

***Ab initio* Calculations: Geometry Optimization and Reaction Path Following of $\text{H}_2 + \text{SiH}_2 \rightarrow \text{SiH}_4$**

1. Introduction

One of the most significant advances in the past thirty years towards our current understanding of the quantum mechanical details of molecules and chemical reactions has been the advent of the digital revolution. The ever increasing capacity for computers to handle large, complex calculations has provided an extremely powerful test on the accuracy of Quantum Theory. Early opponents of quantum theory (such as Albert Einstein and Erwin Schrodinger: of Schrodinger Wave Equation fame!) were skeptical of its success because its foundations are built around a series of postulates, or assumptions, that are not supported by empirical evidence. Rather these postulates are accepted to be true in order for quantum theory to be correct. Fortunately for us, we have eighty years of empirical evidence that supports the postulates of quantum mechanics through the predictive powers demonstrated by *ab initio* calculations.

The uncertainty of quantum theory lies in the approximation methods that we use to calculate the molecular properties that we are interested in (more about this later). Modern high-level *ab initio* calculations can calculate chemical parameters to within a few percent or better depending on the type of calculation used. As chemists, we are primarily interested in molecular properties such as energetics and structure. Common questions are usually of the type: Which structures does the quantum theory predict are the most stable? Which reaction pathways (or transition states) are preferred in a chemical reaction?

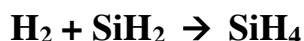
Through most of the 20th Century, these questions were typically posed through experimental evidence where the predictive powers of quantum theory were then tested by its ability to 'predict' the right answer after the experimental evidence was known. As we move into the 21st Century, scientists are shifting the functional paradigm between quantum theory and experiment. It is becoming increasingly the case that quantum theory is used to understand experimental results and not merely prove them after the fact.

As an example consider the relatively simple, but extremely important molecule Melatonin. Melatonin is a neurotransmitter that is produced in our bodies that regulates sleep-related activities and is thought to be instrumental in the production of Serotonin. Erratic Serotonin levels in humans are currently believed to be related to psychotic disorders such as schizophrenia and bi-polar disorder. Following the chemical adage: form follows function, it is crucial that we know the shape of Melatonin if we want a quantitative and mechanistic description of its role in the human body.

Melatonin contains only 24 atoms, but the chemically allowed bond patterns of these 24 atoms leads to literally tens of thousands of realizable structures. The answer cannot be provided by experiment alone. Quantum theory can be used to parse down these possibilities to perhaps 10 likely structures that can be used in conjunction with experiment to reveal the

most likely structure. These comparisons are generally made through experimental spectroscopic properties such as NMR and IR spectra that can be predicted by quantum theory.

While it is far beyond the scope of this lab to have you predict the structure of Melatonin, it is intended to provide exposure to modern ab initio methods as they are used within computational packages such as Gaussian. We will be studying the (somewhat) trivial reaction of:



We will use the Gaussian program to calculate the energies of the reactants, products, and the transition state to elucidate the reaction pathway. In order to interpret your results you will need to know a little about what Gaussian is doing. This will be covered in the methodologies section. The rest of the lab will have the familiar format. Don't worry if you've never seen quantum mechanics, as the purpose of this lab is primarily an introduction to how we use quantum mechanics. Hopefully, this lab will help your understanding of how the writing on the blackboard that you see in the classroom is actually used in modern laboratories. Everything you need to know about this lab is given in Section 2.

2. Methodologies

ALL ab initio calculations are approximations by their very nature. In order for you to have any reliability on the number you get out of the calculation, it is crucial for you to know what the approximations are and why they are used. It is important for you to understand two things right from the beginning: 1) These approximations have NOTHING to do with the validity of quantum theory, and 2) The approximation methods discussed hereafter have the capability to simulate experimental results EXACTLY. The following discussion will hopefully make this seeming contradiction clear: How can approximate methods produce exact answers?

In any calculation, there is a trade-off between calculation time, computer memory limitations, and accuracy. There are good and bad approximations depending on the type of information you want to get out of the calculation. Ultimately, the quality of the information you get out of a calculation is intimately tied to your understanding how these approximations affect the description of the physical system (molecules in our case) you want to model.

