

Conical Intersections: The Seam Space Between the Sciences

Temple University Honors Scholar Thesis

May 2020

Drew Behrendt

Table of Contents

ABSTRACT:	3
FOREWORD:	4
WHAT IS A CONICAL INTERSECTION? PHYSICS AND CHEMISTRY:	5
LIGHT:	6
COLLISIONS:	7
CONICAL INTERSECTIONS AT A GLANCE:	8
THEORY VERSUS EXPERIMENT	9
QUANTUM YIELD AND TIMESCALES OF RELAXATION:.....	11
WHY SHOULD WE CARE? BIOLOGY AND CHEMISTRY:	12
VISION:	12
DNA/RNA RELAXATION:	13
WHY ARE CONICAL INTERSECTIONS INTERESTING? PHYSICS AND MATHEMATICS:	15
OPERATORS AND FUNCTIONS REVIEW:	15
CALCULUS REVIEW:	17
MATRIX REVIEW:	18
DIRAC NOTATION:.....	19
SCHRÖDINGER EQUATION:	20
SOLVING THE SCHRODINGER EQUATION:	22
VARIATIONAL PRINCIPLE:	24
BORN-OPPENHEIMER APPROXIMATION:.....	25
BREAKDOWN OF BORN-OPPENHEIMER APPROXIMATION:.....	27
CONICAL INTERSECTIONS:	28
HOW DO WE STUDY CONICAL INTERSECTIONS? COMPUTATIONAL CHEMISTRY:	31
BASIS SETS:	31
SINGLE-REFERENCE METHODS:.....	32
MULTI-REFERENCE METHODS:	34
BENCHMARKING STUDY -- METHODS OF CALCULATING CONICAL INTERSECTIONS:.....	35
THE SEAM SPACE BETWEEN THE SCIENCES:	38
REFERENCES:	41

Abstract:

When molecules absorb light and become excited, the energy ultimately has to go somewhere; the energy can be lost by radiation, transferred to another molecule, or lost as heat. To predict how molecules interact with light and other matter, theoretical chemists use calculations based on the Born-Oppenheimer Approximation to numerically estimate energies and other properties of interest. Most processes can be explained within the bounds of the approximation; however, the spontaneous nonadiabatic loss of energy as heat cannot. These non-adiabatic processes are driven by conical intersections and play an important role in many known phenomena. Computationally, conical intersections arise out of the breakdown of the Born-Oppenheimer Approximation and the coupling of electronic and nuclear wavefunctions. Physically, conical intersections represent the seam space of degenerate electronic states on the potential energy surface of a molecule. Metaphorically, conical intersections represent the seam space of the research frontiers in biology, chemistry, physics, mathematics, and computer science. The present work is a review of the work in, and application of, each respective field related to conical intersections and a benchmarking study of the most viable current methods used to calculate conical intersections.

Foreword:

This piece is a scientific perspective. I wish to walk the reader through a logic loop I have constructed throughout my undergraduate career about computational chemistry as a whole and to describe why I am dedicating my future career to it. Conical intersections, as will be described in detail, are simply a tool to be used as an exemplar of the complexity, broad applicability, and importance of studying computational chemistry. Personally, I write this piece as a way to document my way of thinking at this point in my life. Broadly, I write this piece dually as a way to convince scientists that this is an important field to concentrate efforts on and to convince any interested party of the effectiveness, beauty, and worthiness of investing in this area of science.

I want the readers of this perspective to be able to understand the significance and connection of every section; however, I also want to convey some more dense mathematics and abstract principles that may only be of interest to those with a background in the sciences. Therefore, I write each section in order -- from most general, fundamental, and easy to understand, to the most complex and abstract material.

I believe there is no such thing as “math people” and “non math people.” There are only levels to which one can understand abstract mathematics: basically, how well you can picture math in your head. Therefore, I am going to attempt to illustrate as many concepts as possible using analogy and explicit real-world examples.

What is a Conical Intersection? Physics and Chemistry:

It is important to start with the basics of energy and chemistry. Chemistry is the study of molecules, and everything in chemistry is driven by relative energy. Energy drives chemistry and chemistry is everything; therefore, energy drives everything. Reactions take the lowest energy pathways, and information on rates and reaction types provides insight on the process of how molecular processes work. Learning how things work is science and learning how molecules work and interact is chemistry.

What makes chemistry so strange and difficult at the fundamental level is that molecules and atoms behave quantum mechanically. Quantum mechanics is intimidating, and while it may be very frustrating to study at times, the fundamentals are accessible to understand, and the logic applies to understanding many worldly situations. The following is an analogy for how it works:

Example 1: Quantum Stairs

Channeling memories from high school physics: your energy can be described by your kinetic (movement) energy and potential energy (think gravity). Potential energy can be visualized as climbing a hill; when you climb higher, you have more gravitational energy available to 'use' at a later time. In classical mechanics, your energy can be anything -- imagine a hill where you can stand at any spot and have any relative potential energy. In quantum mechanics, things become quantized -- energies can only exist at specific values and never in between. Imagine for this a staircase where you can only stand on any single stair; you can only have the potential energy of the stair you stand on at any given time. Molecules are quantized. Energy is the driving factor for everything in molecules, but it can only exist at specific values like steps on a staircase.

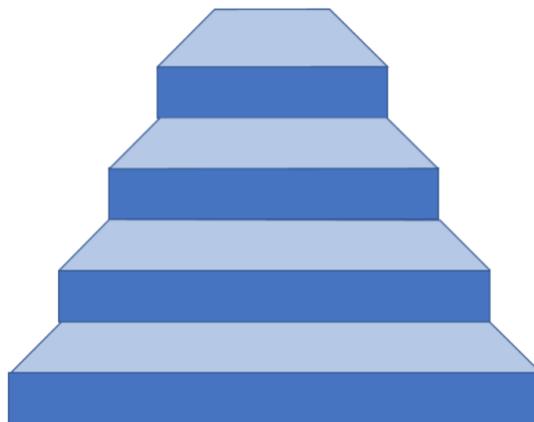


Figure 1. Image of the quantum staircase. Each light step represents a different allowed energy. The shaded dark area is the forbidden region between steps.

But how does the energy go between the levels if it can only ever be one or the other? The answer to this question is at the heart of physical chemistry. Molecules can relax between energy levels in three ways: 1) light; 2) collisions with other molecules; 3) conical intersections.

Light:

A result of the quantization of the energies allowed for a molecule is its *spectra*. Okay, but what is a spectrum?

Light is energy. The energy of a specific photon of light is proportional to its frequency and inversely proportional to its wavelength,

$$E_{\text{photon}} = h\nu = \frac{hc}{\lambda} \quad (1)$$

where E is the energy of a photon, h is Planck's constant, c is the speed of light, ν is the frequency of the light, and λ is the wavelength.

If a molecule is hit by light with the exact energy of two energy levels, it can absorb and excite. Conversely, a molecule can relax and spontaneously move down an energy level by emitting light with the same energy as the difference between the levels. The exact values for the

exact energy levels in any given molecule are unique. If the levels are unique, so too are the differences between the levels. Thus, every molecule has a unique absorption and emission spectrum. A spectrum is just a way to tell the unique energy levels of a molecule.

Relating back to Example 1, imagine the quantum stairs again where you can only be on one step at a time. If you get hit with a photon of light with the same energy as the difference between the rungs, you absorb it and instantaneously are on the next step up. If you relax and jump down a step, you shoot out a photon of light with the exact energy of the difference between the steps.

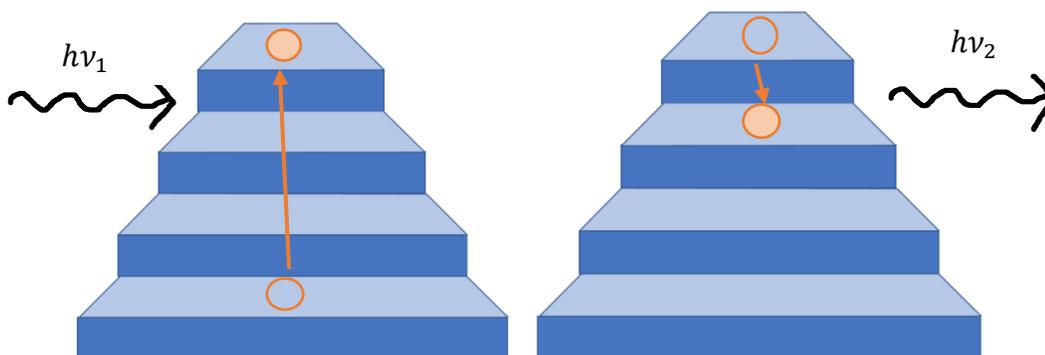


Figure 2. Light exciting a particle up the quantum staircase, and the particle emitting light as it relaxes. Note that the energy of the light can be different for the absorption and relaxation; the transition between each possible step generates the energy spectra for the particle.

