

**The Structure and Vibrational Modes of Propargylene (HCCCH)**  
**Drew Lewis**  
**April 27, 2012**

Supervisor:

Date:

Honors Advisor:

Date:

I grant the Dean's Scholars Program permission to post a copy of my thesis on the University of Texas Digital Repository. For more information, visit <http://repositories.lib.utexas.edu/about>

**The Structure and Vibrational Modes of Propargylene (HCCCH)**

Department: Chemistry

Cannada Andrew Lewis

Signature:

Date:

John F Stanton

Signature:

Date:

# Contents

0.1	What is Computational Chemistry? . . . . .	1
0.1.1	Overview . . . . .	1
0.1.2	Methods . . . . .	2
0.2	HCCCH . . . . .	6
0.2.1	Account of What was Done . . . . .	6
0.2.2	Differences between different levels of theory . . . . .	9
0.2.3	Potential energy surfaces . . . . .	9
0.2.4	Difficulties encountered . . . . .	9
0.2.5	Discuss . . . . .	10
0.3	Computers . . . . .	12
0.3.1	Unix,Linux, and Programming in High Performance Computing . .	12
0.3.2	Problems with Computers . . . . .	13
<b>1</b>	<b>Conclusions</b> . . . . .	<b>15</b>
1.1	Future Work . . . . .	15
1.2	What was learned . . . . .	16
1.2.1	General . . . . .	16
1.2.2	Quantum Chemistry . . . . .	16

## Abstract

The vibrational modes of propargylene were modeled at multiple levels of theory using the CFOUR quantum chemistry package. Then a potential energy plot was produced for two of the unsymmetric modes at both the UHF/CCSD, which allows the  $\alpha$  electron orbitals and the  $\beta$  electron orbitals to be different, and ROHF/CCSD, which requires them to be the same except for any electrons which do not form a pair, levels of theory. The theoretical structure and vibrational modes were used to compare the more accurate representation of propargylene to the classic lewis structure version of the molecule and to discuss how resonance is used to show molecules that can not be accurately depicted using a single lewis structure. These calculations were also used to gain insight into the nature of the different energy levels of propargylene and as a starting point for more refined quantum mechanical calculations, which have not yet been carried out.

# The Introduction

## 0.1 What is Computational Chemistry?

### 0.1.1 Overview

Obviously computational chemistry is the practice of putting computers to use to solve problems in chemistry. In this report only quantum chemistry problems will be discussed although there are other ways of using computers to solve chemistry problems. But what does this mean? To start how do you give a chemistry problem to a computer? While today it is as easy as drawing the structure of the problem you want to study, it was not always this easy. Think about all of the information the computer needs that your drawing leaves out or assumes.

- Where are nuclei situated in space with respect to each other?
- How heavy is each nucleus?
- How many electrons are in the system?
- Do the electrons have to be in pairs or can they be separate?
- Is the structure you drew the lowest energy one?
- You get the point . . .

These are just a few of the things a computer needs to know before any type of calculation can be performed on the system. Luckily others have come before and written programs that will automatically deal with most of the questions above. It is important to note though that while programs will do this automatically it is still possible to find chemical systems which require more care and often require at least parts of the input to be entered by hand or at least a knowledgeable chemist checking the output of the geometry program.

Once you know how to tell the computer what your system looks like what can you do with it? Today, computational chemistry can address a wide range of problems including but not limited to geometry optimizations where the computer finds out what arrangement of the nuclei leads to the lowest energy, vibrational modes and frequencies which are the excitations the molecule undergoes when exposed to infrared radiation(IR), electronic excitations where the molecule gets excited to some higher state often by visible or ultra-violet light, as well as many others. Now you can use packages that someone else has written to do these calculations in a black box manor but it was not always so easy. Before, if you wanted new functionality you would either have to code it yourself, or wait for someone else to do so. It is only recently that commercial packages have gotten to the point where they include more functionality than the majority of chemists can use or understand.

## **0.1.2 Methods**

### **Methods Overview**

Almost all quantum chemical problems cannot be solved analytically. Or in other words it is either too hard to find an answer that can be easily written down, or it is impossible. Thus most of the problems are solved numerically where the answers are just a number generated by the computer. There are several ways to do chemistry problems. Three of

which are given below.