A. Three-Body Problem and Approximation Methods

The need for approximation methods lies in the mathematical frame work we use to describe the universe we live in. So far, human beings have not figured out how to exactly model more than two interacting particles. These particles can be either microscopic (on a molecular level) or macroscopic (on a classical level, or more bluntly, something you can pick up and hold in your hand). Without going into excessive detail, this is called the three-body problem and is purely a failure of our mathematical description of the universe and not the underlying physics.

Fortunately for us, this is not a new problem and there are many different ways to surmount this limitation. There are two general mathematical techniques used to handle the three-body problem: variational and perturbation methods. It is important to realize that these methods are applied to a vast array of physical problems, but the remaining discussion will be limited to how they are applied to quantum mechanics.

The variational method approximates the basis sets in our quantum calculation. Basis sets are a fancy term for the ψ 's in the Schrodinger Wave Equation (SWE). Recall from the particle-in-a-box solutions, the variation in the functional form of the ψ 's was the quantum number n . All possible values of n form a complete basis set: the set of all wavefunctions that are solutions to the particle-in-a-box problem. We'll have more to say about basis sets below. The power of the variational method lies in the variational theorem. This theorem states that no matter what approximate basis set we choose, the final answer will ALWAYS be larger than the true answer. By modifying the functional form and the size of our basis set, our approximate answers will converge to the exact answer. The utility of this result should not be underestimated.

Conversely, perturbation methods approximate the Hamiltonian in our quantum calculation. The Hamiltonian is the H in the SWE. Functionally, the Hamiltonian describes the potential surface that the particle moves on (remember that the derivative of the potential is the force). Perturbation methods divide the Hamiltonian into a part that can be solved exactly and a part that cannot be solved exactly. We won't have much more to say about perturbation methods, but don't interpret that as their not being useful. Typically perturbation methods are more expensive to use with regard to memory and time, but the advantage is that these methods typically converge much faster to the correct answer, particularly if the perturbation is small. **WARNING:** Unlike variational techniques, there is no guarantee that the result from perturbation methods will be larger than the true answer.

Finally, there is an extremely powerful approximation that is specific to molecular systems called the Born-Oppenheimer (BO) approximation. This approximation allows us to treat the nuclei and electrons as separate entities. How is this possible? Recall that the momentum (p) of a particle is linked to its mass (m) and velocity (v) by:

$$p = mv$$

Given that the mass of an electron is nearly $1/2000^{\text{th}}$ that of a proton, the electrons will have significantly less momentum than the nuclei of a molecule. We can therefore consider any change in the nuclei coordinates to have a correspondingly instantaneous change in the coordinates of the electrons, allowing us to mathematically separate the motions of the two particles. Where does the BO approximation break down? One prevalent example is for electrons that are very far from the nuclei (i.e. in Rydberg states). This underscores the important idea that you must have a basic understanding of the physical system you are trying to model. The methods used in this lab rely on the BO approximation being valid. Therefore, you should be suspicious of any results derived from the methods given here if you were trying to model and molecule in a Rydberg state.

B. Construction of the Hamiltonian and Basis Sets: Application to Particle-In-A-Box

The Hamiltonian of any molecular system in one dimension is given by:

$$H = \left(\frac{h}{2m} \right)^2 \frac{\partial^2}{\partial x^2} + V(x).$$

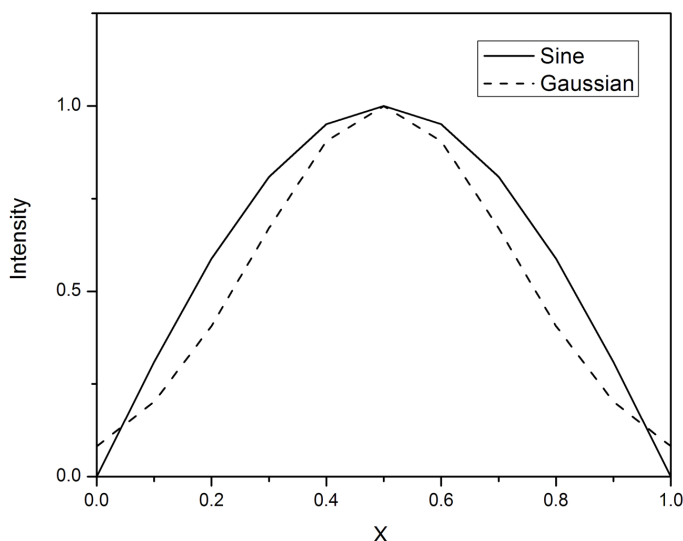
The trick is to define the potential term $V(x)$. For the particle in a box, we simply set $V(x)=0$, leading to an exactly solvable solution. For more physically realistic models, the trick lies in defining the $V(x)$ term. For a simple harmonic oscillator, $V(x)$ is proportional to x^2 , for a Coulombic potential $V(x)$ is proportional to $1/x$. For the remainder of this discussion, we will assume that the functional form of our Hamiltonian is exact for the problem we are trying to solve.