Collisions:

When discussing chemistry, it is crucial to remember that almost nothing in the real world happens in isolation. Molecules are tiny; the ratio of the diameter of one atom to one meter is the same as the ratio of one meter to the distance from the Earth to the Sun. To study an individual molecule in a vacuum is very difficult and involves processes like laser cooling and trapping. So instead, chemists must observe characteristics of a large number of molecules. For a

homogenous group of molecules, each one has the same energy levels and transitions. Therefore, by observing what happens in the bulk, one can gain insight as to how the individual molecules work. So, the example case of a single molecule absorbing a single photon is extrapolated to become many of the same molecules absorbing identical photons at the same time. In the jumbled randomly-moving-mess that are these molecules, an excited molecule can relax and ‘step down a stair’ by colliding with another molecule and transferring the energy. The energy can excite the other molecule or turn into kinetic energy that speeds up the movement of the molecules.

Conical Intersections at a Glance:

Conical intersections are a pathway in which large molecules can transition to a lower level electronic state (step down the quantum stairs) without emitting light or colliding with another molecule. At first, this may sound counter-intuitive with everything previously stated. For this reason, conical intersections were once thought to be incredibly rare; however, it has been found that any molecule with at least three atoms can have one. Extrapolating the quantum stair example, imagine looking down a staircase that is really wide. You can jump down a step right in front of you, however if you go to a direction left or right, imagine the steps change height and converge at some point. This allows you to move down without having to jump.

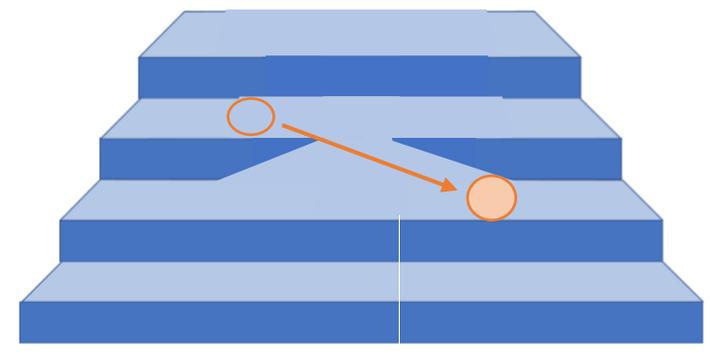


Figure 3. Conical intersection on the quantum steps. No light is emitted upon relaxation.

Theory versus Experiment

Moving forward, it is now of interest to briefly distinguish between theoretical and experimental chemistry. Experiment is simply what one can physically observe molecules doing using various experimental techniques. As mentioned above, it is bulk groups of molecules that are actually observed. Theoretical methods are based on *ab initio* calculations, which means that the energies or characteristics of molecules are modeled purely from mathematics stemming from the size, orientation, and forces in the individual atoms that make the molecule. This is an intimidating thought, and it is an intimidating practice that will be discussed in-depth later. For now, understand that this is extremely complicated and gets more complicated the more pieces that are added. So, molecules are calculated individually in theory; whereas in experiment, molecules are typically only studied in a bulk setting.

Still, the whole purpose of theory is to explain what is actually happening in the world. Therefore, to compare theory and experiment, one must either tailor the experiment to mimic an isolated molecule in a vacuum or tailor the theory to include the effects of the molecule of interest being around many others. Furthermore, there are many concepts that are useful theoretically that are physically impossible to measure directly in the lab for a variety of reasons. To make statements about the efficacy of a theoretical method to describe something that is not experimentally testable, it is necessary to compare that method to an actually possible experiment that tests something adjacent to the property originally desired, as is shown in Example 2.

Example 2: Theoretical Investigation of Pyrazine

The first project I worked on as an undergraduate involved finding possible large molecules that could break apart to form HCN in comets and the interstellar medium.² The source of this molecular formation was unknown, so it was of interest to find molecules that had energy pathways low enough to break apart and form HCN from sunlight. In order to validate the computational methods that predicted these pathways, my group compared experimental spectra to computational spectra. Since the energies we were investigating were on the order of UV light, we compared experimental and computational spectra for pyrazine, one of the molecules we identified as being a possible source of HCN, shown in Figure 4. Because of the high agreement between the two methods, we concluded that the theoretical methods used were a valid way to estimate the energy of the molecule as it transitioned through the reaction.

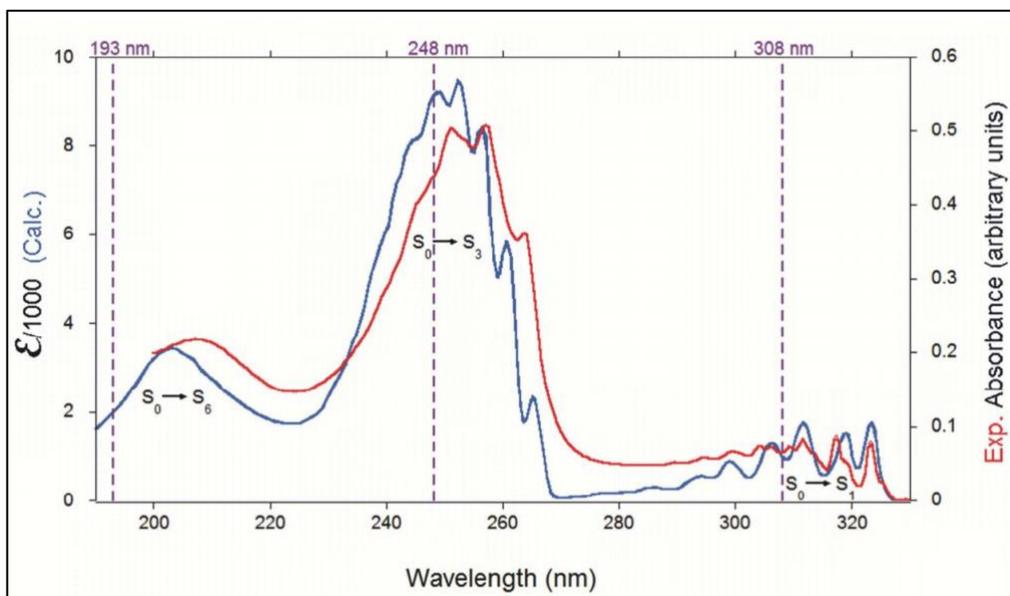


Figure 4. Calculated (blue) UV-Vis spectrum of pyrazine done at the TD-DFT level with APF-D/6-311+G(2d,p) model overlapped with experimental spectrum (red).² Agreement indicates validation of theoretical method.

Quantum Yield and Timescales of Relaxation:

To further complicate the picture of what happens when a molecule absorbs a photon, all three processes mentioned above can and will happen in some combination for any given bulk group of molecules. So, in an experiment, it is essentially impossible to isolate any one of those three processes, however, there are ways to determine the relative ratio of number of molecules that undergo each process.

One such way is known as quantum yield. The idea of quantum yield is fairly simple; one shoots a known number of photons at a sample of molecules and count how many photons the sample emits. The number emitted divided by the number absorbed is the quantum yield. A quantum yield of one means that every molecule emits a photon to relax; a quantum yield of zero means that all of the molecules relaxed through collision or conical intersection.

The second measure for determining exactly what pathway is occurring is by comparing relative timescales of relaxation. Relaxation by conical intersection happens on the order of femtoseconds (10^{-15} s), and collisions happen on the order of nanoseconds (10^{-9} s). So, if it is known that a relaxation process happens super-fast with a low quantum yield, then there is evidence indicating the existence of a conical intersection. To highlight just how different these time scales are, one femtosecond compared to one nanosecond is the same as one second compared to about twelve days; if a conical intersection took the same time as the blink of an eye, then relaxation by collision would take the same time as labor-day weekend. So, while timescale alone may seem like a roundabout way to provide evidence of a process, conical intersections and non-adiabatic transitions are the only way to explain ultrafast molecular relaxation.