- Wave methods
- Density Functional Theory (DFT)
- Molecular Mechanics

Unfortunately doing these problems numerically is currently the only way to get answers to chemically relevant questions for most systems. Even when doing the problems numerically most of the problems are too hard to undertake in a complete sense. For example the nuclei in molecules are never actually stationary in the real world, but attempting to calculate things like the energy and electronic states of a molecule are difficult when the calculation assumes nuclear motion. Another issue that arises when attempting to do calculations is the more massive the nuclei get the larger effects relativity plays in determining electronic structure. These are the two biggest issues quantum chemical calculations attempt to get around.

Starting with nuclear motion, nearly every calculation done today uses an approximation called the Born-Oppenheimer Approximation, named for Max Born and J. Robert Oppenheimer. This approximation assumes that the nuclei are massive enough to ignore their motion because it is very small compared to the motion of the much lighter electrons. Treating nuclei as motionless is in some sense similar to ignoring the gravitational pull of a satellite on the earth. The earth is much more massive than the satellite that the earth's motion due to the satellite is negligible. For most problems the Born-Oppenheimer Approximation is a very good approximation and the nuclear motion can safely be ignored. But sometimes certain problems do not do well under this approximation an example being  $\text{NO}_3$ . Where the nuclear motion does have a large effect on the electronic structure of

the molecule.<sup>1</sup> <sup>I</sup>

The other main approximation is the lack of relativistic treatment. This effect has to do with the theory of special relativity as it relates to speed and mass. Basically, as the electron moves more quickly its relative mass increases. This effect is usually small for molecules with light nuclei (H-F) but begins to increase as the nuclei get more massive until it becomes quite important. The effect can not be ignored for very heavy nuclei such as heavy transition metals and actinides. Because heavier nuclei have a larger positive charge they place their electrons under more force thus increasing the electrons apparent velocity. Once this velocity begins to approach a significant fraction of the speed of light, the effects of relativity can no longer be ignored. The other time relativity cannot be ignored is when extremely accurate predictions are required. For example to effectively model the hyperfine splitting of the hydrogen atom relativity must be used.

Usually the simplest method applied to molecules is called the Hartree Fock method which deals with electron correlation only in an average sense. Basically Hartree Fock assumes that an electron only feels the effects of the average electric field from the other electrons.<sup>2</sup> To improve on Hartree Fock it is necessary to do better than average correlation methods for improvement include perturbation theory and increased correlation correction which can be used separately or together. Couple cluster usually takes the Hartree Fock molecular orbitals and applies the cluster operator to them to better account for electron correlation. Couple cluster has some nice qualities over other methods of improving correlation in that it is size consistent or in other words if you were to calculate the energy of two helium atoms at opposite ends of the universe then their energies would be equal to twice the energy of a single helium atom.<sup>3</sup> The main downside to accounting for correlation is as the number of atoms in the problem becomes larger the more expen-

---

<sup>1</sup>This has not been a complete description of the Born-Oppenheimer Approximation, because that is outside the scope of this introduction.



sive the calculations become. High level theories which accurately account for correlation are often orders of magnitude slower to run than Hartree Fock. On the problem at hand we have used standard couple cluster methods to characterize the vibrational modes and equilibrium geometry of propargylene.

# Research

## 0.2 HCCCH

The molecule consisting of three carbon and two hydrogen atoms has three experimentally detected isomers. Propargylene which is a triplet, cyclopropenylidene which is the smallest cyclic carbene and the most stable isomer, and propdienylidene which is a singlet isomer. The molecules are of interest in organic chemistry, physical chemistry and astrochemistry. The different isomers are thought to exist in the diffuse interstellar bands that exist in space. This report will focus on propynylidene in the  $C_2$  geometry although this is certainly not the only geometry proposed for the structure of the molecule, others include  $C_s$ ,  $C_{2v}$ , and  $C_{2h}$ . Depending on theoretical methods the different geometries appear to be the ground state.<sup>4,5</sup>