The term Basis Set is probably more abstract to you at this point. Let's continue with the PIAB solutions. When we solved the differential equation for the problem we saw that the solutions had the following functional form:

$$\psi = \sqrt{\frac{2}{L}} \sin\left[\frac{n\pi x}{L}\right]$$

where n can take any integer value from $1 \rightarrow \infty$. All possible values of n will then form a complete basis set for the PIAB problem, with each individual value of n being a single basis function in our basis set. Practically, we can never form a complete basis set for any problem because we would have to include an infinite number in our basis set. We typically truncate our basis set to some large number, say 100 or 1000. Be careful, because this truncation will lead to errors in the calculation if we do not include enough basis functions in our basis set. The more basis functions included in the set, the higher the accuracy of the calculation, but the calculation will take longer (and use more memory). It is usually up to the user to decide an appropriate trade-off between calculation time and accuracy.

As a specific example, let's say we want to calculate the wavefunction of the lowest energy (ground) state of the PIAB using the variational method, and our guess for the solution is a function that has the form of a gaussian. We have the advantage that we can compare our guess with the exact answer in the following graph.



We can see that the choice of a gaussian function as our guess is not very good. However, we can add an arbitrary number of gaussian functions to our basis set until we get the desired functional form, in this case about 10. A better guess would be the functional form $(1-x)^n$, which would only need 2 basis functions to form a (nearly) complete basis set for the ground state PIAB.

While choosing a gaussian may seem arbitrary here, there is a very good reason for choosing it. The solutions to the integrals in the SWE involving gaussian functions can be performed very quickly on a computer. The end result is that even though we need more basis functions in our basis set when using a gaussian, the total calculation time is much shorter than using a function that more accurately models our exact solution. The name Gaussian of the computational package we use to perform our *ab initio* calculations should be immediately obvious as to the functional forms used in our basis sets.

C. The Hartree-Fock Method

We will only discuss briefly the details of this method. The method itself is named after its two developers D.R. Hartree and V.I. Fock. While Hartree first proposed this method, it needed some modifications before it could be accurately used. These modifications were developed by Fock, a Russian scientist, who completed the theory while he was imprisoned in a Russian gulag for being accused as a dissident. Rather than solve differential equations as in the SWE, the HF method uses matrices to produce quantum solutions. Matrices may not be the method of choice when solving a problem by pencil and paper, especially since the matrices tend to be huge (1000's x 1000's). Fortunately, this problem is ideally suited to computers and began to be the preferred formulation from the 1950's on when computers began widespread use. Hartree-Fock is a variational method, therefore we know that the answers (the energies) we calculate will be greater than the exact number.

There is a major limitation to the HF method. No matter how big our basis set is, we will never be able to calculate the exact answer using this method. The problem lies with something called Configuration Interaction (CI). Configuration interaction is the influence one electron has on all other electrons in the system (remember BO allows us to ignore the nuclei). This is rooted in the three-body problem and there are other techniques that attempt to account for this, although the calculation time can increase dramatically.

D. Gaussian Nomenclature

As with any computer program, the user must tell the computer what they want to calculate. There is a vast array of methods outside of HF that the user may choose to use. We will restrict ourselves to the HF technique.

Finally, we need to tell gaussian how many basis functions we want to include in our basis set. This is signified by the 3-21G command. Here we are telling Gaussian that we want to use a split valence basis set. Split valence means that we will use different gaussians to handle the core and valence electrons (six gaussian functions in total). The 3-21G basis set is

considered a ‘minimal’ basis set in that it is the smallest basis set we should use to get a reasonable answer. Remember that we can always add more gaussians to our basis set, but the calculation time will increase.

4. Procedure/Experiment

In this experiment you will use the Hartree-Fock variational method to optimize the geometry and determine the energy for the optimized structures of the products, reactants and transition state for the $\text{H}_2 + \text{SiH}_2 \rightarrow \text{SiH}_4$ reaction. You will then use these optimized structures to perform a reaction path following calculation (IRC command in Gaussian) to map the minimum energy trajectory for H_2 approaching SiH_2 to form SiH_4 . The procedural methods will be provided in the computer lab for easy reference.