Why Should We Care? Biology and Chemistry:

Biology is the study of life and the various processes required for it to exist in the way it does. Macroscopically, biology is often focused on cells, organisms, and ecosystems. The study of processes of these large systems does not necessarily involve chemistry directly. For example, to understand genetic heritage and evolution, one does not need to understand DNA on a molecular level. However, microscopically, everything is made of molecules. Therefore, when any biological process is broken down into specific pathways and reactions, the specific chemical processes and molecular interactions become essential to understand fundamentally how the biology happens. In this category of research lies the seam space between chemistry and biology. Conveniently, but by no coincidence, conical intersections are ubiquitous in biology and govern processes that are essential for life as we know it.⁶ For the purpose of this review, two biology pathways are explored: vision and the photo relaxation of DNA/RNA. Both of these molecular transitions have been shown to happen in femtoseconds, proving that there must be a conical intersection involved in the process.

Vision:

The first chemical step in the vision process is the cis-trans isomerization of the molecule retinal.⁷ Light passes through the lens of the eye and through the pupil, eventually hitting the retina where photoreceptors called rods and cones convert the light information to electrical pulses that can be detected and interpreted in the brain. The photoreceptors that absorb the light are nothing more than molecules with specific energy levels as explained previously. The molecule responsible for this absorption is called retinal.

When light hits retinal, the molecule transitions through a conical intersection to ‘flip’ over the double bond as shown in Figure 5. This molecular change causes a change in the greater

protein that contains the retinal, and the light energy is turned into mechanical energy and eventually electronic energy in the form of cell signaling.⁸⁻⁹ This pathway, essential to life as we know it, critically involves a conical intersection; thus, this is proof of the importance of studying conical intersections to biology.

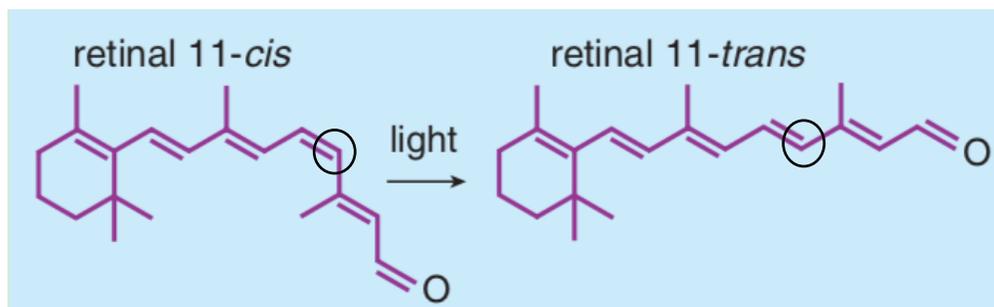


Figure 5. The cis-trans isomerization of retinal.⁸ The changed carbon is circled for emphasis.

DNA/RNA Relaxation:

The DNA and RNA building blocks (Uracil, Adenine, Thymine, Cytosine, and Guanine) are the single molecules that align in a chain to form the sequence of code that contains the information on how to make proteins in organisms. These sequences are essential to store and pass on this information between generations of cells and ultimately organisms. The exact origin of life is a question that many have spent their lives researching and is one that I am in absolutely no position to speculate on. That being said, it is known that some primordial form of DNA and RNA must have existed in the first organisms, at least the first organisms that led to the phylogenetic tree that contains all life currently existing on Earth. These organisms would have not had much, if any, protection from constant and direct light from the sun. The UV-light from the sun that is not captured in the atmosphere of the Earth is high enough energy to excite molecules to the next ‘energy step’, which can cause change and therefore damage to the important life molecules like primordial DNA. It is crucial, therefore, for these DNA/RNA monomers to have a way to energetically relax quickly without damaging itself or surrounding

molecules. By no coincidence, it has been shown that these monomers have conical intersections that lie between the states UV light excites them to and the relaxed ground state.^{4,10} Essentially, a monomer can get hit by UV light and relax within femtoseconds through a conical intersection without creating any damaging radiation. Thus, this again proves the importance of studying conical intersections to fully appreciate the mechanisms of biological processes.

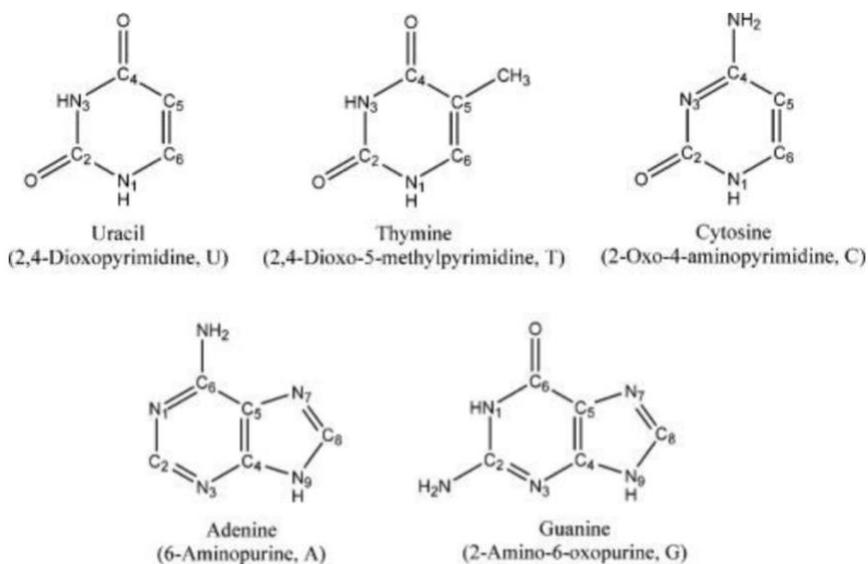


Figure 6. The molecular structure of the DNA and RNA nucleobases.¹¹

Why Are Conical Intersections Interesting? Physics and Mathematics:

This section will start light, but eventually deal with some denser mathematical concepts. For any math subject with which a high school student would say, “why would I ever need to know this?” the answer lies somewhere in quantum mechanics and computational chemistry. As such, there is no way to review the entirety of geometry, algebra, and calculus in this paper, which is required to fully understand the mathematics presented here. This section will focus first on four specific mathematical concepts that are fundamental to theoretical chemistry and second on specific physical situations that arise and result in conical intersections.

Operators and Functions Review:

Operators are fundamental mathematical tools that are very useful when dealing with real-world physics and computer science. The logic behind operators is quite simple yet vastly powerful. As presented here, operators are an idea in the broader subjects of linear algebra and calculus so the concepts can be intimidating. However, breaking it down to the most basic ideas makes the logic accessible. To understand operators, first functions must be reviewed. Everyone is familiar with the function of a line,

$$y = mx + b \tag{2}$$

where y is a function of x , and m and b are constants. Insert a value for x and y will change. It is important to remember that in algebra, letters represent concepts. Any letter can be changed for another letter, as long as all parties understand what the letter means. Math is a universal language because the concepts transcend language, as any letter, symbol, or otherwise can be used in place of any other. With that being said, we can rewrite equation 2 to be,

$$\psi(x) = mx + b \tag{3}$$

where $\psi(x)$ is used to show that ψ is a function of x . Yes, this equation *looks* more complicated than equation 2, but they mean the same thing. Now, even more abstractly, for the present purposes, what ψ equals does not matter; all that matters is that it is a function of *something*. That *something* can extend beyond just one variable. In fact, a function can have multiple variables (think x, y, z) that affect the answer to the function. For example, if I made a function describing where a marble is on a table, the function would depend on both the x distance and y distance from the corner of the table.

Building upon the concept of a function, an operator is essentially an action to a function that produces results of interest. So, for present purposes, an operator is like a function for functions; if you apply an operator to a function, you get the function back times a number. Yes, that is confusing to read but mathematically, this type of operator (used in what is called an “eigenvalue” problem) is represented as,

$$\hat{A}f(x) = af(x) \quad (4)$$

where \hat{A} is an operator (the hat means operator), $f(x)$ is a function (called an eigenfunction), and a is just a number (called an eigenvalue). The crucial concept to understand is that an operator times an eigenfunction returns the function times certain set of numbers: its eigenvalues. If the operator changes the function after applied, that function is not an eigenfunction.¹²

Operators are important in physics because they elicit observable properties in real mathematical systems. If you have a mathematical function that describes a system, then you can use specific operators to find out information about that system. For example, one can describe the movement of electrons using an eigenfunction called the wavefunction. This is a fundamental idea behind quantum mechanics, which will be discussed in detail in following sections.

Calculus Review:

In order to ensure equal footing for all readers, the essentials of derivatives and integrals are briefly presented in this section. A derivative is the rate at which something is changing with respect to something else. The first example every student learns is that the derivative of position with respect to time is velocity, and the derivative of velocity with respect to time is acceleration.

Example 3: Marble Calculus

Let $x(t)$ represent a function of the position of a marble with respect to time (as time changes, the position of the marble changes). The velocity of the marble as a function of time can be represented by the equation,

$$v(t) = x'(t) \quad (5)$$

where the (') represents taking a derivative, so $x'(t)$ is the derivative of $x(t)$.

