

# A Crash Course in Density Functional Theory

Christoph Kowitz



Technische Universität München

*The survey of chemical compounds and their reaction is an important part of chemistry. But modeling these processes is quite cumbersome and a number of methods can be applied to tackle this problem. Density Functional Theory (DFT) is one possibility for that. It provides a much higher accuracy than molecular dynamics with empirical potentials, but it is computationally less expensive than the typical Hartree-Fock method. The coarse outline of a computational algorithm and some post-processing options will be given here. This report and the talk held at the Joint Advanced Spring School in St. Petersburg are strongly based on the PhD thesis of Thomas Grauschopf [1] and the nice book of Koch and Holthausen [2].*

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Basics In Quantum Mechanics</b>	<b>2</b>
2.1	Wave Function . . . . .	2
2.2	Schrödinger Equation . . . . .	2
2.3	Born-Oppenheimer-Approximation . . . . .	3
<b>3</b>	<b>DFT</b>	<b>3</b>
3.1	Basic Principle . . . . .	3
3.2	Kohn-Sham Functions . . . . .	3
3.3	Exchange-Correlation Potential . . . . .	3
<b>4</b>	<b>SCF Cycle</b>	<b>4</b>
4.1	Set of Basis Functions . . . . .	4
4.2	Exchange Correlation Potential . . . . .	5
4.3	Solving the Kohn-Sham Equation . . . . .	5
<b>5</b>	<b>Numerical Effort</b>	<b>5</b>
<b>6</b>	<b>Postprocessing</b>	<b>6</b>

## 1 Introduction

Density Functional Theory (DFT) is a method to successfully describe the behavior of atomic and molecular systems, taking in account their full quantum mechanical interaction. It is used for instance in:

- structural prediction of chemical compounds
- simulation of chemical reactions
- high precision molecular dynamics
- predicting/validating spectroscopic data
- and many more

But still this method is computational rather expensive and for some given problems the computations can take timescales of months. So parallelization is also a big issue in this branch of computational science.

**1** DFT is using taking in account quantum mechanics. But on can also describe large molecular systems by molecular dynamics with empiric potentials. Empiric potentials are represented by analytic functions, which are fitted to experimental data or verified through statistical physics. The probably best known empiric potential is the Lennard-Jones potential

$$L_{ij} = \epsilon \left[ \left( \frac{r_0}{r_{ij}} \right)^{12} - 2 \left( \frac{r_0}{r_{ij}} \right)^6 \right], \quad (1)$$

which is used to model inert gases. As you can see in (1) it is defined as a pairwise potential. That means the total energy of the system will be calculated by

$$L = \sum_{i>j} L_{ij}. \quad (2)$$

**5** But by doing that we actually make an error, because the real physical potential between more than two particles cannot be described by a sum of the pairwise interactions. This is a systematic error coming with the nature

of empirical potentials. Another disadvantage is, that the nature of bonding is not simulated itself, but packed into the potentials. So they are not useful to simulate the chemical bonding of atoms and molecules in detail. But the huge advantage of using this potentials is the low computational effort to compute them. Their derivatives are also relatively easy to derive, so that the interaction forces for molecular dynamics simulations can be implemented and executed quite fast, even for a really large number of objects. But we see, that we need to take in account quantum mechanics to simulate molecular systems properly.

## 2 Basics In Quantum Mechanics

Quantum mechanics bears some fundamental differences to classical mechanics. In classical mechanics a system is determined by the position  $\vec{x}$  and linear momentum  $\vec{p}$  of every particle in the system. If  $\vec{x}$  and  $\vec{p}$  are exactly known at one point in time, all points in the phase space can be calculated from this state (in the past as well as in the future). But unfortunately it is not possible to describe every branch of physics in this matter. For example the strange behavior of electron being diffracted at a small gap: even single electrons seem to be waves and particles at the same time, because even a single electron gets diffracted in that gap, eventhough there is no other electron to interact with. But there are also some other physical effects, which cannot be explained successfully without quantum mechanics. For instance:

- Black Body Radiation
- The Photoelectric Effect
- The detailed structure of atoms

The basic idea of Quantum mechanics is, that it is not possible to know the position  $\vec{x}$  and the linear momentum  $\vec{p}$  with infinite precision at the same time. This fact is known as Heisenberg's Uncertainty Principle:

$$\Delta x \Delta p \geq \frac{\hbar}{2}. \quad (3)$$

In this equation we can see the constant  $\hbar$ , which is defined as  $\hbar = \frac{h}{2\pi}$ .  $h$  is Planck's constant and it represents the smallest possible package of energy. The principle (3) shows us that an exact measurement of  $\vec{x}$  results in a fuzzy measurement in  $\vec{p}$  and vice versa. This is the basic problem of quantum mechanics and leads to a description of physics which is based on probabilities, due to the incomplete knowledge of the state of a system. The state of an object cannot be determined by a measurement anymore, because a measurement is changing the state of a system on these small scales.