The first spectroscopic experiments on propynylidene were in 1965 by Bernheim.<sup>6</sup> ESR was used to understand the triplet state of the molecule. Later in 1972 Chi published the first infrared spectrum of the radical.<sup>7</sup> Then another FTIR spectrum was published by Huang in 1990 and can be seen in Figure 1 with the assignments given in Figure 2.<sup>8</sup>

### 0.2.1 Account of What was Done

The  $C_2$  isomer of HCCCH was optimized at several different levels of theory from Hartree Fock to CCSDTQ couple cluster using both the UHF and ROHF methods. The basis set

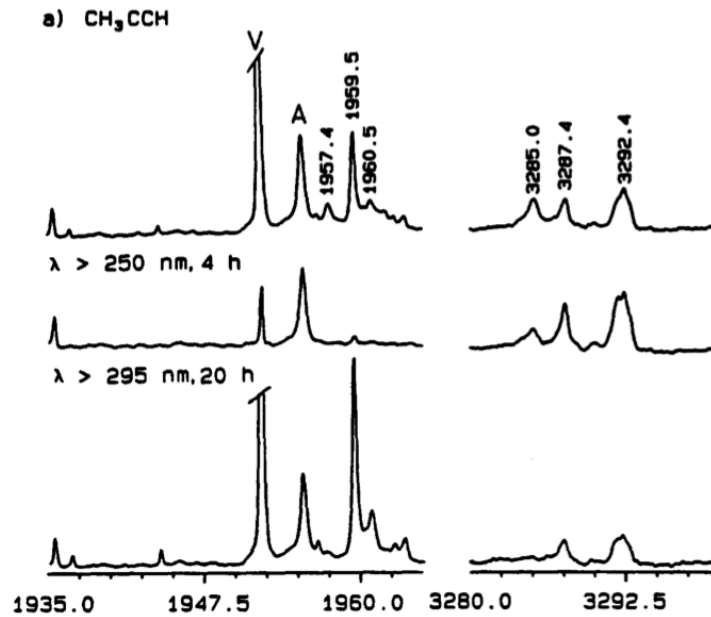


Figure 1: The FTIR spectrum of the  $\text{C}_2$  isomer in a 10K argon matrix taken from Huang et al.

Mode	$\text{C}_3\text{H}_2$	
	Calc.	Obs.
$\nu_1(\sigma_g^+)$	3274.1	
$\nu_2(\sigma_g^+)$	1074.8	
$\nu_3(\sigma_u^+)$	3265.3	3266.0
$\nu_4(\sigma_u^+)$	1610.9	1619.4
$\nu_5(\pi_g)$	302.0	
$\nu_6(\pi_u)$	401.9	402.6
		401.5
$\nu_7(\pi_u)$	245.9	245.9
$\nu_5 + \nu_7$	547.9	547.3

Figure 2: The vibrational frequencies of the  $\text{C}_2$  isomer in a 10K argon matrix from Huang et al.

used was ANO0 because in the past it has given good results when other basis sets have had problems. Specifically the ANO series of basis sets are good for atomic correlation

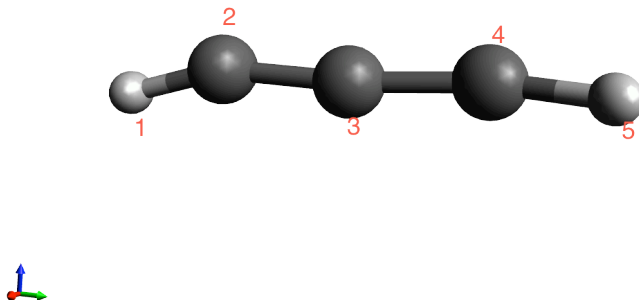


Figure 3:  $C_2$  geometry for triplet propynylidene isomer taken from Seburg et al. Calculated at the CCSD(T)/cc-pVTZ level.

C-C	1.273 Å
C-H	1.06 Å
$\angle$ C-H	162°
$\angle$ C-C	174°
$\sphericalangle$ Dihedral	143°

Table 1: Bond angles and lengths for Figure 3

calculations and have shown good results but have not necessarily been well tested for systems that involve bond formation or charge transfer.<sup>9</sup> Next the frequencies of the vibrational modes of the molecule were calculated. Most of the frequencies matched well, but there were two specifically, which ROHF had a very hard time with. The two asymmetric stretches that would convert the  $C_2$  propargylene into the  $C_s$  structure. The CFOUR quantum chemistry package was used for all calculations. The calculations for all CCSDT and higher order treatments were done using MRCC.

