Evaluation of the Self-Consistent Sternheimer Method in a Localized Basis

Master Project

Julian Zeller (1369369)
j.zeller@student.tue.nl

Eindhoven University of Technology
Department of Mathematics and Computer Science

Supervisor: dr. B. Baumeier

August, 2020

Abstract

Simulating excitation processes in organic materials is an essential part in the development of organic solar cells. These simulations can be performed within the framework of many-body Green's function theory, in this thesis within the *GW* approximation. These simulations are, however, very time consuming. A major bottleneck is the computation of the microscopic polarizability, which includes a sum over all occupied and unoccupied electron states. In this thesis an alternative approach to the *GW* approximation that uses the Sternheimer equation was implemented. The Sternheimer equation is a result from linear perturbation theory and allows the calculation of the linear response of a quantum mechanical system to an external perturbation. The Sternheimer *GW* method avoids the treatment of unoccupied states and could provide a speed up for *GW* calculations. [1] Compared to previous implementations that use plane wave basis sets, this implementation uses localized basis sets that allow for a more efficient treatment of non-periodic systems. Because localized basis sets have different mathematical properties, several topics throughout the implementation process of the *GW* Sternheimer method need special attention. These topics will be discussed throughout this thesis.

While it was possible to reproduce results of standard *GW* methods, the desired speed up was not achieved. This is because the Sternheimer *GW* methods requires multiple numerical integrations that cannot be performed as efficient with localized basis functions as with plane waves. However, the self-consistent Sternheimer cycle, implemented as a part of the *GW* method, can be used to calculate various other properties. [2] [3] As an example of this, the framework to calculate polarizability tensors was implemented and tested as well.

Contents

Chapter 1

Introduction

Computer simulations have become an important part of scientific research. They provide a way to use results from theoretical and experimental research to compute properties that are otherwise hard to obtain. A computational method can thus be used to examine the underlying processes of physical phenomena helping in confirming experimental evidences or enlightening unknown processes, paving the way for new applications. An example is the study of opto-electronic devices, for example solar cells or LEDs. Building on quantum theory, there are several approaches to simulate the processes that need to happen in order for the solar cell to produce electrical energy or an LED to emit the desired light. [4] The basic principle of opto-electronic devices is either the absorption of photons and converting their energy into electricity or using electricity to emit photons. [5]

Structural and atomic properties are the main ingredients to understand the electronic behaviour of materials. [6] In the case of simulating opto-electronic materials the behavior of the electrons is especially interesting. [6] Since electrons are quantum particles they cannot be described using classical mechanics. [7] That is why the principles of quantum mechanics are used. An electron is described as a wave. The square of this wave can be interpreted as a probability amplitude for the position of the electron, and as a result there is no specific position the electron is in, but rather a probability that the electron is in a specific place at a specific time is observed. [7] It will be explained in chapter 2 how the many-body wave function of a quantum mechanical system is a very complicated object to deal with. Consequently, in the context of quantum mechanical simulation, often the wave functions of the individual electrons are examined. [7] A detailed description on how the many-particle problem can be reduced to treating individual single particle wave functions can be found in the next chapter. A region where such an electron wave function is not zero, which means that the electron can be present there, is called an orbital. [7] There are different types of orbitals with different forms and distances from the nucleus. Due to Pauli's exclusion principle each orbital can be occupied by a maximum of two electrons. [7] Each orbital corresponds to an energy level, that depends on

its distance from the nucleus, while different orbitals can have the same energy level. [7]

The state of a quantum mechanical system is obtained as a function of all single electron wave functions. [7] In the context of quantum chemistry one differentiates two types of states the system can be in. The ground state of a quantum mechanical system is the state with the lowest energy level. [7] It is the state an atom or molecule is in when there is no external perturbation. This means, that all electrons are in the lowest orbitals. The highest of these occupied levels is called highest occupied molecular orbital or short HOMO. Electrons in the HOMO level are particularly interesting. They are the least strongly bound by the nucleus and because of that are especially susceptible to external perturbations. [8] That is why they take a prominent role in excitation processes. There are several well tested methods that simulate the ground state of atoms or molecules very well. [9] Excited states on the other hand are states where at least one electron is on a higher energy level. [10] One way to reach an excited state is the aforementioned absorption of a photon. [5] When the photon has the right energy an electron can be elevated to a free orbital with a higher energy level. As mentioned above, electrons on the HOMO level are the easiest to excite since the energy difference to the unoccupied states is the lowest. [6] An important term to mention here is the lowest unoccupied molecular orbital (LUMO), which is the free orbital with the lowest energy level, and, consequently, exciting electrons to this level requires the least amount of perturbation energy. [8] Computing excited states is a lot more challenging since most methods are constructed for the ground state specifically. [10]

In solar cells characteristics of the electronic structure of the molecule or atom determines which photons are absorbed. [6] One important property of a material is the HOMO-LUMO gap, which is the energy difference between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO). [6] Only if a photon has enough energy to bridge the energy gap from occupied to unoccupied states and thus is able to excite an electron it can be absorbed. This creates a so called exciton, that consists of the excited electron and the hole it leaves behind. [11] Absorption alone does not yet lead to electricity being produced. First the electron that has been excited to a higher energy state needs to be separated from the hole. [5] Usually the 'hole' the electron leaves behind is treated as a positive charge because it represents the lack of a negative charge. Thus, another important property of a solar cell is its ability to successfully separate the electron and the hole. [12]

The quality of photovoltaic cell depends largely on the properties of the material used. [13] The rate at which photons are absorbed is important as well as how many of the excitons are separated. Today, most solar cells are made from crystalline semi-conductors like silicon. [14] One major advantage of these cells is that by dividing the cell into two layers and doping one with boron and the other one with phosphorus an internal electric field is present in the silicon crystal. [15] That is because boron has fewer electrons than silicon and phosphorus has more. The internal electric field is able to separate the hole and the excited electron since they are driven to different layers. While these cells are

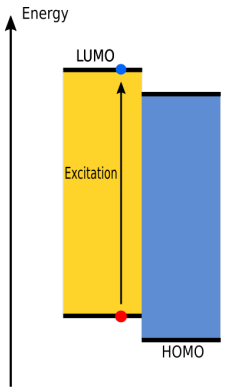

Figure 1.1: This figure shows the schematic structure of a two layer organic solar cell. In the left layer an electron is excited from the HOMO to the LUMO level, leaving behind a hole. The goal of this thesis is to implement an alternative way to obtain the excitation energy of an exciton like this. That in turn allows to find new materials that can be used in organic solar cells.

very efficient once in use, they are very expensive and complicated to produce because high temperatures are needed to form the crystalline structure of silicon. [14] Although they are widely used today, alternatives to these kinds of cells are needed. One of those options is the use of organic polymers. The production of organic solar cells is cheaper and uses less energy, but those cells do not yet reach the efficiency of silicon cells. While silicon cells reach efficiencies of over 20%, organic cells only reach an efficiency of 12%. [16] This is where quantum mechanical simulations come into play. It is possible to calculate energy levels of molecules and simulate the behavior of excitons using for example many-body Green's function theory and the GW approximation. [17] While other methods for excited states simulations exist, in this thesis however the focus lies on the GW approximation presented in chapter 2. These computations can be very helpful for the search of new organic material to use in solar cells and for designing the structure of organic cells. Different materials and arrangements can be simulated and tested for viability in solar cells.

Calculation using the GW approximation can be time consuming and computationally demanding because in addition to the occupied orbitals the unoccupied orbitals of a molecule need to be treated as well. In the standard GW method all possible combinations of occupied and unoccupied need to be examined. [1] This poses a major bottleneck particularly since the number of unoccupied states is significantly larger than the number of occupied states. [10] Therefore, it is favorable to find an approach that avoids the need for unoccupied states. [1] In this thesis such an alternative approach for these kinds of simulations is presented that uses the Sternheimer equation.

The Sternheimer equation will be derived in detail in chapter 3. In summary, it is an equation derived from linear perturbation theory that allows the

computation of the linear response of the electron wave functions to an external perturbation potential. It will be explained in chapter 6 how the Sternheimer equation can be used to avoid the need to treat unoccupied states and replaces that with solving linear systems. The goal of this method is to reduce the computation time of the GW approximation by avoiding the bottleneck of the standard approach. [10]

There already exist implementations of the Sternheimer algorithm that use plane wave basis functions. [10] [2] Plane waves are periodic functions and are therefore very well suited to represent systems with periodic boundary conditions like crystals. While they can also be used for single molecules the boundary conditions can become a problem since the wave functions of the electrons decay to zero towards the boundaries of the system. This leads to plane waves being inefficient in non-periodic systems because a large space around the molecules needs to be examined to avoid self interaction of the molecule through the periodic boundary conditions. Consequently, the plane waves need to be evaluated on a large space which leads to high computational costs. [10]

The goal of this work is to implement the algorithm using localized basis functions that allow cheaper simulations of molecules and atoms. Localized basis functions have their maximum at the nucleus and thus represent the shape of single electron orbital better than a plane wave. [18] The functions used in localized basis sets have different mathematical properties than plane waves. Therefore, several topics need to be addressed when implementing this method with localized basis sets.

1. **Convergence:** The Sternheimer algorithm contains a self-consistent cycle. [1] These cycles often have the problem that convergence is not guaranteed and depends on different parameters. It needs to be investigated how the usage of localized basis sets affects the convergence properties of the self-consistent Sternheimer cycle and what measures can be taken to improve the convergence rate.

2. **Integration:** The Sternheimer method requires an integration over the three spatial directions and over the frequency axis. [19] When using plane wave basis sets, the integrals can be dealt with efficiently using the fast Fourier transform. [20] This is not possible for localized basis sets. The integrals thus have to be evaluated on a discrete grid. It needs to be tested carefully how fine the integration grids need to be in order to calculate accurate results. The finer the grids used are the more time the computations take. It is crucial to find the right balance between computational efficiency and accuracy. Consequently, it needs to be tested how the numerical integrations are affecting the run time of the algorithm and whether the Sternheimer method is also faster for localized basis sets than other GW methods.

3. **Parameters:** The number of parameters that need to be set before starting the Sternheimer calculation also needs attention. As can be seen when the implementation details are discussed in chapter 4, there are several

parameters that need to be chosen. For example, different integration grids can be used, but also the convergence tolerance and factors of an analytical continuation method need to be set beforehand and it needs to be investigated how these parameters affect the results.

4. **Analytic Continuation:** An additional set of parameters is introduced by the use of analytic continuation. Some quantities investigated in this thesis have poles on the real axis. [1] To avoid having to deal with the poles these quantities can be evaluated on the imaginary axis and then projected onto the real axis using analytic continuation. Careful investigations are required to analyse how the continuation used is able to reproduce results on the real axis and how advantageous the calculation on the imaginary axis are compared to the real axis.

5. **Basis expansion:** Lastly, to obtain the desired quantities they have to be expanded in the basis set, in order to calculate them in matrix form. It needs to be examined how the results of the Sternheimer algorithm can be used to compute the needed results when a localized basis set is used.

In the following, first, the fundamentals of quantum mechanical simulation are presented in chapter 2 as well as the conventional approach for excited states simulations using Green's function theory calculations. In chapter 3 the Sternheimer equation is derived from perturbation theory and in chapter 4 the Sternheimer algorithm is presented. Subsequently results of polarisabilty calculations are discussed in chapter 5 to address topics **1**, **3** and **4** without the need of numerical integration. The details of implementing the *GW* algorithm using the Sternheimer methods are presented in chapter 6, dealing with topics **2** and **5**. Finally the results of the implemented method are discussed and assessed in chapter 7.

Chapter 2

Many-Body Theory of Electrons and Nuclei

The focus of this thesis lies on quantum mechanical simulations that describe nuclei, electrons and their interaction. In this chapter an overview over the fundamentals of quantum mechanics is given and different approaches to perform quantum mechanical calculations are presented.

In classical mechanics it is possible to calculate the path of a particle from its position and momentum. A system of particles can be described using Hamiltonian mechanics. Introducing the Hamiltonian function $\mathcal{H}(q_i, p_i, t)$ that depends on the spatial coordinates of the particles q_i, their momenta p_i and the time t. The Hamiltonian function corresponds to the total energy of the system. For example the Hamiltonian function of a non-relativistic particle moving in a potential V is:

$$\mathcal{H}(\mathbf{q}, \mathbf{p}) = \frac{\mathbf{p}}{2m} + V(\mathbf{q}). \tag{2.1}$$

The first term corresponds to the kinetic energy of the particle, the second term to its potential energy. When the potential is constant over time the Hamiltonian function is time independent as well. From the Hamiltonian function the change of the positions and momenta can be calculated from the Hamiltonian equations. The change of the positions is:

$$\frac{d}{dt} q_i = \frac{\partial \mathcal{H}}{\partial p_i}, \tag{2.2}$$

the change of the momentum is:

$$\frac{d}{dt} p_i = -\frac{\partial \mathcal{H}}{\partial q_i}, \tag{2.3}$$

creating a system of differential equations that can be solved to obtain the positions and momenta. For the example above, we obtain:

$$\frac{d}{dt} q_k = \frac{p_k}{m}, \ \frac{d}{dt} p_k = -\frac{\partial V}{\partial q_k}. \tag{2.4}$$

When taking the time derivative of the first equation and multiplying the mass m we can derive the Newtonian equation:

$$\frac{d^2}{dt^2}q_k = F_k = -\frac{\partial V}{\partial q_k}. \tag{2.5}$$

Something similar is done, for example, in molecular dynamics simulations. In each time step the forces on each particle are calculated and from that the position of that particle in the next step. [21]

For quantum mechanical objects this is not possible. Heisenberg's uncertainty theorem states that the position and momentum of a quantum mechanical object, for example an electron, cannot be known at the same time. Thus observing the positions and momenta of electrons as in the classical case is not possible. As an alternative approach a wave function $\Phi(\mathbf{r}, t)$ is introduced to describe a system of quantum mechanical particles like an atom, a molecule or a crystal. [7] Before discussing how to obtain the wave function of a system some practical details are introduced to simplify the notation.

In this thesis atomic units are used in all equations unless stated otherwise. This allows to slim down equations because several frequently used constants are equal to one: $m_e = \hbar = e = c = 4\pi\epsilon_0 = 1$. [22] The constants are the mass of an electron m_e, the elementary charge e, which is the charge of a proton or the absolute of the charge of an electron, the reduced Planck constant $\hbar$, that connects the frequency of a photon to its energy, and the inverse Coulomb constant $4\pi\epsilon_0$. To clarify the dimensions of certain quantities the constants can be reintroduced later.

2.1 The Schrödinger Equation

The central equation in quantum mechanics is the Schrödinger equation:

$$\hat{H}\Phi(t) = i\frac{\partial}{\partial t}\Phi(t). \tag{2.6}$$

It describes the evolution of the wave function over time where $\hat{H}$ is the Hamiltonian of the system and $\Phi(t)$ is the wave function of the system that depends on the time t. The Hamiltonian $\hat{H}$ is an operator that sums over all kinetic and potential energies. It is the quantum mechanical analog to the Hamiltonian function in classical mechanics. When looking at a many body system, like a crystal, an atom or a molecule, the Hamiltonian can be specified as:

$$\begin{aligned}\hat{H} = &-\frac{1}{2}\sum_i \nabla_i^2 - \sum_j \frac{1}{2m_j}\nabla_j^2 - \sum_{i,j}\frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} \\ &+ \sum_{i,k\neq i}\frac{1}{|\mathbf{r}_i - \mathbf{r}_k|} + \sum_{i,k\neq i}\frac{Z_j Z_k}{|\mathbf{R}_j - \mathbf{R}_k|} \\ &= \hat{T}_e + \hat{T}_N + \hat{V}_{eN} + \hat{V}_{ee} + \hat{V}_{NN}.\end{aligned} \tag{2.7, 2.8}$$

The first two terms describe the kinetic energy of the electrons and the nuclei respectively. The last three terms the coulomb interaction between electrons and nuclei, electrons and electrons and nuclei and nuclei. The nuclei-nuclei and electron-electron interactions are repulsive terms because the have the same charge, the nuclei-electron interaction is attractive and thus has a negative sign. In equation 2.7 the lowercase $\mathbf{r}_i$ refers to the position of the electron i and the capital $\mathbf{R}_j$ to the position of the nucleus j. Accordingly electrons are always referred to with the index i and nuclei with the index j. Z_j is the charge of the nucleus j. The operator ∇_n^2 is the Laplacian, the second derivative respective to the three cartesian directions, of the particle n. Since the Hamiltonian depends on the positions of the particles the wave function of the system depends on them as well. [7] The Schrödinger equation is a linear differential equation since there all terms are linear in $\Phi(t, \{\mathbf{r}_i\}, \{\mathbf{R}_j\})$. Therefore, a linear combination of solutions of the Schrödinger equation, the wave functions, is itself a solution of the equation. The wave functions thus form a vector space. Also the standard L^2 scalar product can be applied to the wave functions. This inner product:

$$(\phi\mathbf{r}, \psi\mathbf{r}) = \int d\mathbf{r}\phi(\mathbf{r})\psi(\mathbf{r}) \tag{2.9}$$

can be interpreted as the overlap of two wave functions which is the probability of a system in a state corresponding the wave function $\phi(\mathbf{r})$ transitioning into the state corresponding to $\psi(\mathbf{r})$. A widely used notation in quantum mechanics is the Bra-Ket or Dirac notation. It shortens and clears up notation. An element of the vector space is denoted as a "ket":$|\psi\rangle$. Where as an operator acting on a vector is denoted as a "bra": $\langle\phi|$. A scalar product of two vectors can thus be written as $\langle\phi|\psi\rangle$ and the Hamiltonian acting on a wave function is written as $\hat{H}|\psi\rangle$.

The simplest system to examine is the Hydrogen atom. It consists of one nucleus and one electron. Therefore, the electron electron and nucleus nucleus interaction terms are omitted. For this case the Schrödinger equation is analytically solvable by using the spherical symmetry of the atom to split the three dimensional partial differential equation into three separate equations that can be solved individually. For all cases with more than one electron however, no such solution exists. As will be visible throughout this chapter, the many-body wave function is a very complicated object to deal with.

With no analytic solution available, a possible alternative is to solve the equation numerically. This has proven to be impractical as well. When looking at a system of N particles, the wave function Φ depends on $3N$ coordinates. A naive attempt to solve the equation numerically is to discretize the wave function on a spatial grid. Even when choosing a very coarse grid of 20 grid points in each direction, 20^{3N} values are needed to describe the wave function on the grid. A water molecule has 10 electrons. Therefore, more than 10^{39} values would be required. This is not feasible even on large computers. [9]

For most cases it is sufficient to calculate the wave function of a stationary state. If the Hamiltonian does not explicitly depend on the time, the space and

time dependence of the wave function can be separated: [7]

$$\Phi(t, \{\mathbf{r}_i\}, \{\mathbf{R}_j\}) = \phi(\{\mathbf{r}_i\}, \{\mathbf{R}_j\})\tau(t). \tag{2.10}$$

Inserting this ansatz into the Schrödinger equation and applying separation of variables yields that the time dependent function $\tau(t)$ is always of the form:

$$\tau(t) = \exp\left\{\frac{-iEt}{\hbar}\right\} \tag{2.11}$$

and thus the equation only needs to be solved for $\phi(\{\mathbf{r}_i\}, \{\mathbf{R}_j\})$. Consequently, the time derivative simplifies to a constant E, that is the energy level of the system. The problem is simplified to an eigenvalue problem described using the time-independent Schrödinger equation:

$$\hat{H} \left|\Phi\right\rangle = E \left|\Phi\right\rangle. \tag{2.12}$$

The eigenvectors of the Hamiltonian, Φ_n ,are the stationary solutions of the Schrödinger equation.

Still, all the issues of solving the time dependent Schrödinger equation also apply to the time independent case. As a result, the problem needs to be simplified in order to approximate the wave function of a non-trivial quantum mechanical system.

2.2 The Born Oppenheimer Approximation

A first idea to simplifying the problem is to treat the electrons separately from the nuclei. The Born-Oppenheimer approximation provides a way to achieve exactly that. Using the approximation, the position of the nuclei can be fixed and only the electrons need to be examined. [23] The Born-Oppenheimer approximation is used in a variety of different methods. At the end of this section it will be discussed, when the Born-Oppenheimer equation can be used and it what cases it fails.

The motivation to this approximation is that the mass of the nuclei is approximately four orders of magnitude larger than the mass of an electron. [23] A nucleus with the same kinetic energy as an electron moves therefore significantly slower.