### 2.1 Wave Function

In the previous section we saw, that a system in the quantum world cannot be described by  $\vec{x}$  and  $\vec{p}$ . We need to describe an object in a more probabilistic view. That is

realized through the wave function  $\psi$ . It is in general a complex function and fulfils the following property:

$$\int \psi \psi^* d\vec{x} = \int |\psi|^2 d\vec{x} = 1 \quad (4)$$

But  $\psi$  is a rather abstract notation, because it can be finite or infinite vector  $\vec{\psi}$  or a function with a finite or infinite number of variables  $\psi(x_1, \dots, x_N, \dots)$ . But all this representations of the wave functions must be part of the Hilbert space. This is a complete vector space, where a scalar product is defined. If the scalar product of two members of the Hilbert space is zero, these two elements are orthogonal. This fact will later be an important foundation of DFT. The physical interpretation of the wave function can also be seen in (4). The space integral of the square of the absolute value of  $\psi$  has to be normalized to one, so that it can be interpreted as a probability density. The wave function gives us a clue, where to find a particle with high or with low probability.

### 2.2 Schrödinger Equation

But now we also need a new formalism to deal with the wave function, because we cannot apply classical mechanics on it. The basis formulation of the behavior of a quantum system is done by the general Schrödinger equation:

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{x}, t) = \mathbf{H} \psi(\vec{x}, t). \quad (5)$$

This equation has a time dependency on the lefthand side and is so able to describe the spatial and time evolution of a system. But for the problems we are going to approach with DFT, we only need to take in account the stationary solution of the Schrödinger equation:

$$\mathbf{H} \psi_i(\vec{x}) = E_i \psi_i(\vec{x}). \quad (6)$$

This equation is an eigenvalue problem with the eigenvalues  $E_i$  and the eigenvectors  $\psi(\vec{x})$ . But  $E_i$  is not only an eigenvalue, it also represents the energy of the corresponding wave function  $\psi_i$ . So the state  $\psi_0$  with the lowest energy is the ground state of a system. We can also find the principle of the Hilbert space again in this equation, because all the eigenvectors  $\psi_i$  stand perpendicular on each other as it was required in the previous section.

The  $\mathbf{H}$  on the righthand side is the Hamilton operator, which acts on the wave function of the system. It returns its the kinetic and potential energy of the system. Classically we would compute it by summing up all kinetic and potential energies in the system but for a wave function we derive it by

$$\mathbf{H} = -\frac{1}{2} \Delta + V(\vec{x}). \quad (7)$$

We can see that the kinetic energy of the system is derived by the Laplacian at this position. We can now reformulate (6) to

$$-\frac{1}{2} \Delta \psi_i + V \psi_i = E_i \psi_i. \quad (8)$$

This eigenvalue problem can be solved analytically, but unfortunately only for system with one electron like H, He<sup>+</sup>, Li<sup>2+</sup>...

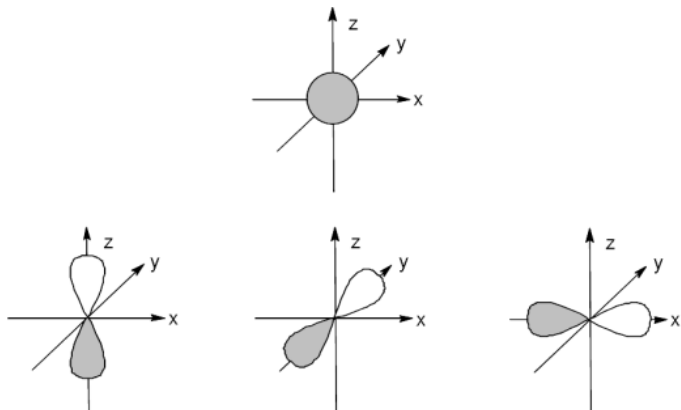


Figure 1: The s and p orbitals of hydrogen. The drawn objects represent the area, in which the probability density to find an electron at this position.