After the completion of the lab you will have:

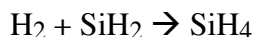
- Energies of the two reactants
- Energy of the transition state
- Energy of the product
- A picture of each reactant
- A picture of the transition state
- A picture of the product
- Bond length and bond angles for each molecule
- The IRC data file, with arrows showing both E_a & ΔE_{RXN}

5. Report

Make sure to include a print out of your results and all appropriate graphs with your final report. Make sure to address all of the following topics:

Discuss your results, in particular with respect to this:

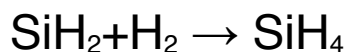
Based on your chemical intuition, do you expect the reaction of:



- to be exothermic or endothermic? Use chemical principles to support your argument.
- Discuss how you would expect the results to change using different basis sets.
- Discuss how you would expect the results to change using a perturbation method.
- Discuss your confidence in these results.
- Of the approximations discussed above, is there a particular approximation that could be removed (or addressed) that would improve the results for the IRC calculation.

Gaussian 03: A Guided Example with Screenshots

These pages describe the use of Gaussian 03 and GaussView to study the chemical reaction:

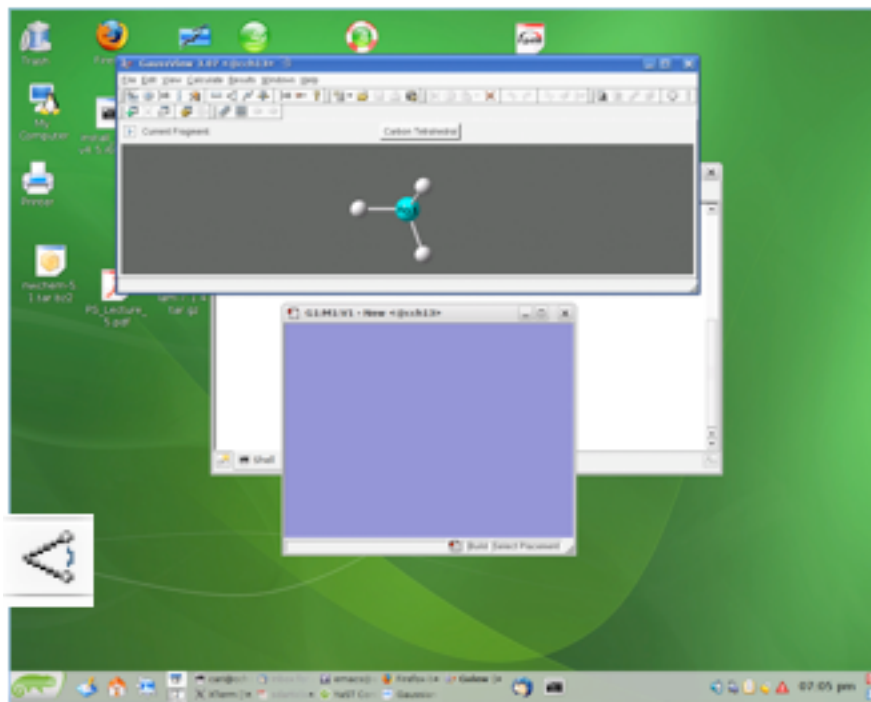


The example consists of several steps. First, we optimize the structure of the two reactants, separately. Then, we optimize the structure of the final molecule. Using the reactants and product structures, we determine the transition state. Finally, we calculate the variation of the energy along the internal reaction coordinate.

STEP 1: Getting Started

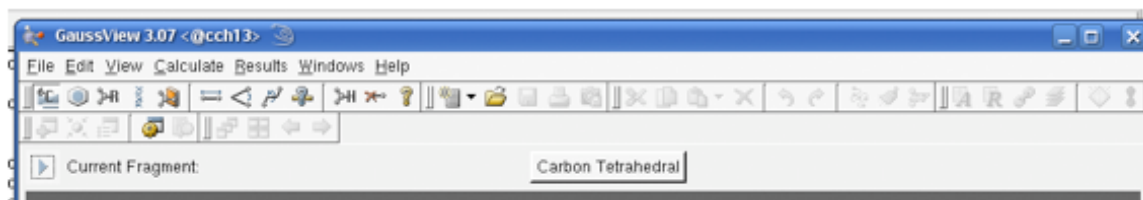
On the PC's, Log-In the **CHM 376** (no password needed).


