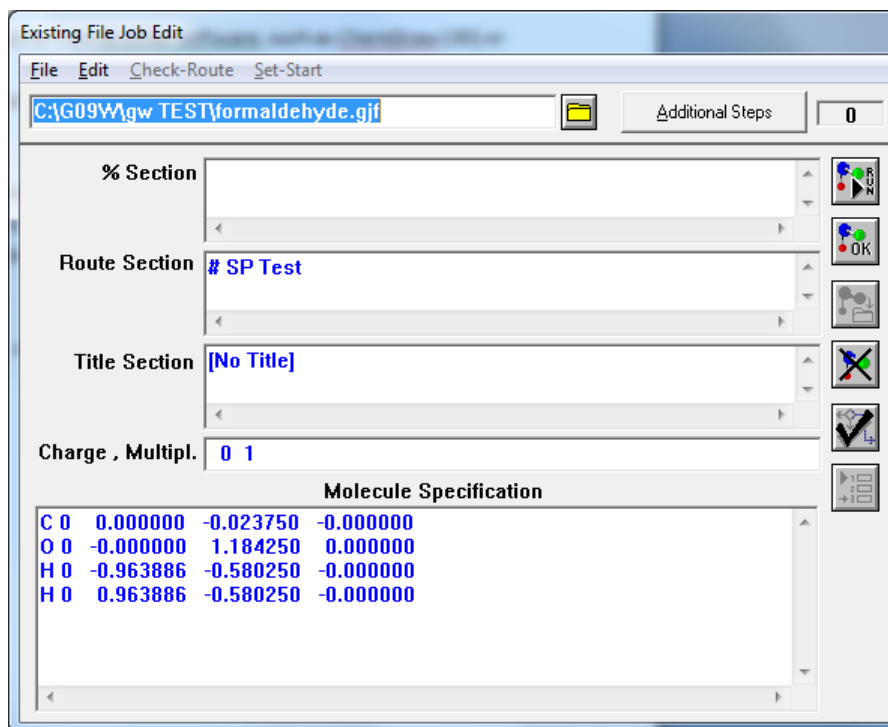


Figure 1: G09 input window

This window appears when you click on File→ Open→ ... in the main G09 window.

- The first line : specifies the path to the file you just loaded.
- Section: specifies the name of the checkpoint file (.Chk)
- Route Section: specifies the basis set, the theoretical model and the type of job you want to perform
- Title Section: specifies the name of the job (for the user's ease only)
- Charge, Multipl.: specifies the charge and spin multiplicity of the current molecule separated by a space.
- Molecule Specification: specifies the atoms and their coordinates ← this is what we want from the ChemDraw and Avogadro input, otherwise, we would need to calculate the coordinates by hand.

More details on each section of the input is available in section (3.0)

1.2 OUTPUT

To run the G09 job, click on the RUN button, on the right panel of the Job Entry window (Figure 1). The output has only one extension: .out. You will be prompted to save the output file before closing G09 program.

1.3 VISUALIZATION

This tutorial will use Avogadro software for visualization of the G09 output. Open the .out file in Avogadro, your input molecule should appear on the view screen. A detailed demonstration of Avogadro is available in the first two parts of the video tutorials.

NOTE If you try to visualize things you have not calculated using G09, the program will freeze!

1.4 COMPUTATION MODEL

The mathematical models used to do the computations are called FEM (Final Element Method) and Symplex method. Using matrices, this method cuts the N-dimensional space into small sub-systems that can be described by N linear equations. These equations can be solved as soon as one of them is solved. Therefore, one must take a guess of the solution and then, recursively solve all the other ones. Once all solutions are obtained, the initial guess can be modified and the calculations repeated. This process is run until the new solution outputs the same result as in the previous iteration. This is called convergence.

Because the initial guess can be very far off the real value, many thousands of iterations are often needed. The basis set selected influences, among other, the quality of the guesses, while the theoretical model influences the type of calculations that the matrices will be subjected to.

The calculation stops as soon as the result converges, however convergence does not mean that the system reached its minimum. It is very likely that the minimum outputted is actually only a saddle point on the potential surface, but the program will not be aware of this. The user must be careful with this, and always perform a check of stability of the system. This is done by perturbing the "stable" system and re-calculating the minimum. If the output is the same, one can assume that the energy obtained is the total minimum.

FOR MORE DETAILS ON THE MATH: [3]

2 FIRST LOOK AT A COMPUTATION

This section will describe how to do a simple calculation in G09 using Avogadro. All the parts of this procedure (and more) are explained more in details in corresponding sections. Let's take a look.

2.1 EXAMPLE 1: SINGLE POINT ENERGY OF WATER

Let's calculate the Single point energy of water. For this, we will first open Avogadro and draw the molecule (Figure 2).

If you think you will reuse this molecule in Avogadro, you can save it with an .cml extension.

NOTE It is good practice to save all parts of your job along the way!

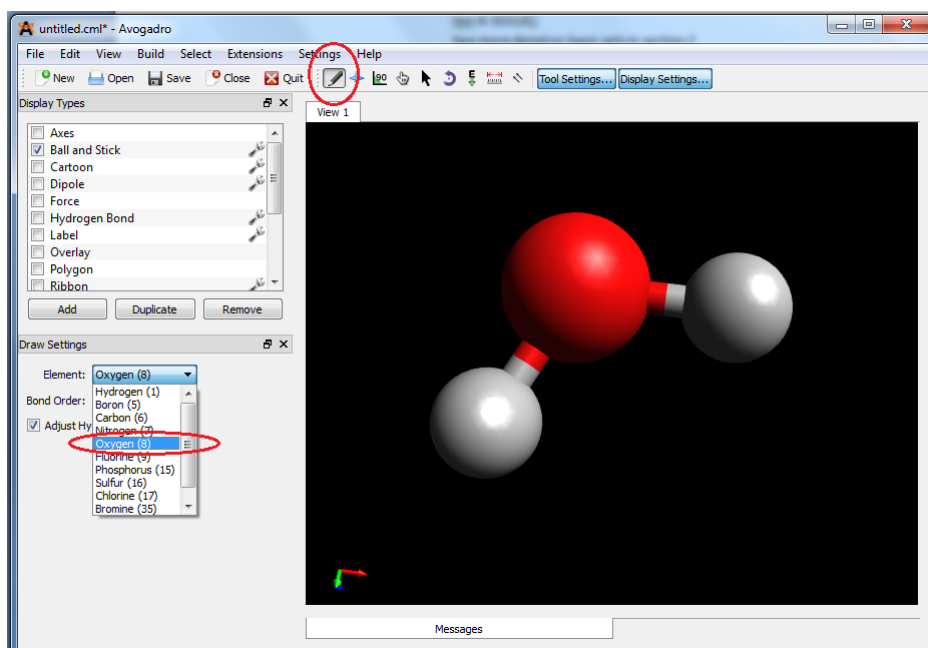


Figure 2: water molecule in Avogadro

Now, find the Extensions button on the top task bar, and select "Gaussain..." under this tab. A new window should appear (Figure 3).

You need to change (A) to "water"; (B) to Single Point energy, and click on the Generate button. Save as waterSP.com. Close this window, minimize Avogadro and open Gaussian. In G09, click on File > Open > waterSP.com.

NOTE If you cannot find your file, select "all Files" instead of Gaussian Input Files (bottom right corner).

Once your input is loaded and all the parameters are set properly, click on the Run button (first on Top Right). This will begin the calculation after prompting you to save the final output. Save as waterSP.out.

The job is over once the message "Processing Complete" is shown as on Figure 4. Don't forget to read the awesome quote at the end of the output! Now, let's analyze the results. To see clearer, you can load the output into the Notepad, by clicking on the magnifying glass button in the top right corner of the window. For a detailed overview of the output, see the video tutorials.

Right before the Population analysis you will see the following:

SCF Done: E(RHF) = -76.0067942514 A.U. after 10 cycles

This is the single point energy (similar to free energy of formation) of water! (units = Hartrees)

The next sections will describe in more details each part of a computation and will show how to do other types of calculations.

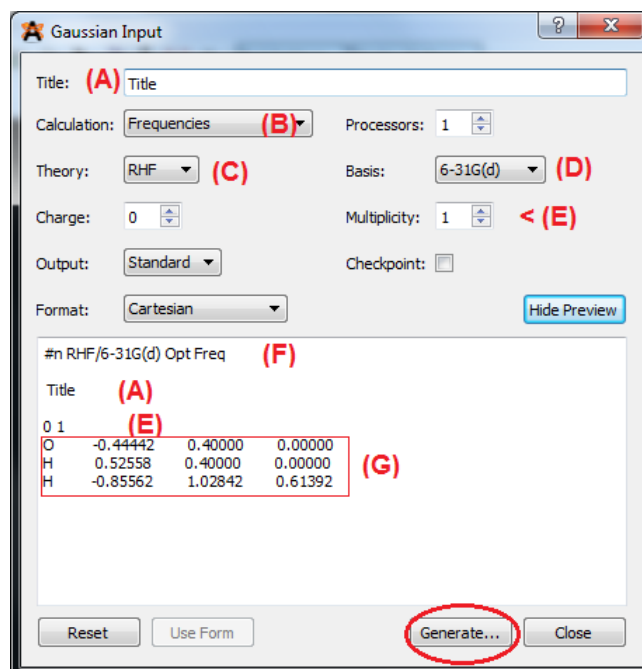


Figure 3: (A) Name of job, (B) type of calculation, (C) theoretical model, (D) Basis set, (E) charge and multiplicity, (F) the route section that will be given to G09. (G) the molecular coordinates.