Table 2: Vibrational modes of propargylene with intensities in parenthesis. The ANO0 basis was used for all calculations. At the time of writing the CCSDTQ calculations were not yet finished.

Mode	UHF	ROHF	UHF/ CCSD	ROHF/ CCSD	UHF/ CCSD(T)	ROHF/ CCSD(T)	UHF/ CCSDT	ROHF/ CCSDT	UHF/ CCSDTQ	ROHF/ CCSDTQ	Experimental
A	3527(6)	3532(8)	3431(5)	3430(5)	3409(2)	3408(5)	3402(5)	3400(5)			
A	1214(0)	1332(0)	1251(0)	1255(0)	1242(0)	1235(0)	1231(0)	1230(0)			
A	519(27)	747(40)	404(32)	410(35)	406(63)	398(42)	421(48)	425(46)			
A	386(21)	551(2)	385(33)	387(33)	400(0)	379(23)	397(12)	400(16)			
A	309(5)	374(18)	283(2)	326(0)	301(2)	247(1)	308(3)	326(2)			
B	3523(38)	3522(26)	3423(67)	3424(68)	3403(0)	3398(58)	3396(58)	3395(58)			3265(71)
B	1654(28)	1445(169)	1599(18)	1684(9)	1611(14)	1450(37)	1620(13)	1661(9)			1621(7)
B	385(13)	443(6)	401(15)	401(12)	398(14)	394(16)	393(13)	392(11)			
B	352(80)	995i(4)	98(77)	302(92)	166(86)	286i(61)	218(90)	292(98)			249(100)
Geometry	UHF	ROHF	UHF/ CCSD	ROHF/ CCSD	UHF/ CCSD(T)	ROHF/ CCSD(T)	UHF/ CCSDT	ROHF/ CCSDT	UHF/ CCSDTQ	ROHF/ CCSDTQ	
R(CC)	1.28838	1.26697	1.28884	1.28782	1.2921	1.29267	1.29412	1.29404	1.294677	1.294668	
R(CH)	1.06287	1.0633	1.07279	1.07293	1.07462	1.07471	1.07518	1.07533	1.075377	1.075409	
∠(CCC)	170.15934	174.160008	172.21121	172.45808	171.80492	171.59154	171.42577	171.47168	171.37057	171.372089	
∠(CCH)	151.83916	148.57903	157.97025	157.60349	156.61724	157.42297	155.61073	155.24991	155.4078	155.31446	
↔ Dihedral	144.34299	136.50012	144.56604	142.54818	143.92416	146.16142	143.92031	143.18758	143.858424	143.704702	

## 0.2.2 Differences between different levels of theory

As is noticeable in Table 2 ROHF has a hard time dealing with the lowest frequency asymmetric mode. This mode would convert the molecule to a structure that has a  $C_s$  symmetry geometry. It was found that the reason ROHF does so poorly in this mode, is due to extremely small eigenvalues during the calculation for SCF. These eigenvalue problems cause difficulties in getting the ROHF single point calculations to converge as can be seen in Figure 4.

## 0.2.3 Potential energy surfaces

The potential energy surfaces were calculated using a grid for the two modes that have geometries similar to the  $C_s$  isomer. Shown in Figure 5 and 6.

