Lab 1: Introduction to Molecular Coordinates
CHE 525/PHY 567, Stony Brook University

1 Introduction

The primary task of electronic structure calculations is to solve approximate forms of the electronic Schrödinger equation, or the Schrödinger equation with the nuclei fixed, in order to calculate observables such as binding energy, electron density, dipole moments, etc. Thus, the starting point for such a calculation is fixing the nuclei. There are two main ways the locations of the nuclei (aka the *geometry*, the '*nucelar coordinates*, or just *coordinates*) are specified, either with *internal coordinates* that specify relations between the atoms, such as bond lengths and angles, or in *Cartesian coordinates*, where the (x, y, z) positions of every atom are specified with respect to some origin. The most common internal coordinate scheme is known as the *Z-matrix*. Here are examples of the coordinates expressed both ways for the optimum geometry of the C_2H_4 molecule we found at the `scf/3-21G` level of theory we found in Lab0.

scf/3-21G C_2H_4 in Cartesian coordinates:

```
C       0.000000      0.000000       0.000000
H       0.000000      0.000000       1.070000
C       1.156144      0.000000      -0.667500
H      -0.943102      0.000000      -0.544500
H       1.156144      0.000000      -1.756500
H       2.099246      0.000000      -0.123000
```

scf/3-21G C_2H_4 in internal (Z-matrix) coordinates:

```
C
C   1 1.3148
H   1 1.0737     2 121.9
H   1 1.0737     2 121.9      3 180.0
H   2 1.0737     1 121.9      3 180.0
H   2 1.0737     1 121.9      3 0.0
```

The unit of length in both representations is Ångstroms (1 Å= 10^{-10} m) in this case. Note that it is important to pay attention to units in general, as geometries may be represented in either Ångstrom or in bohr (aka atomic units or a.u.; 1 bohr = 0.529177 Å) depending on the specific software package being used, and many packages are capable of accepting both units with appropriate input parameters. The units used for angles in the internal coordinate representation is degrees. Both representations above will be interpreted just fine by Psi4, i.e. if you cut paste these coordinates into a `psi4.geometry()` call it will construct a planar C_2H_4 molecule. For the sake of us humans, let's spend a little bit more time discussing what each line means.

The Cartesian-coordinate version is fairly simple. The first element of each row specifies the atom (e.g. 'C' for carbon) and the following three numbers specify the (x, y, z) coordinates of the atom in Ångstroms. Looking at the Cartesian coordinates, it is immediately obvious that it describes a planar molecule since all of the y coordinates are 0, putting all the atoms in the $y = 0$ plane.

In the Z-matrix (internal-coordinate) representation, the first element of each row again specifies a type of atom to add to the molecule, and then the following numbers specify how this atom is located relative to the other atoms already specified in previous lines. For example, the first line of the Z-matrix adds a carbon atom (atom 1). The second line adds another carbon atom 1.3148 Å away from it. The third line adds a hydrogen atom 1.0737 Å away from the first carbon atom (atom 1), with a bond angle (defined between the new CH bond and the existing bond between carbon atoms 1 and 2) of 121.9°. Since two vectors (or three atoms with 2 bonds) define a plane, so far we haven't had to specify a dihedral angle yet. With the 4th line of the Z-matrix, we add another hydrogen atom with similar relations as the hydrogen atom on line 3, but now we additionally specify a dihedral angle (180°) between the plane defined by atoms 4, 1, and 2 (the new atom and the first two atoms specified on this line) and the plane defined by atoms 1, 2, and 3 (the three existing atoms specified on this line). This can be though of in chemically intuitive terms as the angle between the two C-H bonds in a Newman projection along the bond between atoms 1 and 2. The last two lines of the Z-matrix similarly specify two hydrogen atoms bonded to the second carbon atom added on line 2 of the Z-matrix.
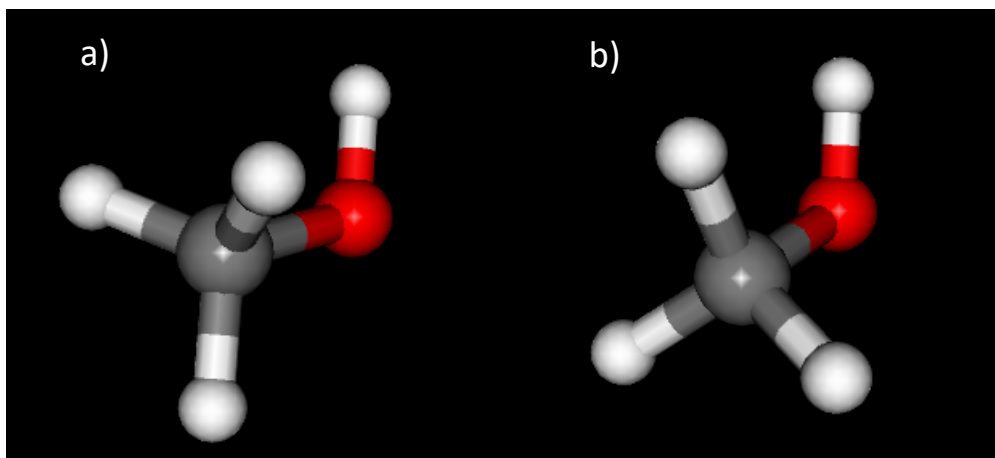
Figure 1: **Two geometries of methanol.** a) normal (staggered) geometry to use for starting unconstrained optimization. b) eclipsed geometry to do the constrained optimization.