3 GAUSSIAN INPUT DETAILS

3.1 LINK 0

This very first line of the input usually contains the name of the checkpoint file to save additional information in. This tutorial will not cover the use of .Chk files, but you can learn about them from the reference manual.

3.2 ROUTE SECTION

This section contains the instructions for a job you want G09 to execute. The input for this section is as follows:

#X Theoretical model/Basis set Type of calculation Options

- # : mandatory sign to begin route section.
- X : specifies the amount of detail you want to acquire: X = T (terse output); X = P (max output); X = N (normal output).
- Theoretical model : keyword that tells G09 which theoretical model to use (ex: RHF) See more detail on theoretical models in section 4.
- Basis set : specifies the basis set to use (ex: 6-31G(d)). See more detail on basis sets in section 5.
- Type of calculation : specifies one or more keywords for G09 jobs to do separated by a space. See more detail on types of calculations in section 6.

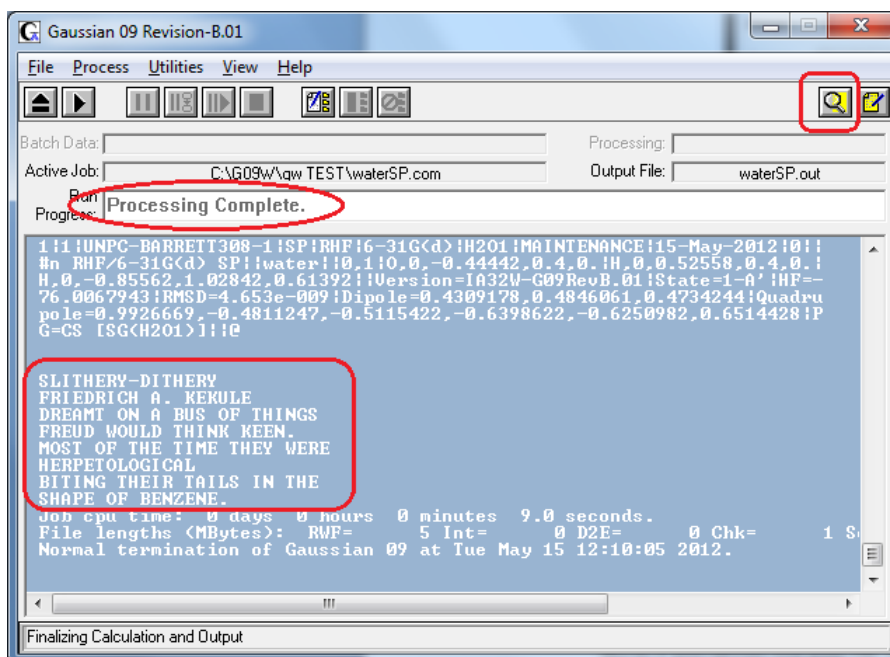


Figure 4: Water SP calculation Job successfully completed

- Options : specifies additional options for this job.

3.3 MOLECULAR STRUCTURE

This section is usually composed of the atoms and Cartesian or Z-matrix coordinates. It is possible to calculate and input the coordinates manually, but it is much easier to obtain them through software such as Avogadro or ChemBio 3D Ultra.

ChemDraw can be used solely to make the input molecule. Once the molecule is created and optimized, save it as a .gjf file. You will have to alter the input the other fields manually once the file will be loaded into G09. ChemDraw provides the coordinates, but nothing else.

Avogadro can easily generate the most popular G09 inputs (Opt, Freq, SP). To see how, refer to the video tutorials (links above) or to the example section. In short, once the molecule is drawn and optimized, click on Extentions (on top)→Gaussian...→.

To create a G09 input, click on Generate... and save as a .com file (shown on Figure 5). This file should then be loaded into G09 for further processing.

4 THEORETICAL MODELS

4.1 WHAT IS A THEORETICAL MODEL?

In short, a theoretical model or method is a way to model a system using a specific set of approximations. These approximations are combined with a calculation algorithm and are applied to atomic orbitals, defined by the basis set (see section 5), in order to compute molecular orbitals and energy. In general, the methods

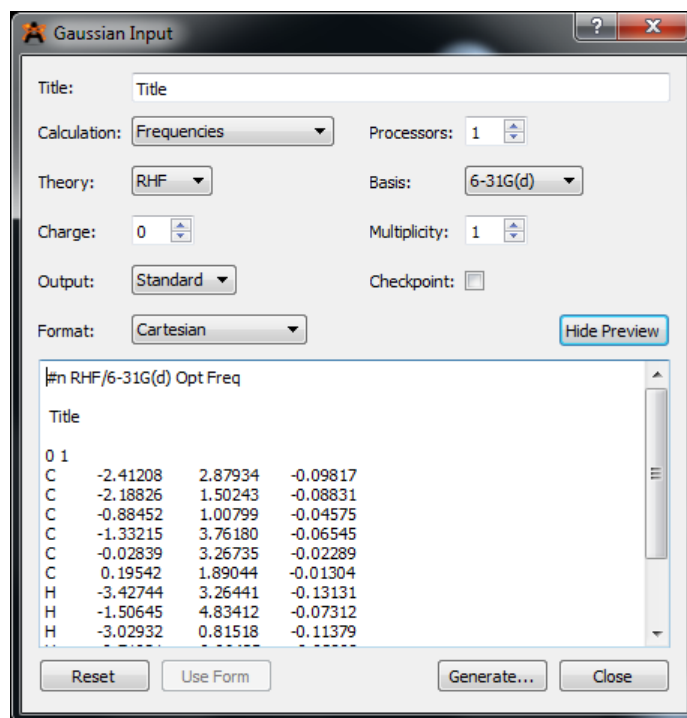


Figure 5: Avogadro input for G09

can be separated into 4 main types: semiempirical, ab initio, density functional, molecular mechanics. The selection of theoretical model depends on the size of the system and on the level of approximation. See "Theoretical Model Flow Chart" (Appendinx) for a quick way to select the right method.

4.2 *Ab initio* METHODS

This type of computation is based only on theoretical principles, using no experimental data. The numerous methods have the same basic approach, but differ in the mathematical approximations used. These are the most popular type of models, despite the fact that the calculations take unbelievably long time.

4.2.1 EXAMPLES OF *Ab initio* METHODS

HF Hartree-Fock is the basic *ab initio* model. It uses the approximation that Coulombic electron-electron repulsion can be averaged, instead of considering explicit repulsion interactions (central field approximation). There are two ways to compute molecular orbitals using HF: UHF (unrestricted) or RHF (restricted). UHF uses a separate orbital for each electron, even if they are paired (used for ions, excited states, radicals, etc.). RHF uses the same orbital spatial function for electrons in the same pair (good for species with paired electrons, no spin contamination).

The major drawback of HF method is the exclusion of electron correlation. The following models start with an HF calculation and then correct for electron repulsion.

MP_n Moller-Plesset perturbation theory are denoted as MP_n (n=2,...,6). In practice, MP2 and MP4 are the only methods used, since the other n's are either too computationally expensive or do not significantly improve the results compared with a lower level of complexity.

CI Configuration Interaction calculations are most often used for excited states. CI can be very accurate, but are also very CPU expensive.

4.3 SEMIEMPIRICAL METHODS

Semiempirical methods use a certain number of experimental data throughout the calculation. For example, bond lengths of a specific type will have a fixed value independently of the system (C=C bond will always be taken as 134 pm, for example). This dramatically speeds up computational time, but in general is not very accurate. Usually, semiempirical methods are used for very big systems, since they can handle large amounts of calculation.

4.3.1 EXAMPLES OF SEMIEMPIRICAL METHODS

ZINDO This method was parametrized to reproduce electronic spectra. It is most often used to compute UV transitions.

AM1 Austin Model 1 is a method that is most often used to model organic molecules.

4.4 DENSITY FUNCTIONAL THEORY

DFT methods are becoming more and more popular because the results obtained are comparable to the ones obtained using *ab initio* methods, however CPU time is drastically reduced. DFT differs from methods based on HF calculations in the way that it is the electron density that is used to compute the energy instead of a wave function.

4.4.1 EXAMPLES OF DFT METHODS

B3LYP This is the most popular DFT model. This method is called to be a hybrid, because it uses corrections for both gradient and exchange correlations.

PW91 Gradient-corrected method.

VWN Based on Local density Approximation.

4.5 MOLECULAR MECHANICS

What is the system you are working with is giant? No panic, you can still model it! This is possible by using molecular mechanics. MM methods approximate atoms as spheres and bonds as springs. They use an algebraic equation for the energy calculation, not a wave function or electron density. The constants in the equation are obtained from experimental data or other calculations and are stored in a data library. The combination of constants and equations is called a force field. These calculations are so simple that you don't even need to perform them in a complicated software such as G09. You can run your calculation right in Avogadro!