## 0.2.4 Difficulties encountered

There are a few difficulties when trying to accurately model the propargylene system. One being the issues with convergence for ROHF as shown in Figure 4, although there

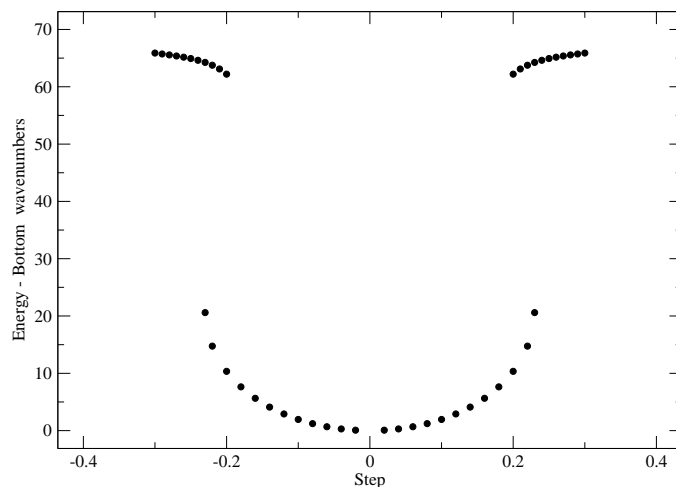


Figure 4: Grid calculation on the lowest energy B mode. The SCF has difficulties converging around the bottom of the well, which combined with symmetry breaking makes this mode hard to modal using ROHF and causes the sudden jump in energy.

are ways around this problem such as using UHF or other more complicated methods.

The other issue is one of structure. The different geometries for propargylene are not that far apart in energy and this sometimes causes problems. In face the  $C_2$  isomer can be viewed in a lewis structure sense as the resonance between the two equivalent  $C_s$  isomers and shown in Figure 7 taken from Seburg et al.<sup>10</sup>

## 0.2.5 Discuss

Essentially what has been done is to model the vibrational modes of propargylene and then to take energy measurements along the problematic normal modes to establish a harmonic potential energy surface for those two modes.

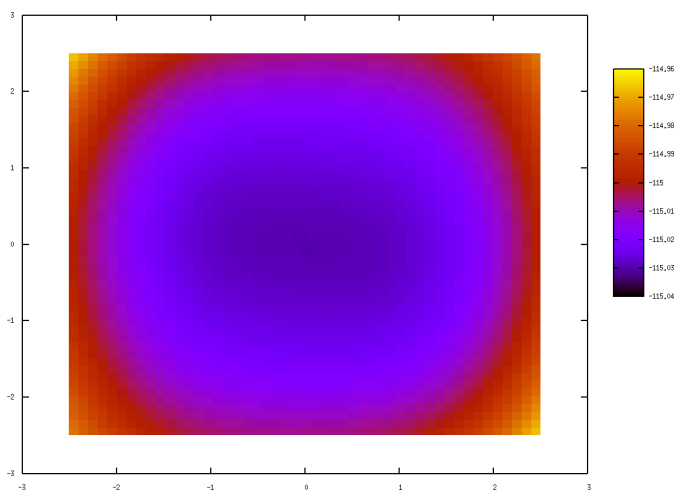


Figure 5: 2D potential energy contour map of the lowest energy B mode as well as the mode that falls around  $1600\text{ cm}^{-1}$  for UHF. The units for energy in the graph are given in Hartrees.

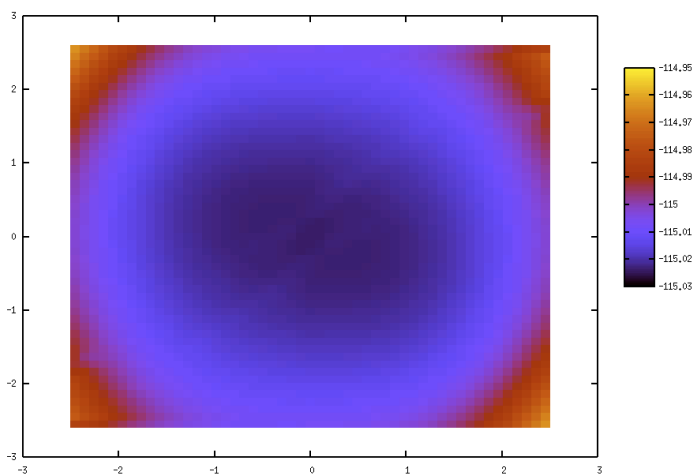


Figure 6: 2D potential energy contour map of the lowest energy B mode as well as the mode that falls around  $1600\text{ cm}^{-1}$  for ROHF. Notice if you look closely there are points missing in the middle due to inability to converge the SCF calculations. The units for energy in the graph are given in Hartrees.