If we take another derivative, we get the acceleration of the marble,

$$a(t) = v'(t) = x''(t) \quad (6)$$

The integral of each function is a little more abstract to comprehend, but the take home message is that an integral is the opposite of a derivative. The notation for position is as follows, if you have a function that describes the $v(t)$ of a marble, then you can get the position at any given time by

$$x(t) = \int_0^t v(t)dt + c \quad (7)$$

where t is the time you want to know where the marble is, and c is a constant of where you are measuring from.

Moving on from the marble example, the concepts of calculus can describe any rate of change and explain many other concepts in physics and beyond.¹²

Matrix Review:

The third mathematical concept that is essential to understanding computational chemistry is matrices. A matrix is a mathematical tool used to solve many equations at the same time. If a system of interest deals with many functions at the same time, and if solving those functions involves calculus, it can get impossible to solve everything just by writing equations and using regular algebra. Thus, matrices and the tools of linear algebra are employed to solve increasingly complex problems like those in computational chemistry. For example, take the two-equation system,¹²

$$3x + 2y = 10 \quad (8a)$$

$$5x - y = 6 \quad (8b)$$

If both of these equations are true at the same time, one could simply solve one for either x or y, substitute into the other equation, and solve for the other variable. When one has the same number of equations as unknowns (2 equations for the 2 unknowns above), the system is always solvable. However, for example, for systems of 6 equations with 6 unknowns substitution is no longer a valid way to solve the problem; it would just take too long. To solve systems like this, matrices are used. Equations 8a and 8b above can be written in matrix form as,

$$\begin{bmatrix} 3 & 2 \\ 5 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 10 \\ 6 \end{bmatrix} \quad (9)$$

By observation, it is easy to see how this matrix is formed from the substituent equations. The specific ways to solve such a matrix are not as important for contextual understanding as just the knowledge that this *can* be solved and solved relatively quickly by a computer. Computers can solve large systems of equations using matrices and linear algebra, and this is essential for computational chemistry.

Dirac Notation:

When dealing with very large systems of equations that involve many variables and differential equations, using typical formulism can get very complicated, fast. This is why in 1939, Paul Dirac invented the language most commonly used in computational chemistry: Dirac Notation. Here is how it works.

Everyone has had to deal with some kind of vector math in their lives. On an xy plane, a vector only must have two points, and is shown by [x,y]. As one learns in algebra, to multiply two vectors, take $i=[a,b]$ and $j=[c,d]$ for example, one takes the dot product of the vectors. The answer to this is a number, known as the inner product space, is simply,

$$\langle i, j \rangle = i \cdot j = a * c + b * d \quad (10)$$

This number represents the total length of the two vectors combined. The notation $\langle i, j \rangle$ stands for the inner product of these two vectors. This inner product space applies to functions as well; for example, take the functions $f(x)$ and $g(x)$, the inner product space is,

$$\langle f(x), g(x) \rangle = \int f(x)g(x)dx \quad (11)$$

Using this idea, Dirac introduced an extension to this where,

$$\langle f(x) | g(x) \rangle = \int f(x)^* g(x)dx \quad (12)$$

The star indicates what is known as the complex conjugate of the equation, which deals with changing the sign on imaginary numbers (yes, imaginary numbers are ubiquitous in computational chemistry, but one must not worry about this to understand the basics.)

What makes this so powerful is that each function can be written as a matrix of many functions, and operators can be used on both matrices in the notation. This combines everything mentioned previously into a simpler notation that its easier to work with and manipulate¹³

Schrödinger Equation:

Armed with the powers of operators, calculus, and matrices, the reader is now prepared to understand one of the most powerful equations invented by mankind: the Schrödinger equation,

$$\hat{H}\Psi = E\Psi \quad (13)$$

where \hat{H} is an operator called a Hamiltonian, Ψ is a specific function called the wavefunction, and E is the energy of the system described by the wavefunction. This is exactly the same type of equation as equation 4, it just has a very specific meaning. Essentially, this equation is a way to determine the energy of atoms and molecules. Energy is everything in chemistry, and this is how theoretical chemists make mathematical predictions about how molecules move.

To understand the significance of this mathematically, one must understand the postulates of quantum mechanics, the set of rules that define how we study quantum systems.^{13,14} Here, three primary postulates are discussed to contextualize the meaning of the Schrödinger equation.

1. The state of a quantum system is described by the wavefunction.

A quantum system is one that can only be explained fully by quantum mechanics like electrons, atoms, and molecules. This means that for an electron, for example, we can define a function Ψ that describes the behavior of the electron. Remember, quantum mechanics is confusing and does not follow classical physics that we see and understand in our daily life. We can never know the both position and momentum of an electron at the same time, so we can only define a probability distribution of where the electron may be. The probability of where to find an electron is defined by the square of the wavefunction,

$$P(x) = |\Psi|^2 \quad (14)$$

This is a very abstract idea, but the important point is that Ψ is the mathematical representation of a quantum system. Just as $x(t)$ is a function described the position of the marble in example 4, $\Psi(x,y,z,t)$ is just a function that describes the distribution of a quantum mechanical system.

2. *For every observable, there is a quantum operator.*

Quantum mechanics is based on operators. The important thing about these operators is that the numerical answers -- the numbers that result from operating on a function (think a in equation 4) -- are real, measurable numbers. There would be no point in having an operator give an answer that is not testable; the whole point of quantum mechanics and physics in general is to mathematically describe what is actually happening in the world. So, what are these measurable numbers, and what are their respective operators? Anything that one can measure physically, such as position, momentum, charge, and energy, have associated operators. In the Schrödinger equation mentioned above, the operator is called a Hamiltonian (\hat{H}), and the observable is the energy of the system. Thus, this postulate simply means that there must be an observable quantity that is explained by the operator equation; otherwise, the answer is nonsense. Using Dirac notation and this postulate, the Schrodinger Equation can be written as,

$$\langle \Psi | \hat{H} | \Psi \rangle = \int \Psi^* \hat{H} \Psi d\tau = E \quad (15)$$

Any observable (like energy) has a quantum operator associated with it to where this equation gives the possible answers to the observable for the quantum system of interest.

3. *The only values ever observed for a quantum operator are the eigenvalues.*

The previous postulate explains that the observable values are the answers to the operator equation; however, it is important to note that for a given eigenfunction, there can be multiple answers. This is why the quantum stair example exists; for a quantum system, the answers to the Schrodinger equation are the allowed energies the system can have. Say, for example, that for a

given quantum system that the allowed values of E are 1, 2, 3, 4, 5... etc. That means that the energy can be exactly one of those numbers, but a fraction -- 1.1 for example -- is not allowed. Thus, this is why the quantum staircase example states that the energy can only exist at specific points, because each of these points are an answer to the Schrödinger Equation.

Solving the Schrodinger Equation:

A primary reason the Schrodinger equation is difficult to solve is that the wavefunctions that describe the system of interest can only be found by solving the equation. Yes, the equation itself can only be found by solving the equation. How in the world would one approach such a problem? Remember, an operator as described can only work if it does not change the function it is operating on; therefore, if we know the operator, there are only certain functions that will work, and the equation can be solved for them. There is a lot more complex mathematics that goes into this process, but the bottom line is that it is possible to solve this equation for simple operators. However, when the operators and the systems of interest become more complex, the equation becomes increasingly difficult to solve until it is no longer possible. A factor that makes the Schrodinger Equation so complex is its operator. Like classical energy, the Hamiltonian operator is a sum of kinetic and potential energy terms. For any given quantum system, the kinetic part is relatively easy to solve. However, the potential energy term can be unsolvable.

The most complex quantum system that can be solved analytically is the hydrogen atom. There are two pieces to a hydrogen atom: a single-proton nucleus with a charge of positive one and one electron with a charge of negative one. The energy of the system depends on the position of the electron as it moves around the nucleus; therefore, the Schrodinger equation must only be solved for the single electron. The kinetic energy of the electron is readily solvable, and the potential energy is too, because it is only a function of the attraction between the proton and

electron. However, as soon as another electron is added to the system (like a helium atom), the potential energy of each electron depends on the position of the other electron. Remember that an electron doesn't have an exact position; it has a wavefunction that defines a probability of where that electron can be. So, in a two-electron system, the operator for each electron depends on the wavefunction for the other, but the wavefunction itself is defined by the operator. Yes, this is incredibly confusing circular logic, which is why quantum mechanics is so difficult to study.