### 2.3 Born-Oppenheimer-Approximation

Now we know the basic rules of quantum mechanics and can now go into further details. The main goal of DFT is to simulate the behavior of molecular systems and in this system we have typically a number of atomic nuclei  $N_n$  and a corresponding and higher number of electrons  $N_e$ . To now be able to solve (6) for this system, the dimensionality increases corresponding to  $3(N_e + N_n)$ , because every particle has one degree of freedom for every direction in space. A really useful way to reduce the dimensionality of the problem is the assumption, that the atomic nuclei are stationary in time. This assumption is based on the different masses of the nuclei and the electrons. Even for a hydrogen atom the nucleus is about 2000 times heavier than the electron bonded to it. And due to that high mass of the nuclei the electrons move on much shorter timescales than the nuclei and one can assume them to be not moving at all. The dimensionality is now only  $3N_e$ .

## 3 DFT

We now proceed with the details of the density functional theory. At first we want to point out some main application areas of DFT. A direct product of DFT is the electronic structure of a molecule. We can instantly find the occupation numbers of the different atomic orbitals and the distribution of electrons in the molecule. One also find the different energy states of a molecule besides the ground state and one can so determine the energies of excited states. Another useful physical value we get out of DFT computations is the energy of a given system. With that one can do geometry optimization of molecules and find reaction pathways. Or it is possible to determine the eigenmodes of a system and with that one has the opportunity to forecast oscillation spectra, which can be directly validated by experiments. And it is also possible to do molecular dynamics, with which one can directly observe the interactions atomic systems. Unfortunately this is only possible on really short timescales. But there is not only DFT to simulate quantum systems.

One can also use the Hartree-Fock method to determine the electron configuration with high precision. Unfortunately this is much more computational expensive than DFT. The advantage of DFT comes is based on the fact, that the main components of it are only based on a scalar function in space, which can be handled rather easy.

### 3.1 Basic Principle

The Hohenberg-Kohn theorem states that the potential  $V$  of a system is only determined by the electron density  $\rho$  (despite an additive constant). This density is the sum of the probability densities of all electrons:

$$\rho(\vec{x}) = \sum_i |\psi_i|^2 = \sum_i \psi_i \psi_i^*. \quad (9)$$

And due to the fact that all other properties of a system just depend on the potential  $V$  and the number of electrons  $N_e$ , they all can be derived from the electron density  $\rho$ . The second Hohenberg-Kohn theorem states that every valid electron density  $\rho$  results in an energy, which is equal or larger than the energy of the ground state.

$$E(\rho) \geq E_0 = E(\rho_0) \quad (10)$$

The electron density  $\rho$  is just a function of  $\vec{x}$  and so we avoid dealing with high dimensional wave function to determine the systems properties.

### 3.2 Kohn-Sham Functions

The expression which connects the density  $\rho$  with the physical behavior of the systems are the Kohn-Sham equations. These equation are derived from the Schrödinger equation (6), one for each electronic spin  $\sigma$ :

$$\mathbf{H}_\sigma \psi_\sigma = (\mathbf{H}_{kin} + \mathbf{V}_{ne} + \mathbf{V}_{coul} + \mathbf{V}_{xc}) \psi_\sigma \quad (11)$$

$$= \epsilon_\sigma \psi_\sigma. \quad (12)$$

The properties in (11) are defined as follows:

$$\mathbf{H}_{kin} = -\frac{1}{2} \Delta \quad (13)$$

$$\mathbf{V}_{ne} = \sum_n \frac{Z_n}{|\vec{x} - \vec{R}_n|} \quad (14)$$

$$\mathbf{V}_{coul} = \int \frac{\rho(\vec{x}')}{|\vec{x} - \vec{x}'|} d\vec{x}' \quad (15)$$

Following the Born-Oppenheimer approximation described in section 2.3, the interaction potential between nuclei and electrons can be described by a simple sum as shown in (14). In this equation  $Z_n$  represents the charge of the nuclei and  $\vec{R}_n$  the position of the nuclei. All four of these properties are operators acting on the functions  $\psi_\sigma$ . These are the Kohn-Sham orbital functions. They represent one electron wave functions, which do not interact.