The central idea is that the electrons follow the movement of the nuclei instantaneously. The initial ansatz is to separat the wave function into

$$\Phi(\{\mathbf{r}_i\}, \{\mathbf{R}_j\}) = \psi_{e[\{\mathbf{R}_j\}]}(\{\mathbf{r}_i\})\psi_N(\{\mathbf{R}_j\}) \tag{2.13}$$

an electronic wave function, that depends parametrically on the position of the nuclei, and the nuclear wave function. To keep the notation short the parametric dependence on all nuclear positions $\{\mathbf{R}_j\}$ is written as $\psi_{e[\mathbf{R}]}$. When fixing the nuclear positions the electronic Hamiltonian simplifies to:

$$\hat{H}_e = -\frac{1}{2}\sum_i \nabla_i^2 - \sum_{i,j} \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} + \sum_{i,k\neq i} \frac{1}{|\mathbf{r}_i - \mathbf{r}_k|} = \hat{T}_e + \hat{V}_{eN} + \hat{V}_{ee}. \tag{2.14}$$

The kinetic energy term of the nuclei can be dropped because the nuclei do not move. The nuclei nuclei Coulomb term provides a constant contribution since it only depends on the fixed nuclei positions. Therefore, it only introduces a constant shift of the eigenvalues and is neglected here. We thus obtain the electronic Schödinger equation:

$$\hat{H}_e \left|\psi_{e[\mathbf{R}]}(\mathbf{r})\right\rangle = E_e \left|\psi_{e[\mathbf{R}]}(\mathbf{r})\right\rangle \tag{2.15}$$

This equation can be used to calculate the electronic part of the wave function. The number of particles as well as the number of terms in the Hamiltonian reduced compared to the original Schrödinger equation.

When inserting the split ansatz of the wave function into the original time independent Schrödinger equation:

$$\hat{H}\psi_{e[\mathbf{R}]}(\mathbf{r})\psi_N(\mathbf{R}) = E\psi_{e[\mathbf{R}]}(\mathbf{r})\psi_N(\mathbf{R}) \tag{2.16}$$

we need to investigate how each term of the Hamiltonian acts on the wave functions. For the kinetic energy term of the electrons we obtain:

$$\hat{T}_e\psi_{e[\mathbf{R}]}(\mathbf{r})\psi_N(\mathbf{R}) = \psi_N(\mathbf{R})\hat{T}_e\psi_{e[\mathbf{R}]}(\mathbf{r}) \tag{2.17}$$

since the operator $\hat{T}_e$ has no dependence on $\mathbf{R}$. The same does not hold for the kinetic energy of the nuclei, here we get:

$$\nabla^2_{\mathbf{R}}\psi_{e[\mathbf{R}]}(\mathbf{r})\psi_N(\mathbf{R}) = \psi_{e[\mathbf{R}]}(\mathbf{r})\nabla^2_{\mathbf{R}}\psi_N(\mathbf{R}) + 2\nabla_{\mathbf{R}}\psi_{e[\mathbf{R}]}(\mathbf{r})\nabla_{\mathbf{R}}\psi_N(\mathbf{R}) + \tag{2.18}$$

$$+ \psi_N(\mathbf{R})\nabla^2_{\mathbf{R}}\psi_{e[\mathbf{R}]}(\mathbf{r}) \tag{2.19}$$

by chain rule.

Inserting this and the electronic Schödinger equation 2.15 back into equation 2.16 yields:

$$\psi_{e[\mathbf{R}]}(\mathbf{r})\nabla^2_R\psi_N(\mathbf{R}) + \psi_N(\mathbf{R})\psi_{e[\mathbf{R}]}(\mathbf{r})\left(E_e + V_{NN}\right) \tag{2.20}$$

$$-\left[\sum \frac{1}{2m_n}(2\nabla_{\mathbf{R}}\psi_{e[\mathbf{R}]}(\mathbf{r})\nabla_{\mathbf{R}}\psi_N(\mathbf{R}) + \psi_N(\mathbf{R})\nabla^2_{\mathbf{R}}\psi_{e[\mathbf{R}]}(\mathbf{r}))\right] \tag{2.21}$$

$$= E\psi_{e[\mathbf{R}]}(\mathbf{r})\psi_N(\mathbf{R}). \tag{2.22}$$

The last term in brackets needs to be approximated because it contains the derivatives of the electron wave function. As long as they are in the equation, we cannot decouple the nuclear part of the equation. It is usually approximated with $p_e^2/2m_n = (m_e/m_n)E_e$. Because m_e/m_n is around 10^{-5}, it can be dropped. This is the actual Born-Oppenheimer approximation that allows the decoupling of the equations. Here we see again the motivation explained in the beginning. The mass ratio between electrons and nucleus is small enough such that we can ignore the coupling term of the nuclear kinetic energy operator and thus the systems can be decoupled with only a small error introduced. What remains is:

$$\psi_{e[\mathbf{R}]}(\mathbf{r})\hat{T}_N\psi_N(\mathbf{R}) + \psi_N(\mathbf{R})\psi_{e[\mathbf{R}]}(\mathbf{r})\left(E_e + V_{NN}\right) = E\psi_{e[\mathbf{R}]}(\mathbf{r})\psi_N(\mathbf{R}). \tag{2.23}$$

Finally we obtain the nuclear Schrödinger equation by dividing out the electronic wave function:

$$(\hat{T}_N + \hat{E}_e + \hat{V}_{NN})\psi_N(\mathbf{R}) = E\psi_N(\mathbf{R}). \tag{2.24}$$

We can now solve the electronic and nuclear Schrödinger equations independently. Usually the nuclei electron interaction in the electronic Schrödinger equation is expressed using a potential $v(\mathbf{r})$ with fixed nuclei. Because this approach neglects nuclear motions like vibrations, the kinetic energy of the nuclei can be introduced again in the nuclear Schrödinger equation. It can be solved to obtain the nuclear motions. [7]

To see when the Born Oppenheimer approximation is applicable, we look at vibronic coupling, a way to improve the approximation by including nuclear coupling effects. [24]

Suppose that the functions $\psi_{k[\mathbf{R}]}(\mathbf{r})$ are the N orthonormal electronic eigenfunctions from the time independent Schrödinger equation. The corresponding matrix formulation of the electronic Hamiltonian:

$$\mathbf{H}^e_{ij} = \left\langle \psi_{i[\mathbf{R}]}(\mathbf{r}) \right| \hat{H}_e \left| \psi_{j[\mathbf{R}]}(\mathbf{r}) \right\rangle = \delta_{ij} E_k(\mathbf{R}) \tag{2.25}$$

is a diagonal matrix. The nuclear Hamiltonian matrix $\mathbf{H}_n$ however is not diagonal as can be seen from equation 2.19. The matrix formulation of the nuclear Hamiltonian, with respect to the electronic wave functions, is:

$$\mathbf{H}^n_{ij} = \delta_{ij} T_n - \sum_{n\alpha} \frac{1}{M_n} \langle \psi_i | P_{n\alpha} | \psi_j \rangle P_{n\alpha} + \langle \psi_i | T_n | \psi_j \rangle . \tag{2.26}$$

With

$$P_{nm}\psi_{i[\mathbf{R}]}(\mathbf{r}) = -i\frac{\partial \psi_{i[\mathbf{R}](\mathbf{r})}}{\partial \mathbf{R}_{n\alpha}};\ \alpha = x, y, z \tag{2.27}$$

and n being the indices of the nuclei.

The term dropped in the Born-Oppenheimer approximation in equation 2.21 represents the off diagonal elements of the nuclear Hamiltonian. Analog to before, the off diagonal elements prevent the decoupling of the Schrödinger equation. Neglecting the off diagonal elements is the same as dropping the term in the Born-Oppenheimer approximation, and thus also leads to the decoupling of the electronic and nuclear wave functions. The off diagonal elements of the matrix formulation satisfy:

$$\langle \psi_i | P_{n\alpha} | \psi_j \rangle = \frac{\langle \psi_i | [P_{n\alpha}, \mathbf{H}_e] | \psi_j \rangle}{E_i(\mathbf{R}) - E_j(\mathbf{R})} \tag{2.28}$$

with the numerator:

$$\langle \psi_i | [P_{n\alpha}, \mathbf{H}_e] | \psi_j \rangle = iZ_n \sum_m \langle \phi_i | \frac{(\mathbf{r}_m - \mathbf{R}_n)_\alpha}{(\mathbf{r}_m - \mathbf{R}_n)^3} | \psi_j \rangle . \tag{2.29}$$

The potential energy surfaces $E_i(\mathbf{R})$ are the electronic eigenvalues obtained from solving the electronic Schrödinger equation for different configurations of

the nuclei. Thus we obtain a potential energy surface for each electron in the system. When the potential energy surfaces are well separated for all electrons:

$$E_0(\mathbf{R}) \ll E_1(\mathbf{R}) \ll E_2(\mathbf{R}) \ll \ldots \ll E_n(\mathbf{R}), \tag{2.30}$$

the Born-Oppenheimer approximation can be trusted because the off diagonal elements of the nuclear Hamiltonian can be neglected. On the other hand, when two potential energy surfaces come too close, the Born-Oppenheimer approximation breaks down since the coupling effects cannot be ignored. Then other methods like vibronic coupling need to be used to account for these effects.

The Born-Oppenheimer approximation is widely used since decoupling the electrons and nuclei decreases the complexity significantly. The problem is separated into two smaller problems. When the approximation is applicable it yields good results and from here on all methods and theories presented work within the Born-Oppenheimer approximation. Additionally the following methods all focus on the electronic structure of the system and the nuclei are assumed to be fixed.

2.3 The Hartree-Fock Method

The Hartree method was one of the first methods introduced to approximate the solution of the Schrödinger equation. [25] The motivation to this methods lies in the variational principle. The correct wave function is the one that minimizes the expected value of the Hamiltonian and thus the ground state energy level. Consequently, any other wave function gives a higher energy value. By testing all possible wave functions the correct one would be easy to identify. Since this is clearly not possible, the goal of the Hartree method is to introduce a way to calculate a trial function that is optimal under the assumptions made for this method.

The fundamental idea of the Hartree method is to break down the many body nature of the N-electron wave function by introducing single particle wave functions. Each electron is represented by an individual wave function $\chi_i(\mathbf{r}_i)$. The pairwise interaction of the electrons is approximated by the interaction with a field potential that summarizes the contributions of all other electrons. The advantage of this method is that instead of solving the N-electron Schrödinger equation the one particle Schrödinger equation can be solved for each electron. This simplifies the calculation significantly.

In the Hartree Method the wave function of the system is assumed to be the product of the single electron wave functions: [26]

$$\Phi(\mathbf{r}_i) = \prod_i \chi_i(\mathbf{r}_i) \tag{2.31}$$

While this ansatz is simple to implement, it fails to reproduce the anti-symmetric property of fermions, such as electrons, that is stated by Pauli's exlusion principle. Switching two electrons should result in a sign change of the wave function. This is not the case when the single particle wave functions are only multiplied.

A more sophisticated approach is used in the Hartree-Fock method, an extension of the Hartree Method. Here the N electron wave function of the system is calculated using a so called Slater determinant: [27]

$$\Psi(\mathbf{r}_i) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{r}_1) & \chi_2(\mathbf{r}_1) & \cdots & \chi_N(\mathbf{r}_1) \\ \chi_1(\mathbf{r}_2) & \chi_2(\mathbf{r}_2) & \cdots & \chi_N(\mathbf{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{r}_N) & \chi_2(\mathbf{r}_N) & \cdots & \chi_N(\mathbf{r}_N) \end{vmatrix} \tag{2.32}$$

When using a Slater determinant, switching two electrons changes the sign of the wave function. [26]

The Hartree-Fock method calculation starts with an initial guess for the one electron wave functions. These guesses can be atomic orbitals of a hydrogen atom or linear combinations of atomic orbitals in the case of molecules. Then the Fock-operators, that are a result of the aforementioned variational conditions, are built from the orbitals as:

$$\hat{F}[\psi_j]_i = \hat{H}_i + \sum_{j=1}^{N/2} 2\hat{J}_{ji} - \hat{K}_{ji}. \tag{2.33}$$

$\hat{F}[\psi_j]_i$ is the Fock-operator built from the wave functions ψ_j for the $i-th$ electron. $\hat{H}_i$ is the single particle Hamiltonian:

$$\hat{H}_i = \frac{1}{2}\nabla_i^2 - \sum_n \frac{Z_n}{|\mathbf{r}_i \mathbf{r}_n|}. \tag{2.34}$$

It consists of the electron's kinetic energy and the nuclei electron Coulomb interaction. J_{ji} is the Coulomb operator that corresponds to the Coulomb interaction between the electron and the electrons in the $j-th$ orbitals. Because each orbital holds two electrons, the operator is multiplied by two. When there is an uneven number of electrons, the single electron orbital needs to be treated outside the sum without the factor two. $\hat{K}$ is the exchange operator that incorporates Pauli's exclusion principle. It is a relic of the Slater determinant used to build the wave function. To obtain the single electron wave functions, the eigenvalue problem:

$$\hat{F}[\psi_j]_i \chi_i = \epsilon_i \chi_i \tag{2.35}$$

is solved. The calculated one electron wave function are themselves orbitals that can be used to construct new Fock-operators. Therefore, the Hartree method is performed in a self-consistent manner. The new one electron wave functions are used to build new Fock operators that in turn are used to calculate new one electron wave functions. This cycle is reaped until self-consistency is reached.

The Hartree-Fock method uses, as described, simplifications, that can for some cases lead to large deviations from experimental observations. One of the main reasons is that the many electron wave function cannot be expressed using only one Slater determinant. The Hartree-Fock method is therefore not

capable to fully incorporate the electronic correlation. There are extensions, called post-Hartree Fock methods, that improve on some of its weaknesses, e.g. Configuration Interaction. The idea of Configuration Interaction is to use more than one electron configuration and thus built the wave function from a linear combination of Slater determinants. While this accounts for electron correlation much better than standard Hartree-Fock methods, Configuration Interaction and other post Hartree-Fock methods are computationally very demanding. [28]

2.4 Density-Functional Theory

Density functional theory (DFT) is an exact theory that allows to investigate the electronic structure of many body systems, like molecules. Traditional density functional theory is specifically constructed for the ground state of a system, and the theory is therefore mainly used to investigate ground state properties.

Compared to other approaches like the Hartree(-Fock) method, it does not require the direct solution of the many electron Schrödinger equation. As already mentioned above, the main problem with solving the Schrödinger equation is its many body nature. It is computationally very demanding to solve the Schrödinger equation for even a small number of atoms. In DFT instead of looking at each electron individually all electrons are treated simultaneously by observing the electron density. The ground state energy can be uniquely expressed by that electron density that depends on three spatial coordinates only. [9] This reduces the complexity of the many body problem from 3N electron coordinates to only three spatial coordinates of the electron density through the use of functionals.

2.4.1 The Hohenberg-Kohn Theorem

The central idea of DFT is to map the many particle problem of the Coulomb interaction onto the electron density $n(\mathbf{r})$. The density is defined as:

$$n(\mathbf{r}) = N \int d\mathbf{r}_2 \cdots \int d\mathbf{r}_N \, \Phi^*(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N)\Phi(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N) \tag{2.36}$$

with $\Phi(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ the many electron wave function.

At the core of DFT stands the theorem of Hohenberg and Kohn, which states that for a fixed potential $v(\mathbf{r})$, that can for example be the potential of the nuclei, the ground state energy of a quantum mechanical system can be expressed through a unique functional of the electron density $n(\mathbf{r})$. [29] Instead of working with the wave function, that depends on $3N$ variables, we can look at the electron density that only depends on 3 variables. The resulting ground state energy functional has the following form:

$$E[n] = T[n] + \int v(\mathbf{r})n(\mathbf{r})d\mathbf{r} + \frac{1}{2}\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + G[n]. \tag{2.37}$$

The first part corresponds to the kinetic energy of the electrons, the second term to the contribution from the external field $v(\mathbf{r})$, the third term corresponds to the Hartree energy and the fourth term $\hat{G}[n]$ corresponds to contributions from the electron-electron interactions not covered from the terms before. This functional is often referred to as the exchange-correlation functional. The exact form of this functional is not known for non-trivial cases. Therefore, it needs to be approximated. It is important to note that DFT is an exact theory. The only approximation lies in the estimation of the exchange-correlation functional. [9]

When looking at this result it is not immediately clear how the number of variables can be reduced from $3N$ to three while still obtaining the same information. To understand this apparent contradiction, we need to look at the variational principle using the density. For a given ground state density $n_0(\mathbf{r})$, the corresponding ground state wave function $\Phi(\mathbf{r}_i)$ must not only reproduce the ground state density, it also needs to minimize the energy. [9] The ground state wave therefore must satisfy the requirement:

$$E_{v,0} = \min_{\Phi \to n_0} \langle \Phi | \hat{E} | \Phi \rangle \tag{2.38}$$

$\Phi \to n_0$ stands for all wave functions that reproduce the ground state density n_0. From all the wave functions that reproduce the ground state density, the ground state wave function is the one that also minimizes the energy. [9] Using an electron density $n(\mathbf{r})$ different from the ground state density, we obtain a different wave function Ψ. When looking at the functional

$$E_v[n] = \min_{\Phi \to n} \langle \Phi | \hat{E} | \Phi \rangle , \tag{2.39}$$

according to the variational principle the resulting energy for an arbitrary density is greater or equal than the ground state energy:

$$E_v[n] \geq E_v[n_0] = E_{v,0}. \tag{2.40}$$

Thus, it is always possible to set an upper bound to the ground state energy even if the ground state density is not exactly known. The knowledge of the ground state density combined with the energy minimization requirement therefore allows us to uniquely express the ground state wave function as a functional of the electron density.

2.4.2 The Kohn-Sham Equations

There are several approaches to practical DFT calculations. An obvious idea would be to directly minimize the total Energy with respect to the density. The correct density is the one that minimizes the energy functional. In practice this is not optimal. A more popular approach is the Kohn-Sham approach. [30] It does not exclusively rely on electron density calculations, but instead introduces non-interacting single orbital wave functions.

The idea remains to minimize the ground state energy functional:

$$\frac{\delta E[n]}{\delta n(\mathbf{r})} = \frac{\delta T[n]}{\delta n(\mathbf{r})} + \frac{\delta V[n]}{\delta n(\mathbf{r})} + \frac{\delta U_{\mathrm{H}}[n]}{\delta n(\mathbf{r})} + \frac{\delta G[n]}{\delta n(\mathbf{r})} = 0 \tag{2.41}$$

with respect to the electron density. $V[n]$ and $U_H[n]$ are the external field and Hartree energy functionals respectively. The ansatz introduced by Kohn and Sham is to use non-interacting particles that move in a potential v_s. The equation thus simplifies to:

$$\frac{\delta E[n]}{\delta n(\mathbf{r})} = \frac{\delta T[n]}{\delta n(\mathbf{r})} + \frac{\delta v_\mathrm{s}[n]}{\delta n(\mathbf{r})} = 0, \tag{2.42}$$

when the potential v_s is chosen to be:

$$v_s[\mathbf{r}] = v[\mathbf{r}] + v_\mathrm{H}[\mathbf{r}] + v_\mathrm{xc}[\mathbf{r}], \tag{2.43}$$

the sum of the external, Hartree and exchange-correlation potential. This are exactly the potentials corresponding to the energy functionals in equation 2.37. The density $n(\mathbf{r})$ that solves equation 2.42 is the same density that solves equation 2.41. [9] The many electron problem is thus reduced to a problem of non-interacting electrons moving in a potential. In practice the computation starts by calculating an electron density using non-interacting electron states:

$$n(\mathbf{r}) = 2\sum_{i=1}^{n_\mathrm{occ}} \psi_i^*(\mathbf{r})\psi_i(\mathbf{r}). \tag{2.44}$$

The number of occupied orbitals is n_{occ} and the factor 2 accounts for spin degeneracy.

Spin is a property of quantum mechanical objects. Although no real rotation is involved it can be described as an angular momentum. In quantum mechanics the term degeneracy is used if two or more different quantum states correspond to the same energy level. Electrons are fermions and therefore obey the Pauli exclusion principle. It states that two electrons in an atom cannot be in the same quantum state. Thus each atomic orbital can be occupied by a maximum of two electrons, one with up spin and one with down spin. From now on we use the degeneracy factor of two to avoid the need to explicitly treat the spin of the electrons.

The density obtained is then used to construct the Kohn-Sham Hamiltonian:

$$\hat{H}^\mathrm{KS} = -\frac{1}{2}\nabla^2 + \hat{v}(\mathbf{r}) + \hat{V}^\mathrm{H}(\mathbf{r}) + \hat{V}^\mathrm{XC}(\mathbf{r}) \tag{2.45}$$

where

$$V^\mathrm{H}(\mathbf{r}) = \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{2.46}$$

is the Hartree potential. Using the non-interacting particles reduces the complexity of the Hamiltonian because no pairwise interactions need to be treated. This reduces the cost of solving the Schrödinger equation drastically.

$$\hat{H}^\mathrm{KS}\psi_n(\mathbf{r}) = \epsilon_n\psi_n(\mathbf{r}) \tag{2.47}$$

These single particle Schrödinger equations are also known as the Kohn-Sham equations. We now have a self-consistent procedure similar to the one from the

Hartree-Fock method. The electron density depends on the single particle wave functions, which are calculated from the non interacting Schödinger equation, that is constructed from the Kohn-Sham Hamiltonian that in turn depends on the electron density. This cycle can be iterated until convergence is reached. A brief description on how the Kohn-Sham density functional theory can be implemented using localized basis functions can be found in chapter 4.

2.4.3 The Local Density Approximation

The last piece missing is the exchange-correlation functional. There are several different approaches to approximating this functional. One of them is the local density approximation.