4.5.1 EXAMPLES OF MM METHODS

UFF Universal Force Field is the method used in Avogadro by default. It can be used on organic and inorganic molecules.

MMFF Merck Molecular Force Field is another general-purpose model, used mainly with organic systems.

5 BASIS SETS

5.1 WHAT IS A BASIS SET AND WHY IS ITS SELECTION IMPORTANT?

A basis set is a set of wave functions that describes the shape of atomic orbitals (AOs). The molecular orbitals (MOs) are computed using the selected theoretical model by linearly combining the AOs (LCAO). Not all theoretical models require the user to choose a basis set to work with. For example, PMn (n=3,...,6) models use an internal basis set, while ab initio or density functional theory require a basis set specification. The level of approximation of your calculation is directly related to the basis set used. The choice to make is a trade-off between accuracy of results and CPU time.

5.2 A BIT OF THEORY: SLATER VS GAUSSIAN

Both Slater Type Orbitals (STOs) and Gaussian Type Orbitals (GTOs) are used to describe AOs. STOs describe the shape of AOs more closely than GTOs, but GTOs have an unbeatable advantage: they are much easier to compute. In fact, it is faster to compute several GTOs and combine them to describe an orbital than to compute one STO! This is why combinations of GTOs are commonly used to describe STOs, which in turn, describe AOs. Yes, a bit complicated, but the computers don't mind. There are other differences between STOs and GTOs, but they will not be covered here.

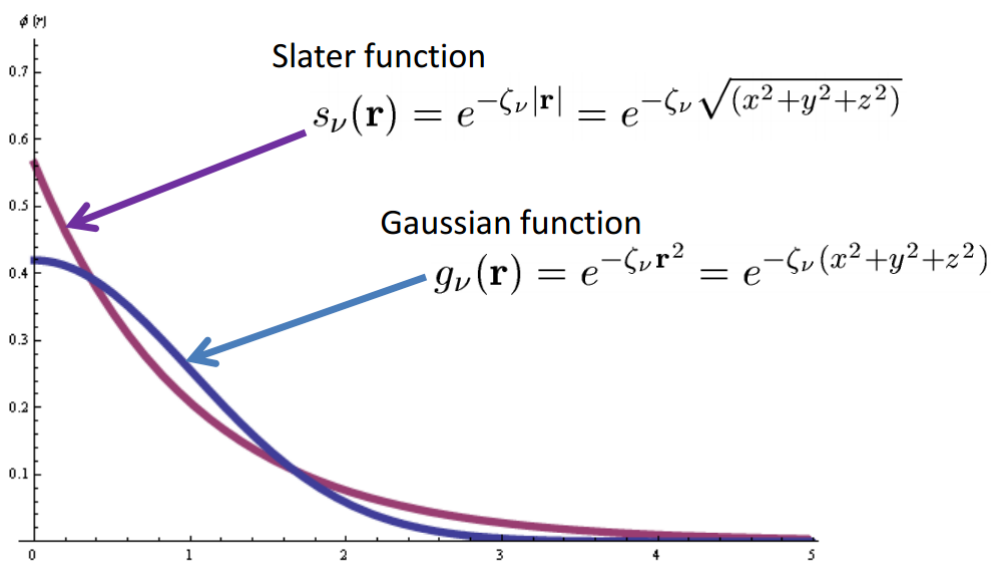


Figure 6: Slater VS Gaussian Type functions [4]

5.3 TYPES OF BASIS SETS AND NOTATION

5.3.1 MINIMAL

These basis sets use only one function for each AO. STO-nG (n=2,...,6) means that n GTOs are used to describe one STO, and only one STO is used to describe an AO (single Zeta). Usually n < 3 gives too poor results, so STO-3G is called the

minimal basis set. Minimal basis sets are used for either qualitative results, very large molecules or quantitative results for very small molecules (atoms).

5.3.2 SPLIT VALENCE

These basis sets are also called Pople basis sets and allow to specify the number of GTO's to use for core and valence electrons separately (size adjustable). These are double Zeta (2 functions per AO) or triple Zeta. The notation is as follows: K-LMG, where

- K = number of sp-type inner shell GTOs
- L = number of inner valence s- and p-type GTOs
- M = number of outer valence s- and p-type GTOs
- G = indicates that GTOs are used

POPLE basis sets are usually employed for organic molecules:

- 3-21G : 3 GTOs for inner shell, 2 GTOs for inner valence, 1 GTO for outer valence
- 4-31G
- 4-22G
- 6-21G
- 6-31G
- 6-311G : 6 GTOs for core orbital, 3 GTOs for inner valence, 2 different GTOs for outer valence (triple zeta)
- 7-41G

POLARIZED Pople basis sets can be modified to obtain an approximation that better describes the system you are working with. This can be done, for example, by letting the AOs distort from original shape (get polarized under the influence of the surroundings). Polarization can be added as * or (d).

- (d) or * type : d-type functions added on to atoms other than Hydrogens and f-type functions added on to transition metals
- (d,p) or ** type : p-type functions added on to Hydrogens, d-type functions added on to all other atoms, f-type functions added on to transition metals

EX: 6-31G(d) or 6-31G**

DIFFUSE Pople basis sets can also be modified by letting the electron move far away from the nucleus, creating diffuse orbitals. This modification is useful when working with anions, excited states and molecules with lone pairs. Diffuse functions can be added as + or ++ in front of the G.

- + : diffuse functions added on to atoms other than Hydrogens
- ++ : diffuse functions added on to all atoms

EX: 6-31+G(d) or 6-31++G(d)

5.3.3 CORRELATION-CONSISTENT

All of the basis sets described until now were optimized at a Hartree-Fock level. However, it is legitimate to doubt that this optimization might not be the best for correlated computations. Thom Dunning created a set of basis sets optimized using correlated (CISD) wavefunctions. They are denoted as cc-pVXZ, where:

- cc = indicates that it is a correlation-consistent basis
- pV = indicates that it is a polarized valence basis
- XZ = indicates the zeta number (X= D for double, T for triple, Q for quadruple, 5,6,7)

The prefix aug- can be used to add diffuse functions.

EX: cc-pVDZ or aug-cc-pVTZ

5.3.4 USEFUL TIPS FROM DAVID SHERRILL [5]

- a bigger basis set is not always the best! (ex: cc-pVQZ is overkill for RHF)
- STO-3G should be used only on very big systems
- anions must have diffuse functions
- usually cc-pVDZ is equivalent or worse than 6-31G(d,p)
- usually cc-pVTZ is better than 6-311G(d,p) or similar
- convergence of ab initio methods is very slow

5.3.5 COMPARISON BETWEEN POPLER AND CC BASIS SETS

These are approximately equivalent:

6-31G \longleftrightarrow cc-pVDZ

6-311G \longleftrightarrow aug-cc-pVDZ

6-31+G(d) \longleftrightarrow cc-pVTZ

6-311+G(d) \longleftrightarrow aug-cc-pVTZ

6-31++G(d,p) \longleftrightarrow cc-pVQZ

6-311++G(d,p) \longleftrightarrow aug-cc-pVQZ

These are only a few of the available possibilities. See more details on basis sets in G09:[6].

6 TYPES OF CALCULATION

Gaussian 09 is a powerful software that can perform a multitude of calculations on a given molecule. Here is a short list of the most useful types of jobs. If you need more details, refer to the Gaussian Reference Manual. There is a hard copy available in the lab office, as well as an online version available at [7].

NOTE If working with an unknown system (i.e. no reference values), you should ALWAYS perform a stability check. See section 5.4.

6.1 GEOMETRY OPTIMIZATION

KEYWORD: Opt

6.1.1 WHAT IS IT?

Geometry optimization is a name for the procedure that attempts to find the configuration of minimum energy of the molecule. The procedure calculates the wave function and the energy at a starting geometry and then proceeds to search a new geometry of a lower energy. This is repeated until the lowest energy geometry is found. The procedure calculates the force on each atom by evaluating the gradient (first derivative) of the energy with respect to atomic positions. Sophisticated algorithms are then used at each step to select a new geometry, aiming for rapid convergence to the geometry of the lowest energy. In the final, minimum energy geometry the force on each atom is zero. It is important to recognize that this procedure will not necessarily find the global minimum, i.e. the geometry with the lowest energy. By its nature, a successive search for a minimum finds a local minimum but not necessarily the lowest. In fact, the optimization procedure stops when it finds a stationary point, i.e. a point where forces on the atoms are zero, and this may be also a saddle point (i.e. a transition structure). This will occur particularly if we restrict the symmetry of the molecule and do not allow the program to search the full space of molecular configurational degrees of freedom. [...] It is always a good idea to start a procedure of a geometry optimization calculation with a small basis set and a relatively poor method before we move to the basis set and method of choice for a particular problem. You can then start the final geometry optimization from the geometry selected by the simpler and less accurate approach. We can even use the second derivatives of the energy with respect to atomic coordinates (obtained during the optimization process or from a frequency run 1) to further improve the geometry optimization.[8]

See also: [9]

6.1.2 FOR YOUR JOB

Unfortunately, neither Avogadro's geometrical optimization, nor ChemDraw's, can be fully trusted, so for any job you do, you should optimize your molecule before performing any calculation. It is important to optimize geometry in the drawing software too, before generating G09 input coordinates. Do this by running an MM geometry optimization calculation directly in Avogadro. Some small molecules don't require this step, but if you want to be safe, do it anyway. The downside of optimization is that for bigger molecules 2 things may happen.