### 3.3 Exchange-Correlation Potential

The interaction of these wave functions will be handled by the special potential  $V_{xc}(\rho)$ , which is the exchange correlation potential. This potential shall catch the error we

make by simulating the multi-electron system by a sum of non-interacting one electron wave functions. Unfortunately this potential cannot be derived analytically yet and we have to stick to approximations in order to calculate the the proper electron distribution. Despite this potential the whole DFT is absolutely correct, so that better and better approximations for  $V_{xc}$  lead to better and better results for the simulation of microscopical systems. A large part of DFT implementation is the determination of the exchange correlation potential, because one can use iterative methods to approach it. By starting with a initial guess for  $V_{xc}$  one can determine  $E_{xc}$  by solving the Kohn-Sham equations (11). From there we get an approximate  $\rho$  and with that we can calculate  $E_{xc}$  by

$$E_{xc} = \int \rho \epsilon_{xc} d\vec{x}, \quad (16)$$

where  $\epsilon_{xc}$  is the exchange correlation energy functional, which will be discussed later. And with this energy  $E_{xc}$  we can now find a new approximation for  $V_{xc}$  by a variational derivative:

$$V_{xc} = \frac{\delta E_{xc}}{\delta \rho} \quad (17)$$

So we have the following dependency:

$$V(\rho(\psi)) \Leftrightarrow \psi(V).$$

So we need to push the electron density  $\rho$  to be self consistent with its own potential. And the algorithm we roughly described to do that is the self-consistency cycle (SCF cycle).

## 4 SCF Cycle

The cycle we roughly described before contains of the following main points, from which we will describe some in detail:

1. choose set of basis functions  $\mathbf{w}$
2. set initial  $\rho$
3. calculate  $V_{class} = V_{coul} + V_{ne}$
4. determine  $V_{xc}$
5. build Hamilton matrix  $\mathbf{H}$
6. solve Kohn-Sham equations:  
 $\mathbf{H}_\sigma \psi_\sigma = (\mathbf{H}_{kin} + V_{class} + V_{xc}) \psi_\sigma$
7. determine occupation numbers  $n_i$  of the orbitals  $\psi_i$
8. calculate  $\rho = \sum n_i \psi_i^* \psi_i$
9. if not converged: GOTO 3  
else: calculate energy  $E$  and do postprocessing

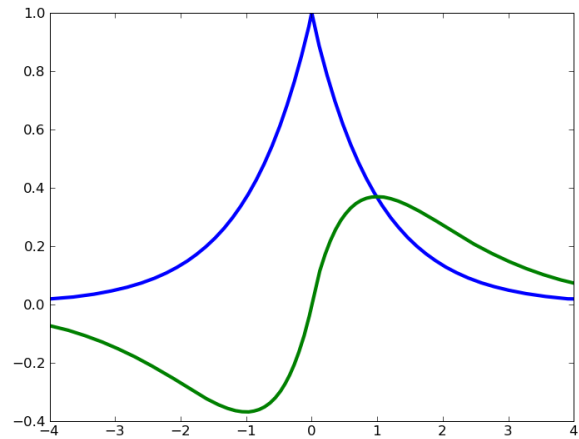


Figure 2: The Slater function in one dimension with  $w(x) = e^{-\alpha|x|}$  in blue and for  $w(x) = x e^{-\alpha|x|}$  in green.

### 4.1 Set of Basis Functions

This basis functions  $\mathbf{w}$  are used as a basis set for solving the Kohn-Sham equations. A linear combination of this functions will later represent the electron density of the system. These basis functions have to be orthogonal to allow a solution of the Kohn-Sham equation. They should be also LCAO (Linear Combination of Atomic Orbitals), because the correct solution will mostly be a superposition of this kind of orbitals. Due to that fact, the basis function should contain a exponentially decreasing term, so that an orbital is really vanished in far distance from its own nucleus. The real atomic orbitals do also decrease in the same matter, so that it is natural to choose the basis functions also with this behavior. For that one can use the Slater functions:

$$w(\vec{x}) = f(\vec{x} - \vec{x}') e^{-\alpha|\vec{x} - \vec{x}'|}. \quad (18)$$

As one can see, it provides the exponentially decaying part with the parameter  $\alpha > 0$ . With this parameter one can steer the locality of the basis function. The function  $f(\vec{x} - \vec{x}')$  can be any function in spatial coordinates as spherical harmonics or cartesian polynomials. Its shape and direction should be determined by the magnetic quantum number. In figure 2 we can see two examples for Slater functions. The clear advantage of this of Slater functions is the correct exponential decay. But due to the fact that we have to integrate complicated products of basis functions and potentials later, we seek for an easy evaluable analytical solution. Unfortunately this is not feasible with the Slater functions.