The local density approximation (LDA) estimates the exchange-correlation energy of the system with the exchange-correlation energy of a homogeneous electron gas. Therefore, it can be written as:

$$\hat{E}_{\mathrm{XC}}[n(\mathbf{r})] = \int \epsilon_{\mathrm{XC}}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r}. \tag{2.48}$$

The function $\epsilon_{\mathrm{XC}}(\mathbf{r})$ is the energy per electron at point $(\mathbf{r})$. [31] This is a very simple approximation since variations in the charge density are not accounted for. Regardless, the LDA has proven to be reliable for a large variety of systems.

The LDA can be improved by including the gradient of the electron density to account for changes in the density. These functionals are known as general gradient approximation functionals. The exchange-correlation functional can be further improved by including the Laplacian of the electron density or include results of the exact exchange energy from Hartree-Fock calculation. The latter are known as hybrid functionals. Also linear combinations of different functionals exist to balance out advantages and disadvantages of the used functionals.

Because these various different approximations for the exchange-correlation functional exist, the choice of the exchange-correlation functional is crucial for the results of a DFT simulation. Often a compromise between accuracy and computational cost has to be found. However, more accurate results are not guaranteed when for example using a GGA functional instead of an LDA functional. [31] [32]

2.4.4 Analysis and Limits

While Kohn-Sham DFT yields good results for many different applications, it also has its drawbacks. The Introduction of non-interacting particles is an essential part in reducing the complexity of the problem, but as a result the Kohn-Sham wave functions have little physical meaning because they represent artificial non-interacting particles and not real electrons. A good example where this causes problems is Koopman's theorem. It states that in Hartree-Fock calculations the first ionization energy is the negative orbital energy of the HOMO level. [33] The theorem assumes that removing an electron from the highest occupied orbital does not change the orbitals. The accuracy of ionization

energies calculated this way is directly linked to the quality of the Hartree-Fock calculation and yields results that qualitatively agree with experimental values. [34] This is not possible in the same way for Kohn-Sham orbitals. While there exists a version of this theorem for Kohn-Sham orbitals, the resulting ionization energies are often very different from experimental values. [35] This is due to the approximation of the exchange-correlation functional and the nature of the artificial wave functions.

Due to its construction, DFT also fails to produce reliable results for excited states calculation. That is why other theories like the many body Green's function theory are applied to compute properties of excited systems.

2.5 Green's Function Theory and the GW Approximation

As mentioned above, excited states calculations prove to be a difficult task since the widely used density functional theory is limited to ground state calculations. The many body Green's function theory provides a way to treat excited states by introducing creation and destruction operators which can add a particle, for example an electron, to the system or remove it. First the Green's function is introduced and some simple applications are shown. We then focus on the GW approximation which uses the single particle Green's function G as well as the screened Coulomb interaction W to approximate the self energy Σ of a system.

2.5.1 The Green's Function

Green's functions are widely used for solving partial differential equations (PDEs). [36] They where first introduced by George Green. For excited states calculations we state the single particle Green's function: [20]

$$G(\mathbf{r}, t, \mathbf{r}', t') = \langle N | \hat{T} \left[\hat{\psi}(\mathbf{r}, t) \hat{\psi}^\dagger(\mathbf{r}', t') \right] | N \rangle . \tag{2.49}$$

In this equation $\hat{T}$ is the time ordering operator. In the case $t > t'$ it orders the events in a way that t' happens before t. ψ and $\psi^\dagger$ are the creation and destruction field operators for electrons. They are defined as:

$$\hat{\psi}(\mathbf{r}', t') = \sum_n \phi_n(\mathbf{r}) c_n(t) \tag{2.50}$$

and

$$\hat{\psi}^\dagger(\mathbf{r}', t') = \sum_n \phi_n(\mathbf{r}) c_n^\dagger(t), \tag{2.51}$$

with ϕ_n being single particle wave functions, c_n the annihilation operator and $c_n^\dagger$ the creation operator. Further, N is the ground state wave function for the

N particle system.
Using the Heaviside function Θ, equation 2.49 can be expanded to

$$\begin{aligned} G(\mathbf{r},t,\mathbf{r}',t') = &- i\Theta(t-t') \langle N| \hat{\psi}(\mathbf{r},t)\hat{\psi}^\dagger(\mathbf{r}',t') |N\rangle \\ &+ i\Theta(t'-t) \langle N| \hat{\psi}^\dagger(\mathbf{r}',t')\hat{\psi}(\mathbf{r},t) |N\rangle . \end{aligned} \tag{2.52}$$

The Heaviside function is zero for negative arguments and one for positive arguments. Therefore, the the first bracket is evaluated if $t > t'$ and the second if $t \leq t'$. For the first case, this equations describes the insertion of an electron at time t' and coordinates $\mathbf{r}'$ and the removal of said particle at t and $\mathbf{r}$. In the second case the equation represents the removal of an electron and therefore the creation of a hole at time t, which is filled again at time t'.
The time dependence of the field operators can be expressed using the single particle eigenvalues ϵ_n:

$$\hat{\psi}(\mathbf{r}',t') = \sum_n \phi_n^*(\mathbf{r}) c_n e^{-i\epsilon_n t}. \tag{2.53}$$

Inserting this into equation 2.49 yields

$$\begin{aligned} G(\mathbf{r},t,\mathbf{r}',t') = &- i\Theta(t-t') \sum_{\epsilon_n > \epsilon_f} \phi_n(\mathbf{r})\phi_n^*(\mathbf{r}') e^{-i\epsilon_n(t-t')} \\ &+ i\Theta(t'-t) \sum_{\epsilon_n < \epsilon_f} \phi_n(\mathbf{r})\phi_n^*(\mathbf{r}') e^{-i\epsilon_n(t-t')} \end{aligned} \tag{2.54}$$

ϵ_f is the energy level of highest occupied electron state. The first term of the equation only takes unoccupied states into account since it represents the addition of an electron into an unoccupied state. The opposite holds for the second term since only the occupied states are used. This is because in this case an electron has to be removed from an occupied state. For $\mathbf{r}' \to \mathbf{r}$ and $t' \to t^+ = t + \delta$ the Green's function becomes the charge density of the system as creation and destruction times and positions coincide. [20]

This is an example on how the Green's function can be used to treat a one particle excitation. The Green's function theory can also be expanded to treat the excitation of more than one particle. In this thesis however we look at the *GW* approximation that uses the single particle Green's function to approximate the self-energy of a system.

2.5.2 The Screened Coulomb Interaction

To define the screened Coulomb interaction, first, the dielectric function has to be established. Therefore, we look at the potential that is acting on the electrons: [20]

$$V(\mathbf{r},t) = \phi(\mathbf{r},t) - i \int d\mathbf{r}' v(\mathbf{r},\mathbf{r}') G(\mathbf{r}',\mathbf{r}',t,t^+). \tag{2.55}$$

In this case ϕ is the a perturbation potential acting on the system. The term $iG(\mathbf{r}', \mathbf{r}', t, t^+)$ is the same as the single particle density $n(\mathbf{r}')$, as the arguments of the Green's function coincide.

The inverse dielectric function is then defined as the variation of the potential with respect to the perturbation potential: [20]

$$\epsilon^{-1}(\mathbf{r}, \mathbf{r}', t, t') = \frac{\delta V(\mathbf{r}, t)}{\delta \phi(\mathbf{r}', t')}. \tag{2.56}$$

Putting both equations together we obtain

$$\epsilon^{-1}(\mathbf{r}, \mathbf{r}', t, t') = \delta(\mathbf{r} - \mathbf{r}')\delta(t - t') + \int d\mathbf{r}'' v(\mathbf{r}, \mathbf{r}') \frac{\delta n(\mathbf{r}'', t)}{\delta \phi(\mathbf{r}', t')} \tag{2.57}$$

this form for the inverse dielectric function. [20] It contains the variation of the charge density with respect to the perturbation. Since there is a change in the charge density, the coulomb interaction between two points is also changed. Therefore, we obtain the changed Coulomb interaction W also called the screened Coulomb interaction:

$$W(\mathbf{r}, \mathbf{r}', t, t') = \int d\mathbf{r}'' dt'' v(\mathbf{r}, \mathbf{r}'')\delta(t - t'')\epsilon^{-1}(\mathbf{r}', \mathbf{r}'', t', t''). \tag{2.58}$$

2.5.3 The Self-Energy

Using this knowledge about the Green's function and the screened Coulomb interaction, one can derive a self-consistent system of five equations, called Hedin's equations. [17] For these equations a different notations is used to minimize the number of arguments of each functions. We introduce $1 = (\mathbf{r}, t)$, $2 = (\mathbf{r}', t')$ and so on. Further the polarizability:

$$P(1, 2) = \frac{\delta n(2)}{\delta V(1)} \tag{2.59}$$

and the vortex function:

$$\Gamma(1, 2; 3) = \frac{\delta G^{-1}(1, 2)}{\delta V(3)} \tag{2.60}$$

are needed for Hedin's equations. These equations will however not be derived here as the focus lies on the alternative approach presented in chapter 6. Combining all five quantities yields: [17]

$$\Sigma(1, 2) = i \int d4d3W(1^+, 4)G(1, 3)\Gamma(3, 2; 4) \tag{2.61}$$

$$W(1, 2) = \int d3\epsilon^{-1}(1, 3)v(3, 2) \tag{2.62}$$

$$\epsilon(1, 2) = \delta(1, 2) - \int d3v(1, 2)P(3, 2) \tag{2.63}$$

$$P(1,2) = -i \int d4d3G(1,3)\Gamma(3,4;2)G(4,1^+) \tag{2.64}$$

$$\Gamma(1,2;3) = \delta(1,2)\delta(1,3) + \int d4d5d6d7 \frac{\delta\Sigma(1,2)}{\delta G(4,5)} G(4,6)G(7,5)\Gamma(6,7,3) \tag{2.65}$$

These equations can be solved self-consistently by starting with a guess and then solving the equations iteratively to calculate the self-energy of the system. The self-energy can then be used to calculate the quasi particle correction of the DFT eigenvalues and eigenfunctions. The quasi particle energy is obtained by solving the fixed point problem: [37]

$$\Sigma_n(\omega^*) + \varepsilon_n^{KS} - V_n^{xc} = \omega^*. \tag{2.66}$$

The quasi particle energy is the energy needed to remove an electron from a specific orbital and are thus very important properties of a molecule when it comes to the development of solar cells. [37]

In practice solving Hedin's equation is computationally very demanding. To perform a single iteration of Hedin's equation we start by applying the GW approximation that approximates the vertex function in a very simple way:

$$\Gamma(1,2;3) \approx \delta(1,2)\delta(1,3). \tag{2.67}$$

This is turn reduces the calculation of the polarizability to:

$$P(1,2) = -iG(1,2)G(2,1^+). \tag{2.68}$$

From the polarizability the dielectric functions can be calculated which then can be used to calcualte the screened Coulomb interaction. Using again the GW calculation, the self-energy is then obtained as:

$$\Sigma(1,2) = iG(1,2)W(1^+,2). \tag{2.69}$$

When starting from DFT calculations the polarizability is calculated using the Adler-Wiser expression: [38] [39]

$$P(\mathbf{r},\mathbf{r}';\omega) = 2\sum_{nm} \frac{f_n - f_m}{\epsilon_n - \epsilon_m - \omega} \phi_m^*(\mathbf{r})\phi_n(\mathbf{r})\phi_m(\mathbf{r}')\phi_n^*(\mathbf{r}'). \tag{2.70}$$

In this case, f_n is one if n is an occupied state, and zero otherwise. Here the problem with this method of calculating the self energy mentioned in the Introduction is visible. It is necessary to sum over all possible combinations of occupied and unoccupied states, n and m. This can amount to a lot of terms needing to be computed. The dielectric function can then be calculated as seen before:

$$\epsilon(\mathbf{r},\mathbf{r}';\omega) = \delta(\mathbf{r},\mathbf{r}') - v(\mathbf{r},\mathbf{r}')P(\mathbf{r},\mathbf{r}';\omega) \tag{2.71}$$

and from that the screened Coulomb interaction:

$$W(\mathbf{r},\mathbf{r}';\omega) = \epsilon^{-1}(\mathbf{r},\mathbf{r}';\omega)v(\mathbf{r},\mathbf{r}'). \tag{2.72}$$

The Green's function is obtained as:

$$G(\mathbf{r}, \mathbf{r}'; \omega) = \sum_n \frac{\phi_n(\mathbf{r})\phi_n^*(\mathbf{r})}{\epsilon_n - \omega \pm i\eta}, \tag{2.73}$$

where η is an infinitesimal. Finally, combining the Green's function and the screened Coulomb potential, the self energy is calculated as:

$$\Sigma(\mathbf{r}, \mathbf{r}'; \omega) = i \int G(\mathbf{r}, \mathbf{r}'; \omega - \omega') W(\mathbf{r}, \mathbf{r}'; \omega') e^{i\omega'\delta} d\omega'. \tag{2.74}$$

The self energy obtained through a single iteration of Hedin's equations like this is known as the G_0W_0.

As can be seen from equation 2.70 this calculation of the self-energy scales badly with the number of unoccupied states. Thus, in this thesis an alternative approach to calculate the screened Coulomb potential is presented that avoids the need of including unoccupied states. Therefore, the Sternheimer equation is introduced and derived in the next chapter.

Chapter 3

The Sternheimer Equation in Linear Perturbation Theory

3.1 The Sternheimer equation

In this chapter the Sternheimer equation will be derived using linear perturbation theory, and it is explained how perturbation theory is linked to screened Coulomb interaction and, consequently, to the self energy. The Sternheimer equation:

$$(\hat{H} - \varepsilon_i) \left| \psi_i^{(1)} \right\rangle = -\hat{P}_C V_{\text{ext}} \left| \psi_i \right\rangle \tag{3.1}$$

is a result from linear perturbation theory and was first used by RM Sternheimer in 1954. [2] [40] In this chapter the equation will be derived step by step and all symbols and operators will be explained. We start from the time independent Schrödinger equation:

$$\hat{H} \left| \psi_i \right\rangle = \varepsilon_i \left| \psi_i \right\rangle . \tag{3.2}$$

The equation is perturbed by adding a linear term in the following way:

$$\hat{H}(\lambda) \left| \psi_i(\lambda) \right\rangle = \varepsilon_i(\lambda) \left| \psi_i(\lambda) \right\rangle \tag{3.3}$$

with

$$\hat{H}(\lambda) = \hat{H} + \lambda \hat{H}^{(1)} \tag{3.4}$$

$$\psi(\lambda) = \psi + \lambda \psi^{(1)} \tag{3.5}$$

$$\varepsilon(\lambda) = \varepsilon + \lambda \varepsilon^{(1)} \tag{3.6}$$

and λ being the strength of the perturbation. The perturbation to the Hamiltonian $\hat{H}^{(0)}$ can be interpreted as an external potential V_{ext} acting on the system. Substituting this into equation 3.3 yields:

$$\left(\hat{H} + \lambda \hat{H}^{(1)} \right) \left| \psi_i + \lambda \psi^{(1)} \right\rangle = \left(\varepsilon_i + \lambda \varepsilon_i^{(1)} \right) \left| \psi_i + \lambda \psi^{(1)} \right\rangle \tag{3.7}$$

Multiplication expands the equation to:

$$\hat{H}\left|\psi_i\right\rangle + \lambda\hat{H}\left|\psi_i^{(1)}\right\rangle + \lambda\hat{H}^{(1)}\left|\psi_i\right\rangle + \lambda^2\hat{H}^{(1)}\left|\psi_i^{(1)}\right\rangle = \tag{3.8}$$
$$\varepsilon_i\left|\psi_i\right\rangle + \lambda\varepsilon_i\left|\psi_i^{(1)}\right\rangle + \lambda\varepsilon_i^{(1)}\left|\psi_i\right\rangle + \lambda^2\varepsilon_i^{(0)}\left|\psi_i^{(1)}\right\rangle.$$

To simplify this equation, we choose to only take the linear response of the system into account. We therefore neglect the terms quadratic in λ. This is known as the linear approximation which has proven to be a reasonable assumption. [2]

When we subtract the time independent Schrödinger equation 3.2, we end up with the terms linear in λ:

$$\hat{H}\left|\psi_i^{(1)}\right\rangle + \hat{H}^{(1)}\left|\psi_i\right\rangle = \varepsilon_i\left|\psi_i^{(1)}\right\rangle + \varepsilon_i^{(1)}\left|\psi_i\right\rangle. \tag{3.9}$$

Since λ is not zero because this case would just be the unperturbed system, we can divide it out. Factoring out the wave functions leads to:

$$(\hat{H} - \varepsilon_i)\left|\psi_i^{(1)}\right\rangle = -(\hat{H}^{(1)} - \varepsilon_i^{(1)})\left|\psi_i\right\rangle. \tag{3.10}$$

This looks already similar to the Sternheimer equation

$$(\hat{H} - \varepsilon_i)\left|\psi_i^{(1)}\right\rangle = -\hat{P}_C\hat{H}^{(1)}\left|\psi_i\right\rangle, \tag{3.11}$$

where the the operators

$$\hat{P}_{\text{occ}} = \hat{P}_C = \sum_{n\in\text{occ}}\left|\psi_n\right\rangle\left\langle\psi_n\right| \tag{3.12}$$

and

$$\hat{P}_{\text{unocc}} = \hat{P}_V = \sum_{n\in\text{unocc}}\left|\psi_n\right\rangle\left\langle\psi_n\right| \tag{3.13}$$

are the projection operators of the occupied states and the unoccupied states respectively. They satisfy the following equation because of the orthonormality of the wave functions:

$$1 = \hat{P}_V + \hat{P}_C. \tag{3.14}$$

The right hand side of the Sternheimer equation can therefore be expanded to:

$$(-1 + \hat{P}_V)\hat{H}^{(1)}\left|\psi_i\right\rangle = (-\hat{H}^{(1)} + \hat{P}_V\hat{H}^{(1)})\left|\psi_i\right\rangle. \tag{3.15}$$

Using the expansion from equation 3.13 we obtain

$$(-1 + \hat{P}_V)\hat{H}^{(1)}\left|\psi_i\right\rangle = -\hat{H}^{(1)}\left|\psi_i\right\rangle + \sum_{n\in\text{unocc}}\left|\psi_n\right\rangle\left\langle\psi_n\right|\hat{H}^{(1)}\left|\psi_i\right\rangle. \tag{3.16}$$

The orthonormality constrains of the wave functions yield the first order eigenvalues of the first order correction of the Hamiltonian:

$$\varepsilon_i^{(1)} = \left\langle\psi_i\right|\hat{H}^{(1)}\left|\psi_i\right\rangle. \tag{3.17}$$

Therefore, we obtain:

$$-\hat{P}_C\hat{H}^{(1)}\left|\psi_i\right\rangle = -\hat{H}^{(1)}\left|\psi_i\right\rangle + \sum_{n\in\text{unocc}} \varepsilon_i^{(1)}\left|\psi_n\right\rangle \delta_{ni}. \tag{3.18}$$

Which then simplifies to:

$$-\left(\hat{H}^{(1)} - \varepsilon_i^{(1)}\right)\left|\psi_i\right\rangle. \tag{3.19}$$

That is exactly the right hand side of equation 3.10 with $\hat{H}^{(1)} = \hat{V}_{\text{ext}}$. So the Sternheimer equation was derived from linear perturbation theory.

To account for time dependent perturbations, the Sternheimer equation is extended using the ansatz: [41]

$$\psi_v^{(1)}(\mathbf{r}) \longrightarrow e^{\pm i\omega t}\psi_v^{(1)}(\mathbf{r}, \pm\omega) \tag{3.20}$$

and

$$\hat{V}_{\text{ext}}(\mathbf{r}) \longrightarrow \hat{V}_{\text{ext}}(\mathbf{r}, \pm\omega), \tag{3.21}$$

in combination with Fourier transformation, introducing the time t and frequency ω. Inserting these equations into the time dependent Schrödinger equation:

$$i\frac{\partial\psi_n}{\partial t} = \hat{H}\psi_n, \tag{3.22}$$

we obtain

$$(\hat{H} - \varepsilon_n \pm \omega)\psi_n^{(1)}(\mathbf{r}, \pm\omega) = -\hat{P}_C\hat{V}_{\text{ext}}(\mathbf{r}, \pm\omega)\psi_n^{(0)}(\mathbf{r}) \tag{3.23}$$

for all occupied states n. The Sternheimer equation can now be used to calculate the response of the wave function and therefore the electron density of a given system to an external perturbation potential that depends on a frequency ω. In the context of density functional theory an interesting property to observe is the electron density of the perturbed system. It can be calculated from the perturbed wave function analogously to unperturbed density in equation 2.36:

$$\Delta n(\mathbf{r}) = N\int d\mathbf{r}_2 \cdots \int d\mathbf{r}_N\, \Delta\Phi^*(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N)\Phi(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N) \tag{3.24}$$

Because the change of the electron density is in itself a perturbation of the system, the Sternheimer equation needs to be solved self-consistently. The perturbation potential is updated using the response of the electron density and the equation is solved again using the new perturbation. The details of this process are explained in the implementation part 4 of this thesis.

