Computational Exercise 1

Today we'll be getting used to some of the data analysis software on the CRC machines, and we'll be fitting approximate potential energy surfaces for three model problems in which we have pre-calculated the quantum mechanical electronic energies. The three model problems are.

1. model bond: N_2 molecules [using DFT: B3LYP/6-311++G(d,p)]
    Energy data stored in `/afs/crc.nd.edu/group/gezelter/650/N2.dat`

2. model torsion: ethane hindered rotations [also using DFT]
    Energy data stored in `/afs/crc.nd.edu/group/gezelter/650/eth.dat`

3. model non-bonded interaction: Ar-Ar dispersion [CCSD/6-311++G(d,p)]
    Energy data stored in `/afs/crc.nd.edu/group/gezelter/650/Ar2.dat`

The first few sections of the exercise are to show you how to connect to the CRC machines.

Preliminaries:

1. You'll need an X-windows program on your laptop. X-windows allows the remote computer you are using to display things on your local desktop.

    - If you have a Mac, you can download a free one here: www.xquartz.org

    - If you have a Windows machine, OIT provides a downloadable X-Windows here: https://oit.nd.edu/software-downloads/x-windows-client/ , or you can install the popular Xming: https://sourceforge.net/projects/xming/

    - If you have a Linux machine, X-windows (and terminals and ssh) should be built in.

2. You'll also need access to a command line terminal and the secure shell "ssh" command. On a Mac, this is built-in:

    - Drag `Applications / Utilities / Terminal` into your dock. We'll be using the terminal a lot this semester, so we'll want to have it handy.

    - Click the Terminal icon in the dock.

3. On a Windows machine you can download the PuTTY commands: www.putty.org for free. This will install a terminal program (PuTTY) as well as all of the back-end programs to make ssh function. We've also heard that MobaXterm (contains both X11 and ssh) is good.

From here on out, I'm going to be assuming most of you have figured out ssh and the terminal. If you are having trouble on any of this, this is a good time to consult with friends with the same kind of machine…

Connecting to CRC resources:

1. To connect to the CRC computers and redirect the X windows connection back to your local terminal, type the following command:

   `ssh user@newcell.crc.nd.edu -Y`

   replace `user` with your own Notre Dame NetID.

2. The CRC software is controlled by modules. This is a simple system that lets you load only the software you are going to use. To use the xmgr module, you can type:

   ```
   module load xmgr
   xmgrace &
   ```

3. Make sure you can use xmgrace to plot some data: `Data → Import → ASCII...` and navigate until you reach `/afs/crc.nd.edu/group/gezelter/650/N2.dat` Alternatively, you can run `xmgrace` directly on the data file:
   `xmgrace /afs/crc.nd.edu/group/gezelter/650/N2.dat`

4. Click the "OK" button to import the data file, and then "Cancel" to get rid of the import window. Your window should now have an intermolecular potential plotted in it. This is the electronic energy as N_2 molecule is stretched and compressed. The units along the interatomic (x) axis are Angstroms, and along the energy (y) axis are Atomic Units (Hartree).

Fitting a curve in xmgrace:

1. First we need to change the energy units to something more common (kcal / mol), and shift the curve so that the zero of energy is at the *bottom of the well*. With the N_2 data still on screen, go to `Data → Transformations → Evaluate expression...`

2. Select the Source to be Graph: G0, Set: G0.S0 and the Destination to Graph: G0, Set: G0.S0. In the formula box, type:
   `y = y * 627.509`
   Then, click: `Apply`

3. Click the Autoscale `AS` button to replot the data on a better scale.

4. Use the magnifying glass tool to zoom in on the minimum of the potential, which should be at around -68749.4. Follow the procedure in point 2, to ***shift*** *(not scale)* the potential up by 68749.4 kcal / mol. That is:
   `y = y + 68749.4`
   Click Autoscale and verify that your potential has a minimum energy of 0.

5. Go to `Data → Transformations → Non-linear curve fitting...`

6. Select the Source to be Graph: G0, Set: G0.S0 and leave the Destination set *unselected*. In the formula box, type:
   `y = 0.5 * A0 * (x-A1)^2`
   (i.e. a harmonic fit) and set Parameters to 2.

Pick an initial value of 1.1 for A1, and 100 for A0. Click Apply. You should see a terrible fit.

7. Change the formula in the box to:
`y = A0*(1 - exp(-A2*(x-A1)))^2`
(i.e. a Morse oscillator fit) and set Parameters to 3.
Pick an initial value of 1.1 for A1. Click Apply. You should see a much better fit. In the Grace Console window, you should be able to find the best fitting bond length and dissociation energy.

Limiting the region for a fit:

1. We'd like to tailor our fits to only the region around the minimum. With the N_2 data still on screen, go to `Edit → Regions → Define...` Click the Define button, and start drawing a box around the minimum. Use the right mouse (or option-click) to close the region. This has now defined region 0 for the fit. (Note that if option-click doesn't work, check the Preferences (Input) for your X-windows program and make sure it is set to emulate a 3-button mouse)

2. Go back to the fit window, click the Advanced tab and restrict the fit to region 0. Re-do the morse oscillator fit. You should have a much better fit near the minimum.