For any system larger than a single hydrogen atom, the Schrodinger equation is impossible to solve. That means that all of the large molecules of interest are impossible to calculate exactly by themselves, and to simulate a molecule in the actual world surrounded by millions of other atoms quantum-mechanically is almost unfathomable. Thus, to use these mathematical concepts in any meaningful way we must use approximations and other methods to generate scientific estimates of what the answer likely is.

It is important to remember here that theory is not all doom and gloom because humans have gotten very good at approximating these systems. Think back to Example 3 with the simulation of the pyrazine spectrum; Pyrazine is a relatively large molecule, and the calculated spectrum clearly reflects the actual spectrum of the molecule. This is why it is ever important to compare theory with experiment, because it is just as important to know what theory cannot do, as what it can.

Variational Principle:

One of the cooler mathematical tools computational chemists use to solve the Schrodinger equation for larger systems is the variational principle. Essentially, what this principle states is that when using non-exact mathematics in the manner presented below, the answer we get is always an over-estimation.^{13,14} Therefore, the lowest possible answer is the best guess. Mathematically, this is represented by,

$$W = \frac{\langle \phi | \hat{H} | \phi \rangle}{\langle \phi | \phi \rangle} \geq E \quad (16)$$

where W is the best guess for the energy, and ϕ is an approximate wavefunction.

We cannot know the exact wavefunction for any given system larger than the hydrogen atom because the equation is simply not solvable (remember the answer to the equation is a necessary part of the equation itself). Therefore, we repeatedly ‘guess’ what the wavefunction could be and keep optimizing it until it converges on the smallest number possible. The most basic computational method to find energies of molecules, known as Hartree-Fock, is based on this principle. Hartree-Fock is the method to which all others are compared, as will be discussed in detail in the computational methods section.

In theory, if we used an infinite number of guesses and added them together, we would get the actual wavefunction. Unfortunately, we cannot actually do infinite calculations, so we must ‘cut off’ the approximation at some point. The point at which we cut off the calculation is one way in which computational methods can vary, as will be discussed later.

Born-Oppenheimer Approximation:

The Born-Oppenheimer Approximation is the foundation on which computational chemistry is built, as it enables calculations to be performed for molecules. As described above, the Schrodinger equation cannot even be solved for a helium atom, and the molecules that are of interest to study for practical purposes consist of up to hundreds of atoms! Uracil, for example, consists of 12 atoms, over half of which are more complex than helium. To calculate the energy of a singular nucleus, one must only calculate the energy of the electrons as they move around a relatively stationary nucleus. For a molecule, the nuclei move relative to each other and the electrons move around them. What a mathematical mess!

To find the energy of a whole molecule, the molecule itself is the quantum system that must be studied through the Schrodinger equation. That means the wavefunction must contain all of the electrons and all of the nuclei in the system. This is represented mathematically by $\Psi(\mathbf{r}, \mathbf{R})$, where \mathbf{r} represents the coordinates of the electrons and \mathbf{R} represents the nuclear coordinates. It is simply out of the question to solve these as one big system; this is why we need the Born-Oppenheimer Approximation.¹⁵ The approximation is that the electrons move so fast relative to the nuclei that we can separate their wavefunctions,

$$\Psi(\mathbf{r}, \mathbf{R}) = \chi(\mathbf{R})\psi(\mathbf{r}; \mathbf{R}) \quad (17)$$

where $\chi(\mathbf{R})$ is the nuclear wavefunction that does not depend on the electrons at all and $\psi(\mathbf{r}; \mathbf{R})$ is the electronic wavefunction that only depends parametrically on the nuclei. In simpler terms, the electron energy can be calculated independently at different nuclear positions. Then we solve the electronic-only Schrödinger equation,

$$\widehat{\mathcal{H}}^e \psi(\mathbf{r}; \mathbf{R}) = E^e \psi(\mathbf{r}; \mathbf{R}) \quad (18)$$

where E_e is the electronic energy in the system, and the Hamiltonian term only accounts for electronic contributions. Substituting this back into the equation for the whole system gives,

$$(\hat{\mathcal{H}}^{nuc} + E^e)\chi(R) = E\chi(R) \quad (19)$$

where the remaining Hamiltonian term accounts for the nuclear energies, and the answer is the total energy of the system.

So, for a simple molecule like dihydrogen, which is just two bonded hydrogen atoms, one can calculate the energy as the bond length changes and the nuclei move. This creates what is called a potential energy surface (PES), shown in Figure 7.

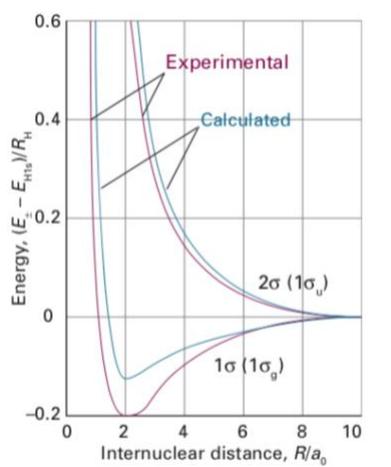


Figure 7. Potential energy surface for the dihydrogen molecule.¹

Remember, again, that there are multiple solutions to the Schrodinger equation for each system. For a molecule, this means that there are multiple potential energy surfaces that are different in energy. The different potential energy surfaces are called electronic states. These are the states that form the steps on the quantum staircase for molecular energy. When a photon of light hits a molecule that matches the energy difference between these potential energy surfaces, the molecule can absorb it and excite as explained previously. Understanding the depth and complexity of the mathematics to explain this phenomenon only adds proof to the awesome beauty and power behind the predictive abilities of computational chemistry.

Breakdown of Born-Oppenheimer Approximation:

The Born-Oppenheimer approximation is not perfect, and there are mathematical situations in which it breaks down. To be specific, the approximation breaks down when two potential energy surfaces have the exact same energy for a given nuclear configuration. This is because with a small change in where the nuclei are, there can be a large change in where the electrons are; therefore, we cannot decouple the electronic and nuclear wavefunctions.

Mathematically, this complicates everything mentioned in the previous section, and the once relatively simple equation 19 becomes,³

$$\left(\widehat{\mathcal{H}}^n + \frac{1}{2\mu} K_{II}(R) + E_I^e(R) \right) \chi_I(R) - \sum_{J \neq I} \frac{1}{2\mu} (K_{IJ}(R) + 2\mathbf{f}_{IJ}(R) \cdot \nabla) \chi_J(R) = E \chi_I(R) \quad (20)$$

where I and J denote the two states of interest, and the K and f terms represent the derivative coupling between the states. What is important here is not that the reader understands this equation but understands that this equation is really intense and depends on the derivative coupling term, f_{IJ} .

$$\mathbf{f}_{IJ} = \langle \Psi_I | \nabla \Psi_J \rangle \quad (21)$$

As mentioned previously, imagine a quantum staircase with really wide steps that change height from left to right where two steps actually become the same height at a specific point. This is not actually possible for a two-atom system like HCl, only for molecules with at least three atoms. In larger atoms, the potential energy surface is a function of every bond and angle between the atoms. For a (non-linear) molecule with N atoms, there will be 3N-6 dimensions of the potential energy surface. This is very confusing and impossible to imagine; luckily, mathematics is not limited to what we can picture in our heads. When these 3N-6 dimensional

spaces have the same energy, the intersection between the PES also has this large dimension. An intersection of many dimensions is called a seam in mathematics.

Conical Intersections:

A conical intersection is a way to physically picture the seam space between electronic states. Again, there is no way to picture a seam space of more than three dimensions, so defining a conical intersection is how computational chemists convey characteristics of this seam. To do this, first a point on the seam must be chosen. Because the seam covers a range of positions for the nuclei the point on the seam where the energy is minimal, (minimum energy crossing point/MECP), is chosen. Note, a ‘point’ in this context actually refers to a specific molecular geometry. In one dimension, this seam space is a line as shown in Figure 8. for a generic system.

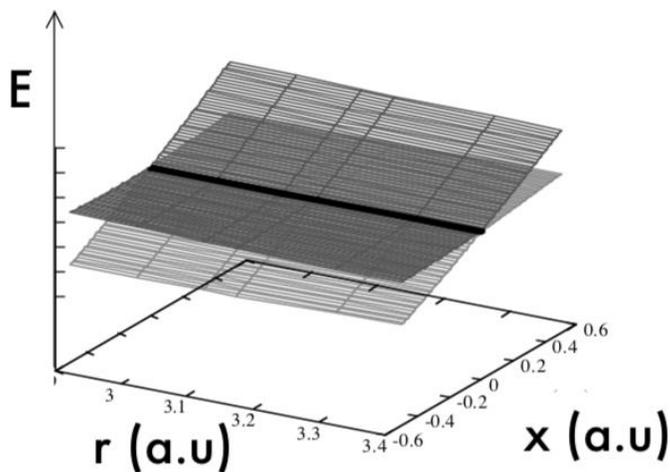


Figure 8. Image of a hypothetical one-dimensional seam space formed by intersecting PES.³

The characteristics of the conical intersection are defined by the area around the seam as much as the seam itself. Thus, the molecular motions that make the energies of the states different is how one can characterize the intersection. There are two important vectors representing molecular motion that can be used to define a plane to visualize many dimensional conical intersections. The magnitudes of these vectors characterizes the region around the seam.