The Sternheimer equation has a lot of practical applications. [2] [42] It allows the calculation of the response of the system to an external perturbation, under the assumption that the linear approximation mentioned above holds. Depending on the chosen perturbation it is possible to calculate various properties. As mentioned in the previous chapter, the Sternheimer equation can be used to calculate the self energy of a system. Depending on the initial perturbation chosen, other quantities such as the polarizability or the inter atomic forces can

be calculated from the linear response of the electron density calculated via the Sternheimer equation. Further, the solutions of the Sternheimer equation depend only on the occupied states of the quantum mechanical system. This is a great advantage of this strategy compared to the approach outlined in chapter 2. As mentioned in the Introduction the sum over all state combinations in the Adler-Wiser expression for the polarizability is a major bottleneck in the standard GW approach. Due to its construction, the Sternheimer equation does not need unoccupied states and thus has the potential to reduce the computational workload. [1]

For practical calculations, the Sternheimer equation can be very ill conditioned. It is therefore favorable to add αP_V to the Hamiltonian on the left hand side. [1] This does not change the solution since it only depends on the occupied states. Adding the projection operator onto the unoccupied states therefore does not effect the results but leads to a better conditioned left hand side of the equation.

This equation can now be used to formulate an algorithm that allows the calculation of various properties of functional perturbation theory.

3.2 Analogies to Density Functional Perturbation Theory

In this section the perturbation theory used before to derive the Sternheimer equation will be put into the context of density functional theory and the GW approximation explained in the previous chapter.

Density functional perturbation theory, or short DFTP, is, as the name suggests, an extension of classic DFT. The theory looks at how the electron density described by DFT responds to external potentials. In most cases only the linear response of the electron density is accounted for. We now investigate the case in which the perturbation is a point test charge introduced into the system. In the dielectric approach the total potential experienced by an external test charge is expressed in term of the external perturbation using the inverse dielectric matrix $\epsilon^{-1}(\mathbf{r}, \mathbf{r}')$: [43] [2]

$$\Delta V_{\text{test}}(\mathbf{r}) = \int d\mathbf{r}' \epsilon^{-1}(\mathbf{r}, \mathbf{r}') \Delta V(\mathbf{r}'). \tag{3.25}$$

Alternatively, the response of the electron density can be expressed through the electron polarizability $\chi(\mathbf{r}, \mathbf{r}')$:

$$\Delta n(\mathbf{r}) = \int d\mathbf{r}' \chi(\mathbf{r}, \mathbf{r}') \Delta V(\mathbf{r}'). \tag{3.26}$$

When looking at DFT, the independent-electron polarizability $\chi_0(\mathbf{r}, \mathbf{r}')$ can be defined as the electron density response to the a perturbation of the Kohn-Sham potential:

$$\Delta n(\mathbf{r}) = \int d\mathbf{r}' \chi_0(\mathbf{r}, \mathbf{r}') \Delta V_{\text{KS}}(\mathbf{r}'). \tag{3.27}$$

The independent-electron polarizability can be expressed in terms of the Kohn-Sham wave functions:

$$\chi_0(\mathbf{r},\mathbf{r}') = \sum_{n,m} \frac{f_n - f_m}{\epsilon_n \epsilon_m} \phi_n^*(\mathbf{r})\phi_m(\mathbf{r})\phi_m^*(\mathbf{r}')\phi_n(\mathbf{r}'), \tag{3.28}$$

with the summation going over all occupied states n and unoccupied states m and f_n being the occupancy of a state. The variation of the Kohn-Sham potential can be linked to external perturbation via the kernel:

$$K(\mathbf{r},\mathbf{r}') = \frac{e^2}{|\mathbf{r}-\mathbf{r}'|} + \frac{\delta^2 E_{xc}}{\delta n(\mathbf{r})\delta n(\mathbf{r}')}, \tag{3.29}$$

which expresses the potential that acts on the electron density when redistributing in response to a point charge. The first term is the classical Coulomb potential, the second term corresponds to quantum effects. The Kohn-Sham variation is thus:

$$\Delta V_{\mathrm{KS}}(\mathbf{r}') = \Delta V(\mathbf{r}') \int d\mathbf{r}' K(\mathbf{r},\mathbf{r}')\Delta n(\mathbf{r}'). \tag{3.30}$$

From this result the electron polarizability can be calculated as:

$$\chi(\mathbf{r},\mathbf{r}') = \chi_0(\mathbf{r},\mathbf{r}') + \int d\mathbf{r}_1 d\mathbf{r}_2 \chi_0(\mathbf{r},\mathbf{r}_1)K(\mathbf{r}_1,\mathbf{r}_2)\chi(\mathbf{r}_2,\mathbf{r}'). \tag{3.31}$$

The inverse dielectric matrix can be obtained from the electron polarizability with:

$$\epsilon^{-1}(\mathbf{r},\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}') - \int d\mathbf{r}_1 \frac{e^2}{|\mathbf{r}-\mathbf{r}_1|}\chi(\mathbf{r}_1,\mathbf{r}). \tag{3.32}$$

The approach is in many ways equivalent to the Sternheimer approach. Its advantage is that all relaxation processes introduced by the point charge can be expressed through the dielectric matrix. The drawback is that this method is only valid for point charges and not for non-local perturbation potentials. The Sternheimer method is more versatile as it can deal with various different perturbations, as will be explained in the next chapter.

An additional advantage of the Sternheimer method can be seen when looking at the calculation of the independent-electron polarizability. Here all states, occupied and unoccupied, need to be treated. The Sternheimer method avoids this and uses occupied states only.

Chapter 4

Implementation

Before discussing the implementation of the Sternheimer algorithm, we need to introduce a numerical solution strategy for the Kohn-Sham density functional theory because the starting point for Sternheimer calculations are the results from DFT calculations like the Kohn-Sham wave functions and the Kohn-Sham Hamilitonian. In practice often basis functions are used as a way to represent the wave functions introduced by the Kohn-Sham approach of DFT.

4.1 Basis Functions

To approximate the wave functions in DFT calculations commonly a set of basis functions called a basis set is used. The wave functions are expressed through a linear combination of basis functions:

$$\psi_i(\mathbf{r}) = \sum_{j=1}^{M} \mathbf{X}_{ij}\chi_j(\mathbf{r}) \tag{4.1}$$

Where M is the number of functions $\chi_j(\mathbf{r})$ in the used basis set and $\mathbf{X}$ the corresponding coefficient matrix. There are different approaches to which types of function to use in a basis set. [10]

One option is to use plane waves of form:

$$\chi(\mathbf{r}) = \sum_{\mathbf{k}} u_{\mathbf{k}} e^{i\mathbf{kr}}. \tag{4.2}$$

These sets work very well for periodic systems such as crystal structures since the basis functions themselves are periodic. Further some very common integrals over the real space can be treated efficiently using fast Fourier transform. For non-periodic systems however, plane wave basis sets are a poor choice because of their periodic character. To avoid self-interaction, large boxes need to be used leading to a large space on where the plane waves need to be evaluated on, which is very inefficient. [10] That is why in this thesis localized basis sets are used.

Slater orbitals are another type of basis functions:

$$\chi(\mathbf{r}) = \mathbf{r}^n e^{-\alpha|\mathbf{r}|}. \tag{4.3}$$

They approximate the form of atomic orbital wave functions really well since the have a sharp peak at $\mathbf{r} = 0$ and they decay with $\exp\{-|\mathbf{r}|\}$. That is because Slater type orbitals are the analytic solution of the Hydrogen atom Schrödinger equation. [44] Therefore, Slater-type orbitals are physically very well founded as they originate from physical results. They require however numerical integration because they are not differentiable for $\mathbf{r} = 0$.

A compromise for this are Gaussian basis sets:

$$\chi(\mathbf{r}) = \mathbf{r}^n e^{-\beta\mathbf{r}^2}. \tag{4.4}$$

Gaussian basis functions are cheaper to integrate than Slater orbitals, but do not reproduce the properties of the wave functions as well. [45] [46] Another advantage of Gaussian basis functions is that the linear combination of Gaussian basis functions is itself a Gaussian function. Therefore, the functions used in the implementation are all of the same type.

Additionally several integrals that are very common in DFT implementations, like overlap integrals or kinetic energy term, can be computed analytically when using Gaussian basis functions. [47] This reduces the computation time significantly, as no costly numerical integration has to be performed.

The software package votca-xtp which was used for this thesis uses Gaussian basis sets. Therefore, all calculation in this thesis use them as well. In the following some key properties of Gaussian basis functions are explained.

4.1.1 Properties of Gaussian Basis Functions

An alternative way to formulate a Gaussian basis function is as Cartesian Gaussians:

$$G_{ikl}(a, \mathbf{r}, \mathbf{A}) = x_A^i y_A^j z_A^k e^{-a\mathbf{r}_A^2}. \tag{4.5}$$

Where $\mathbf{r}_A = \mathbf{r} - \mathbf{A}$ with $\mathbf{A}$ being the center of the Gaussian. [48]

The goal is to represent the different electron orbitals of an atom with Gaussian basis functions. To account for different types of atomic orbitals, the exponents are chosen independently for each Cartesian coordinate. For example, an s orbital is represented by G_{000}, the three p orbitals, one in each direction, are represented by G_{100}, G_{010} and G_{001} respectively. Due to this construction of the Gaussian basis functions, they can be split into the contribution of each direction: [46]

$$G_{ikl}(a, \mathbf{r}_A) = G_i(a, x_A)G_j(a, y_A)G_k(a, z_A) \tag{4.6}$$

with

$$G_i(a, \mu_A) = \mu_A^i e^{-a\mu_A^2},\ \mu = x, y, z. \tag{4.7}$$

Therefore, we can treat each contribution separately and put them back together in the end.

Differentiating such a Gaussian results in a linear combination of two Gaussians:

$$\frac{\partial G_i(a,\mu_A)}{\partial \mu} = iG_{i-1}(a,\mu_A) - 2aG_{i+1}(a,\mu_A) \tag{4.8}$$

Another important property is, as mentioned above, that the product of two Gaussian functions is another Gaussian:

$$e^{-a\mu_A^2}e^{-b\mu_B^2} = e^{-\alpha X_{AB}^2}e^{-p\mu_P^2}. \tag{4.9}$$

With $e^{-\alpha X_{AB}^2}$ being a prefactor independent of μ and $e^{-p\mu_P^2}$ being the product Gaussian with the exponential factor $p = a + b$ and center $P_\mu = \frac{aA_\mu + bB_\mu}{p}$. The exponent of the prefactor can be given in term of the Gaussian factors as $\alpha = \frac{ab}{a+b}$ and $X_{AB} = A_\mu - B\mu$. The center of the product Gaussian lies between the centers of the factor Gaussians. The cost of an integral over Gaussian basis functions can be reduced significantly by contracting a product of Gaussians into a single Gaussian. [48]

4.2 Implementation of Kohn-Sham Density Functional Theory

The foundations of every Sternheimer calculation are the results of an underlying density functional theory calculation. To build the Sternheimer equation, the Kohn-Sham Hamiltonian and wave functions are needed. That is why in this section a summary of how the Kohn-Sham approach to density functional theory is implemented is presented.

As explained in section 2.4 the DFT workflow is done in a self consistent manner. The computation starts with an initial guess for the Kohn-Sham orbitals. One way is, as discussed, using non-interacting particle orbitals as a starting point for the DFT calculation. The orbitals are obtained by solving the non-interacting Schrödinger equation.

The orbitals are used to set up the overlap matrix:

$$\mathbf{S}_{ij} = \langle \chi_i | \chi_j \rangle \,, \tag{4.10}$$

the non-interacting Hamiltonian matrix $\mathbf{H}_n$ and the initial guess for the electron density matrix:

$$\mathbf{D}_{kj} = \sum_{i=1}^{N} \mathbf{X}_{ik}\mathbf{X}_{ij}. \tag{4.11}$$

$\mathbf{X}$ is the coefficient matrix that constructs the orbitals from the basis functions as shown in equation 4.1.

From the density matrix the exchange-correlation matrix:

$$\mathbf{V}_{\mathrm{xc},ij} = \langle \chi_i | \, \hat{V}_{\mathrm{xc}} \, | \chi_j \rangle \tag{4.12}$$

that depends on the exchange-correlation functional chosen and the electron repulsion matrix:

$$\mathbf{V}_{\mathrm{H},ij} = \langle \chi_i | \hat{V}_{\mathrm{H}} | \chi_j \rangle = \int \int d\mathbf{r} d\mathbf{r}' \chi_i(\mathbf{r}) \chi_j(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}'|} n(\mathbf{r}'). \tag{4.13}$$

are calculated. The integral used to calculate the electron repulsion matrix is a very complicated object to deal with. It can be written as:

$$\langle \chi_i | \hat{V}_{\mathrm{H}} | \chi_j \rangle = \sum_{kl} \mathbf{D}_{kl} (ij|kl). \tag{4.14}$$

with

$$(ij|kl) = \int d\mathbf{r} d\mathbf{r}' \ \frac{\chi_i(\mathbf{r}) \chi_j(\mathbf{r}) \chi_l(\mathbf{r}') \chi_l(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{4.15}$$

being called a four center integral. A very similar integral needs to be evaluated in the Sternheimer algorithm. That is why an efficient approximation of the integral can be found in section 4.3.12. Once all matrices needed to build the Kohn-Sham Hamiltonian:

$$\mathbf{H}_{\mathrm{KS}} = \mathbf{H}_{\mathrm{n}} + \mathbf{V}_{\mathrm{xc}} + \mathbf{V}_{\mathrm{ee}} \tag{4.16}$$

are constructed, the Kohn-Sham equation in matrix form is obtained as:

$$\mathbf{H}_{\mathrm{KS}} \mathbf{X}_i = \epsilon_i \mathbf{S} \mathbf{X}_i. \tag{4.17}$$

Diagonalizing the Kohn-Sham Hamiltonian yields the Kohn-Sham energies ϵ_i and the coefficient Matrix $\mathbf{X}$ that can be used to construct the new Kohn-Sham wave functions. They can in turn be used to construct a new electron density matrix that is used to construct a new Kohn-Sham Hamiltonian. This cycle is repeated until convergence.

This section only scratches the surface of DFT implementations. There are a lot of ways to optimize the convergence properties of the self-consistent cycle and the computation time of the integrals. However, it should give an overview over how the matrices used in the next section are constructed.

In the next section the self-consistent Sternheimer algorithm will be presented. It builds directly on the results of the DFT calculations as the Kohn-Sham Hamiltonian, the wave functions, the density matrix as well as the overlap matrix and the coefficients are needed to construct the Sternheimer equation.

4.3 The Self-Consistent Sternheimer Method

As discussed in the previous chapter 3, the Sternheimer equation can be used to calculate different properties, such as the polarizability or the screened Coulomb potential in the context of density functional perturbation theory. In this section the algorithm used to compute these properties is presented. Also several of the important topics outlined in the Introduction are addressed. The algorithm is

commonly referred to as the self consistent Sternheimer method. In this section the method is explained step by step. Further some additional algorithms are presented that are required to perform the procedure or provide some sort of improvement. This chapter mostly focuses on the calculation of the polarizability tensor [3] as an example on how the Sternheimer method can be used, but changes needed for other computations are noted when they arise. It is noted here that computing the polarizablity tensor does not involve an integration over space or frequency. Therefore, issues arising from numerical integrations will be addressed in chapter 6 where the Sternheimer *GW* method is discussed.

Before discussing the algorithm itself the Sternheimer equation needs to be transformed into matrix form. [19] We start by expanding the wave function in the basis functions:

$$\psi_v(\mathbf{r}') = \sum_{i\in\text{occ}} = c_{vi}\chi(\mathbf{r}') \tag{4.18}$$

$$\psi^{(1)\pm}_{v[\omega]}(\mathbf{r}') = \Delta\psi^{\pm}_v(\mathbf{r}') = \sum_{i\in\text{occ}} \Delta c^{\pm}_{vi[\omega]}\chi(\mathbf{r}'). \tag{4.19}$$

Inserting this into the dynamic Sternheimer equation 3.23 and projecting both sides onto the basis functions we obtain:

$$\sum_{ij} \langle\phi_i|\, \hat{H} - \omega\hat{1}\, |\phi_j\rangle\, \Delta c^v_i = -\sum_{nm} \langle\phi_n|\, \hat{1} - \hat{P}_{occ}\Delta V\, |\phi_m\rangle\,. \tag{4.20}$$

This simplifies to

$$(\mathbf{H} - (\varepsilon_v \pm \omega)\mathbf{S})\Delta c^{\pm}_{v[\omega]} = -(\mathbf{1} - \mathbf{S}\boldsymbol{\rho}^T)\Delta\mathbf{V}_{[\omega]}c_v \tag{4.21}$$

the matrix formulation of the Sternheimer equation. [19] With $\mathbf{S}$ being the orbital overlap matrix, $\boldsymbol{\rho}$ being the density matrix, $\mathbf{H}$ the Hamiltonian matrix, $\Delta\mathbf{V}$ the perturbation potential and c_v the coefficients of the respective wave functions.

The Sternheimer method starts with a standard DFT calculation to obtain the Kohn-Sham wave functions and the Kohn-Sham Hamiltionian of the system. We can then use the Matrix formulation of the Sternheimer equation from above:

$$[\mathbf{H} - (\varepsilon_v \pm \omega)\,\mathbf{S}]\Delta\mathbf{c}^{\pm}_v = -\left[\mathbf{1} - \mathbf{S}\boldsymbol{\rho}^T\right]\Delta\mathbf{V}\mathbf{c}_v. \tag{4.22}$$

The results of the DFT calculation allow the construction of the Sternheimer equation. Inserting the Kohn-Sham Hamiltonian $\mathbf{H}$, the overlap matrix $\mathbf{S}$, the density matrix $\boldsymbol{\rho}$, the Kohn-Sham coefficients c_v and eigenvalues ϵ_v, the only pieces missing are the perturbation matrix that is chosen according to the wanted quantity and the frequency ω. [19] Consequently, the self-consistent Sternheimer method can now be set up.

4.3.1 Frequency Grid

The Sternheimer equation can be solved for different frequencies. Often the property is wanted as a function of the frequency, so the Sternheimer equation

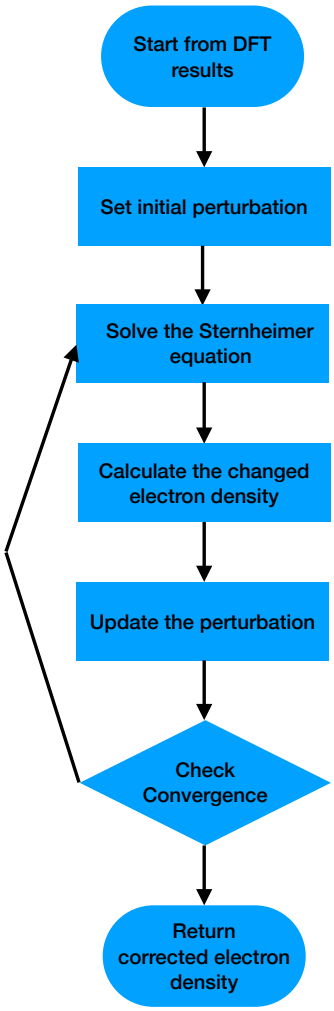

Figure 4.1: This flowchart describes the workflow of the self-consistent Sternheimer equation. In the following, each step is explained in detail.

needs to be evaluated on a grid of frequencies. Since the matrix on the left hand side of the equation can be ill conditioned for real frequencies, the Sternheimer equation is solved for complex frequencies of form $\omega_{\text{real}} + i\kappa$. [3] For values of κ between 3-5eV, the system is better conditioned and converges faster. [19] The approximation of the values on the real axis is performed using analytic continuation, in this case the Padé approximation. [49] How many frequencies are evaluated depends on the desired quantity. In the following chapter 5 an analysis of different frequency grids can be found.

4.3.2 Perturbation

Depending on what property needs to be calculated, the initial perturbation $\Delta\mathbf{V}$ needs to be chosen accordingly.

It is possible to use a variety of different perturbation potentials and calculate the response of the system. For polarizability calculations the perturbation is initialized as an electric field $V_{\text{in}} = \mathbf{r}\cdot\mathbf{E}(\omega)$. [3] For GW calculations the potential is chosen as the bare Coulomb potential $V_{\text{in}} = v(\mathbf{r}, \mathbf{r}')$. [1] It is explained later how to calculate the stated properties from the solution of the Sternheimer equation.

The Matrix values of the perturbation are determined by expansion in the basis functions: [19]

$$V_{\text{in},ij[\omega]} = \int d\mathbf{r}\phi_i(\mathbf{r})V_{\text{ext}}(\mathbf{r},\omega)\phi_j(\mathbf{r}). \tag{4.23}$$

4.3.3 Spatial Grid

Depending on the perturbation chosen, the Sternheimer equation sometimes needs to be solved on a spatial grid. For the polarizability tensor this is not the case since the electric field does not depend on $\mathbf{r}$. The solution only depends on ω:

$$[\mathbf{H} - (\varepsilon_v \pm \omega)\,\mathbf{S}]\Delta\mathbf{c}^{\pm}_{v[\omega]} = -\left[\mathbf{1} - \mathbf{S}\boldsymbol{\rho}^T\right]\Delta\mathbf{V}\mathbf{c}_v. \tag{4.24}$$

For GW calculations on the other hand, the perturbation directly depends on $\mathbf{r}$. [19] Therefore, the solution of the Sternheimer equation depends on $\mathbf{r}$ as well:

$$[\mathbf{H} - (\varepsilon_v \pm \omega)\,\mathbf{S}]\Delta\mathbf{c}^{\pm}_{v[\omega,\mathbf{r}]} = -\left[\mathbf{1} - \mathbf{S}\boldsymbol{\rho}^T\right]\Delta\mathbf{V}_{[\mathbf{r}]}\mathbf{c}_v. \tag{4.25}$$

The Sternheimer equation, in that case, has to be solved on a spatial grid to obtain the response of the wave function for each point in space. The details on how the grid can be set up and how the spatial integration is implemented can be found in chapter 6 because for now the focus lies on the polarizability calculation.