So another appropriate set of basis functions are Gauss like functions:

$$w(\vec{x}) = f(\vec{x} - \vec{x}') e^{-\alpha|\vec{x} - \vec{x}'|^2}. \quad (19)$$

These functions provide the possibility to calculate integrals and products of them analytically, so that a computational intensive numerical integration doesn't have to be

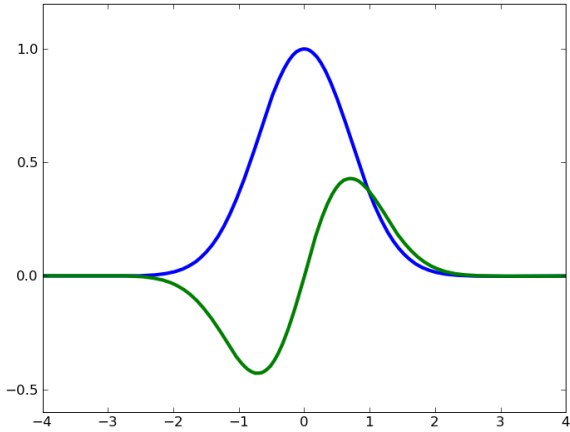


Figure 3: A simple Gaussian function with  $w(x) = e^{-x^2}$  in blue and a Gauss like function with another function  $f(x) = x$  in green.

executed. Unfortunately they do not provide the right exponential decay, so that one is forced to circumvent that by superposition of more than one Gauss like function. These superposition shall provide the peak in the middle as well as the slower exponential decay. In figure 3 we can clearly see that it is decaying much faster than the Slater function in figure 2.

## 4.2 Exchange Correlation Potential

As we mentioned before, the exchange correlation potential and energy have an important role in DFT. The crucial property to determine is the exchange correlation functional  $\epsilon_{xc}$ . The choice of the functional strongly influences the results of the computations. And this choice strongly depends on the problem you want to solve. There are a lot of functionals known in science and each is useful for a special problem. But all these functionals can be divided in a few main groups.

The easiest way to approximate this potential is the local density approximation (LDA):

$$E_{xc} = \int \rho \epsilon_{xc}(\rho) d\vec{x} \quad (20)$$

This is the simplest possible approximation and just using the electron density at a certain position  $\vec{x}$ . But it is still useful for the determination of equilibrium structures, harmonic frequencies and charge moments of molecules. The next more complicated approximation is the local spin-density approximation (LSDA).

$$E_{xc} = \int \rho \epsilon_{xc}(\rho_{\uparrow}, \rho_{\downarrow}) d\vec{x} \quad (21)$$

The difference to LDA is the different evaluation of  $\epsilon_{xc}$ . It now depends explicitly on the two different spins, which has the advantage, that the final result is not necessarily symmetric. This is useful for systems with an odd number of electrons.

The most general approach for the exchange correlation functional is the general gradient approximation (GGA). In this approximation one also take the gradient of the local density into account.

$$E_{xc} = \int \rho \epsilon_{xc}(\rho_{\uparrow}, \rho_{\downarrow}, \nabla \rho_{\uparrow}, \nabla \rho_{\downarrow}) d\vec{x} \quad (22)$$

The GGA functionals are most commonly used right now, but they also produce a lot more computational effort and they are much harder to implement.  $\epsilon_{xc}$  is mostly a rather complicated function. But with GGA one gets really precise results, especially for bonding energies.

## 4.3 Solving the Kohn-Sham Equation

Now we have all the ingredients we need to build the Hamilton matrix  $\mathbf{H}$ .

$$\mathbf{H}_{ij} = \int w_i \mathbf{H} w_j d\vec{x} \quad (23)$$

and solve the eigenvalue problem of the Kohn-Sham equations (11). If we now diagonalize the Hamilton matrix  $\mathbf{H}$ , we get as a result a matrix  $\Psi$  filled with the eigenvectors of the problem. We can now construct the orbitals of the system by linearly combining the basis functions with this eigenvectors, where as each reconstructed orbital has an energy corresponding eigenvalue, which is the energy.

$$\psi_i = \sum_n w_n \Psi_{in} \quad (24)$$

We can now also assign the occupation numbers of the orbitals by just filling them up, starting from the lowest energy. As soon as we have done this it is possible to calculate the new electron density  $\rho$  by

$$\rho = \sum_i n_i \psi_i^* \psi_i, \quad (25)$$

where  $n_i$  is either one for an occupied orbital or zero for an unoccupied orbital. We now can start with the SCF cycle again until it has converged.