The first vector, \mathbf{g}_{IJ} , represents the difference between the gradients of the two states,³

$$x = \mathbf{g}_{IJ} = \mathbf{g}_I - \mathbf{g}_J \quad (22)$$

where each individual \mathbf{g} vector is the gradient (derivative of energy with respect to every dimension) of the potential energy surface. This value will result in a three-dimensional vector for every atom, but this is multiplied using inner product spaces to find a single value for the magnitude of the whole vector.

The second vector, \mathbf{h}_{IJ} , is a little harder to understand physically. What is important about this vector contextually is that it is related to the derivative coupling vector mentioned previously, \mathbf{f}_{IJ} . The vector is defined in the y direction as,³

$$y = \mathbf{h}_{IJ} = \langle \Psi_I | \nabla \hat{\mathcal{H}} | \Psi_J \rangle = \mathbf{f}_{IJ} * \Delta E_{IJ} \quad (23)$$

In equation 23, the energy difference (ΔE_{IJ}) between the states is zero at the conical intersection.

Using the two vectors in the x-y plane allows one to plot the conical intersection visually. This is done using equation 24 and results in an image like Figure 9,

$$E_{IJ} = \sqrt{(\mathbf{g}x)^2 + (\mathbf{h}y)^2} + \text{"tilt"} \quad (24)$$

where the tilt is an additional parameter related to the slope of the intersection that can be reviewed for those looking to characterize conical intersections in real systems.³

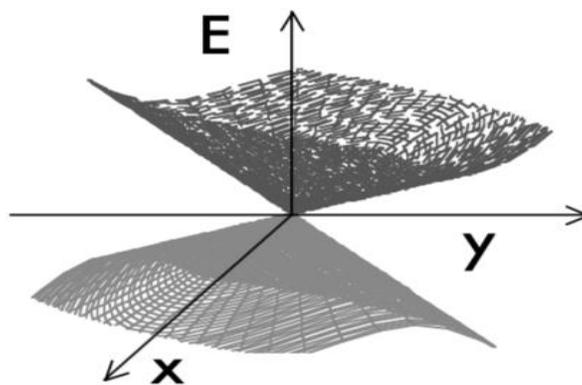


Figure 9. A representative image of a conical intersection.³

These conical intersections arise from the mathematics in calculating the potential energy surfaces of polyatomic molecules. In theory, there can be an infinite amount of these in any given molecule forming subspaces (see Figure 8), or none at all. What is of interest to the practical theorist are conical intersections that lie on important parts of potential energy surfaces for important molecules. In Uracil, for example, there are conical intersections that lie near minima of the excited electronic states. What that means analogously is that if uracil is excited by light to a higher quantum step in Example 1, then it will spontaneously fall to the lower part of the step that touches the next one down and “fall” to the lower state without emitting light. What makes conical intersections interesting is that this mathematical phenomenon actually governs life processes.

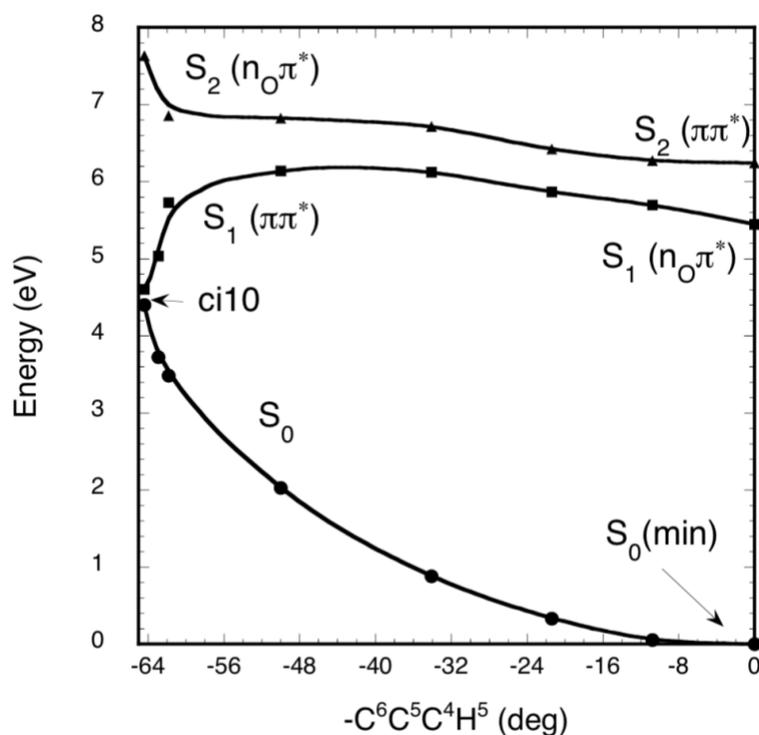


Figure 10. Image of the PES that create a conical intersection in Uracil.³ The S_n denotes the level of each singlet state. With analogy to the quantum staircase, the S levels are the steps and “ci10” is the seam where the steps meet.

How Do We Study Conical Intersections? Computational Chemistry:

As shown thus far, computational chemistry relies on foundational physics and mathematics to study the molecular processes that are relevant to explaining material science and molecular biology. The math becomes so difficult when dealing with these systems that all of the calculations must be done using computers. Herein lies the challenge that many active researchers in computer science and chemistry take on: how to get the most accurate answer with the least amount of computational cost. For conical intersections specifically, there are two factors to whether a method can characterize a conical intersection: 1) Analytical gradients and derivative coupling must be available for each state (\mathbf{g} and \mathbf{h} vectors), and 2) The states of interest must be treated equivalently.

Basis Sets:

The Schrödinger equation (equation 13) looks relatively simply when written in the most basic form; there are only four terms; two are the same and one is the answer. The complexity comes through the application and utilization of it in calculations. One primary consideration when performing a calculation is called the basis set. Remember that we cannot ever know the exact wavefunction, so we optimize a combination of guesses until we get the lowest answer possible. The basis set is essentially the combination of ‘guesses’ given to the computer to optimize. These ‘guesses’ are not random; we can solve the equation for the hydrogen atom exactly, so the ‘guesses’ are actually the answers to the equation for hydrogen. Answers to the one electron wavefunction are called orbitals. (The average chemist is very familiar with the practical application of orbitals). By using linear combinations (summations) of these atomic orbitals, we can generate optimized molecular orbitals.¹⁵

A basis set is the predetermined set of orbitals to be used in a calculation. The larger the basis set, the larger the set of orbitals that can be used to optimize, and the more accurate the answer will be. However, more orbitals to optimize means a bigger matrix for the computer to solve, thus creating a higher computational cost. What is often functionally done with basis sets is to use smaller ones to ensure everything else about a calculation works, then ultimately use and report answers using the largest basis set possible to get the most accurate possible answer.

Single-Reference Methods:

Chemically, the different electronic states that have potential energy surfaces correspond to the actual physical location of the electrons around the nuclei in a molecule. The lowest electronic state is called the ground state, and this is the state that molecules exist in under normal conditions. The higher energy excited states are not indefinitely stable, and a molecule will always eventually decay from them back to the ground state. Chemically, in an excited state the electron is physically moved to a different orbital and can have different characteristics (here the reader is spared the gory detail of what transitions are allowed and the character of different states). In the ground state for most stable molecules, all the electrons are paired up with one of opposing spin (electrons can have either 'up' or 'down' spin). This is known as a "singlet" state.

The single reference methods considered in the present work along with a brief summary of the distinguishing features is listed below:

1. CIS – (Configuration Interaction Singles) The most basic excited state method possible.

The ground state is calculated using the Hartree-Fock method (based on the variational principle using a specific basis set). CIS then uses the optimized orbitals to allow electron excitations to generate the excited states. This is the least accurate yet cheapest method.