4.3.4 Solving the Sternheimer Equation

Once the initial perturbation potential is known, the Sternheimer equation can be solved for $\Delta c_v(\mathbf{r}, \omega)$. While different solvers can be used to compute the solu-

tions of the Sternheimer equation, the complex bi conjugate gradient algorithm has proven to be a good method for this task. [1] [50] It can also be used to initialize a multi shift solver that is introduced later on.

For small real parts of ω the right hand side of the Sternheimer equation is changed to $(\mathbf{H} + \alpha\mathbf{S}\boldsymbol{\rho}) - (\epsilon \pm \omega)\mathbf{S}$. [1] This is the matrix formulation of adding the projection operator introduced in the theory section. While this does not effect the solution, the system is better conditioned using this left hand side.

4.3.5 Calculating the Change in Electron Density

The solution of the Sternheimer equation can be used to calculate the change in the Kohn-Sham electron density Δn caused by the perturbation potential. This change is a result of the electrons interacting with the perturbation potential.

The following formula is used to determine the reaction of the electron density: [1]

$$\Delta n_{ij}(\omega) = 2 \sum_{v,\sigma=\pm} c_{v,i} \Delta c^{\sigma}_{v,j}(\omega). \tag{4.26}$$

The reaction of the electron density is calculated from the coefficients of the KS wave function and the first order correction of the coefficients calculated from the Sternheimer equation. This is an important step because from the response of the electron density, the perturbation potential can be updated. [1]

4.3.6 Updating the Perturbation Potential

The change of the electron density is in itself a perturbation to the original system, that has to be taken into account. That is why the Sternheimer equation is solved in a self-consistent manner. To include the potential induced by the change of the electron density the Hartree and exchange-correlation potentials are calculated and added to the perturbation potential: [3]

$$\Delta V_{ij}(\omega) = V_{\mathrm{ext},ij[\omega]} + V_{\mathrm{H}(ij)}(\omega) = V_{\mathrm{ext},ij[\omega]} + \sum_{kl} (ij|kl)_{\mathrm{H,xc}} \Delta n_{kl}(\omega). \tag{4.27}$$

With

$$(ij|kl)_{\mathrm{H,xc}} = \int\int d\mathbf{r} d\mathbf{r}' \phi_i(\mathbf{r}) \phi_j(\mathbf{r}) (v_{\mathrm{c}}(\mathbf{r}, \mathbf{r}') + f_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}')) \phi_k(\mathbf{r}') \phi_l(\mathbf{r}') \tag{4.28}$$

being a four point integral. In this implementation the integral is split up into the sum two integrals:

$$(ij|kl)_{\mathrm{H}} + (ij|kl)_{\mathrm{xc}} = \int\int d\mathbf{r} d\mathbf{r}' \phi_i(\mathbf{r}) \phi_j(\mathbf{r}) (v_{\mathrm{c}}(\mathbf{r}, \mathbf{r}')) \phi_k(\mathbf{r}') \phi_l(\mathbf{r}') \tag{4.29}$$

$$+ \int\int d\mathbf{r} d\mathbf{r}' \phi_i(\mathbf{r}) \phi_j(\mathbf{r}) (f_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}')) \phi_k(\mathbf{r}') \phi_l(\mathbf{r}') \tag{4.30}$$

In comparison to the Random Phase Approximation, that is often used to calculate the linear screening of electrons, this method takes exchange-correlation

interaction into account. Thus this approach goes beyond what is done in the standard approach and includes additional term leading to increased physical accuracy.

To speed up the computation an auxiliary basis is used to approximate the first integral via three point integrals. This method is explained later on in section 4.3.12. The second integral is evaluated numerically on a spatial grid. To reduce computation time, the second integral can be computed once at the start of the method and then saved. Hence, the integral does not have to be evaluated in every step. This is especially important because the evaluation of the integrals is the most time consuming part of the cycle. As seen in figure 4.2, presaving the integral reduces the computation time for a small basis sets by a factor of more than 20. Although the implementation uses the symmetries of the integral to save memory, for a higher number of basis functions the time saved per cycle is reduced. When exceeding 120 basis functions, it is more efficient to evaluate the integral in every step. Another reason why this solution is not feasible for large basis sets is the amount of memory it requires. The memory needed scales with N^4 where N is the number of basis function. To reduce the computation time for larger molecules, the auxiliary basis approximation mentioned above could be used for the f_{xc} integral as well.

When the integration is complete, the updated perturbation is inserted back into the Sternheimer equation and the equation is solved again. This self consistent cycle is continued until the change of the perturbation potential in one step is below a predefined threshold. In the implementation a history of the last perturbation potentials is stored and compared to the updated potential. If the difference is not below the defined threshold, the perturbation is inserted into the Sternheimer equation again and the cycle is repeated once more.

4.3.7 Convergence Properties

The first problem that needs to be investigated when implementing the self-consistent Sternheimer method using localized basis sets is the convergence properties of the self-consistent cycle.

The self-consistent cycle is not guaranteed to converge. Simply using the newly calculated perturbation potential resulted in the cycle diverging for most test cases. There are several ways to improve the convergence of this fixed point iteration. The crucial point is updating the perturbation potential. Calculating the full four center integral instead of approximating it improves the convergence because even small errors introduced can hinder the convergence of the cycle. In this case calculating the full integral is impractical for larger systems since it scales badly with the system size.

Another method that was tested is moving the frequency grid points away from the real axis. While this improves the convergence, the approximation of the real frequency values via analytic continuation is not accurate enough for this method to be feasible. The error introduced by the large shift of the frequencies is too large. For Benzene, shifts of $\kappa = 25\mathrm{eV}$ are necessary to guarantee convergence. As discussed later, only shifts up to 5eV lead to reasonable

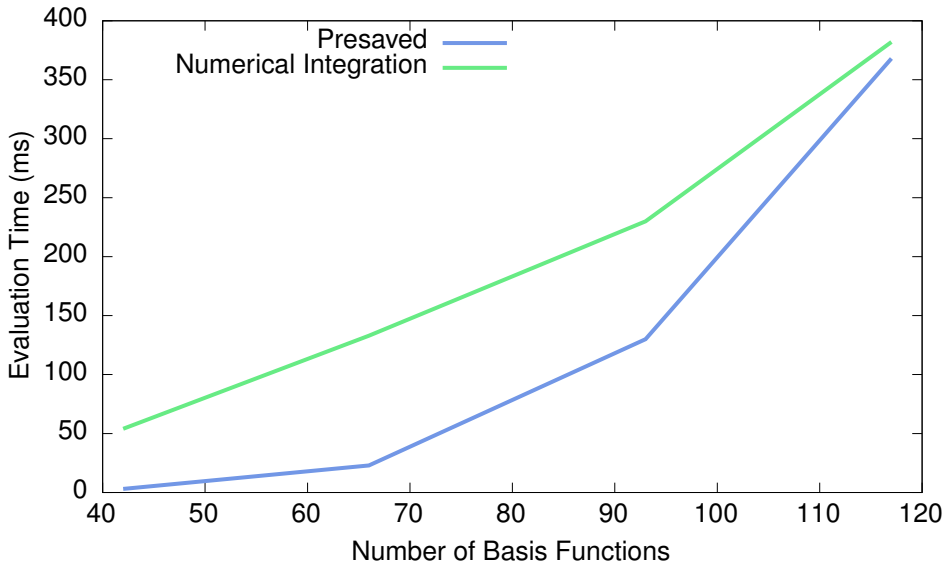

Figure 4.2: The plot shows the computation time needed to evaluate the f_{xc} integral 4.30 and to update the potential, depending on the number of basis functions used. When presaving the integral and updating the potential using the stored integral, less time is needed when using less than 120 basis functions. The scaling with the number of basis functions is worse however for this method. When using more basis functions, computing the integral in every step is more efficient.

results. As a result, a different method is needed to improve the convergence of the self-consistent Sternheimer method, independent of the frequencies chosen.

Instead of blindly using the new potential, a mix of the new potential and potentials calculated in previous steps is used. [51] A very simple method of doing this is choosing the new input potential to be linear combination of the last two potentials:

$$\Delta V_{\text{in}} = \alpha \Delta V_{\text{new}} + (1-\alpha)\Delta V_{\text{old}} \tag{4.31}$$

This simple method already leads to a considerable improvement though it did not lead to convergence in every test case. Another downside of the method is that α has to be chosen according to the system and the frequency ω.

Therefore, a more sophisticated method was implemented. The Anderson method provides a way to mix the last M potentials calculated into a new input potential. [52] Consider

$$\Delta V_{\text{out}}^{(m)} = F[\Delta V_{\text{in}}^{(n)}] \tag{4.32}$$

where F stands for the Sternheimer algorithm described above that calculates the new potential from an old potential. And $\Delta V_{out/in}^{(in)}$ the input and output potentials in the mth iteration of the self-consistent cycle Sternheimer respectively. In the Anderson method the mixing coefficients are chosen in a way that minimizes the quadratic difference between the latest output and inputs potentials.

We define

$$D_m = \Delta V_{\text{out}}^m - \Delta V_{\text{in}}^m, \tag{4.33}$$

the difference between input and output potential in each iteration. Now the last M input and output potentials are mixed separately:

$$\tilde{\Delta V}_{\text{in}} = \Delta V_{\text{in}}^N + \sum_{m=1}^{M} \vartheta \left(\Delta V_{\text{in}}^{N-m} - \Delta V_{\text{in}}^N \right), \tag{4.34}$$

$$\tilde{\Delta V}_{\text{out}} = \Delta V_{\text{out}}^N + \sum_{m=1}^{M} \vartheta \left(\Delta V_{\text{out}}^{N-m} - \Delta V_{\text{out}}^N \right). \tag{4.35}$$

Where N is the latest iteration and the coefficients ϑ are calculated by solving the linear system:

$$A\vartheta = c, \tag{4.36}$$

with

$$A_{mj} = Re \left\langle D_N - D_{N-m} | D_N - D_{N-j} | D_N - D_{N-m} | D_N - D_{N-j} \right\rangle \tag{4.37}$$

and

$$c_m = Re \left\langle D_N - D_{N-m} | D_N | D_N - D_{N-m} | D_N \right\rangle. \tag{4.38}$$

Lastly, the new input potential is the mix of the input and output potentials calculated above:

$$V_{\text{in}}^{N+1} = \alpha \tilde{V}_{\text{out}} + (1-\alpha)\tilde{V}_{\text{in}}. \tag{4.39}$$

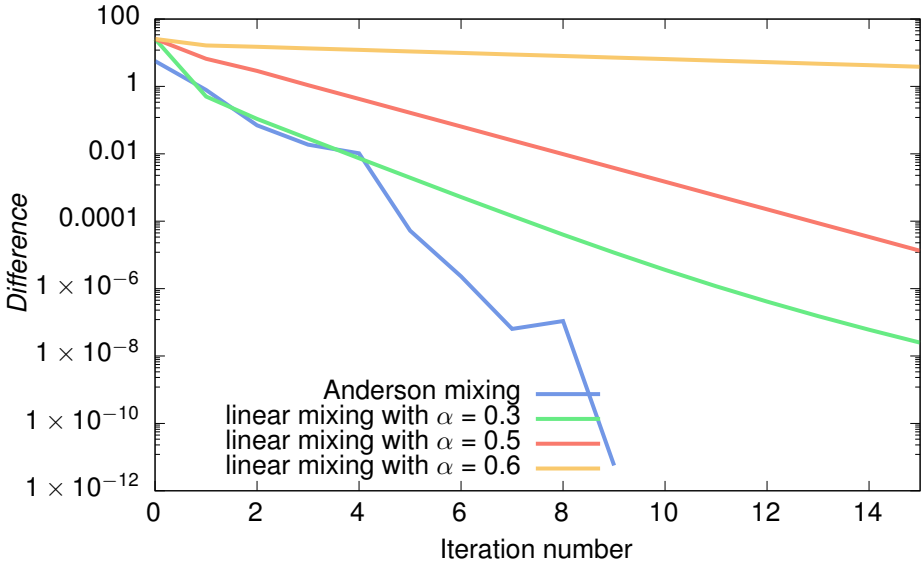

Figure 4.3: The graph shows the norm off the difference of the self-consistent perturbation potential for Benzene. One can see that the fastest convergence is reached using the Anderson Mixing. Using linear mixing does not necessarily lead to convergence but does depend on the mixing constant α. Choosing the right constant is not straight forward because it depends on the system and frequency. The Anderson mixing is thus the preferred method for the self-consistent Sternheimer algorithm.

Again, the factor α can be chosen freely, but in this case, the impact of the value chosen is not as important as in the linear mixing case.

Using the Anderson mixing of order $M \geq 10$, the self-consistent Sternheimer method converged for all test cases used. Also the number of iterations needed until convergence was reduced compared to linear mixing.

In can be concluded that by introducing the Anderson mixing scheme, the convergence of the self-consistent Sternheimer cycle ensured. Consequently, topic **1** can be regarded as solved.

In the following it will be explained how the results of the Sternheimer equation can be used to calculate the polarizability tensor.

4.3.8 Calculating the Polarizability Tensor

Once the self-consistent Sternheimer cycle is converged, the results like the corrected density or the converged perturbation potential can be used to calculate the desired property. One example that is given hear is the calculation of the polarizability tensor the system. [3]

To calculate the polarizability tensor, the self-consistent Sternheimer method

is used to calculate the response of the electronic structure of system, for example a molecule, to an external electric field. Under the assumption that the positions of the nuclei are fixed, the electric field induces an electric dipole $\mathbf{p}(t)$. [3] It is given by:

$$\mathbf{p}(t) = -e \int d\mathbf{r}\ \mathbf{r}\ n(\mathbf{r}, t). \tag{4.40}$$

Fourier transforming $\mathbf{p}(t)$ and $n(\mathbf{r}, t)$ as $p(\omega)$ and $n(\mathbf{r}, \omega)$, the polarizability tensor $\overleftrightarrow{\alpha}(\omega)$ can be defined as the first order variation of the dipole with respect to the external electric field:

$$\overleftrightarrow{\alpha}(\omega) = \frac{\partial \mathbf{p}(\omega)}{\partial \mathbf{E}(\omega)}. \tag{4.41}$$

Putting the two equations together, we obtain:

$$\overleftrightarrow{\alpha}_{\mu\nu}(\omega) = -e \int d\mathbf{r}\ r_\mu \frac{\partial n(\mathbf{r}, \omega)}{\partial \mathbf{E}(\omega)_\nu}. \tag{4.42}$$

Where r_μ points into one of the three spatial directions $\mu = x, y, z$.

The variation of the electron density with respect to the electric field is exactly the Δn calculated during the Sternheimer cycle. [3] By expanding it with the basis functions:

$$\Delta n(\mathbf{r}, \omega) = \sum_{ij} \Delta n_{ij}(\omega) \phi_i(\mathbf{r}) \phi_j(\mathbf{r}) \tag{4.43}$$

we obtain a matrix formulation of Δn, that can be inserted into equation 4.42.

The polarizability tensor can then be calculated from the change of the electron density:

$$\overleftrightarrow{\alpha}_{\mu\nu}(\omega) = -\sum_{ij} \Delta n^\nu_{ij}(\omega) \int d\mathbf{r} \phi_i(\mathbf{r}) \mathbf{r}_\mu \phi_j(\mathbf{r}). \tag{4.44}$$

With $\Delta n^\nu(\omega)$ being the converged change in electron density resulting from an external electric field in direction ν and the integral being the dipole integral in direction μ. [3]

In the result section of this thesis the isotropic average of the polarizability tensor is used to compare results. It is defined as:

$$\overline{\alpha} = (\overleftrightarrow{\alpha}_{xx} + \overleftrightarrow{\alpha}_{yy} + \overleftrightarrow{\alpha}_{zz})/3, \tag{4.45}$$

the average of the diagonal entries.

To calculate the polarizability tensor for a given frequency, several parameters need to be chosen that are not necessarily known beforehand. For example the optimal shift along the imaginary frequency axis can depend on the system. The closer the frequencies are to the real axis, the more accurate the Padé-approximation is. But when the frequencies are too close to the real axis, the matrices become ill conditioned and the self consistent cycle becomes unstable.

Further the absolute value of the polarizabilty depends on the strength of the electric field applied to the system.

4.3.9 Kramers-Kronig

The Sternheimer method calculates the real part of the response function. For some cases it is also desirable to investigate the imaginary part of the function. The Kramers-Kronig relations allow the computation of the imaginary part of a function from its real part if it satisfies this symmetry:

$$\phi^{(1)}(-\omega^*) = \phi^{(1)*}(\omega). \tag{4.46}$$

This identity follows directly from the construction of the dynamic Sternheimer equation in equation 3.20. [41] We therefore obtain $Im(\phi(\omega))$ with: [41]

$$Im(\phi^{(1)}(\omega)) = -\int_{-\infty}^{\infty} d\xi \frac{Re(\phi^{(1)}(\xi))}{\xi - \omega} \tag{4.47}$$

4.3.10 The Padé Approximation

The only thing left is to project the calculated values onto the real axis. Since the Sternheimer equation is solved for frequencies shifted away from the real axis, a continuation onto the real axis is needed.

The N-point Padé approximation is an analytic continuation of the following form: [49]

$$C_N(\omega) = \frac{a_1}{1 + \frac{a_2(\omega-\omega_1)}{1+\ldots\frac{a_N(\omega-\omega_{N-1})}{1}}} \tag{4.48}$$

For each frequency and value, a coefficient $a_i = g_i(\omega_i)$ is calculated using the recursive formula: [53]

$$g_p(\omega) = \frac{g_{p-1}(\omega_{p-1}) - g_{p-1}(\omega_p)}{(\omega - \omega_{p-1})g_{p-1}(\omega_p)} \tag{4.49}$$

with initial starting value:

$$g_1(\omega_i) = u_i, \tag{4.50}$$

where u_i is the calculated value at ω_i.

The N-Point Padé approximation can be evaluated using another recursive algorithm: [53]

$$C_N(z) = \frac{A_N(z)}{B_N(z)} \tag{4.51}$$

with

$$A_{n+1}(z) = A_n(z) + (z - z_n)a_{n+1}A_{n-1}(z), \tag{4.52}$$

$$B_{n+1}(z) = B_n(z) + (z - z_n)a_{n+1}B_{n-1}(z) \tag{4.53}$$

and $A(0) = 0$, $A(1) = a_1$, $B(0) = B(1) = 1$.

The use of the Padé Approximation allows us to solve linear systems that are ill conditioned or even singular, by adding an imaginary part to the frequency ω. This reduces the numerical error of the solver and speeds up the solution

process. For some cases finding a solution would even be impossible without introducing a shift.

The Padé approximation also speeds up the convergence of the self consistent Sternheimer cycle since the solution of the linear system is better for well conditioned system. Smaller numerical errors lead to a more stable cycle.

The accuracy of the Padé Approximation depends mainly on two factors, the number of input points and how close these points are to the evaluation points. Therefore, it is important, as shown above, to choose the grid size and the shift carefully.

One of the drawbacks of this recursive implementation is that the number of grid points is limited. Since the values A and B directly depend on the previous value, they tend to become very large or small which in some cases can lead to a overflow error. This error is caught in the implementation. Nevertheless, for a large range of frequencies splitting up the range into smaller batches is recommended.

Further this implementation can only approximate scalar valued function. Therefore, each entry of the tensor is approximated independently. A matrix valued approximation was implemented but was ultimately not used since it was less accurate due to the error induced from the matrix inversion needed during the calculations.

Since during the calculation several divisions are needed, the algorithm always checks the denominator to avoid division by zero. This can lead to an input point being rejected. But for a large enough grid size, this does not effect the overall accuracy.

To account for the symmetry of the polarizability function, for each tensor calculated four points are added to the Padé Approximation: [3]

$$f(\omega_{\text{real}} + i\kappa) = \alpha, \tag{4.54}$$

$$f(\omega_{\text{real}} - i\kappa) = \alpha^*, \tag{4.55}$$

$$f(-\omega_{\text{real}} + i\kappa) = \alpha, \tag{4.56}$$

$$f(-\omega_{\text{real}} - i\kappa) = \alpha^*. \tag{4.57}$$

This scheme ensures that the symmetries

$$\alpha(\omega^*) = \alpha^*(w) \tag{4.58}$$

and

$$\alpha(\omega) = -\alpha(w) \tag{4.59}$$

are preserved.

4.3.11 The Multi Shift Algorithm

The Multi Shift algorithm is a method designed to speed up the calculation of the solution of a linear system that are similar to an already solved seed system

$Ax = b$. A linear system of form $Ax + \omega I = b$ can be solved by replacing expensive matrix vector operation with vector operation.