3. Change the formula back to `y = 0.5 * A0 * (x-A1)^2` with Parameters set to 2, and re-do the harmonic fit. Zoom and tighten the region if necessary until you get a reasonable fit near the bottom of the well. What's the best fitting bond length and force constant?

Trigonometric functions in fits:

1. Start up xmgrace and plot: `/afs/crc.nd.edu/group/gezelter/650/eth.dat`

2. Shift the y axis so that the 0° torsion is at the zero of energy (`y = y + 79.8561900875`)

3. Scale the y axis to so that you are using kcal / mol instead of Hartrees (`y = y * 627.509`)

4. Scale the x axis so that we are in units of radians, not degrees (`x = x * 3.1415/180.0`)

5. Fit the data to the functional form `y = A0*(1-cos(A1*x +A2))`

6. If the fit does not appear to be converging, try inserting different initial values and refitting the data.

Difficult fits:

1. Start up xmgrace and plot: `/afs/crc.nd.edu/group/gezelter/650/Ar2.dat`

2. Shift the y axis so that the infinitely separated dimer is the zero of energy:
`y = y + 1053.61385084`

3. Scale the y axis to so that you are using kcal / mol units instead of Hartree:
`y = y * 627.509`

4. Zoom in and find the well.

5. Fit the data to the functional form `y = 4 * A0 * ((A1/x)^12 - (A1/x)^6)`

6. Fix the problem with the fit you get - that is, look more closely at the well and see if your fit gets this feature correct. If not, use a smaller region.

7. The standard form for the Lennard-Jones potential for Argon has A0 = 0.2381 kcal / mol and A1 = 3.41 angstroms. This potential is known to give good liquid state behavior. We'll create this function as follows:

    a.   `Edit → Data Sets...`

    b.   Within this window do: `Edit → Create new → By formula`

    c.   `Start at: 2.9, Stop at: 10, Length: 100,`
          `y = 4*0.2381*((3.41/x)^12 - (3.41/x)^6)`

Homework questions:

1. The spectroscopic frequency of a bond can be obtained from the reduced mass of the oscillator, $\mu = m_1 m_2 / (m_1 + m_2)$, and the second derivative of the potential at the bottom of the well:

$$\tilde{\nu} = \frac{1}{2\pi c} \sqrt{\frac{v''(r_e)}{\mu}}$$

    What is the frequency in *wavenumbers* (cm^{-1}) of the best fitting harmonic potential for the N_2 bond? How does this compare to the experimental value?

2. What is the dissociation energy (in kcal / mol) for the best fitting Morse potential for N_2? How does this compare to the experimental value?

3. What was your barrier (in kcal / mol) for hindered rotation in ethane? The Amber force field uses a functional form of
$$V(\phi) = 0.15(1 - \cos(3\phi)) \text{ kcal mol}^{-1}$$
for HC-CT-CT-HC torsions. *Explain* why the Amber torsion energy parameter is so much smaller than your hindered rotation barrier.

4. If you fit to a Lennard-Jones potential to the Argon dimer data, what values of σ (in Angstroms) and ε (in kcal/mol) are the best choices? What is the corresponding dissociation energy? What is the distance of closest approach before the energy becomes repulsive? What is the distance of the energy minimum? Qualitatively, is Lennard-Jones a good fit? Can you come up with ideas as to why the standard Lennard-Jones potential for Argon is so different from the quantum data? Is the problem with the fit or with something else? I would like a thoughtful answer to this question.

5. One of the primary uses of empirical potentials is to provide classical forces on the atoms as they move around. Consider two atoms that are interacting via a radial pair potential, $v(r_{ij})$, where the distance between the two atoms is given by:

$$r_{ij} = \left| \vec{R}_j - \vec{R}_i \right| = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2 + (z_j - z_i)^2}$$

Use the chain rule to obtain analytical expressions for the components of the force vector on atom i,

$$\vec{F}_i = - \begin{pmatrix} \partial v / \partial x_i \\ \partial v / \partial y_i \\ \partial v / \partial z_i \end{pmatrix}$$

in terms of the first derivative of the radial potential, $v'(r) = dv/dr$.

6. For a bend potential, this is a bit more difficult. Given the formula for the bend angle,

$$\cos \theta_{ijk} = \frac{\vec{R}_{ij} \cdot \vec{R}_{jk}}{\left| \vec{R}_{ij} \right| \left| \vec{R}_{jk} \right|}$$

Use the chain rule to find the forces on atom j (the central atom) in terms of the derivative of the bend potential; $v'(\theta_{ijk}) = dv/d\theta_{ijk}$. What happens to the forces when the bond angle goes to 180°? Do you think this is a problem? Why or why not?

7. Using xmgrace, create publication-ready plots of your fits for the N_2 bond, ethane torsion, and Argon dimer potentials comparing them to the quantum mechanical data. A publication-ready plot is a plot you would be proud to show to your advisor. This means your axes should be *labeled* and have *units*, raw data should be distinguished from the fits using a *legend* and different symbols or line styles, the plot should be *titled*. Figure out how to do this in xmgrace, and add these plots to your solutions.