2. TDDFT – (Time Dependent Density Functional Theory) This method has gained much popularity in recent years due to its similarity to CIS in computational cost, yet distinct superiority to CIS in accuracy for excited state energies.¹⁶ This method is similar in principles to CIS, but includes “electron correlation effects” which encapsulates much of the weird quantum nature of many electron systems. This is done by approximating the potential term of the Hamiltonian from electrons to be the density distribution of where they could be instead of the tiny moving point charges they are.
3. EOM-CCSD – (Equation of Motion- Coupled Cluster Singles and Doubles) This theory is known as the “gold standard” for single method computational methods. This method is defined by the treatment of the wavefunction; this method can characterize multi-configurational wavefunctions even though it is single reference. Because of this, EOM-CC is an attractive method to study multi-state systems.^{17,18}

Single reference methods use the ground state as the basis for all calculations; the ground state is optimized first, then the excited states are calculated using the same molecular orbitals. Fundamentally, this means that the ground state is treated differently than the excited states. Therefore, using traditional single reference methods, one cannot accurately calculate conical intersections that involve the ground state.

Fortunately, much progress has been made in recent years to implement analytical gradients and derivatives for commonly used single reference methods in recent years. These methods are able to calculate \mathbf{g} and \mathbf{h} vectors between excited states relatively reliably.

In order to use these methods to characterize the ground state, a different reference must be used. Chemically, if an electron is excited and its spin is flipped, it is in what is called a “triplet” state. The triplet state is higher in energy than the ground state, and it is different

enough that it can be used as the reference instead of the ground state. Thus, using a “spin-flip” single reference method can theoretically be used to characterize conical intersections with the ground state.²⁰ There are mathematical issues when using this method, thus it is of interest to investigate how well these methods compare to the more accurate multi-reference methods. These methods are discussed and compared in the following sections.

Multi-Reference Methods:

Multi-reference methods are the most accurate computational methods; however, they are by far the most difficult for both the computer and the computational chemist. For present purposes, two multi-reference methods are investigated, and one builds on top of the other to get the most accurate answer possible. The first method is called CASSCF (Complete Active Space Self-Consistent Field). This method simultaneously optimizes the molecular orbitals and the relative combinations of them for each state of interest. A crucial factor of this method is that the ground and excited states are treated equivalently. However, the user must define which orbitals to include in the “active space” of the calculation. If all possible orbitals were used, the answer would be perfectly accurate but would take infinite time. The user-dependence of the orbitals used is the reason this method is more difficult on the user; if the wrong active space is chosen, the answer will be wrong.³

To go beyond the CASSCF method, one must use what is known as the MRCI method. This method takes something called “dynamical correlation” into account, which makes the answers ultimately more accurate. In practice, this method uses the orbitals generated in a CASSCF calculation to factor in excitations of electrons to all the possible orbitals as defined in the active space.³ This is the most accurate method to calculate conical intersections and is used as the reference to which all the other methods are compared.

Benchmarking Study -- Methods of Calculating Conical Intersections:

To effectively compare the different methods of calculating conical intersections, one must study both a conical intersection between excited states and involving the ground state. As mentioned previously, traditional single reference methods cannot characterize ground state conical intersections, so the spin flip variety of each method must be tested independently. The nucleobase uracil has been proven to have important intersections between the first and second excited singlet states (S_1 - S_2), and between the ground and first excited singlet states (S_0 - S_1).⁴ Therefore, uracil is a quality representative larger molecule on which to compare the different computational methods. The present study is a benchmarking of the five discussed computational methods on the two representative conical intersections in uracil; for the ground state calculations, the spin flip variety of the single reference methods was used. The intersections were optimized to the MECP using MRCI,⁴ then this geometry was used for the other methods.

MRCI is the reference calculation method, as it is the most accurate way to calculate conical intersections. The ΔE_{IJ} measurement shows how close to degeneracy the two states of interest are at the optimized geometry. Because the methods are fundamentally different, the geometry will not be an exact conical intersection for each method; a lower ΔE_{IJ} indicates that the point is closer to the actual conical intersection. Fortunately, both the \mathbf{g} and \mathbf{h} vectors are not explicitly dependent on energy, though the values for these can become distorted if the point of interest is substantially separated from the seam. For these vectors, an accurate result is one that is similar to that of MRCI, keeping in mind the measured geometry does not perfectly lie on the seam for the methods other than MRCI. The two representative conical intersections in uracil were measured using all five methods and are tabulated in Tables 1 and 2.

Table 1. Comparison of theoretical methods, in order from least to most accurate, on the representative S₁-S₂ conical intersection in uracil.⁴

Method	Basis Set	ΔE_{IJ} (eV)	$\ f_{IJ}\ $ (au)	$\ h_{IJ}\ $ (au)	$\ g_{IJ}\ $ (au)
CIS	6-31G*	1.838	0.0515	0.035	0.26
TDDFT	6-31G*	0.1355	10.4	0.052	0.19
EOM-EE-CCSD	6-31G*	0.6797	4.69	0.117	0.35
CASSCF	6-31G*	0.3307	2.01	0.024	0.34
MRCI	cc-pVDZ	0	x	0.035	0.15

Table 2. Comparison of theoretical methods, in order from least to most accurate, on the representative S₀-S₁ conical intersection in uracil.⁴

Method	Basis Set	ΔE_{IJ} (eV)	$\ f_{IJ}\ $ (au)	$\ h_{IJ}\ $ (au)	$\ g_{IJ}\ $ (au)
SF-CIS	6-31G*	0.1065	0.425	0.002	0.010
SF-TDDFT	6-31G*	0.3710	2.90	0.039	0.157
EOM-SF-CCSD	6-31G*	0.02188	20.3	0.016	0.034
CASSCF	6-31G*	0.3376	5.83	0.070	0.202
MRCI	cc-pVDZ	0	x	0.064	0.097

The first consideration impacting the efficacy of different methods of calculating conical intersection is whether or not the g and h vectors can be computed at all; this was true for all methods studied. However, for the CIS and TDDFT, the derivative coupling vector f was calculated, and h was extrapolated from it. For CCSD, CASSCF, and MRCI, the h vector was calculated by the program and the f vector was extrapolated from it. Because of the 0-energy

difference for MRCI (this is perfectly on the seam), the derivative coupling extrapolating would involve dividing by zero and is therefore not available.

Due to the fact that the geometry for the intersection was not optimized for each individual method, there are limitations to the extrapolations that can be made from the data. For example, in Table 1, EOM-EE-CCSD gives a larger energy gap between the states, meaning this point is further from the seam, and this results in a higher g for this geometry. While this may be an implicit overestimation by method, one cannot rule out the possibility that the gradient difference just increases rapidly away from the seam.

Of particular note are the trends that are contrary to expectations based on energy difference, as it is these trends that indicate the true shortcomings/ advantages of a given method.

For coupled cluster, EOM-SF-CCSD gives a very low energy differences yet underestimates both g and h vectors substantially; this indicates that there are some complications that arise when using the spin flip variety of CCSD. Conversely, the EOM-EE-CCSD overestimates both the g and h vectors substantially, indicating that more work must be done to ensure coupled cluster is accurately representing the derivative coupling between states.

For TDDFT, the results are the most accurate of the single reference methods, and comparable to CASSCF, even for the conical intersection with the ground state. This is an impressive feat and potentially very important because TDDFT is so computationally cheap that it could be the perfect balance between cost and accuracy for many in the electronic structure field. From the limited results presented here, TDDFT is more accurate in some respects than CASSCF; this speaks to the importance of including dynamical correlation in these calculations and to the immense progress that has been made in developing this method in recent years.

The Seam Space Between the Sciences:

The final section of this paper is a purely subjective analysis of the current state of the field of computational chemistry and why it is so important.

Thus far, this piece has involved quantum chemistry of excited states: an area of research fittingly known as electronic structure theory. However, not even all computational chemists explicitly use quantum mechanics. For large molecular systems like proteins, DNA as a whole, and other biologically relevant processes, purely quantum calculations are impossible due to time constraints. A technique used to calculate properties of interest for these systems is known as QM/MM (quantum mechanics/molecular mechanics).²¹ At the most basic level, this method uses quantum mechanics to simulate a specific region of interest yet uses classical mechanics to simulate the environment it takes place in. The possible applications of this is endless, as it is a way to actually simulate molecules moving in solution; literally everything is molecules moving in solution.