In the Start Menu, open **GaussView**, which is a graphical interface to interact with **Gaussian 03**. Your screen should look like:





The top panel is the main *GaussView* interface, the lower (blue) panel is the *MolGroup* window.


Now, let's examine some of the features of the main GaussView interface:



Clicking on  will pop-up a periodic table, where you can select the element needed to draw your molecule in the MolGroup window.

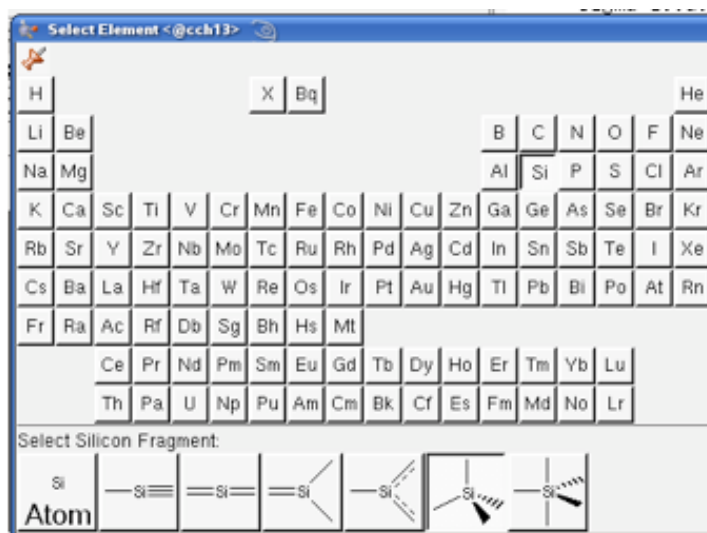
 allows you to modify a bond, and measure its length.

 allows you to modify an angle, and determine its value.

You can delete atoms using .

Note that when you select an element from the periodic table, you also have to select the fragment that you want to use:

Then, SiH_4 could be directly drawn in one step, though it is more fun to build it.



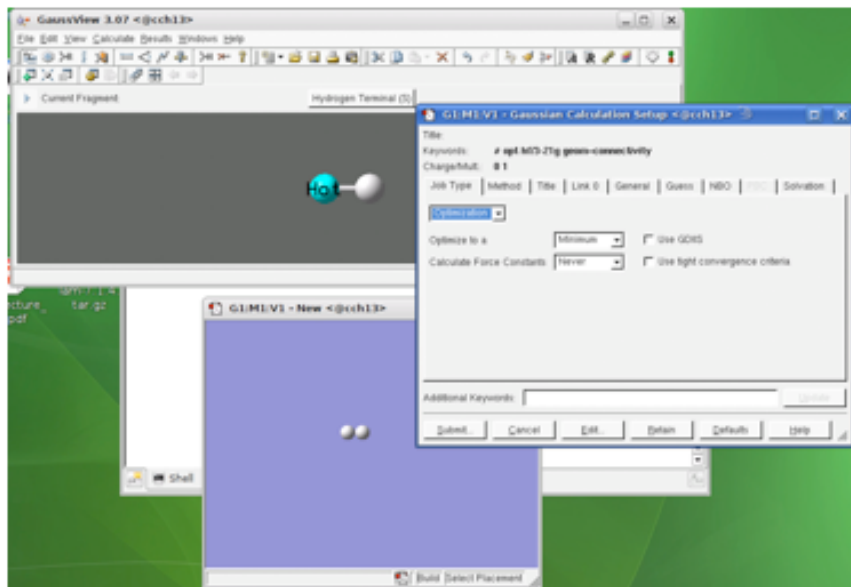
STEP 2: Reactants

In this step you will optimize the structure of the reactants. That means to find the molecular geometry that minimizes the energy of the system.