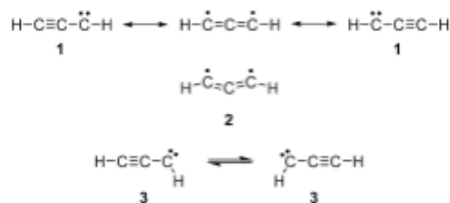


Figure 7: The top conversions shows the resonance nature of the  $C_2$  isomer as a mix of the two equivalent  $C_s$  structures. Taken from Seburg et al.

## 0.3 Computers

### 0.3.1 Unix, Linux, and Programming in High Performance Computing

Almost all high performance computing is performed either on Linux or Unix operating systems. These systems are open source and have a rich background in scientific computing application. Another reason their use is so wide spread is their stability. Unix servers can go for months at a time with out needing to be shut down, which has not always been the case for other operating systems. When calculations can take upwards of months it is important that the system be stable. Finally a large amount of work has been done to help these operating systems scale to computers with a very large number of processors and memory. Thus almost all super computers use these operating systems and ultimately one must be familiar with them to effectively use super computing resources.

One skill that is very important in using the unix or linux operating systems is the command line interface (CLI), which is a way of telling the computer what to do without having to point and click on things, but instead typing in commands. The interface is at first very awkward for the younger generation who did not grow up with CLI, but once learned it offers several benefits over a graphical user interface (GUI).

- Far more control and customization
- Quick navigation through large file trees



- Requires only small amounts of system resources so many different users can be on the same node or processor
- Ability to SSH into server, or basically log in from a different computer over the internet. It is not actually necessary to use a CLI when doing this, but lag and bandwidth are much smaller issues
- Ability to quickly navigate through a large number of files as well as edit multiple files at the same time using scripts

For anyone who seriously wants to get into high performance or scientific programming learning the CLI for Unix or Linux is almost a necessity.

Then there is the problem of programming. Many engineers use Matlab and other fields use Python or even Excel to carry out their calculations, but that is because their calculations are done in a matter of hours or more likely minutes. Often times in high performance scientific computing calculations can take days to upwards of months to run. Due to these long run times it is worthwhile to program them in more efficient languages, which today means either Fortran or C++. These languages unlike a lot of their slower counter parts are compiled languages which are translated into machine code by a compiler. Fortran and C++ are both capable of being significantly faster than Matlab, python, and much much much faster than excel. The down side is they can be harder to learn and use which is why general computing languages such as python have a wide following, but when speed means the difference between one month and two months the extra difficulty in programming in a fast language is worth it.

### **0.3.2 Problems with Computers**

Computers do have several limitations, which unfortunately quantum chemistry runs into fairly often. The first is quantum chemistry calculations often need access to large

amount of memory and when their memory runs out they are forced to write to disk. Writing to disk causes calculations to run substantially slower than they could otherwise run. There is not an easy way around this problem except to purchase more memory which can get very expensive. Sometimes it is possible to solve the problems in a less memory intensive manor, but these methods are often either hard to program or provide access to less information than the memory intensive methods.

Another issue is the problem of scalability, or the ability to divide calculations across multiple processors. Today the majority of speed increase in computing is coming from parallelization or the increase in the number of processors that a computer contains. But parallelization is only beneficial if the program is written in such away as to take advantage of the parallelization and even then a calculation can never be faster than the single slowest step. This is an issue for quantum chemistry because writing programs to take advantage of parallel architectures is harder than writing them for single processors. While quantum chemists may write many programs, they are often not professional programmers meaning as programming becomes harder less people are capable of writing these programs. Finally there is the issue in chemistry of problems tending to scale very poorly or in other words as the system grows the problem begins to take an excessive amount of time or memory or both. One example is a CCSDTQ calculation, which scales with  $n^{10}$ . This will quickly outpace the ability of computers to increase in power and so while over time faster computers will make more problems possible, there will still always be problems that are too expensive to computer and high levels. The only way around this issue is better methods, not more computing power. Increase in computing power will always benefit quantum chemistry, but it is not a fix all. Without better programs and new approaches adding extra computing power will only take computational chemistry so far.