In the capitalist world we live in, the objective way to measure whether a science is valuable is whether or not there are companies making money doing it. An exemplar of this is a company fittingly called Schrödinger. The business model of this company is to provide a computational chemistry platform that lets drug development companies find properties of molecules without having to actually synthesize them first. Obviously, this is not perfect, and simulation is not a replacement for drug trials and testing. However, if as a drug company you have 20,000 molecules that might work as a drug and you can rule out 18,000 of them by doing calculations, then the calculations essentially save the money that would have gone into testing those 18,000 drugs. That is an immense value when extrapolated to the expansive field that is pharmaceutical research.²²

What makes computational chemistry such an interesting field for the future is that the predictive power of the calculations grows with computing power and advancements. A long-known qualitative phenomenon known as Moore's law states that computational power (the number of transistors that can be fit on a computer chip) doubles every two years. While this will not hold forever and there are already signs that this relationship is slowing, if raw computational power is increasing exponentially, so too will the size of systems that can be calculated. This means that molecular simulations and computational chemistry will get more precise and useful in the future just by the work that is being done by computer scientists.

In addition to the growing raw computational power of traditional computers, there are other promising fields that will promote the efficiency and efficacy of computational chemistry: namely machine learning and quantum computing.

Machine learning is here; we are using it now and getting results that we never could have otherwise. While this is an entire field itself, the applications to computational chemistry increases with its progress. Theory, at heart, is about explaining observed data, and machine learning is an amazing tool to help with mass amounts of data, calculated or experimental. In a paper published March 6, 2020, a team of researchers proved the ability to use machine learning and computational NMR spectra to identify molecules. NMR is a technique used to identify and characterize organic molecules; it is a spectra-based process that relies on the unique energy levels of molecules to gain insight about their character. For simple molecules, one can observe the spectrum and reliably know what the molecule is. For larger molecules, subtle differences in spectra can indicate major structural differences. What used to be a purely human logic game of identifying molecules based on spectra can now be done by an advanced machine learning program in conjunction with computational chemistry. That is an advancement that was literally

made a month prior to the writing of this piece and may have been thought to be impossible by many just a few years ago.²³

Quantum computing is something that is a little further off from being useful in practice yet holds immense promise for the field of computational chemistry as it advances. A quantum computer is fundamentally different from a regular computer. A regular computer functions on bits, which are either 0 or 1; this generates the binary language that computers perform logic with. Quantum computers use what is called a quantum bit (qubit), which is a superposition of 0 and 1. Superposition is a fundamental element of quantum mechanics and is the reason why there is uncertainty in quantum measurements. Computationally, taking advantage of qubits and superposition allows quantum computers to solve certain problems in minutes that would take traditional computers thousands of years to solve. Application-wise, there exists a possibility of being able to map a quantum system like those discussed onto these qubits to be able to run calculations that were once thought to be impossible. This is purely speculation and the limits of this lie as much in the foundations of quantum mechanics and philosophy as it does in computer science; however, the thought of a world where we can map simple quantum systems onto a quantum computer and run simulations coordinated with a traditional computer (essentially ridiculously powerful QM/MM simulations) is truly awesome.

In conclusion, computational chemistry relies on decades of progress in mathematics, new and budding physics, and state of the art computer science to perform predictive chemistry tasks to explain important biological processes to ultimately make the world better through knowledge. That is why I have chosen to dedicate my career to it, and why I believe it is so important that everyone be able to understand the value in pursuing it going forward.

References:

- [1] Atkins, P. W., et al. *Atkins' Physical Chemistry*. Oxford University Press, 2019.
- [2] Wilhelm, Michael J., et al. "UV Photolysis of Pyrazine and the Production of Hydrogen Isocyanide." *The Journal of Physical Chemistry A*, vol. 122, no. 46, 2018, pp. 9001–9013., doi:10.1021/acs.jpca.8b09179.
- [3] Matsika, Spiridoula. "Conical Intersections in Molecular Systems." *Reviews in Computational Chemistry*, 2007, pp. 83–124., doi:10.1002/9780470116449.ch2.
- [4] Matsika, Spiridoula. "Radiationless Decay of Excited States of Uracil through Conical Intersections." *The Journal of Physical Chemistry A*, vol. 108, no. 37, 2004, pp. 7584–7590., doi:10.1021/jp048284n.
- [5] Matsika, Spiridoula, and Pascal Krause. "Nonadiabatic Events and Conical Intersections." *Annual Review of Physical Chemistry*, vol. 62, no. 1, 2011, pp. 621–643., doi:10.1146/annurev-physchem-032210-103450.
- [6] Levine, Benjamin G., and Todd J. Martínez. "Isomerization Through Conical Intersections." *Annual Review of Physical Chemistry*, vol. 58, no. 1, 2007, pp. 613–634., doi:10.1146/annurev.physchem.57.032905.104612.
- [7] Wald, G. "Molecular Basis of Visual Excitation." *Science*, vol. 162, no. 3850, 1968, pp. 230–239., doi:10.1126/science.162.3850.230.
- [8] Kolb, Helga. "How the Retina Works." *American Scientist*, vol. 91, no. 1, 2003, p. 28., doi:10.1511/2003.1.28.
- [9] Mathies, R., et al. "Direct Observation of the Femtosecond Excited-State Cis-Trans Isomerization in Bacteriorhodopsin." *Science*, vol. 240, no. 4853, 1988, pp. 777–779., doi:10.1126/science.3363359.
- [10] Matsika, Spiridoula. "Three-State Conical Intersections in Nucleic Acid Bases." *The Journal of Physical Chemistry A*, vol. 109, no. 33, 2005, pp. 7538–7545., doi:10.1021/jp0513622.
- [11] Roca-Sanjuán, Daniel, et al. "Ab Initio Determination of the Ionization Potentials of DNA and RNA Nucleobases." *The Journal of Chemical Physics*, vol. 125, no. 8, 2006, p. 084302., doi:10.1063/1.2336217.
- [12] Goode, Stephen W., and Scott Annin. *Differential Equations and Linear Algebra*. Pearson, 2017.
- [13] Sakurai, Jun John, and Jim Napolitano. *Modern Quantum Mechanics*. Cambridge University Press, 2019.

- [14] Levine, Ira N. *Quantum Chemistry*. Pearson, 2014.
- [15] Szabo, Attila, and Neil S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, 2006.
- [16] Levine, Benjamin G., et al. "Conical Intersections and Double Excitations in Time-Dependent Density Functional Theory." *Molecular Physics*, vol. 104, no. 5-7, 2006, pp. 1039–1051., doi:10.1080/00268970500417762.
- [17] Krylov, Anna I. "Equation-of-Motion Coupled-Cluster Methods for Open-Shell and Electronically Excited Species: The Hitchhiker's Guide to Fock Space." *Annual Review of Physical Chemistry*, vol. 59, no. 1, 2008, pp. 433–462., doi:10.1146/annurev.physchem.59.032607.093602.
- [18] Krylov, Anna I. "Equation-of-Motion Coupled-Cluster Methods for Open-Shell and Electronically Excited Species: The Hitchhiker's Guide to Fock Space." *Annual Review of Physical Chemistry*, vol. 59, no. 1, 2008, pp. 433–462., doi:10.1146/annurev.physchem.59.032607.093602.
- [19] Faraji, Shirin, et al. "Calculations of Non-Adiabatic Couplings within Equation-of-Motion Coupled-Cluster Framework: Theory, Implementation, and Validation against Multi-Reference Methods." *The Journal of Chemical Physics*, vol. 148, no. 4, 2018, p. 044103., doi:10.1063/1.5009433.
- [20] Zhang, Xing, and John M. Herbert. "Spin-Flip, Tensor Equation-of-Motion Configuration Interaction with a Density-Functional Correction: A Spin-Complete Method for Exploring Excited-State Potential Energy Surfaces." *The Journal of Chemical Physics*, vol. 143, no. 23, 2015, p. 234107., doi:10.1063/1.4937571.
- [21] Brunk, Elizabeth, and Ursula Rothlisberger. "Mixed Quantum Mechanical/Molecular Mechanical Molecular Dynamics Simulations of Biological Systems in Ground and Electronically Excited States." *Chemical Reviews*, vol. 115, no. 12, 2015, pp. 6217–6263., doi:10.1021/cr500628b.
- [22] Abel, R.; Mondal, S.; Masse, C.; Greenwood, J.; Harriman, G.; Ashwell, M.A.; Bhat, S.; Wester, S.; Frye, L.; Kapeller, R.; Friesner, R.A. Accelerating drug discovery through tight integration of expert molecular design and predictive scoring, *Current Opinion In Structural Biology* (2017) 43, 38-44 DOI: 10.1016/j.sbi.2016.10.007
- [23] Howarth, Alexander, et al. "DP4-AI Automated NMR Data Analysis: Straight from Spectrometer to Structure." *Chemical Science*, 6 Mar. 2020, doi:10.3897/bdj.4.e7720.figure2f.