1. The optimization does not converge → error, the process stops (see section 10.1)
2. It may take long time. So no stress, you have enough time to take a coffee break. Probably more than one.

This is why, it is useful to save the optimized structures in case you need to perform several calculations on the same molecule and avoid redoing the Opt. To make sure that the structure obtained is stable you should do the following:

1. perform a Stable=Opt calculation → see section 6.4
2. do a vibrational analysis (Freq) to make sure there is no imaginary frequencies → see section 6.3

6.1.3 WHAT INFORMATION DO YOU GET OUT OF THIS CALCULATION?

1. Atomic coordinates of optimized molecule
2. Optimized Parameters: atomic distances and angles
3. HOMO/LUMO eigenvalues (Hartrees)
4. Mulliken atomic charges (short)
5. Dipole moments

MORE DETAILS HERE: [10]

See example 2 in section 9.

6.2 SINGLE POINT ENERGY

KEYWORD: SP

6.2.1 WHAT IS IT?

This procedure simply calculates the energy, wave function and other requested properties at a single fixed geometry. It is usually done first at the beginning of a study on a new molecule to check out the nature of its wave function. It is also frequently carried out after a geometry optimization, but with a larger basis set or a more accurate method than is possible with the basis set and method used to optimize the geometry. Thus, for very large systems the geometry may be optimized at the Hartree-Fock level with the (relatively small) 3-21G basis set, but energy differences between isomers are then explored with more accurate methods and larger basis sets.[8]

See also: [9]

6.2.2 FOR YOUR JOB

This keyword is the default one, and is assumed if no keyword is specified in the route section. G09 was found to make systematic errors in the frequencies and Zero-Point Energies. The scaling factors vary according to the different theoretical models and basis sets used for the calculation. Table 1 (Scaling factors for Freq and ZPE/Thermal calculations), which can be found in the Appendix, lists the recommended correction factors for these calculations.[10, 11]

6.2.3 WHAT INFORMATION DO YOU GET OUT OF THIS CALCULATION?

1. Single Point Energy (Hartrees)
2. Orbital symmetries
3. Mulliken atomic charges (short)
4. Dipole moments

+See examples 1 and 4 in section 9.

6.3 FREQUENCIES AND THERMOCHEMISTRY

KEYWORD: Freq

6.3.1 WHAT IS IT?

This keyword will indicate to G09 to calculate the vibrations of the molecule in question, along with other parameters that you can choose to include or erase from the output. Each vibrational analysis also includes a thermochemistry section by default. If you want to learn how G09 computes these values, please read [12-14].

If you generate the input for G09 using Avogadro, then the Opt keyword will be automatically added in front of Freq, and the calculation will run in 2 steps.

Because the molecular frequencies depend on the second derivative of the energy with respect to the nuclear positions, you should make sure that the theoretical model you chose for the calculation computes second derivatives (ex: HF, B3LYP, MP2, CASSCF). *Ab initio* and DFT theoretical models are known to give the best results (based on experimental values).

NOTE: You must use the same combination of theoretical model and basis set for both the Opt and Freq calculations.

6.3.2 FOR YOUR JOB

Most likely, you would do a vibrational analysis to assign the peaks of an experimentally obtained spectrum to specific vibrations. Visualization using Avogadro software can be extremely helpful, however as a rule of thumb, the following can be used: bond stretches > bond bends > bond torsions (in order of decreasing energy).

SCALING The frequencies generated must be scaled according to the theoretical method and basis set used. A short list of the scaling factors is collected in Table 1 of the Appendix, but a complete list can be found at [11]. Note that the scaling factors for frequencies and for zero point energies are different! You can input the scale factor to be used for thermochemistry analysis using the Scale keyword as an option (ex: "Scale=0.95").

STABILITY The stability of your molecule can easily be checked by verifying whether there exists imaginary (negative) frequencies. Imaginary frequencies indicate instability of molecular geometry. Also, knowing that each molecule should have either 3N-6 or 3N-5 vibrational modes, you can notice that there are more frequencies than supposed to be! These are the translational and rotational modes and their intensities should close to zero (+/- 10). Big intensities for these frequencies is another way to verify the stability of you molecule. See more on p.72 of [10].

IMAGINARY FREQUENCIES Negative frequencies indicate instability in the molecule or, in other words, saddle points on the potential energy surface. A stable molecule should have no imaginary frequencies, a transition state should have one (1st order saddle point), while more than one imaginary frequency means that there is a problem with your molecule's geometry.

ANHARMONIC FREQUENCIES By default, G09 computes frequencies based on harmonic oscillator approximation (second order derivative with respect to the nuclei movement). If you wish to obtain more accurate results, you might want to consider using anharmonic corrections, i.e. computing higher-order derivatives. For this, add "Anharmonic" keyword into the route section.

RAMAN If you wish to add Raman intensities to the output, add "Raman" keyword to the Route section.

THERMOCHEMISTRY By default, several thermodynamic values are computed during a vibrational analysis. You can change the thermochemistry parameters (temperature, pressure, ect) by writing Freq=ReadIsotopes in the Route section. The values obtained are based on ideal gas assumptions.

See examples in the video tutorial Part 5.

6.3.3 WHAT INFORMATION DO YOU GET OUT OF THIS CALCULATION?

1. Atomic coordinates of optimized molecule
2. Optimized parameters: atomic distances and angles
3. HOMO/LUMO eigenvalues (hartree)
4. Mulliken atomic charges (short)
5. Dipole moments
6. Single Point energy
7. Harmonic frequencies (wavenumbers)
8. Reduced masses (amu)
9. Force constants
10. IR intensities
11. Raman intensities (not default)
12. Thermochemistry
 - (a) Temperature
 - (b) Pressure
 - (c) Isotopes used
 - (d) Molecular mass
 - (e) Thermal energy: E (Thermal)
 - (f) Constant volume molar heat capacity (CV)
 - (g) Entropy (S)
 - (h) Free Energy (sum of electronic and thermal Free Energies)
 - (i) Enthalpy (sum of electronic and thermal Enthalpies)

6.4 STABILITY CHECK

KEYWORD: Stable

6.4.1 WHAT IS IT?

The stability calculation determines whether the wavefunction computed for the molecular system is stable or not: in other words, whether there is a lower energy wavefunction corresponding to a different solution of the SCF equations. If the wavefunction is unstable, then whatever calculation you are performing is not being done on the expected/desired state of the molecule. The stability of SCF solutions for unknown systems should always be tested. Stability considerations apply and may be tested for in calculations using DFT methods as well.[10]

SCF = self consistent field

DFT = density functional theory

6.4.2 IN G09

The `Stable` keyword causes the program to compute a wavefunction as usual and then to determine if the resulting determinant is a local minimum with the specified degrees of freedom taken into consideration. The default is to test for all instabilities but not to re-optimize the wavefunction. If `Stable=Opt` is specified, by default the wavefunction is allowed to be unrestricted if necessary.[15]

6.4.3 FOR YOUR JOB

The output will then contain a message telling you whether the system is stable or not under the "stability analysis":

- Stable geometry message
 - The wavefunction is stable under the perturbations considered.
- Unstable geometry messages
 - The wavefunction has an RHF → UHF instability.
 - The wavefunction has internal instability.

6.5 MOLECULAR ORBITALS AND POPULATION ANALYSIS

KEYWORD: `Pop=X`; `X=None`, `Reg`, `Full`, `NBO`, ...

6.5.1 WHAT IS IT?

POPULATION ANALYSIS

Population analysis is a mathematical way of partitioning a wave function or electron density into charges on the nuclei, bond orders, and other related information. These are probably the most widely used results that are not experimentally observable. Atomic charges cannot be observed experimentally because they do not correspond to any unique physical property. In reality, atoms have a positive nucleus surrounded by negative electrons, not partial charges on each atom. However, condensing electron density and nuclear charges down to partial charges on the nucleus results in an understanding of the electron density distribution. These are not integer formal charges, but

rather fractions of an electron corresponding to the percentage of time an electron is near each nucleus. Although this is an artificial assignment, it is very effective for predicting sites susceptible to nucleophilic or electrophilic attack and other aspects of molecular interaction. These partial charges correspond well to the chemist's view of ionic or covalent bonds, polarity, and so on. Only the most ionic compounds, such as alkali metal halides, will have nearly whole number charges. Organo-metallics typically have charges on the order of +/- 0.5. Organic compounds often have charges around +/-0.2 or less. [1]

PARTIAL CHARGES :

The Mulliken analysis is the most common population analysis method, it is also one of the worst and is used only because it is one of the oldest and simplest. [...] In the Mulliken analysis half the overlap population is assigned to each contributing orbital, giving the total population of each AO. Summing over all the atomic orbitals on a specific atom gives us the gross atomic population. [16]

WARNING! Mulliken method is extremely basis set sensitive! Usually, the smaller the basis set, the better (opposite for energy calculations).