The foundation of the Multi Shift method is the complex bi conjugate gradient algorithm. [50] It is an extension of the well known conjugate gradient algorithm that allows to compute solutions of non symmetric complex linear systems. Both algorithms are Krylov subspace methods that rely on projecting the solution vector on a Krylov subspace that is expanded in each step. It can be shown that the associated subspaces of the seed system: $\{b, Ab, A^2b, ...\}$ and the shifted system $\{b, (A + \omega I)b, (A + \omega I)^2 b\}$ form the same linear space. [50]

Therefore, it is possible to store and reuse the coefficients:

$$\alpha_n = \frac{\langle \tilde{r}_n | r_n \rangle}{\langle \tilde{p}_n | A p_n \rangle} \tag{4.60}$$

and

$$\beta_n = -\frac{\langle A^* \tilde{p}_n | r_{n+1} \rangle}{\langle \tilde{p}_n | A p_n \rangle} \tag{4.61}$$

from the BiCG algorithm. Here r_n and $\tilde{r}_n$ are the biorthogonal residuals and p_n and $\tilde{p}_n$ the search directions. The matrix vector operations performed here are the most expensive part of the algorithm.

One obtains the new residuals and coefficients for the shifted system using the following formulas:

$$r_{n,\omega} = \frac{r_n}{\pi_{n,\omega}} \tag{4.62}$$

$$\alpha_{n,\omega} = \frac{\pi_{n,\omega}}{\pi_{n+1,\omega}} \alpha_n \tag{4.63}$$

$$\beta_{n,\omega} = \left(\frac{\pi_{n,\omega}}{\pi_{n+1,\omega}} \right)^2 \beta_n \tag{4.64}$$

where π is calculates recursively:

$$\pi_{n+1,\omega} = (1 + \omega \alpha_n) + \frac{\alpha_n \beta_{n-1}}{\alpha_{m-1}} (\pi_{n,\omega} - \pi_{n-1,\omega}). \tag{4.65}$$

These equations avoid matrix vector multiplication. Therefore, the Multi Shift Solver reduces the computational costs of solving shifted linear system. [50]

For the Sternheimer equation it is straight forward to bring it into a form that is suitable for the Multi Shift solver. Multiplying the inverse density matrix S^{-1} yields:

$$[\mathbf{S}^{-1} - (\varepsilon_v \pm \omega)\mathbf{1}] \Delta \mathbf{c}_v^{\pm} = -\left[\mathbf{S}^{-1} - \boldsymbol{\rho}^T\right] \Delta \mathbf{V} \mathbf{c}_v. \tag{4.66}$$

This is exactly the form needed for the Multi Shift method. Once the system is solved for $\omega = 0$, the Sternheimer equation can be solved more efficiently for other frequencies.

4.3.12 Four Center Integration

When updating the perturbation potential a four point integral has to be evaluated. It is noted here that this integration is not the spatial integration mentioned as topic **2** in the Introduction, but rather an integral of shape:

$$(ij|kl)_{\mathrm{H}} = \int\int d\mathbf{r}d\mathbf{r}'\phi_i(\mathbf{r})\phi_j(\mathbf{r})(v_{\mathrm{c}}(\mathbf{r},\mathbf{r}'))\phi_k(\mathbf{r}')\phi_l(\mathbf{r}'). \tag{4.67}$$

Integrating over four sets of basis functions is a very costly procedure and scales badly with the N^4, with N being the number of basis functions. [45] There are N^2 different combinations of this product $\phi_i(\mathbf{r})\phi_j(\mathbf{r})$. Fortunately the functions in this set of products are highly linear dependent and can therefore be approximated using a reduced number of functions. This reduced set of around $3-5N$ functions is called an auxiliary basis set. We call the functions $\xi_\nu(\mathbf{r})$ in this set the auxiliary basis functions.

The four center integral can thus be approximated using two and three center integrals: [54]

$$(ij|kl)_{\mathrm{H}} \approx \sum_{\nu,\mu}(ij|\nu)(\nu|\mu)^{-1}(\mu|kl), \tag{4.68}$$

where

$$(\nu|\mu) = \int\int d\mathbf{r}d\mathbf{r}'\frac{\xi_\nu(\mathbf{r})\xi_\mu(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{4.69}$$

is the two center repulsion integral and $(\nu|\mu)^{-1}$ its inverse. $(ij|\nu)$ is the corresponding three center integral:

$$(ij|\nu) = \int\int d\mathbf{r}d\mathbf{r}'\frac{\phi_i(\mathbf{r})\phi_j(\mathbf{r})\xi_\nu(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{4.70}$$

This approximation leads to a significant reduction in computational load since the approximation scales only with N^3. The quality of the approximation can be improved by using a larger auxiliary basis set.

In the next chapter the self-consistent Sternheimer method will be tested by calculating the polarizability tensor for different parameter sets.

Chapter 5

Polarizability Results

For this thesis the self-consistent Sternheimer algorithm was implemented for the simulation code votka-xtp. Additionally algorithms for the Padé Approximation and the Anderson mixing were implemented. Before goining into details on how the Sternheimer cycle can be used in *GW* calculations, the self consistent Sternheimer method was tested by calculating polarizability tensors using the method described in the previous chapter. To test and analyse the Sternheimer method, the framework to compute the polarizability tensor was also added to the code. The calculation of the polarizability tensor does not require an integration over space or frequency other than the four center integral from section 4.3.12. Thus these calculations provide a good way to see how the Sternheimer method behaves when changing different parameters like the shift of the Padé-approximation or the basis set used without the need to investigate the numerical integrations as well. This will help to get a deeper understanding of several topics stated in the Introduction. In this chapter topic **3** is discussed in detail as several different parameter sets are tested and compared to investigate their respective effect on the calculations. Additionally topic **4** is addressed because different input configurations of the Padé approximation are used and discussed as well.

All calculations were performed for Benzene to compare the data to calculations from reference [3].

The algorithm was tested for different parameters. First the frequency grid was altered. Several numbers of grid points were tested to see how many are needed to achieve a sufficient accuracy.

The plot in figure 5.1 shows that even with very few grid points the first peak is still approximated very well. For frequencies greater than 9eV the graphs differ a lot. Using the Sternheimer method, it is fairly simple to reproduce the first peak, while everything beyond that is very sensitive to the parameters.

Looking at the isotropic average for different imaginary shifts of the frequencies in figure 5.2, a similar pattern is visible. The first peak is reproduced for all shifts but for high frequencies there are no similarities. The first peak is very useful since it corresponds to the energy gap between the highest occupied

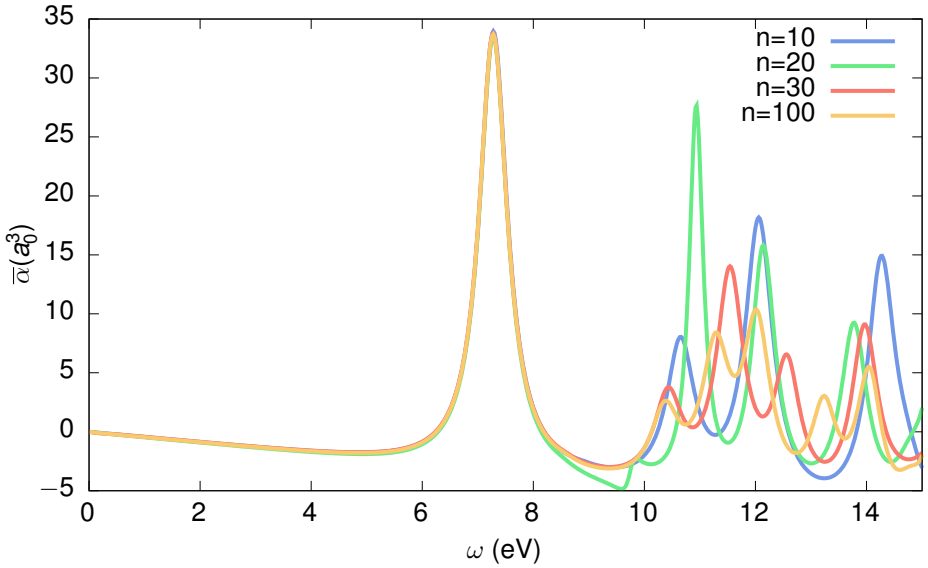

Figure 5.1: The graph shows the isotropic average of the polarizability tensor of benzene for different frequency grids. The first peak is reproduced for even small grids with only ten grid points. For frequencies larger than 10 a higher number of grid points is needed for better accuracy.

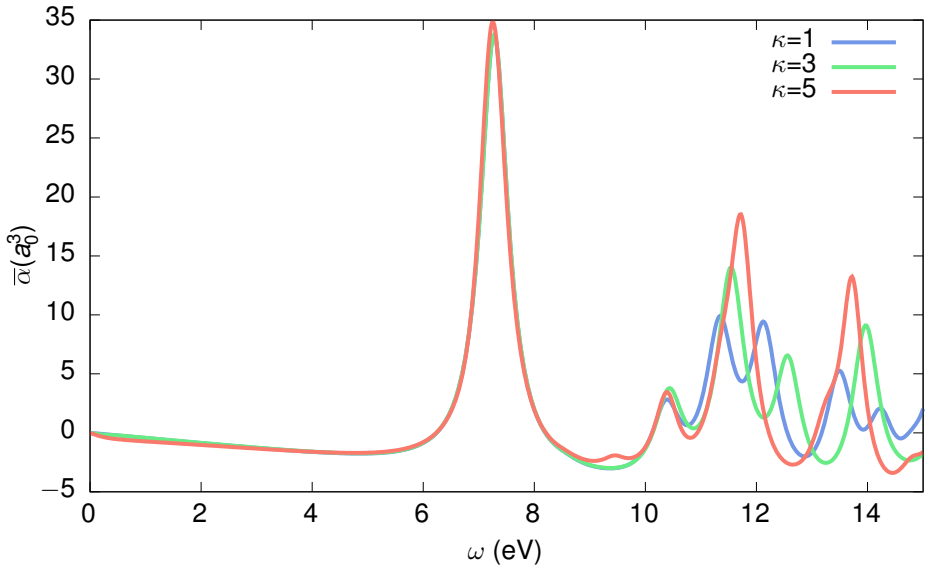

Figure 5.2: The graph shows the isotropic average of the polarizability tensor of benzene for different values of κ in the frequency grid. Again, the first peak is very similar for all different shifts and for higher frequencies the graphs differ a lot. This is due to a loss of accuracy for higher values of κ. When the grid points are further away from the real axis, the Padé approximation is less accurate.

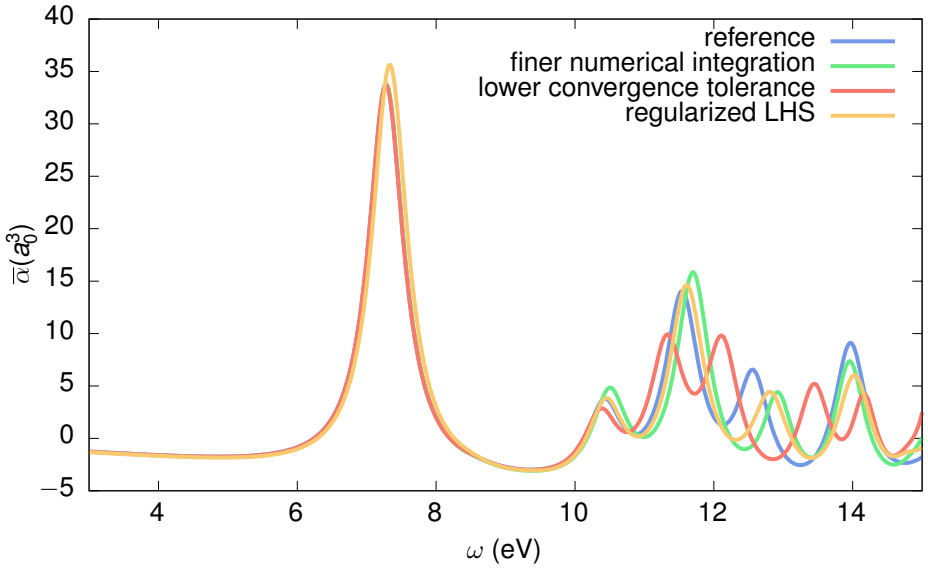

Figure 5.3: The graph shows the isotropic average of the polarizability tensor of benzene for different parameters. Changing the accuracy of the numerical integration for the four center integrals needed to update the perturbation potential does not have a large effect on the results. Also changing the tolerance of the convergence does not effect the output drastically. Using the projection operator onto the unoccupied states to regularize the left hand side of the Sternheimer equation has a small effect on the first peak of the graph.

orbital and the lowest unoccupied orbital. The specific parameters of the grid do not have a large influence on the result. Therefore, to calculate this property, the grid can be chosen freely as long as there are enough grid points and the shift is around 3eV.

This matches with the results from other tests seen in figure 5.3. Changing the accuracy of the numerical integration of the four center integrals calculated to update the perturbation potential in the self-consistent Sternheimer cycle does not effect the first peak of the isotropic average. Neither does changing the tolerance criterion for convergence. Only the values for higher frequencies differ. It is therefore possible to choose a coarse grid for numerical integration which saves around $\frac{2}{3}$ of computation time. Improving the regularization of the left hand side of the Sternheimer equation by adding the projection operator onto the occupied states manifold changes the results slightly but for small frequencies it is necessary to so for the Sternheimer equation to be numerically solvable.

Using different basis sets indicates how the algorithm behaves for a different number of basis functions. More basis functions usually increase the accuracy of the computation, however there is no guarantee that this is the case. Figure

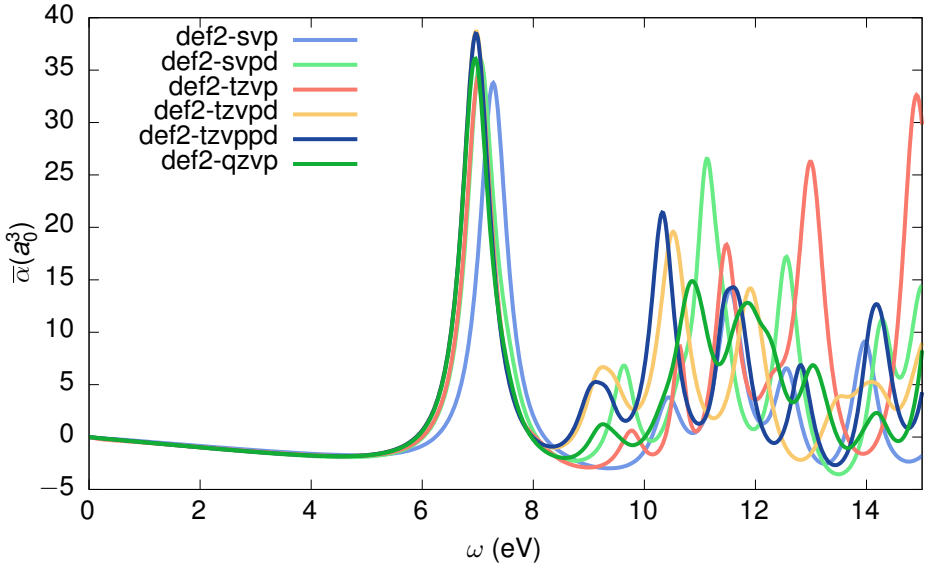

Figure 5.4: The graph shows the isotropic average of the polarizability tensor of benzene using different basis sets. All used basis sets are Karlsruhe basis sets with def2-svp being the smallest set and def2-qzvp being the largest one. The first peak shifts slightly towards smaller frequencies for larger sets. For frequencies larger than 10eV there are no similarities between different basis sets to be seen.

5.4 shows that increasing the size of the basis sets leads to a shift of the first peak towards smaller frequencies. For even larger sets only the absolute value of the isotropic average changes. For frequencies larger than ten the graphs differ heavily. This indicates that the Sternheimer methods is not well suited to calculate high frequency polarizabilities.

To summarize all the tests explained above it is clearly visible that the first peak is reproduced by all parameter sets tested while the results differ for frequencies larger than 10eV. Conclusively the Sternheimer method is very robust respective to changes in the parameters discussed in this chapter when it comes to calculating the first spike of the isotropic average. This is to be expected because the first peak is a very distinct feature of the polarizability tensor as in the reference used to compare the results to [3] only the first peak could be reproduced from experimental data. So while the effects of changing the grid, the Padé parameters, the convergence tolerance and other parameters can be seen in the results for higher frequencies, the pronounced features are mostly unaffected by the changes. The largest changes can be observed when changing the basis set. This can be explained by the fact that while the parameters are affecting the mathematical tools and methods used in the Sternheimer method changing the basis sets also affects the underlying DFT calculations and thus the single particle wave functions and the electron density used as the starting point for the cycle.

To conclude on topic **3**, different parameters always need testing to find suitable values. In the case of the self-consistent Sternheimer method the parameter sets shown in this chapter prove to be reliable and robust in regards to changes in the magnitudes depicted here. Regarding topic **4** the Padé-approximation proves to be a reliable method to avoid treating poles on the real axis, and enforce symmetric properties onto the observed function. It is important that enough input points are used and that the distance to the targeted evaluation points is small enough, but as the examples shown here indicate already ten gridpoint on a range of more than 25eV is enough to produce good results while the shift that is introduced can be chosen freely as long as it is under 5eV.

Chapter 6

GW Sternheimer Implementation

As mentioned above, the self-consistent Sternheimer method can be used for *GW* calculations as well. [1] The goal is to construct the self-energy:

$$\Sigma(\mathbf{r}, \mathbf{r}', \omega) = \frac{i}{2\pi} \int d\omega' G(\mathbf{r}, \mathbf{r}', \omega + \omega') W(\mathbf{r}, \mathbf{r}', \omega'). \tag{6.1}$$

In the Introduction it was already mentioned as topic **5** that it needs to be examined how the self-energy, the screened Coulomb potential and the Green's function can be expressed in a Gaussian basis set. In this chapter it is first explained how the screened Coulomb interaction and the Green's function can be obtained, and subsequently how the integrations can be performed to evaluate the self-energy.

6.1 The Screened Coulomb Interaction and the Green's Function

The Screened Coulomb interaction $W(\mathbf{r}, \mathbf{r}', \omega)$ can be calculated directly using the self consistent Sternheimer method as already mentioned in chapter 4. When using the bare Coulomb potential $v(\mathbf{r}, \mathbf{r}')$ as the initial perturbation potential, the resulting self consistent potential is the screened Coulomb potential. [1] The only change to the polarizability calculation is, as explained previously, that the Sternheimer equation needs to be solved not only on a frequency grid but also on a spatial grid because the bare Coulomb potential depends on two spatial variables. One spatial dependency is integrated out by expansion in the Gaussian basis:

$$W^{[r]}_{\mu\nu}(\omega') = \int d\mathbf{r}' \chi_\mu(\mathbf{r}') W_{[\mathbf{r}]}(\mathbf{r}', \omega') \chi_\nu(\mathbf{r}'). \tag{6.2}$$

The remaining spatial dependence is treated using a spatial grid thus evaluating the self-consistent Sternheimer cycle for different points in space.

The other quantity needed is the Green's function

$$G(\mathbf{r}, \mathbf{r}'; \omega) = \sum_n \frac{\psi_n(\mathbf{r})\psi_n^*(\mathbf{r}')}{\omega - \epsilon_n + i\eta_n}. \tag{6.3}$$

The poles of the Green's function have a physical meaning. They are the excitation energies of the electrons. [55] Because the propagation direction is different for occupied and unoccupied states, the poles corresponding to occupied states lie on the opposite site of the axis compared to the poles corresponding to the unoccupied states. That is why the infinitesimal η_n is positive for occupied states and negative for unoccupied states. An example of the pole structure of the Green's function can be seen in figure 6.1.