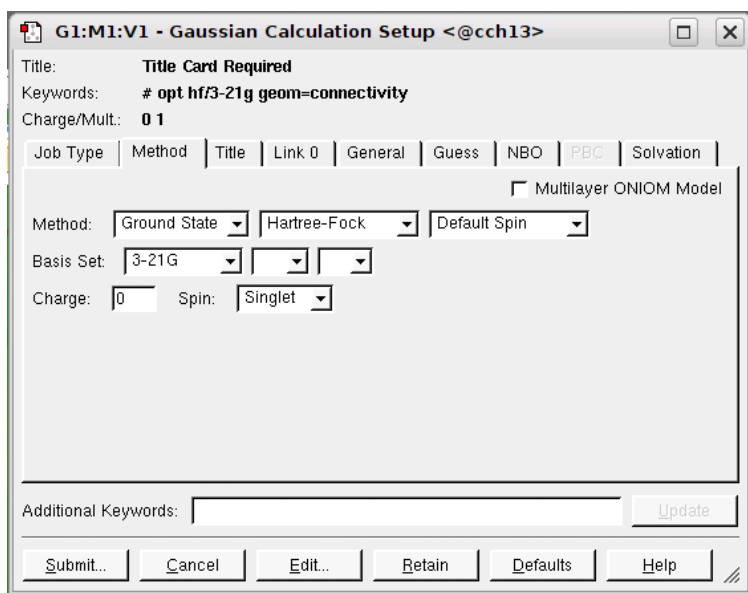
In GaussView, click on the “Element Fragment” button until the periodic table appears. Choose **H** to get the Hydrogen Atom. Click on the blue MolGroup window to add this H atom. Repeat for second hydrogen atom. Use “Delete Atom” to remove any extra unwanted atoms.

Click on “Modify Bond”, then (in the blue screen) click the first H, then the second, then choose a single bond, then OK. Click on the “Clean” button (resembles a broom).

Practice a few minutes drawing different molecules the MolGroup window. Once you understand the process, draw a H_2 molecule in the new MolGroup window. Then, click the **Calculate** menu and select **Gaussian**. The **Gaussian Calculation Setup** window will pop-up. In the **Job Type** tab select **Optimization**, and optimize to “Minimum”. The diagram on the next page shows this screen.

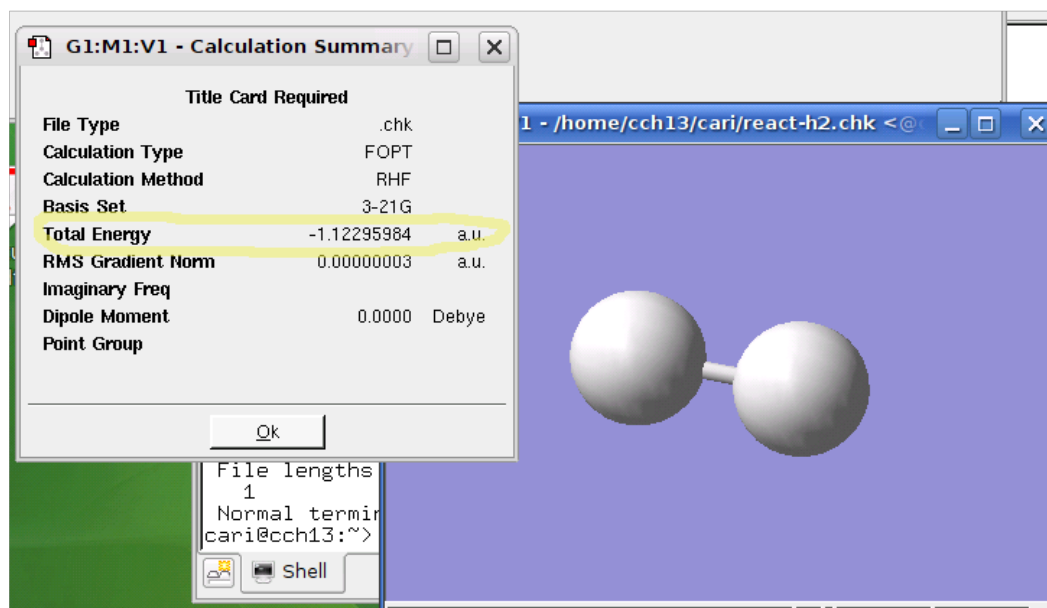


Explore the other tabs, but you do not need to change anything in them. The **Method** tab tell us that we intend to do a Hartree-Fock calculation, using the 3-21G basis set:



Click **Submit**, select an adequate name (e.g. *reac-h2*) and run the program. It will finish in a few seconds. Once the optimization is done, GaussView will offer you to open the results

file. Accept the offer and open **REAC-H2** (results are in All Capitals → may need to set “Files of type”=All Files). In the GaussView main window, select **Results** → **Summary**. The Summary window gives the energy of the system, among other information. This screen is shown on the following page. Record the resulting energy in your lab book.