6.5.2 FOR YOUR JOB

Pop=X keyword

X=NONE no orbital information displayed

X=REG HOMO-5 up to LUMO+5 orbital information displayed

X=FULL all orbitals information displayed

X=NBO Mulliken analysis is replaced by Natural Bond-Order analysis

X=MK, CHEL, OR CHELPG produce charges fit to electrostatic potential (ESP)

G09 uses Mulliken analysis by default. Remember that if you need orbitals' energies, select a bigger basis set, if you need partial charges, select minimum basis set. It is fine to use the obtained results from Mulliken analysis for a qualitative overview of partial charges, but these values are not reliable and cannot be compared with literature ones! If you need quantitative results, you should do for example a Natural Bond-Order analysis (NBO) or use Electrostatic charges (ESP).

NBO How is NBO analysis different from Mulliken?

NBOs are an orthonormal set of localized "maximum occupancy" orbitals whose leading N/2 members (or N members in the open-shell case) give the most accurate possible Lewis-like description of the total N-electron density.[17]

NBO uses eigenfunctions of the first order reduced density matrix, localized and orthogonalized.

ESP This method uses electrostatic potential to compute charges on nuclei. ESP is usually the best to describe charge interactions with other species, but require significant amount of CPU time.

6.5.3 WHAT INFORMATION DO YOU GET OUT OF THIS CALCULATION?

1. Atomic coordinates of optimized molecule
2. Optimized parameters: atomic distances and angles
3. Selected number of orbitals eigenvalues (hartree)
4. Orbitals' population
5. Atomic (partial) charges (full analysis)
6. Dipole moment
7. Single Point energy
8. Electrostatic potential derived charges (with ESP keyword)

6.6 UV-VIS AND ELECTRONIC TRANSITIONS

KEYWORD: TD or CIS or ZINDO

6.6.1 WHAT IS IT?

Most molecules have bound higher energy excited electronic states in addition to the ground electronic state E_0 . These states may be thought of as arising from the promotion of one of the electrons from the occupied orbital in the ground state to a vacant higher energy orbital. The excitation of an electron from the occupied orbital to a higher-energy orbital occurs when a photon with the energy that matches the difference between the two states interacts with the molecule. The classical Franck-Condon principle states that because the rearrangement of electrons is much faster than the motion of nuclei, the nuclear configuration does not change significantly during the energy absorption process. Thus, the absorption spectrum of molecules is characterized by the vertical excitation energies. [18]

See also: [19]

6.6.2 FOR YOUR JOB

Depending on what your system is, you need to choose to work with either vertical or adiabatic excitation energies. Vertical excitations are obtained by taking the difference between the excited state potential energy curve and the ground state potential energy curve using the optimized geometry of ground state. This is done automatically for you using the methods listed below. Vertical excitations energies are used for a system that has no time to get to equilibrium before returning to ground state. Adiabatic excitations energies are obtained by first optimizing the geometry of the excited electronic state and then taking the difference between excited and ground energies. This is only possible if the excited states are long lived.

ZINDO This is a semiempirical method that was optimized to compute electronic transitions. Also applicable to inorganic systems.

CIS / CIS(D) Single-excitation configuration interaction is the easiest way to obtain excited states energies. CIS calculation uses the optimized MOs (for ground state) and calculates the energy to move one electron to one of the virtual orbitals. CIS(D) also includes correlation of excited states, and is a quick and relatively good method. (Always use CIS(D) instead of CIS)

TD This Time Dependent calculation is based on a DFT computational method. TD-DFT obtains the wave functions of MOs that oscillate between ground state and the first excited states.

You can select the number of excited states to compute (ex: TD=NStates=5). The default is 3. You can also select what type of transitions to look for (ex: TD=Singlets).

For more details of TD: [20].

6.6.3 WHAT INFORMATION DO YOU GET OUT OF THIS CALCULATION?

1. Ground to excited state transition
2. Excitation energies and oscillator strengths
3. Electronic Circular Dichroism (ECD)
4. Selected number of orbitals eigenvalues (hartree)
5. Orbitals' population
6. Mulliken Analysis (short)
7. Dipole moments

6.7 POTENTIAL ENERGY SURFACE

KEYWORD: Scan

6.7.1 WHAT IS IT?

The potential energy surface (PES) is a central concept in computational chemistry. A PES is the relationship, mathematical or graphical, between the energy of a molecule (or a collection of molecules) and its geometry. The Born-Oppenheimer approximation says that in a molecule the nuclei are essentially stationary compared to the electrons. This is one of the cornerstones of computational chemistry because it makes the concept of molecular shape (geometry) meaningful, makes possible the concept of a PES, and simplifies the application of the Schrodinger equation to molecules by allowing us to focus on the electronic energy and add in the nuclear repulsion energy later. [21]

See also: [22]

6.7.2 FOR YOUR JOB

The PES Scan is somewhat complicated to setup since you need to specify which coordinate to "freeze" for the calculation. Please refer to Chapter 8 in [10].

6.8 SOLVATION EFFECT

KEYWORD: SCRF

6.8.1 WHAT IS IT?

It is crucial to remember that by default G09 uses gas phase for any calculation. Most probably, you will be interested to compute molecular properties in solution. For this purpose, you have a choice between implicitly or explicitly adding solvation effects. Explicitly adding solvent molecules in your system, by drawing them in Avogadro, is extremely computationally expensive. As an alternative, you can use the SCRF keyword, which will create a solvent cavity around your molecule and perform an given calculation.

See also: [23]

6.8.2 FOR YOUR JOB

It is important to first optimize the geometry in the solvent of choice before doing any other calculation. You can also compute the excited states in solvent to see how solvation affects the transition energies. For a list of solvents that are available in G09 check the reference manual.

EX1: To compute geometry optimization in methanol, using RHF/6-31G(d), the Route section will look like this:

```
#n RHF/6-31G(d) Opt SCRF=(Solvent=Methanol)
```

EX2: To compute transition states in methanol, using RHF/6-31G(d), the Route section will look like this:

```
#n RHF/6-31G(d) TD SCRF=(Solvent=Methanol)
```

6.9 OTHER MOLECULAR PROPERTIES

6.9.1 POLARIZABILITY

KEYWORD: Polar

6.9.2 FORCES ON NUCLEI

KEYWORD: Force

6.9.3 MOLECULAR VOLUME

KEYWORD: Volume

6.9.4 NMR ANALYSIS

KEYWORD: NMR

6.9.5 ELECTROSTATIC POTENTIAL AND ELECTRON DENSITY

KEYWORD: cubegen

7 FOR A SUCCESSFULL ANALYSIS

This section will describe in short the main steps to perform a good study of a given system. There are many ways to go about doing computations, but a systematic approach will give you more chances to stay on track and obtain reasonable results.

7.1 MAIN STEPS OF A SUCCESSFULL COMPUTATIONAL STUDY

7.1.1 STEP 1: DEFINE THE BOUNDRIES

You should start by defining how much time, energy and CPU are available for your study. There are some steps that could be skipped if results must be obtained in little time. The choice of theoretical models and basis sets will also heavily depend on time and CPU resources. If your computer is not powerful enough, you will need much more computational time than when a more powerful computer is used.

7.1.2 STEP 2: SET UP YOU COMPUTER

To speed up computations, it is strongly recommended to turn off all the programs (except G09) on your PC. Programs like Skype, open folders, Internet browsers, etc., all demand CPU power. Shut them off! If you know how to, also turn off all the other services (if you're not sure, skip this). Make sure your computer does not go into sleeping or hibernating mode automatically after some period of time, this might kill a G09 job.

7.1.3 STEP 3: DEFINE A GOOD WAY TO NAME THE JOBS

It is crucial to have a strict notation for all of your G09 jobs. For example, you could use something like this:

```
trans_azo_UFF(MP2_6-31G(d))_Opt
```

This notation puts the name of the molecule followed by the MM method that was used initially to optimize the geometry, followed by the theoretical model, followed by the basis set and finally by the type of calculation performed.

7.1.4 STEP 4: DEFINE A PARENT MOLECULE

(skip if not working with library of molecules)

Starting with a parent molecule will aid you to choose the right theoretical model and basis set. Run several calculations using the same theoretical model, but with increasingly better basis sets, compare with literature values. Then restart with a different model. This will show you which combination outputs the smallest error, and therefore suites the best for this case.

7.1.5 STEP 5: SUCCESFULL SEQUENCE OF CALCULATIONS

You should perform the following sequence of calculations for any molecule:

1. Geometry Opt in Avogadro using MM (UFF or MMFF).
2. Geometry Opt and Freq calculation in G09 .
3. Check for imaginary frequencies (none for ground state, 1 for TS, etc).
4. If frequencies are alright, continue, otherwise restart the geometry opt.
5. Compute any property that you need.
6. Re-opt and re-compute frequencies in solvent (if needed).
7. Re-compute any property that you need in solvent (if needed).

7.1.6 STEP 6: READ THE OUTPUT

You can use a script called "G09\output\interpreter" that can be found on the desktop of the Gaussian workstation (PC). This program lists all the .out files found in C:\G09W\IO\COMPUTED\BATCH\done directory, asks which file you want to work with and lets you extract only the pertinent information out of it. This avoids scrolling up and down to find the right values, and saves you lots of time. Don't forget that the .out file must be copied into the directory mentioned above, otherwise the list will be empty!