So when splitting the sum into a sum over occupied states v and unoccupied states c we obtain:

$$G(\mathbf{r}, \mathbf{r}'; \omega) = \sum_{v \in \text{occ}} \frac{\psi_v(\mathbf{r})\psi_v^*(\mathbf{r}')}{\omega - \epsilon_v - i\eta} + \sum_{c \in \text{unocc}} \frac{\psi_c(\mathbf{r})\psi_c^*(\mathbf{r}')}{\omega - \epsilon_c + i\eta} \tag{6.4}$$

We continue to separate the Green's function, by adding and subtracting the sum over the occupied states only:

$$G(\mathbf{r}, \mathbf{r}'; \omega) = \sum_{v \in \text{occ}} \frac{\psi_v(\mathbf{r})\psi_v^*(\mathbf{r}')}{\omega - \epsilon_v - i\eta} + \sum_{c \in \text{unocc}} \frac{\psi_c(\mathbf{r})\psi_c^*(\mathbf{r}')}{\omega - \epsilon_c + i\eta} \pm \sum_{v \in \text{ occ}} \frac{\psi_v(\mathbf{r})\psi_v^*(\mathbf{r}')}{\omega - \epsilon_v + i\eta} \tag{6.5}$$

By rearranging the fractions we can split the Green's function into an analytic and a non-analytic term:

$$G(\mathbf{r}, \mathbf{r}'; \omega) = G^A(\mathbf{r}, \mathbf{r}'; \omega) + G^N(\mathbf{r}, \mathbf{r}'; \omega) \tag{6.6}$$

The analytic term is the sum over all states with a negative η:

$$G^A(\mathbf{r}, \mathbf{r}'; \omega) = \sum_n \frac{\psi_n(\mathbf{r})\psi_n^*(\mathbf{r}')}{\omega - \varepsilon_n + i\eta}. \tag{6.7}$$

The left over term form the non-analytic term:

$$G^N(\mathbf{r}, \mathbf{r}'; \omega) = \sum_{v \in \text{occ}} \frac{\psi_v(\mathbf{r})\psi_v^*(\mathbf{r}')}{\omega - \varepsilon_v - i\eta} - \frac{\psi_v(\mathbf{r})\psi_v^*(\mathbf{r}')}{\omega - \varepsilon_v + i\eta}. \tag{6.8}$$

For real frequencies when η goes to zero the formula simplifies to:

$$G^N(\mathbf{r}, \mathbf{r}'; \omega) = 2\pi i \sum_{v \in \text{occ}} \delta(\omega - \varepsilon_v)\psi_v(\mathbf{r})\psi_v^*(\mathbf{r}'), \tag{6.9}$$

where $\delta(\omega - \varepsilon_v)$ is the Dirac delta distribution. The non-analytic Green's function G^N can be calculated directly, from the formula states above. Later in this

chapter when the frequency integration methods are described it will become clear that in some cases complex frequencies are needed. In these cases the complete formula from equation 6.8 is needed. To calculate the analytic part G^A both sides of equation 6.7 are expanded with $(\hat{H}-\omega)$ to obtain:

$$(\hat{H}-\omega)G^A_{\mathbf{r},\omega} = \sum_n \frac{\psi_n(\mathbf{r})[\hat{H}-\omega]\psi_n}{\omega-\varepsilon_n+i\eta}. \tag{6.10}$$

Using the orthonormality of the basis functions the equation simplifies to:

$$(\hat{H}-\omega)G^A_{\mathbf{r},\omega} = -\delta_{[\mathbf{r}]}(\mathbf{r}'). \tag{6.11}$$

This is a linear system similar to the Sternheimer equation that can be solved to obtain the analytic part of the Green's function. It is important to note that there is no need for self-consistency for the Green's function because the right hand side of the equation is known and not dependent on the solution.

6.2 The Self-Energy

We have all the pieces to calculate the self-energy. We start by projecting the self-energy on the Kohn-Sham states:

$$\Sigma_{nm}(\omega) = \int d\mathbf{r}d\mathbf{r}'\ \psi^{\mathrm{KS}}_n(\mathbf{r})(\Sigma^c(\mathbf{r},\mathbf{r}',\omega)+\Sigma^{ex}(\mathbf{r},\mathbf{r}'))\psi^{\mathrm{KS}}_m(\mathbf{r}'). \tag{6.12}$$

The self-energy is split into two parts: the Coulomb part:

$$\Sigma^c(\mathbf{r},\mathbf{r}',\omega) = \frac{i}{2\pi}\int d\omega' G(\mathbf{r},\mathbf{r}',\omega+\omega')(W(\mathbf{r},\mathbf{r}',\omega)-v(\mathbf{r},\mathbf{r}')) \tag{6.13}$$

and the exchange part:

$$\Sigma^{ex}(\mathbf{r},\mathbf{r}') = -\sum_{v\in\mathrm{occ}} \psi^*_v(\mathbf{r})\psi_v(\mathbf{r}')v(\mathbf{r},\mathbf{r}'). \tag{6.14}$$

The exchange part needs no integration and can be calculated directly from the Coulomb potential. The Coulomb part needs a closer look. For simplicity reasons the notation $W^c_{[\mathbf{r}]}(\mathbf{r}',\omega) = W_{[\mathbf{r}]}(\mathbf{r}',\omega)-v(\mathbf{r}-\mathbf{r}')$ is used. The Kohn-Sham states are known from DFT calculations and can be expressed in the Gaussian basis functions using the Kohn-Sham coefficients c^n_μ:

$$\Sigma_{nm}(\omega)^c = \frac{i}{2\pi}\sum_{\mu\nu} c^n_\mu \left(\int_{-\infty}^{\infty} d\omega' I_{\mu\nu}(\omega,\omega')\right) c^m_\nu \tag{6.15}$$

With $I_{\mu\nu}(\omega,\omega')$ being the double spatial integral over the Green's function and the screened Coulomb interaction. We express one of those integrals numerically to obtain:

$$I_{\mu\nu}(\omega,\omega') = \sum_{\mathbf{r}\in\text{ grid}} w_{[\mathbf{r}]}\chi^{[\mathbf{r}]}_\mu \int d\mathbf{r}' G_{[\mathbf{r}]}(\mathbf{r}',\omega+\omega')W^c_{[\mathbf{r}]}(\mathbf{r}',\omega')\chi_\nu(\mathbf{r}'). \tag{6.16}$$

With $w_{[\mathbf{r}]}$ being the integration weight corresponding to the grid position $\mathbf{r}$. Further, the Green's function is expanded in the Gaussian basis as well:

$$G_{[\mathbf{r}]}(\mathbf{r}', \omega + \omega') = \sum_{\alpha\beta} G_{\alpha\beta}(\omega + \omega')\chi_{\alpha}^{[\mathbf{r}]}\chi_{\beta}(\mathbf{r}'). \tag{6.17}$$

Inserting this into the integral 6.16 yields:

$$I_{\mu\nu}(\omega, \omega') = \sum_{\mathbf{r}\in \text{ grid}} w_{[\mathbf{r}]}\chi_{\mu}^{[\mathbf{r}]} \sum_{\alpha\beta} \chi_{\alpha}^{[\mathbf{r}]} G_{\alpha\beta}(\omega + \omega') \int d\mathbf{r}' \chi_{\beta}(\mathbf{r}') W_{[\mathbf{r}]}^{c}(\mathbf{r}', \omega')\chi_{\nu}(\mathbf{r}'). \tag{6.18}$$

The integral is exactly the matrix representation of the screened Coulomb potential from equation 6.2. Replacing the integral leads to:

$$I_{\mu\nu}(\omega, \omega') = \sum_{\mathbf{r}\in grid} w_{[\mathbf{r}]}\chi_{\mu}^{[\mathbf{r}]} \sum_{\alpha\beta} \chi_{\alpha}^{[\mathbf{r}]} G_{\alpha\beta}(\omega + \omega') W_{\beta\nu}^{[\mathbf{r}]c}(\omega'). \tag{6.19}$$

Putting the $I_{\mu\nu}(\omega, \omega')$ back into equation 6.15 for the self-energy we obtain an expression for the self energy:

$$\Sigma_{nm}^{c}(\omega) = \frac{i}{2\pi} \sum_{\mu\nu} c_{\mu}^{n} \left(\int_{-\infty}^{\infty} \sum_{\mathbf{r}\in \text{grid}} w_{[\mathbf{r}]}\, \chi_{\mu}^{[\mathbf{r}]} \sum_{\alpha\beta} \chi_{\alpha}^{[\mathbf{r}]} G_{\alpha\beta}(\omega + \omega') W_{\beta\nu}^{c,[\mathbf{r}]}(\omega') d\omega' \right) c_{\nu}^{m} \tag{6.20}$$

In practice to evaluate this expression, the evaluation frequency ω is fixed. To perform the frequency integration, a set of integration frequency grid points ω_n' is chosen. For each of these frequencies the self-consistent Sternheimer cycle is evaluated on all spatial grid points. Consequently, for each frequency point ω_n' a numerical integration over all spatial grid points is performed. This is repeated for all ω_n' so that the frequency integral can be numerically approximated. The procedure can be repeated for all chosen frequencies ω. As already noted in the Introduction as topic **2**, the two integrals that need to be approximated form the bottleneck of the Sternheimer GW approach with localized basis sets.

6.3 Numerical Integration

The spatial integration can be performed using a three dimensional rectangle quadrature:

$$\int_{\text{box}} f(\mathbf{r}) d\mathbf{r} \approx \sum_{n=1}^{N} w_n f(\mathbf{r}_n) \tag{6.21}$$

Where 'box' is the area around the molecule that can be determined from the nuclear coordinates and N is the number of grid points. The integration area can be bounded because the Gaussian basis function decays as e^{-r^2}. Because of that, this integral is not as problematic as the frequency integral. Also to ensure reasonable computing times the grid size should be kept as small as possible.

That is why the placement of the integration points needs special attention. For this thesis a cubic, equidistant grid as well as a atom centered spherical grid as used in reference [56] were used. A discussion of both methods can be found in the next chapter 7.

The frequency integration poses more problems because the integration needs to be performed over the whole frequency axis. Additionally the poles of the Green's function and screened Coulomb potential make numerical integration difficult. [1] In figure 6.1 the structure of the poles can be seen. Integrating along the real axis can pose problems because the poles lie very close to it which can result in very unstable results. That is why in this chapter several different integration methods and integration paths are presented. First three different quadratures for the integration over ω' will be explained.

All methods used are Gaussian quadratures. These types of quadratures aim to optimize the numerical approximation of the integral by splitting up the integrated function into a weight function $w(\omega)$ and a polynomial $\Psi(\omega)$ and approximating the integration as a sum:

$$\int_a^b f(\omega)d\omega = \int_a^b \Psi(\omega)w(\omega)d\omega \approx \int_a^b p_n(\omega)w(\omega)d\omega = \sum_{i=1}^{n} \Psi(\omega_i)w_i. \tag{6.22}$$

with n being the integration order. The polynomial p_n is chosen in a way that it approximates the function at the grid points and can be integrated exactly. The difference in the integration methods used comes from the integration interval $[a, b]$, the grid points ω_i and the weight function $w(\omega)$. The three methods used are the Gauss-Legendre, the Gauss-Hermite and the Gauss-Laguerre quadrature.

The Gauss-Legendre can be used to compute an integral of form:

$$\int_{-1}^{1} f(\omega)d\omega \approx \sum_{i=1}^{n} f(\omega_i)w_i \tag{6.23}$$

With the grid point ω_i being the roots of the Legendre polynomial $P_n(\omega)$ and the weights are given by:

$$w_i = \frac{2}{(1-\omega_i^2)P_n'(\omega_i)^2}. \tag{6.24}$$

This choice of weights and grid points is optimal in the way that a polynomial of order $2n-1$ can be integrated exactly, which is the maximum degree possible. This integration method can be adapted to arbitrary intervals by scaling the grid points to the desired interval:

$$\int_a^b f(\omega)d\omega \approx \frac{b-a}{2} \sum_{i=1}^{n} f(\frac{b-a}{2}\omega_i + \frac{a+b}{2})w_i. \tag{6.25}$$

In the case of an infinite integration interval a cutoff c needs to be introduced so the integral is approximated by a finite integral: [57]

$$\int_{-\infty}^{\infty} f(\omega)d\omega \approx \int_{-c}^{c} f(\omega)d\omega \tag{6.26}$$

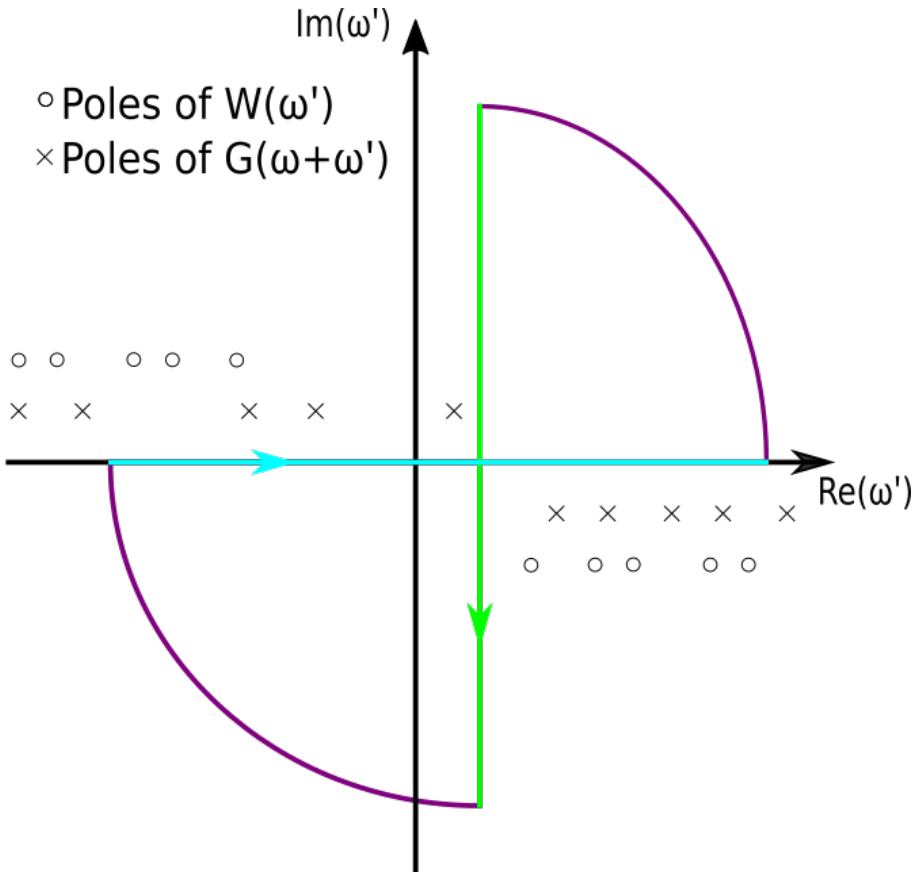

Figure 6.1: In this figure the location of the poles of $G(\omega + \omega')$ and $W(\omega')$ illustrated. The poles of both functions are located close to the real axis. That poses a problem when integrating along the real axis. The path shown is an alternative approach. The value of the purple paths vanishes when the radius goes to infinity and thus the value of the green path must be minus the value of the cyan path because the complete integral equals zero as it contains no poles. Thus, instead of computing the integral along the real axis, it can be evaluated on a vertical line that is further away from the poles. This should improve the stability of the numerical integration.

The Gauss-Hermite integration does not need a cutoff. When the function being integrated decays fast enough which for this case is e^{-x^2}, it is possible to approximate the integral on a finite interval with:

$$\int_{-\infty}^{\infty} e^{-\omega^2} f(x) d\omega \approx \sum_{i=1}^{n} w_i f(\omega_i) \tag{6.27}$$

with ω_i being the roots of the nth Hermite polynomial H_n and w_i the corresponding weights:

$$w_i = \frac{2^{n-1} n! \sqrt{\pi}}{n^2 (H_{n-1}(x_i))^2}. \tag{6.28}$$

The Gauss-Laguerre quadrature approximates integrals of the form:

$$\int_{0}^{\infty} e^{-\omega} f(\omega) d\omega. \tag{6.29}$$

The integration weights and points are obtained in the same way as before, only this time the Laguerre polynomials are used. When applying this quadrature to the integration at hand, we need to assume symmetry of the functions integrated. This is not the case when integrating along the real frequency axis so when using the Gauss-Laguerre quadrature, the integration interval needs to be transformed.

Transforming the integration interval is preferable even for the other quadrature methods as the Green's function and the screened Coulomb potential, as mentioned above, are not well behaved on the real axis. [1]

Because there are two frequency dependencies, ω and ω', there are several different transformations to take into account. A simple way is to keep the evaluation points of the self-energy ω_i on the real axis and only changing the integration points ω'_j by rotating the integration interval by 90 degrees and thus integrating along a vertical line parallel to the imaginary axis. This can be achieved by looking at a contour integral that consists of an integral along the real axis and an integral along a vertical line along the imaginary axis connected by two quarter circles as depicted in figure 6.1. The values of the circular parts vanish as the radius of the quarter circles approaches infinity because the functions $G(\mathbf{r}, \mathbf{r}', \omega')$ and $W(\mathbf{r}, \mathbf{r}', \omega')$ decay as $|\omega'|^{-1}$ and $|\omega'|^{-2}$ respectively. [58] Because the contour integration does not contain any poles, the vertical integration must be the same as the horizontal integration, only with a different sign, leading to this variations of the self-energy formula 6.1:

$$\Sigma(\mathbf{r}, \mathbf{r}', \omega) = \int_{-\infty}^{\infty} d\omega' G(\mathbf{r}, \mathbf{r}', \omega + i\omega') W(\mathbf{r}, \mathbf{r}', i\omega'). \tag{6.30}$$

The frequencies $i\omega$ are purely imaginary, and thus the integration is performed parallel to the imaginary axis. The evaluation frequency ω of the self-energy is kept real. Because the integrants satisfy: [58]

$$G(\mathbf{r}, \mathbf{r}', \overline{\omega}') = \overline{G(\mathbf{r}, \mathbf{r}', \omega')} \tag{6.31}$$

and

$$W(\mathbf{r}, \mathbf{r}', \overline{\omega}') = \overline{W(\mathbf{r}, \mathbf{r}', \omega')}, \tag{6.32}$$

this integration is suitable for the Gauss-Laguerre quadrature and can be written as:

$$\Sigma(\mathbf{r}, \mathbf{r}', \omega) = 2 \int_0^\infty d\omega' G(\mathbf{r}, \mathbf{r}', \omega + i\omega') W(\mathbf{r}, \mathbf{r}', i\omega'). \tag{6.33}$$

It is important to note that this approach is only viable if the path does not contain any poles. Depending on ω the path can contain poles because changing the value of ω shifts the poles of the Green's function parallel to the real axis. This can result into the poles above and below the axis overlapping so that a path containing no poles is no longer possible. In that case the residues of the poles need to be taken into account according to Cauchy's integral theorem. A different approach is approximating the integral over the real axis by shifting the integration interval into the complex plane resulting in: [1]

$$\Sigma(\mathbf{r}, \mathbf{r}', \omega) = \int d\omega' G(\mathbf{r}, \mathbf{r}', \omega + \omega' + i\kappa) W(\mathbf{r}, \mathbf{r}', \omega' + i\kappa). \tag{6.34}$$

That way the integration interval is moved away from the poles. The further away the interval is shifted the more stable the integration becomes, but the approximation becomes also less accurate.

An alternative approach is to avoid the evaluation of the self-energy on the real axis all together and evaluate it on the imaginary axis:

$$\Sigma(\mathbf{r}, \mathbf{r}', i\omega) = \int d\omega' G(\mathbf{r}, \mathbf{r}', i\omega + i\omega') W(\mathbf{r}, \mathbf{r}', i\omega'). \tag{6.35}$$

and use analytic continuation in form of the Padé-approximation to map the self-energy onto the real axis. A detailed analysis of all three different quadratures and the different intervals can be found in the next chapter 7.

One final approach that was tested is motivated from the split of the Green's function into an analytic and non-analytic part as seen in equation 6.6. The non-analytic part of the Green's function contains all the poles and is thus harder to integrate numerically. Therefore, the self-energy:

$$\Sigma(\mathbf{r}, \mathbf{r}', \omega) = \frac{i}{2\pi} \int d\omega' G(\mathbf{r}, \mathbf{r}', \omega + \omega') W(\mathbf{r}, \mathbf{r}', \omega') \tag{6.36}$$

is, consequently, separated into a part corresponding to the analytic part of the Green's function:

$$\Sigma^{\mathrm{COH}}(\mathbf{r}, \mathbf{r}', \omega) = \frac{i}{2\pi} \int d\omega' G^A(\mathbf{r}, \mathbf{r}', \omega + \omega') W(\mathbf{r}, \mathbf{r}', \omega') \tag{6.37}$$

called the Coulomb hole part (COH) and the screened exchange part (SEX) corresponding to the non-analytic Green's function:

$$\Sigma^{\mathrm{SEX}}(\mathbf{r}, \mathbf{r}', \omega) = \frac{i}{2\pi} \int d\omega' G^N(\mathbf{r}, \mathbf{r}', \omega + \omega') W(\mathbf{r}, \mathbf{r}', \omega') \tag{6.38}$$

$$= \frac{i}{2\pi} \int d\omega' G^N(\mathbf{r}, \mathbf{r}', \omega + \omega') W_c(\mathbf{r}, \mathbf{r}', \omega'). \tag{6.39}$$

The Coulomb hole part is evaluated in the same way as the Coulomb part of the self-energy. The Coulomb potential can be left out since G^A is an analytic function in the upper half of the complex plain, therefore, the integration with the frequency independent Coulomb potential $v(\mathbf{r}, \mathbf{r}')$ vanishes and we are left with $W_c(\mathbf{r}, \mathbf{r}', \omega)$. Because only the analytic part of the Green's function is evaluated the numerical integration should be more stable. Next, the formula for the non-analytic part of the Green's function is inserted into equation 6.39:

$$\Sigma^{\mathrm{SEX}}(\mathbf{r}, \mathbf{r}', \omega) = -\sum_{v \in \mathrm{occ}} \int d\omega' \delta(\omega + \omega' - \epsilon_v) \psi_v(\mathbf{r}) \psi_v^*(\mathbf{r}') W(\mathbf{r}, \mathbf{r}', \omega'). \quad (6.40)$$

From the delta function we obtain $\omega' = \epsilon_v - \omega$ and thus:

$$\Sigma^{\mathrm{SEX}}(\mathbf{r}, \mathbf{r}', \omega) = -\sum_{v \in \mathrm{occ}} \psi_v(\mathbf{r}) \psi_v^*(\mathbf{r}') W(\mathbf{r}, \mathbf{r}', \epsilon_v - \omega). \quad (6.41)$$

The screened exchange part can be evaluated without numerical integration over the frequency axis. The formula can be split up further by using $W = W_c + v$ resulting in:

$$\Sigma_c^{\mathrm{SEX}}(\mathbf{r}, \mathbf{r}', \omega) = -\sum_{v \in \mathrm{occ}} \psi_v(\mathbf{r}) \psi_v^*(\mathbf{r}') W_c(\mathbf{r}, \mathbf{r}', \epsilon_v - \omega). \quad (6.42)$$

and

$$\Sigma^{ex}(\mathbf{r}, \mathbf{r}') = -\sum_{v \in \mathrm{occ}} \psi_v^*(\mathbf{r}) \psi_v(\mathbf{r}') v(\mathbf{r}, \mathbf{r}'), \quad (6.43)$$

the frequency dependent screened exchange part and the frequency independent exchange part.