You can study the optimal geometry using the **Inquire** button. After clicking the inquire button, click on each of the two atoms, and the distance between them will be displayed in the lower part of the MolGroup window. This will have units of Angstroms. The same tool is useful to obtain angles and dihedral angles of molecules. Record the bond lengths and angles.

Save the picture of your optimized molecule:

In the MolGroup window rotate your molecule (left click) to select a good viewpoint.

Click on the molecule, then from the main menu **EDIT** → **Image Capture** (this copies the image).

Paste this image into a separate Word or Python document.

Repeat the same procedure to optimize and obtain the energy of the other reactant, **SiH₂**.

Do not forget to save a picture of the optimized molecule.

Your resulting optimized structure should have an energy of -288.4842893 a.u.

Create the SiH₂ molecule. Save as “reac-SiH2.com” and submit.

Calculate → Gaussian

Job Type → **Optimization** Optimize to “**Minimum**”. Calculate Force constants “**Never**”.

Press “Submit”. Open “**REAC-SIH2**”.

Results → Summary

(Total Energy for SiH₂ = -288.4842893 a.u.)

Conversion factor: **Energy a.u. = 1 Hartree = 627.5095 Kcal/mol**

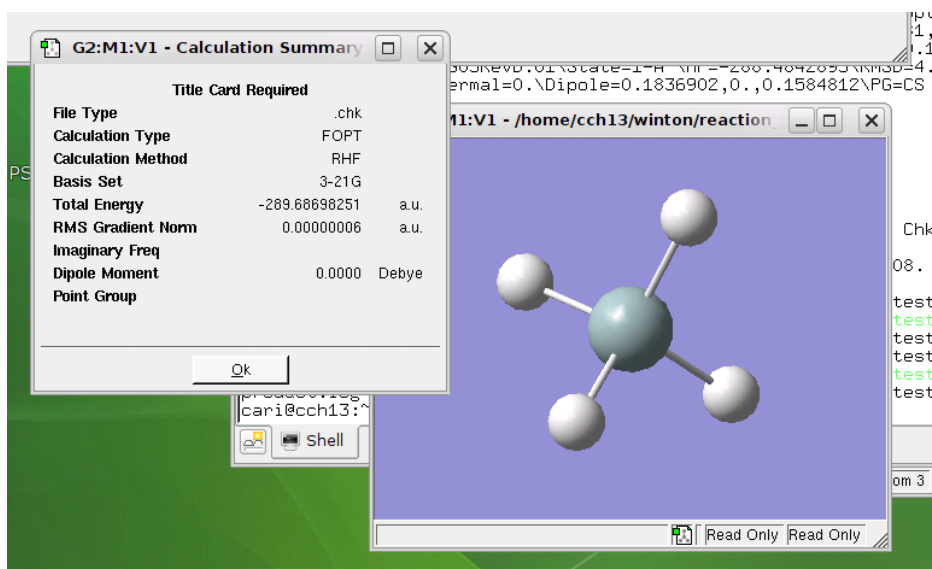
STEP 3: Product

The product of our reaction is SiH₄. Create this molecule, but do make sure it is tetrahedral, and not planer. Using the procedure described in the previous section, optimize the structure of this molecule.

Hint 1: Select the most convenient Si fragment from the table of elements.

Hint 2: Run an optimization job.

Hint 3: Once the optimization is done, check the energy of the system:



Save a picture of the optimized structure.