8 BATCH FILES: AVOID WAITING AROUND

This section will explain how to use batch file to speed up you computational project.

8.1 WHAT ARE BATCH FILES?

A Batch file is a text file that contains a specially formatted list of G09 jobs that will be completed automatically one after the other without the need for the user to physically launch the calculation. Such a file has an extension .bcf recognized by G09. When you create a batch file, you list all the jobs you want G09 to do and the names of the corresponding output files, then open the .bcf file in G09, and start the computation! That easy!

8.2 HOW TO CREATE A .BCF FILE?

To make a new batch file, you fist must prepare the input files separately using Avogadro. Ideally, your input files should be independent of each other (to avoid errors that will be carried on). This means that the output of one of the files in the batch should not be used as input for some other file in the same batch file.

Once you create the input (.com) files, save them all in the same folder named "Batch". If you don't have this folder already, you should create it in the G09 directory.

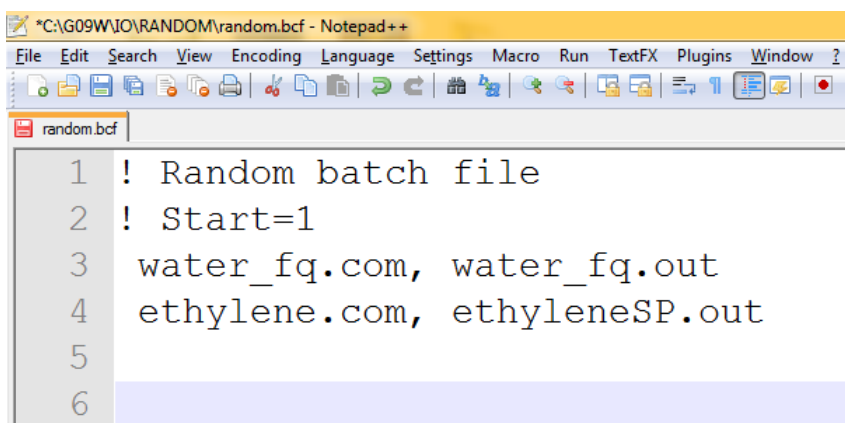
Now, open a blank (new) file in Notepad. The first two lines of this file must be exactly as follows:

! Name of the batch file

! Start=1

Then follow the list of input files that will be in this batch file followed by the name of their corresponding output files, separated by a comma. Each pair of input and output names must be on a new line. (See Figure 7)

When you have listed all the files you want to include in this batch, save this file with .bcf extension in the Batch folder. Open your batch file in G09. This will open a new window called "Edit Batch Control List" where your jobs should be listed. Make sure they are the right ones by verifying that the output files correspond to the input files. When you are ready, click on Exit in the right bottom of this window. This will bring you back to the main G09 screen. To start the calculations, select "Process" on the top task bar and click on "Begin Processing". Your batch file should start executing.



```
*C:\G09W\IO\RANDOM\random.bcf - Notepad++
File Edit Search View Encoding Language Settings Macro Run TextFX Plugins Window ?
random.bcf
1 ! Random batch file
2 ! Start=1
3 water_fq.com, water_fq.out
4 ethylene.com, ethyleneSP.out
5
6
```

Figure 7: Example of Batch File

8.3 BATCH FILE GENERATOR

For your convenience, I wrote a little script called "Batch_file_generator.py" that can be found on the desktop of the Gaussian workstation (PC). This program can be run by double clicking on the icon. You will be given a list of .com files found in the C:\G09W\IO\COMPUTED\BATC directory. Using these files, a batch file will be automatically generated for you and saved into the same directory as mentioned above. You will need to start the batch file in G09W yourself. The .out files will be saved into the same directory.

9 EXAMPLES

Here are some detailed examples that use Avogadro and G09. If you wish to see more examples, refer to [10] or watch the video tutorials listed in section 1.

9.1 EXAMPLE 1: SP OF H₂O

See section 2.

9.2 EXAMPLE 2: OPT OF ETHANOL

Let's optimize the geometry of ethanol, and compare the results obtained by Avogadro, and the one obtained from the Opt in G09. For this, we will first open Avogadro and draw the molecule. Now, click on Extensions > Optimize Geometry. (shortcut: Ctrl+Alt+O) Save the result as ethanol.cml. This will be our result (1).

Now, generate the G09 input by selecting Extensions → Gaussian...; change the title to "ethanolOpt", and the calculation to "Geometry Optimization", click on generate and save as "ethanolOpt.com". Open "ethanolOpt.com" in G09, and run the job. Save the output as "ethanolOpt.out".

Finally, in Avogadro, open "ethanolOpt.out" and compare this result (2) with what you obtained initially, using the Avogadro Optimization (result 1).

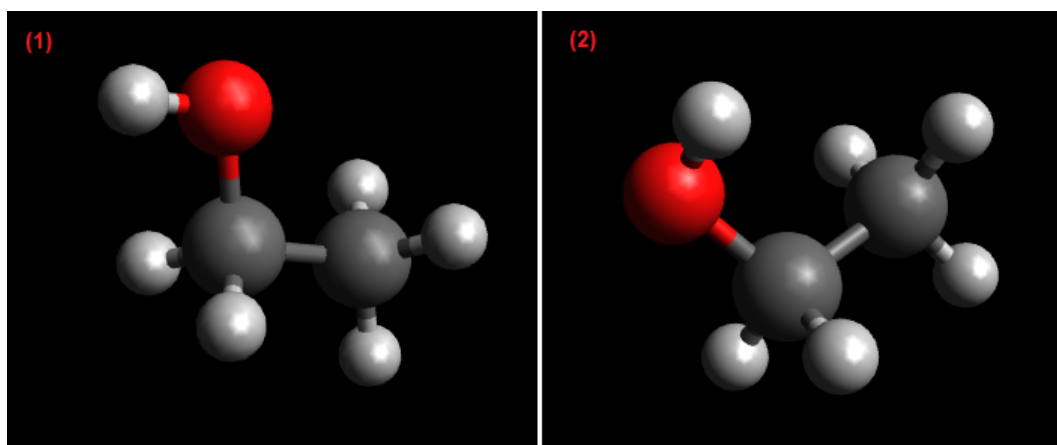


Figure 8: Optimization of ethanol: (1) Avogadro's Opt structure (2) Gaussian final Opt structure

9.3 EXAMPLE 3: MOLECULAR ORBITALS CALCULATION AND VISUALIZATION OF HF

Let's calculate the molecular orbitals of hydrogen fluoride. First, draw the molecule in Avogadro, optimize its geometry and generate the G09 input file. To see how these steps are done in details, refer to examples 1 and 2. The change now comes in the Route section: you have to change the calculation to "Pop=Reg". This command will display the higher five occupied and the lowest five virtual molecular orbitals and other pertinent information. If you wish to see more orbitals, type "Pop=Full". Your G09 job should look like the window in Figure 9.

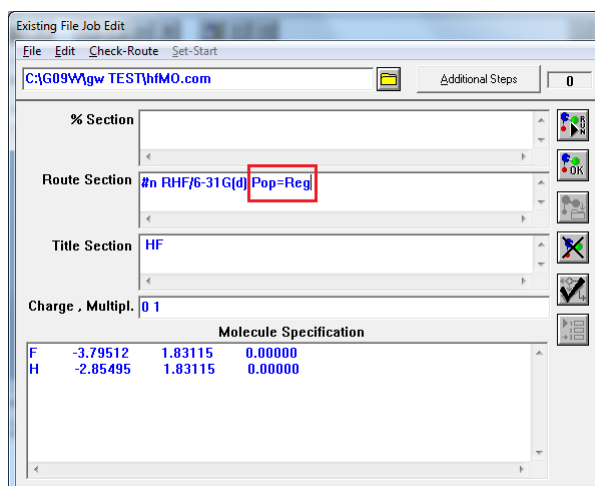


Figure 9: Route section for MO computation of HF molecule

After "Pop=Reg" command, add a space and type "FormCheck". This command is necessary in order to create the file, containing the information about the molecular