2 Using Molden to generate Z-matrices

Both forms of the coordinates become difficult for humans to deal with as molecules get larger. There is where using software such as Molden can be extremely powerful. Log in to Seawulf using ssh -X, as you did in Lab0, and start a session of Molden with

```
module load molden/6.2
molden
```

Then use Molden to generate Z-matrices for the following molecules:

- water in its normal geometry

- water that is constrained to have an HOH angle of 170°

- methanol in its normal geometry with the hydrogens of the methyl group staggered with respect to the H atom bonded to the oxygen, as illustrated in figure 1.

- methanol in a geometry where the hydrogen atom bonded to the oxygen eclipses one of the hydrogen atoms of the methyl group, as illustrated in figure 1.

This may take some messing around with Molden to figure out how to do this, but once you get the hang of it starts to become intuitive. To open the Z-matrix editor, click "ZMAT Editor." To add your first atom, click "Add Line," and then select the element to be added by clicking on the periodic table. To add a second atom, click "Add Line" again, select an element from the periodic table, and then click on the existing atom in the main Molden window to indicate that this new atom should be bound to the first one. Add additional atoms by a similar procedure, clicking on successive atoms to indicate the three atoms that define the bond distance, bond angle, and dihedral angle for each new atom. Don't worry about getting the bond distances and angles to be exactly what are in the literature. The goal here is just to produce something close enough to the minimum energy geometry such that Psi4's optimize function can find the minimum. Molden's default bond distances should be fine. A simple way to go from water to methanol is to use Molden's "Substitute atom by fragment" feature to swap out one of the hydrogens for a methyl group. Note that you can directly change bond lengths, angles, etc. by clicking in the box that displays that particular coordinate and editing it directly. A Google search will bring up more detailed tutorials for how to use Molden's Z-matrix editor that you may find useful.

Save your Z-matrices for each molecule in separate files using Molden's 'Write Z-matrix' feature. In principle, you can make any of the coordinate output formats work with Psi4 with appropriate editing. However, the 'Gamess-US' and 'Cartesian XYZ' seem to work the most seamlessly with Psi4.

Example python scripts to carry out simple calculation on a similar small molecule (OCS) are provided in the coarse git repository. (You can pull down the latest version with the command `git pull origin master`). Specifically, `Lab1starter_OCS.py` does a simple geometry optimization of the OCS molecule, and `Lab1starter_OCS.py` does a constrained optimization, with the OCS angle frozen at its initial value. Before proceding, read carefully through the provided files. Detailed comments (indicated by '#') explain which parts of the code do what. A more 'pythonic' example (making use of more features of the Python programming language) that performs both calculations in a single run can be found in `Lab1starter_pythonic.py`.

To perform your own calculations, you can use these as examples, and modify as needed. For example, you will need to replace the geometry of OCS with the coordinates of water/methanol that you created using Molden. If you exported in XYZ format, you can also create a molecule object by having Python read in the raw text from the file generated by Molden and the inserting the read string into the `psi4.geometry()` function (see `Lab1starter_OCS.py`). If you exported a Z-matrix, you can incorporate the coordinates as you see in `Lab1starter_OCSbent.py`. The two approaches achieve the same goal, so it's a matter of personal preference which approach you choose. You should find that the the energy calculated with the molecule defined by either coordinate representation is the same, so long as the geometry is the same.

3   Performing Geometry Optimizations and Constrained Geometry Optimizations

For each of your molecules, use Psi4 to perform the following tasks:

1. Calculate the energy of all four systems at the initial geometries generated with Molden using the Hartree-Fock method (`scf` in Psi4) with the `3-21G` basis set.

2. Calculate the energy of all four systems at the initial geometries generated with Molden again using `scf` but now with the larger `6-31G*` basis set.

3. Optimize the geometry of all four molecules using the `scf/6-31G*` combination. For the case of straight water and eclipsed methanol, you will have to do a *constrained* optimization. The starter codes gives you an example of how to do this and you can also consult the documentation for setting options in the `optking` module of Psi4 online.

When making your Python program that does all of this, remember that computers are mindless idiots good at doing repetitive tasks. Indeed, the calculations going on behind the scenes in Psi4 are perhaps the epitome of boring repetitive tasks. So you may like to *automate* your code using for loops such that you are not repeating the same code over and again. Some example of this appear in `Lab1starter_pythonic.py`.

4   Deliverables

- Present figures showing the geometries of all four systems after optimization and also report the energies calculated with the `scf/6-31G*` method. You can use Molden for this but there are many other softwares that can take a geometry and give nicer pictures. For example, figure 1 of this writeup was generated with Molekel. Also provide the coordinates (you can find this in the Molden file). Clearly label and explain things such that your results can be reproduced from the information provided.

- Compare the results of your calculations 1 and 2 for all four systems in a table. Here give the *absolute* energy of the system provided by Psi4 (i.e. the total energy required to seperate all electrons and nuclei in the molecule to infinite separation). Which basis set gives a higher energy and by how much? Explain why one basis set might give a higher energy than another.

- Usually we are more interested in energy *differences* as the geometry of a molecule is changed than the total electron binding energy. From your optimization and constrained optimization results, calculate the energy differences between a) fully relaxed water vs. water constrained to have a 170° bend angle and b) fully relaxed methanol vs. methanol constrained in the eclipsed geometry at the `scf/6-31G*` level of theory. Report this energy in some familiar set of units from chemistry (e.g. kcal/mol or cm^{-1}) rather than in Hartree. For methanol compare your result to the value of the 370 cm^{-1} ground state torsional barrier derived from high-resolution spectroscopy [O. Boyarkin et al. *J. Chem. Phys.* **23**, 11359 (1999)].