# Chapter 1

## Conclusions

### 1.1 Future Work

For the future, the CCSDTQ calculations need to be finished in order to establish the “right” answer to the question. The right answer being correct in the limit of the harmonic oscillator approximation along with it essentially being correct for the ANO0 basis set. To deal with the issue of convergence for ROHF other techniques could be applied such as using reference wave functions that do not stem from Hartree Fock, and other methods also exist.

Finally what could ultimately be done is to treat the problem in a non Born-Oppenheimer manner. By focusing on the modes that couple together, we should be able to more accurately reproduce the spectroscopic features of propargylene using only ab initio methods.

## 1.2 What was learned

### 1.2.1 General

During the course of the research plan I learned about Unix and CLI, about programming and the limits of computational methods due to factors such as computer memory and disk space. This particular molecule demonstrates a fundamental chemistry concept: a single lewis structure cannot accurately represent a molecule and thus resonance is required to represent many molecules. This molecule also provides a situation in which the Born-Oppenheimer approximation might be holding back the calculations from getting better accuracy, although better examples of this phenomena exist such as  $\text{NO}_3$ .

### 1.2.2 Quantum Chemistry

I learned about quantum chemical methods such as couple cluster and Hartree Fock. These methods are wave function methods and provide good accuracy for small molecules as well. Also the orbitals provide meaningful insight into the real orbitals of molecules. For bigger molecules it is sometime necessary to us DFT which while providing accurate answers to certain questions such as the electron density and ground state energies can be difficult to interpret in a chemistry sense. For example trying to qualitatively plot DFT molecular orbitals is a mistake. Finally I learned that while strong efforts are being made to make quantum chemistry black box so that an organic chemist can use quantum chemistry without having to know about all of the potential pitfalls, it still has quite a ways to go before "hard" systems will become robust enough to use a black box approach. The danger comes when chemists who do not fully understand quantum chemistry try and use a commercial package to work on systems that have difficult to model behavior. In these situations it is often easy to obtain numbers from those programs that do not make

any sense, but it takes a trained eye and a certain level of intuition about the system to know which numbers are and are not reasonable. This is why there is an effort to make programs more robust so that non-computational chemists can use them without having to worry about the specific details of the implementation.

# Bibliography

- (1) Born, M.; Oppenheimer, J. R. *Ann. Physik* **1927**, *84*, 457.
- (2) McQuarrie, D. A., *Quantum Chemistry*, Second; University Science Books: 2008.
- (3) Crawford, D.; Schaefer, H. F. In *Reviews in Computational Chemistry*; John Wiley and Sons: 2000; Vol. 14; Chapter 2, p 136.
- (4) Seburg, R. A.; Patterson, E. V.; Stanton, J. F.; McMahon, R. J. *J. Am. Chem. Soc.* **1997**, *119*, 5847–5856.
- (5) Stanton, J. F. *Faraday Discuss.* **2011**, *150*, 331–343.
- (6) Bernheim, R. A.; Kempf, R. J.; Gramas, J. V.; Skell, P. S. *J. Chem Phys* **1965**, *43*, 196–200.
- (7) Chi, F. K. *Disseratation East Lansing, MI* **1972**.
- (8) Huang, J. W.; Graham, W. R. M. *J. Chem. Phys.* **1990**, *93*, 1583–1596.
- (9) Neese, F.; Valeev, E. F. *J. Chem. Theory Comput.* **2011**, *7*, 33–43.
- (10) Seburg, R. A.; Patterson, E. V.; McMahon, R. J. *J. Am. Chem. Soc.* **2009**, *131*, 9442–9455.

# About the author <sup>I</sup>

Drew Lewis started his career as an experimental chemist working in the world of inorganic. He eventually decided that wet lab was not for him and joined the group of Dr. John Stanton decidedly a very nice person. Drew has always been interested in computers, and often thought he should have majored in computer science. But even though it took a while to meander his way there he eventually made it close to computer science via computational chemistry. In the fall of 2012 he will go to Virginia Tech to work on method development for quantum chemistry in the group of Dr. Edward Valeev.

---

<sup>I</sup>I think every one should include this section. It's nice to be able to get a short bio about former DS students and see where they went.