orbitals, which will be plotted later. Press RUN and start the calculation. When the process is done, the output will contain several results, among which you can find the sequence shown on Figure 10.

```

hfMO.out - Notepad
File Edit Format View Help

Population analysis using the SCF density.
*****
Orbital symmetries:
  occupied (SG) (SG) (SG) (PI) (PI)
  virtual  (SG) (SG) (SG) (PI) (PI) (SG) (PI) (PI) (DLTA)
  (DLTA) (SG) (SG)
The electronic state is 1-SG.
Alpha occ. eigenvalues -- -26.27930 -1.57238 -0.73584 -0.62619 -0.62619
Alpha virt. eigenvalues --  0.21047  1.07341  1.33089  1.49191  1.49191
Alpha virt. eigenvalues --  1.75292  2.09228  2.09228  2.15811  2.15811
Alpha virt. eigenvalues --  2.93583  4.06424
Molecular orbital coefficients:
  1      2      3      4      5
  0      0      0      0      0
Eigenvalues -- -26.27930 -1.57238 -0.73584 -0.62619 -0.62619
1 1  F  1S  0.99441 -0.22543 -0.05825  0.00000  0.00000
2 1  F  2S  0.01759  0.52617  0.10857  0.00000  0.00000
3 1  F  2PX 0.00000  0.00000  0.00000  0.00000  0.65946
4 1  F  2PY 0.00000  0.00000  0.00000  0.00000  0.00000
5 1  F  2PZ -0.00093 -0.06298  0.55022  0.00000  0.00000
6 1  F  3S  0.00861  0.44698  0.26506  0.00000  0.00000
7 1  F  3PX 0.00000  0.00000  0.00000  0.00000  0.49088
8 1  F  3PY 0.00000  0.00000  0.00000  0.49088  0.00000
9 1  F  3PZ 0.00042 -0.04196  0.37247  0.00000  0.00000
10 1  F 4XX -0.00464  0.02118  0.00033  0.00000  0.00000
11 1  F 4YY -0.00464  0.02118  0.00033  0.00000  0.00000
12 1  F 4ZZ -0.00499  0.04726 -0.05755  0.00000  0.00000
13 1  F 4XY 0.00000  0.00000  0.00000  0.00000  0.00000
14 1  F 4XZ 0.00000  0.00000  0.00000  0.00000 -0.02740
15 1  F 4YZ 0.00000  0.00000  0.00000 -0.02740  0.00000
16 2  H  1S  0.00056  0.10716 -0.25473  0.00000  0.00000
17 2  H  2S -0.00038  0.00467 -0.12071  0.00000  0.00000
  6      7      8      9      10
  V      V      V      V      V
Eigenvalues --  0.21047  1.07341  1.33089  1.49191  1.49191
1 1  F  1S  0.07881  0.02318 -0.11563  0.00000  0.00000
2 1  F  2S  0.00498 -0.18640 -1.82825  0.00000  0.00000
3 1  F  2PX 0.00000  0.00000  0.00000  0.00000 -0.94698
4 1  F  2PY 0.00000  0.00000  0.00000 -0.94698  0.00000
5 1  F  2PZ 0.26479  0.36955 -0.44336  0.00000  0.00000
6 1  F  3S -0.96947  0.22140  3.83647  0.00000  0.00000
7 1  F  3PX 0.00000  0.00000  0.00000  0.00000  1.04273
8 1  F  3PY 0.00000  0.00000  0.00000  1.04273  0.00000
9 1  F  3PZ 0.47301  0.21358  0.34373  0.00000  0.00000
10 1  F 4XX 0.08080 -0.13432 -0.86277  0.00000  0.00000
11 1  F 4YY 0.08080 -0.13432 -0.86277  0.00000  0.00000
12 1  F 4ZZ 0.05679  0.13740 -0.88216  0.00000  0.00000
  
```

Figure 10: G09 output file for MOs of HF molecules

This is a lot of numbers! Most often, you will simply need the HOMO/LUMO energy gap, so you need to find in here the HOMO and LUMO eigenvalues, which correspond to their energies. The HOMO is designated by the last orbital that is labeled by an "O", while the LUMO is the first orbital labeled by a "V". In this case, the HOMO is orbital 5, and the LUMO is orbital 6. Notice, that the HOMO has a negative eigenvalue, while the LUMO has a positive value.

THE HOMO/LUMO GAP The difference in the eigenvalues of the LUMO and the HOMO will give you the HOMO/LUMO gap (units are Hartrees by default).

NOTE: If you actually need the HOMO/LUMO Gap energy you shouldn't take simply the difference between the HOMO and the LUMO eigenvalues. For these

cases, you should compute electronic transitions.

Now, let's visualize these orbitals using Avogadro. This is when the "FormCheck" command comes in. FormCheck creates a new file called "Test.FChk", which is automatically saved in a folder named "Scratch" (C:\G09W\Scratch). If you want to see the molecular orbitals of HF, follow these steps:

1. Locate file Test.FChk as shown on Figure 11.

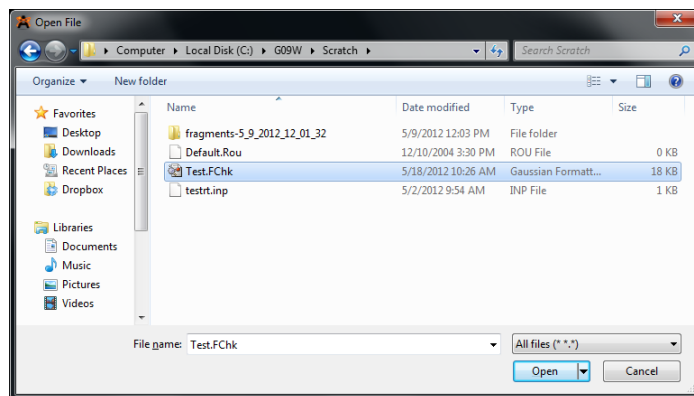


Figure 11: Test.FChk file in the scratch directory

2. Open it in Avogadro, you should see the HF molecule.

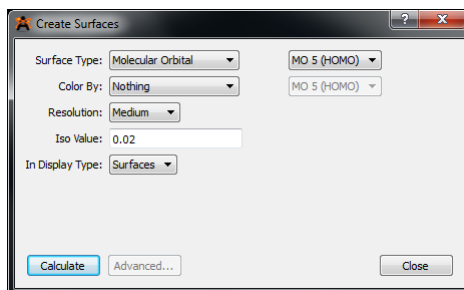


Figure 12: Create surfaces window

3. Click on Extensions → Create Surfaces. A new window should appear (Figure 12). Select "Molecular Orbitals" as Surface Type (in the drop-down menu), and choose the MO you want to visualize (in the right menu).
4. Click on "Calculate". If you click too quickly and do not let the program a second to think, it might freeze!

You should be able to see something that looks like Figure 13 that shows the LUMO graphical representation of HF molecule.

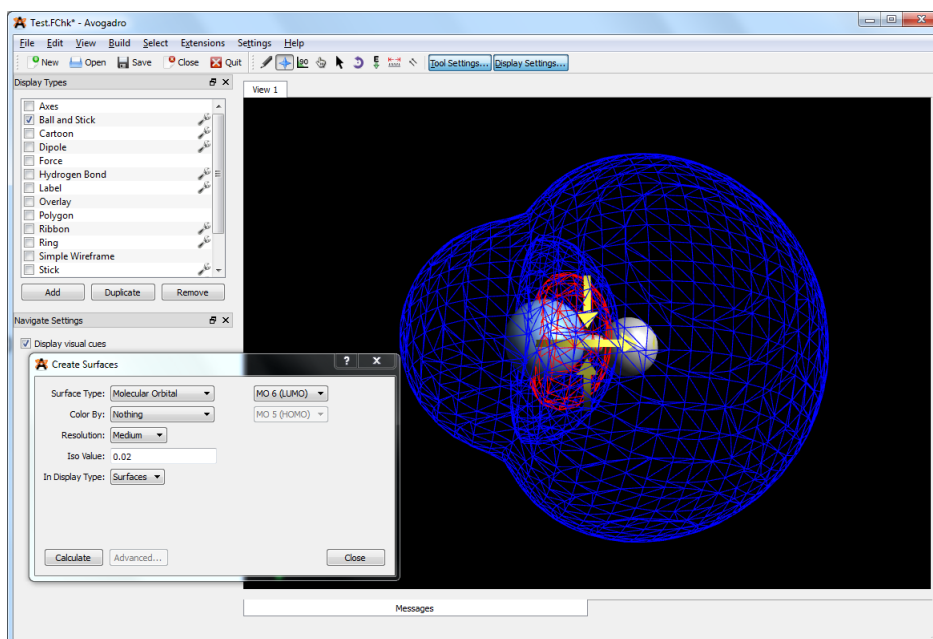


Figure 13: Visualization of the LUMO of HF

9.4 EXAMPLE 4: THE ENERGY OF STEREOISOMERS AND SCALING FACTORS (BUTENE)

Let's compare the output of a geometry optimization for the cis and trans isomers of butene. First, draw cis-butene in Avogadro, optimize its geometry. Generate the G09 input file selecting Geometry Optimization, using B3LYP model with a 6-31G(d) basis set. To see how these steps are done in details, refer to example 1 and 2. Save as cis_butene.com. Repeat the same for trans-butene. Save as trans_butene.com.

Run the jobs in G09. Save as their respective names with the .out extension. You can visualize the molecules in Avogadro, but we will focus on their SCF energies:

Cis-butene: SCF Done: E(RB3LYP) = -157.224311885 A.U. after 12 cycles

Trans-butene: SCF Done: E(RB3LYP) = -157.226083227 A.U. after 12 cycles

We have to use the correction factor listed in Table1 for the method and basis set (B3LYP/6-31G(d)) we used for the calculation: 0.9804. Therefore the real energies are actually:

Cis-butene: SCF Done: E(RB3LYP) = -154.1427154 Hartrees

Trans-butene: SCF Done: E(RB3LYP) = -154.144452 Hartrees

As proved by the energies, the trans configuration is more stable than the cis one.

10 TROUBLESHOOTING

10.1 CONVERGENCE CANNOT BE ACHIEVED?

SCF=Tight is the default for G09. You should retry your computation after including SCF=QC in the Route section.