STEP 4: Transition State

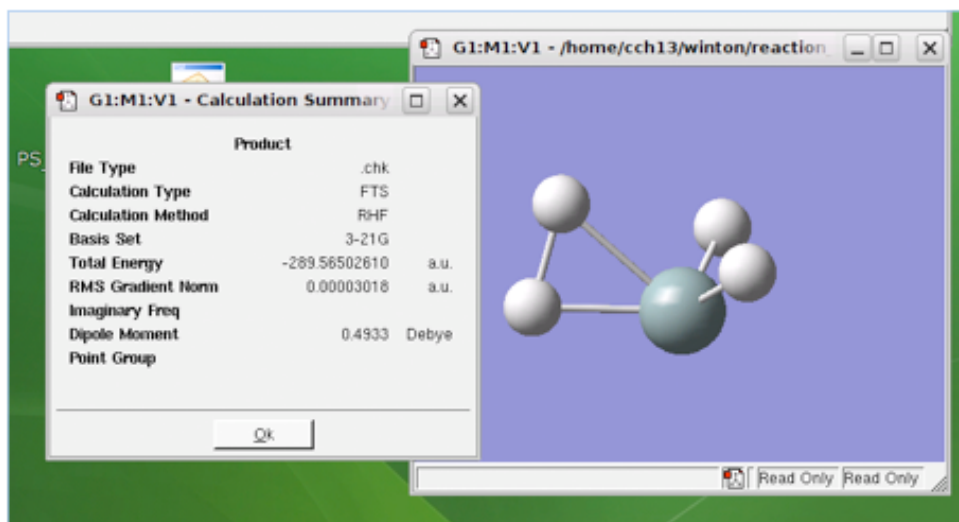
The transition state is an intermediate state during the reaction. The transition state is a saddle point in the energy landscape of the system, which means that the energy surface *curves up* in one direction, but *curves down* in all the other directions.

Gaussian has specific algorithms to locate the transition state of a reaction. One of them is the **QST2** method, which requires as input the structure of the reactants and the product in the same input file. The construction of this input file is somehow delicate, and we provide you with tested input file.

Locate the input file on your computer: **reac-ts.gjf** (located in your lab time folder on the desktop). Optimize to a **TS (QST2)**.

Load the input file with GaussView. Explore the file, and submit the calculation.

An example of the transition state is shown on the next page (it may look different).



Compare the energy of the transition state with the energy of the product, and to the sum of the energies of the reactants.

Save a picture of the transition state structure.

STEP 5: Internal Reaction Coordinate (IRC) Calculation

The energy profile along the internal reaction coordinate (IRC) is important information to describe the kinetic of the reaction.

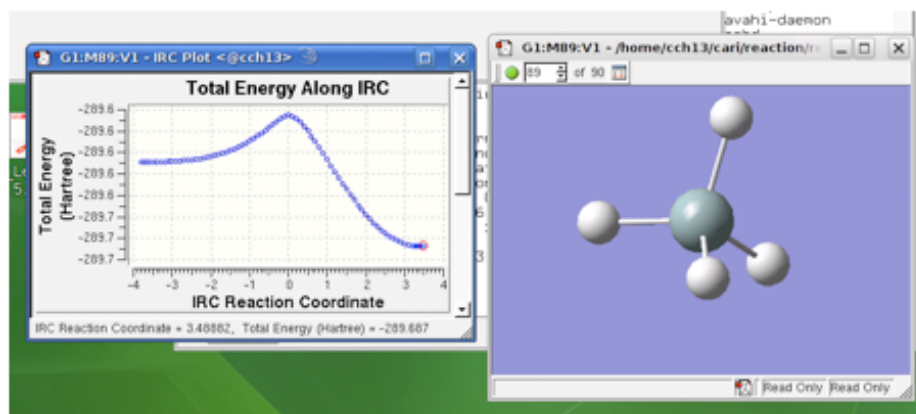
Download the input file to your computer: **reac-irc.gjf**

Load the input file with GaussView. Explore the file, then submit the calculation.

This time it will require several minutes to finish the calculation.

The input for the IRC calculation is based on the structure of the transition state. From there, Gaussian 03 can follow the reaction path toward the product molecule (forward direction) and also toward the reactant parts (reverse direction). In the blue MolGroup window, set the frame to 0 (out of 90), then click and hold on the Up arrow to visually see the attachment of the two hydrogens.

Once the calculation is complete, load the corresponding result file (all capital letters). In the Results menu, select **IRC** to generate the IRC plot.



Note on the IRC plot there is a red circle on one of the points. The molecular geometry represented in the MolGroup window corresponds to the state indicated by the red circle. Click on different points of the IRC curve to position the red circle in different places along the IRC curve.

Explore the correspondence between the IRC curve and the molecular structure.

Save the IRC data into a file:

Right click, Save data, select the filename `irc.txt`

We will then open the data file in Excel to create a printable copy of the IRC Plot.

We will need to use Excel to do some mathematical manipulations. We first want the “lowest” value of energy (which should be the Product SiH_4) to be set to **zero**, then multiply all of the energies by 627.5095 so that the energy is converted into Kcal/mol.

$$y = (y + \text{“abs. energy of lowest molecular structure”}) * 627.5095$$

↑
must be positive value!

FILE → Print