## 5 Numerical Effort

This is a short section, which shall show you roughly, where most of the computational power is used.

One big issue is the diagonalization of the Hamilton matrix  $\mathbf{H}$ . It is a process which takes  $\mathcal{O}(N^3)$  operations for a  $N \times N$  matrix. One can circumvent this problem by exploiting symmetries in the simulated system. With that the number of needed operation can be reduced dramatically.

The next big issue is the SCF cycle itself. It contains on the one side the diagonalization of the matrix, as well as the computation of the exchange correlation potential  $V_{xc}$ , which is a complicated integral and variational derivative. Also the integration of the basis functions with that and other potentials is rather expensive for problems with higher dimensions.

But after the SCF cycle some other computations can take place, which can be also rather expensive. We will

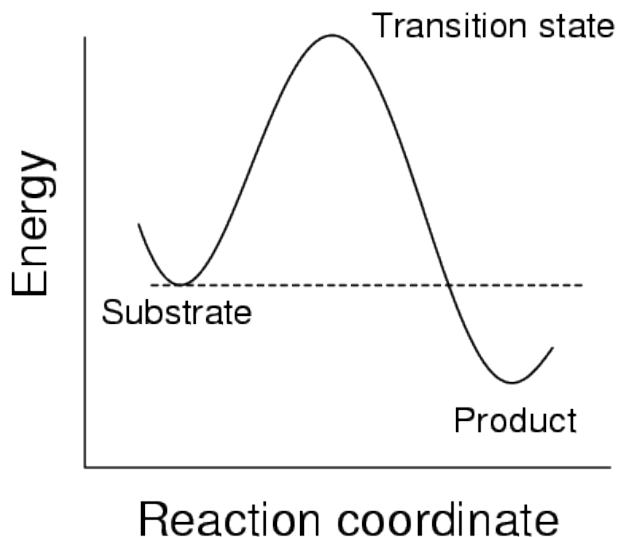


Figure 4: The typical mode of action of a typical chemical reaction. The system is moving in direction of the reaction coordinate over the transition state to reach an energetic or entropic favorable state. The amount of energy needed to reach the top of the transition state is the activation energy.

discuss one of these methods shortly in the next section. This computational intensive method still needs a lot of time, even on modern CPU's. So its parallelization is absolutely crucial and one has to be able to implement all these methods effectively in parallel, to be able to compute the problems coming with modern biology and chemistry.

## 6 Postprocessing

As mentioned before the post processing of the energy data found by the SCF cycle is also consuming a lot of the computational resources. As an example for that we will shortly describe one possible application field of DFT. In computational chemistry it is of special interest to find the reaction coordinate of chemical reaction. This reaction coordinate is an abstract notion for the changes in the electronic and spatial configuration of the initial chemical compounds. During this reaction the energy of the system typically changes. One needs a initial activation energy to push the system over a hill of potential energy (transition state) and after overcoming this hill the reaction moves on until it has reached the final state (compare figure 4) in which all the substrates reacted to the products. In many cases it is not necessary to find the exact reaction coordinate of a system, but the transition state. By knowing this state and its energy one can easily determine reaction parameter for macroscopic systems. But finding this transition state is still hard to do, because it is a first order saddlepoint of the potential energy surface (or landscape) in a  $3N - 6$  dimensional space. There are techniques to search this transition state, but these methods also involve a rather large number of energy calculation, including also the gradient and second derivatives.

In this potential energy landscape one is also able to use the second derivatives of the potential energy (the Hessian matrix respectively) to find eigenmodes of a molecule. For that one needs to diagonalize the Hessian of the potential energy function of a certain configuration. The eigenvalues in that diagonal and the corresponding eigenvectors give a good impression of all vibrational modes of a molecule. Following this, it is also possible to derive the oscillation spectra and compare them to experimental results. And this is again a opportunity to validate the initial DFT calculations.

*We have seen the power and the limitations of density functional theory. We hope that short report helped to understand the general idea and some theoretical aspects of it.*

## References

- [1] T. Grauschopf, *Eine Implementierung der Dichtefunktionalmethode mit einer hybriden Finite-Elemente-Basis*. Dissertation, TU München, 1999.
- [2] W. Koch and M. C. Holthausen, *A Chemists Guide to Density Functional Theory*. Weinheim: Wiley-VCH, second ed., 2001.