10.2 MY JOB FROZE, WHAT CAN I DO TO AVOID RESTARTING FROM SCRATCH?

Check out the in the Gaussian reference manual whether the "Restart" option is available for the type of calculation the you were doing. This option will read the parameters from the checkpoint file (aka you can use this only if you made a .Chk file for the job that froze), and continue on.

10.3 A DOUBLE BOND DISAPPEARED FROM THE STRUCTURE AFTER COMPUTATION?

G09 has a visualization bug that sometimes happens after an "Opt Freq" calculation: the double bond looks like a single bond, but no hydrogens are added to compensate the valency of atoms involved in that bond. This is a visual problem, and if you continue on with our computations of some property, you will obtain results as if you molecule was of good geometry (i.e. with all correct bonds).

10.4 MY BATCH FILE DOES NOT RUN PROPERLY, WHY?

Usually problems with batch files are due to 3 things:

1. Name of batch too complicated: try to keep it simple, no special characters, no commas, no underscores. Something like batchAzo2.bcf will work perfectly, while something like Azo_cmpds[2].bcf will not run.
2. Names of input files are incorrect: make sure they correspond to the files existing in the current directory! You should use the Batch file generator script provided (on desktop) to create batch files without errors.
3. Name of input files contain special characters: **COMMAS ARE NOT ALLOWED IN FILE NAMES!**

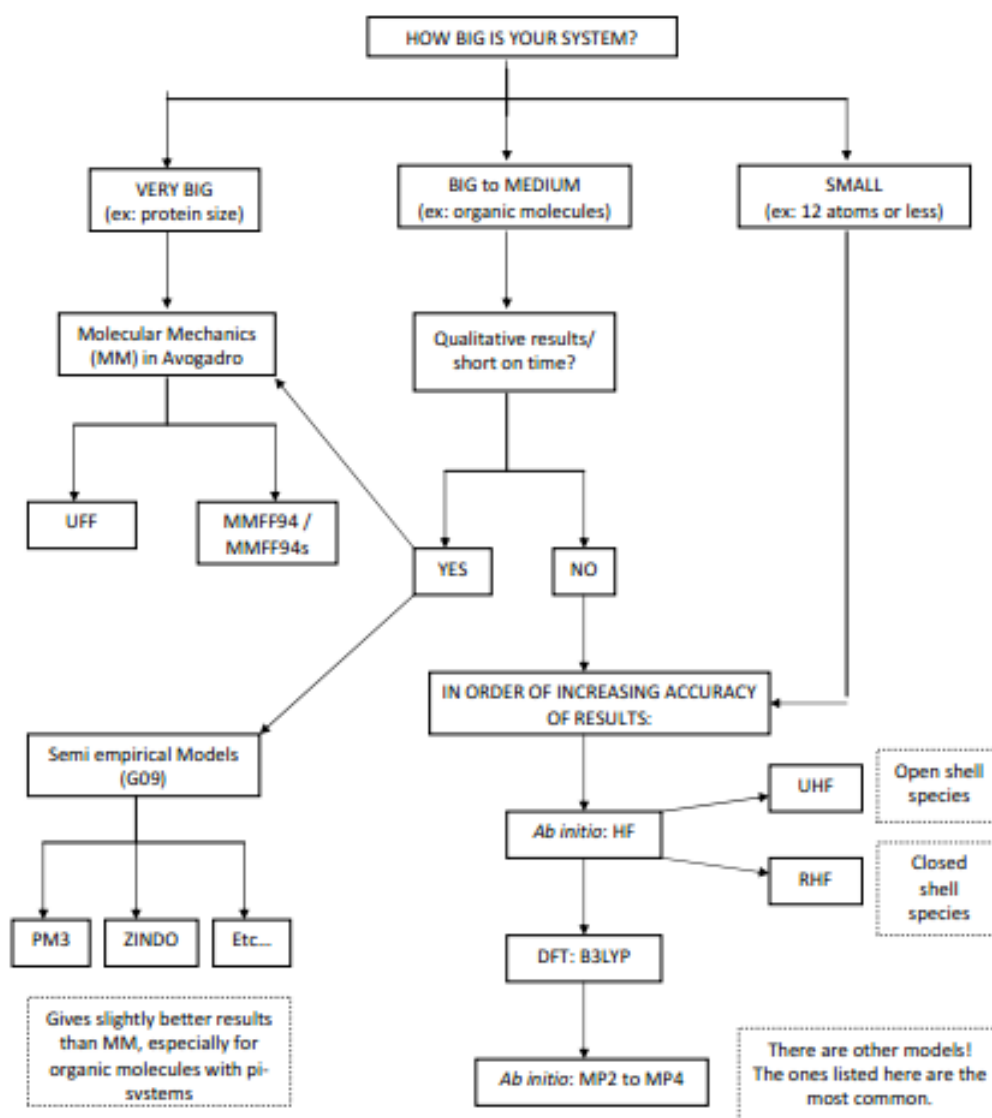
11 APPENDIX

Very good online notes are available on the web. I found computational courses by Dr Ephraim Eliav particularly good for beginners! Google him =>

TABLE 1:
Scaling factors for Freq and ZPE/Thermal calculations

Method	Basis set	Scaling Factor	
		Freq	ZPE/Thermal
HF	3-21G	0.9085	0.9409
HF	6-31G(d)	0.8929	0.9135
MP2	6-31G(d)	0.9427	0.6460
SVWN	6-31G(d)	0.9833	1.0079
BLYP	6-31G(d)	0.9940	1.0119
B3LYP	6-31G(d)	0.9613	0.9804

HOW TO CHOOSE THE RIGHT THEORETICAL MODEL?



DISCLAIMER: THE ACCURACY OF RESULTS ALSO HEAVILY DEPENDS ON THE BASIS SET USED!

12 REFERENCES

1. Young, D. Computational Chemistry: A Practical Guide for Applying Techniques to Real World Problems; John Wiley & Sons, inc: USA, 2001.
2. Jensen, F. Introduction to computational Chemistry, 2nd ed.; John Wiley & Sons, inc: England, 2007.
3.
 - <http://math.nist.gov/mcsd/savg/tutorial/ansys/FEM/>
 - http://en.wikipedia.org/wiki/Finite_element_method
 - <http://http://www.britannica.com/EBchecked/topic/430575/optimization/235508/The-simplex-method>
 - <http://www.math.cornell.edu/~web401/matt.simplex.pdf>
4. C.-K. Skylaris, CHEM3023: Spins, Atoms and Molecules, Lecture 6: Basis Sets. (http://www.southampton.ac.uk/~compchem/other/CHEM3023/chem3023_L6.pdf)
5. C. David Sherrill, Basis Sets in Quantum Chemistry (<http://vergil.chemistry.gatech.edu/courses/chem6485/pdf/basis-sets.pdf>)
6. http://www.gaussian.com/g_tech/g_ur/m_basis_sets.htm
7. http://www.gaussian.com/g_tech/g_ur/g09help.htm
8. Prof. Ephraim Eliav, Computational Chemistry Lab 2011 (<http://www.tau.ac.il/~ephraim/complab.html>)
9. David Lai Gwai Cheung, Structures and Properties of Liquid Crystals and Related Molecules from Computer Simulation (http://cmt.dur.ac.uk/sjc/thesis_dlc/thesis.html)
10. Foresman, J.B.; Frisch, A. Exploring Chemistry with Electronic Structure Methods, 2nd ed.; Gaussian, Inc: Pittsburgh, PA, 1996.
11. Computational Chemistry Comparison and Benchmark DataBase: Precomputed vibrational scaling factors (<http://cccbdb.nist.gov/vibscalejust.asp>)
12. Joseph W. Ochterski, Ph.D, Vibrational Analysis in Gaussian (http://www.gaussian.com/g_whitepap/vib/vib.pdf)
13. Joseph W. Ochterski, Ph.D., Thermochemistry in Gaussian (http://www.gaussian.com/g_whitepap/thermo/thermo.pdf)
14. http://www.gaussian.com/g_tech/g_ur/k_freq.htm
15. http://www.gaussian.com/g_tech/g_ur/k_stable.htm
16. Dr Patricia Hunt, Year 2, Lecture 7: Molecular Orbitals and Population Analysis (http://www.huntresearchgroup.org.uk/teaching/teaching_comp_chem_year4/L7_bonding.pdf)

17. Theoretical Chemistry Institute of the University of Wisconsin, What Are NBOs (and Other "Natural"-Type Orbitals)? (http://www.chem.wisc.edu/~nbo5/web_nbo.htm)
18. Dr. Kalju Kahn, Description of Electronically Excited States (http://web.chem.ucsb.edu/~kalju/chem126/public/elspect_theory.html)
19. Chemistry at Sheffield Hallam University, UV-Vis. Absorption Spectroscopy - Theoretical Principles (<http://teaching.shu.ac.uk/hwb/chemistry/tutorials/molspec/uvvisab1.htm>)
20. http://www.gaussian.com/g_tech/g_ur/k_td.htm
21. E. Lewars, "Computational Chemistry", Kluwer, New York, 2003.
22. http://www.gaussian.com/g_tech/g_ur/k_scan.htm
23. http://www.gaussian.com/g_tech/g_ur/k_scrf.htm