Consequently, the integration over the non-analytic Green's function can be avoided, which should lead to more stable results.

The representation of the self-energy in Gaussian basis functions along with the integration methods stated, can now be used to calculate the self energy of system using the self-consistent Sternheimer cycle:

$$\Sigma_{nm}(\omega) = \quad (6.44)$$
$$\frac{i}{2\pi} \sum_{\mu\nu} c_\mu^n \left[\sum_{\omega' \in freq.} w_{\omega'} \sum_{\mathbf{r} \in grid} w_{[\mathbf{r}]} \chi_\mu^{[\mathbf{r}]} \sum_{\alpha\beta} \chi_\alpha^{[\mathbf{r}]} G_{\alpha\beta}(\omega + \omega') W_{\beta\nu}^{c,[\mathbf{r}]}(\omega') \right] c_\nu^m$$

Topic **5** is thus dealt with, and the only problem left is topic **2** the performance of the numerical integration. In the next chapter the Sternheimer GW method is tested in particular respective to the integration methods used.

Chapter 7

GW Results

The Sternheimer *GW* method was used to calculate the self-energy $\Sigma(\omega)$ of hydrogen and water molecules. The simple structure of these molecules allows for fast calculation so that different parameter sets can be tested. While the same parameters investigated in the polarizability section can be tested here again, this section is focused on the numerical integration aspects of calculation. As presented in the previous chapter to obtain the self-energy of a system a numerical integration over the space as well as the frequencies needs to be performed.

7.1 Spatial Integration

First, the spatial integration is analysed. This is a very costly operation especially since it has to be performed for each step of the frequency integration. The most basic idea to perform numerical integration is to set up an regular equidistant grid and evaluate the screened Coulomb potential on all these grid points. The corresponding weights can be calculated as $w = d^3$ with d the distance between the grid points. Although this is very easy to implement it is not the most efficient way to perform the numerical integration. To increase accuracy d needs to be decreased. But since the number of grid points scales with d^3, the number of grid points can become too large to allow computations in reasonable time. Taking the structure of the observed system into account can lead to less grid points being needed to achieve the same accuracy. This motivates a different arrangement of the integration grid points. Arranging the points in a spherical fashion around the nuclei leads to more grid points being close to the nuclei and less points being further away from the nuclei. Grid points are arranged around each nucleus and overlapping cells are merged. The exact process of how the grid points and weights are calculated can be found in. [56] The computation of the spatial integral takes a lot of time. Although each iteration of the self-consistent Sternheimer cycle is really cheap for small molecules, as investigated in figure 4.30, the sheer amount of cycles needed for

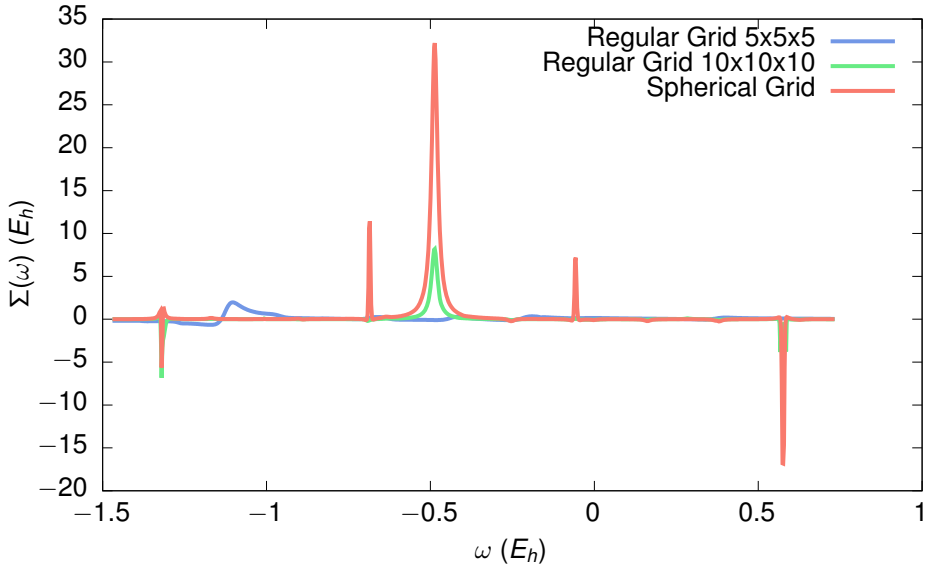

Figure 7.1: The plot shows the real part of the self-energy diagonal matrix element of the HOMO level of a hydrogen molecule. The grid with 5 points in each direction does not reproduce the results of the finer grids, whereas the spherical grid and the 10 step grid produce similar results. The peaks of the spherical grid graph are however much more pronounced although overall less grid points where used.

a *GW* calculation leads to high computation times. For example, the smallest grid tested, a $5 \times 5 \times 5$ equidistant grid, contains 125 grid points. A simple integration scheme for the integration over the frequency ω' needs 12 grid points, and to plot the self energy in dependence of the frequency ω 50 calculation are needed, resulting in $12 * 125 * 50 = 75000$ Sternheimer cycles being calculated in a simple hydrogen *GW* calculation. That is why it is important to keep the number of integration points as low as possible.

From the test performed it can be observed that even a comparably very coarse grid is able to produce stable results. The graphs of the 10 step grid and the spherical grid are very similar but the spherical grid uses less grid points. Consequently, it can be assumed that the spherical grid provides a good accuracy to complexity ratio. Still, even the coarse grids used in these tests, lead to very high computation times. That is also why finer grids where not tested as even for small molecules the computational workload is too high.

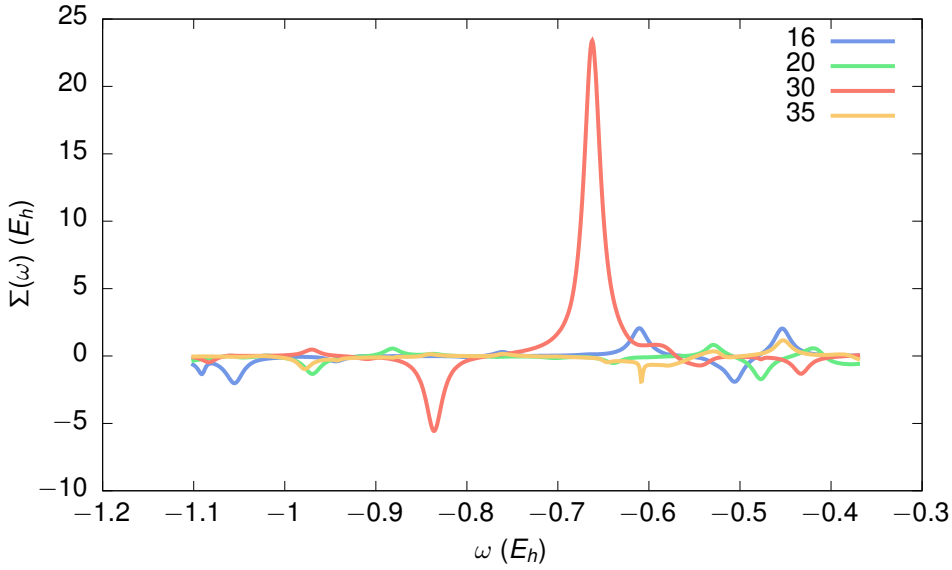

Figure 7.2: The plot shows the real part of the self-energy diagonal matrix element of the HOMO level of a water molecule. Different integration orders of the Gauss-Hermite quadrature where used to integrate along the real axis to produce these results. It can be observed that changing the quadrature order drastically changes the results. Consequently, integrating along the real axis is not a feasible option.

7.2 Frequency Integration

The integration over ω' poses a different challenge because the integration needs to be performed over the whole frequency axis. Thus the three different Gaussian quadratures outlined in the previous chapter are used to see which performs best. Additionally the Green's function and the screened Coulomb potential have poles close to the real frequency axis, as was explained in figure 6.1.

Consequently, already mentioned as topic **2** in the Introduction the numerical integration over the real frequency axis is a major challenge of the implementation of the Sternheimer method. During the course of this thesis several different approaches to resolve the two problems noted where tested in different combinations. In the implementation chapter 6 three different quadrature methods were outlined: the Gauß-Legendre, Gauß-Laguerre and Gauß-Hermite quadrature methods. All methods address the problem of integrating over an infinite interval in different ways. In chapter 6 also different ways to treat the integration interval where presented. In the following tests of the different methods are presented and discussed.

First, the purely real and purely imaginary approaches where tested. From the plots in figures 7.2 and 7.3 of the tests it can be observed that the frequency

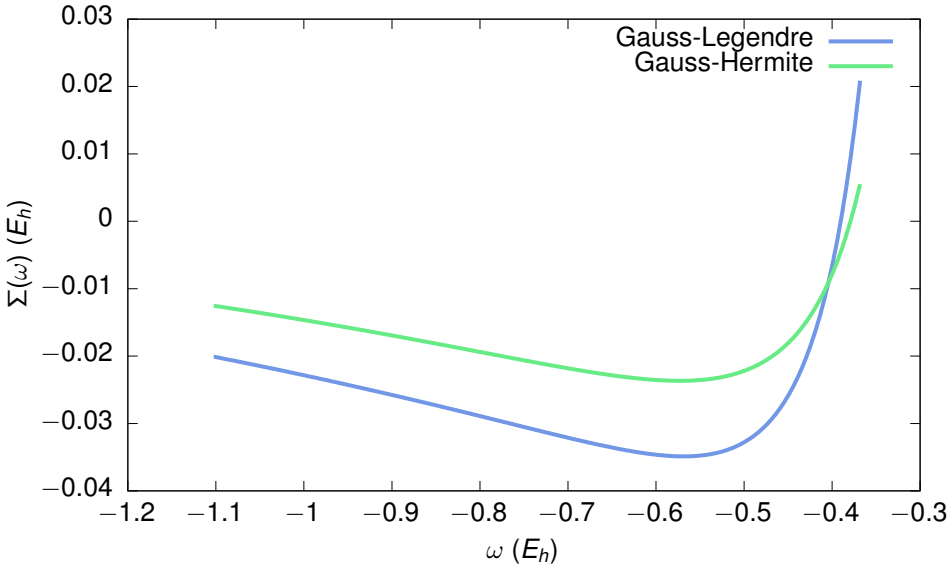

Figure 7.3: The plot shows the real part of the self-energy diagonal matrix element of the HOMO level of a water molecule. For this test, imaginary frequencies combined with the Padé-approximation where used to produce the results. The structure of the self-energy cannot be reproduced by this approach as the values are way to small and in comparison to results from the standard approach show, that this integration method is also not feasible.

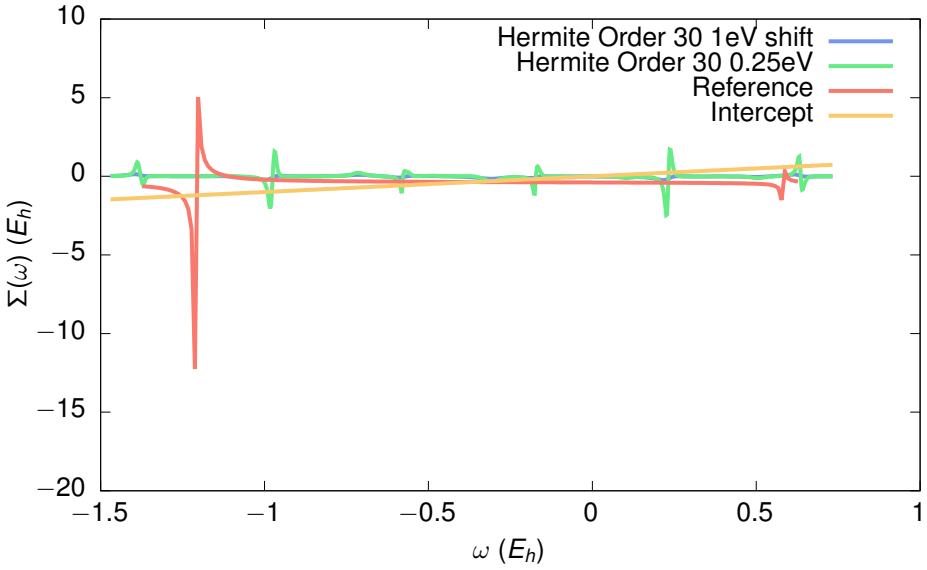

Figure 7.4: The plot shows the real part of the self-energy diagonal matrix element of the HOMO level of a hydrogen molecule. These results where produced using the shifted approach. It can be observed that this ansatz is much more reliable. An increased shift flattens the peaks. The reference data was produced using votca:xtp's G0W0 calculation, that uses a standard *GW* approach. The intercept with the yellow line gives the quasi particle energy of the HOMO level as the intercept gives the fixed point of the quasi particle equation 2.66. It can be observed that the Sternheimer method is able to reproduce the intercept well for both shifts. However, a larger shift leads to a more stable result.

integration is indeed a major issue of the *GW* Sternheimer implementation. Integrating along the real axis only is not feasible. The results are very unstable and even small changes in the integration order or changing the quadrature method lead to vastly different results. Consequently, results obtained that way are not reliable. Using purely imaginary grid points is also not a viable option because the Padé approximation is not able to reproduce the structure of the self-energy on the real axis from the purely imaginary input points. The structure is too smooth and barely visible and all values are very close to zero. Reference data was not included in these plots because the results produced by these approaches are too unreliable to be comparable. Thus, the focus from here on lies on the two other transformation methods: shifting the interval by an imaginary term or integrating along a vertical line and employing the COHSEX approach.

The shifted and vertical approaches produced much more reliable results as can be seen from plots 7.5 and 7.4. Changing parameters of the integration does not affect the results as badly as before. Consequently, it is preferable to use

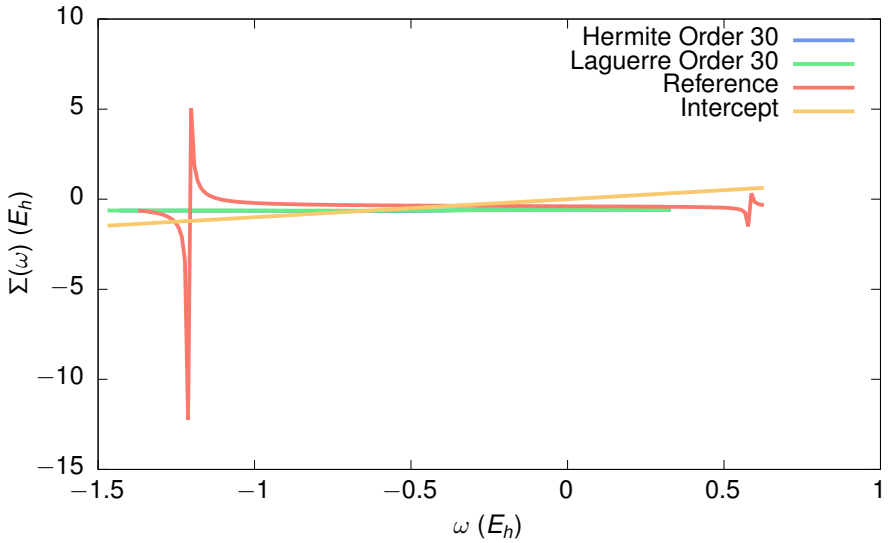

Figure 7.5: The plot shows the real part of the self-energy diagonal matrix element of the HOMO level of a hydrogen molecule. These results where produced using the vertical integration approach. The structure of the reference is not reproduced. However, the intercept that gives the quasi particle energy is very close to the reference. Also the Gauss-Laguerre quadrature could be used here because of the vertical integration. It returns the same results as the Gauss-Hermite quadrature.

one of these two frequency integration approaches. Also the Sternheimer *GW* method is able to reproduce the quasi particle energy with a small error. The quasi particle energy is obtained by computing the fixed point of the quasi particle equation 2.66 which is represented by the intercept with the yellow line in the figures. By convention the intercept that is not part of a pole is chosen. [37] Possible reasons for this error are the numerical integrations performed as particularly the spatial grid is very coarse. The structure of the reference graph is not reproduced. Most likely the attempts to avoid integrating close to the poles also leads to the poles of the self-energy to be smoothen out as well. In topic **2** in the Introduction it was already discussed that a major difference of going from plane waves to localized basis function is the numerical integrations that need to performed. The numerical integrations are very costly to do, particularly the spatial integration because it needs a lot of grid points in combination with the integration being needed to be performed for all combinations of ω_i and ω'_j . Therefore, it is not possible to perform the spatial integration on a finer grid making it a likely source of errors. Further, the vertical integration does not compensate for poles being present inside the integration path presented in figure 6.1. While the results suggest that this does not have a great effect because both vertical and shifted integration lead to similar results, it still is a source of error, that needs to be mentioned. Although the frequency integration was vastly improved by transforming the integration interval, numerical integration, particularly over an infinite interval, can still lead to errors.

The COHSEX approach also leads to a more stable numerical integration along the frequency axis compared to the standard approach. Because the non-analytic part of the Green's function is treated separately, there are no poles of the Green's function close to the real axis that are hard to handle for the numerical integration applied. Thus no transformation of the integration interval needs to be applied to avoid the poles. The COHSEX approach, consequently, provides a way to stabilize the integration without the need to approximate the integral by shifting the interval or by using vertical integration path that needs knowledge of the poles position in order account for the residues of the poles inside the integration path depicted in figure 6.1. However, the results are not as smooth as the results from the transformation approaches. But the COHSEX approach reproduces the structure of the self-energy better that the transformation approaches, that fail to reproduce the poles. However, the quadrature order still needs to be chosen carefully as in the test cases only quadratures of order 30 or higher were reliably able to reproduce the results of the reference.

The best results were achieved by combining the two approaches. Using the COHSEX splitting as well as integrating along a vertical line leads to remarkable stability of the integration as well as an accurate reproduction of the reference data. The first pole can be reproduced with vertical integration because the non-analytic part of the Green's function is treated separately. When not using the COHSEX approach the pole is not visible because the integration path is transformed to be away from the poles of the Green's function and so the pole of the self-energy cannot be reproduced. Additionally the combination of the vertical integration and the split ansatz lead to a very smooth results and an

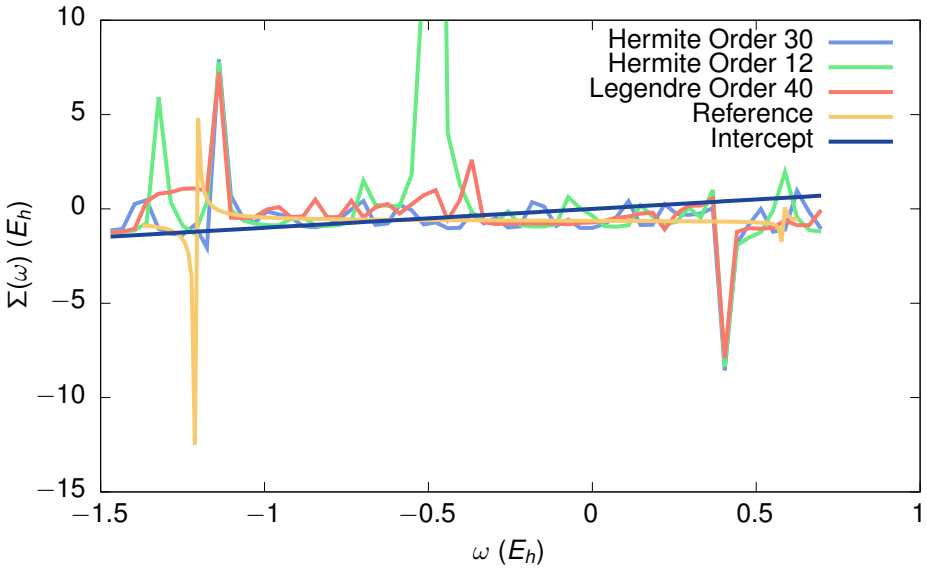

Figure 7.6: The plot shows the real part of the self-energy diagonal matrix element of the HOMO level of a hydrogen molecule. The results where computed using the COHSEX splitting approach and numerical integration along the real axis. It can be observed that compared to figure 7.2 the results are more stable when changing the quadrature. However, when the quadrature order is too small the accuracy of the results suffers as can be seen from the graph of the Hermite quadrature of order 12.

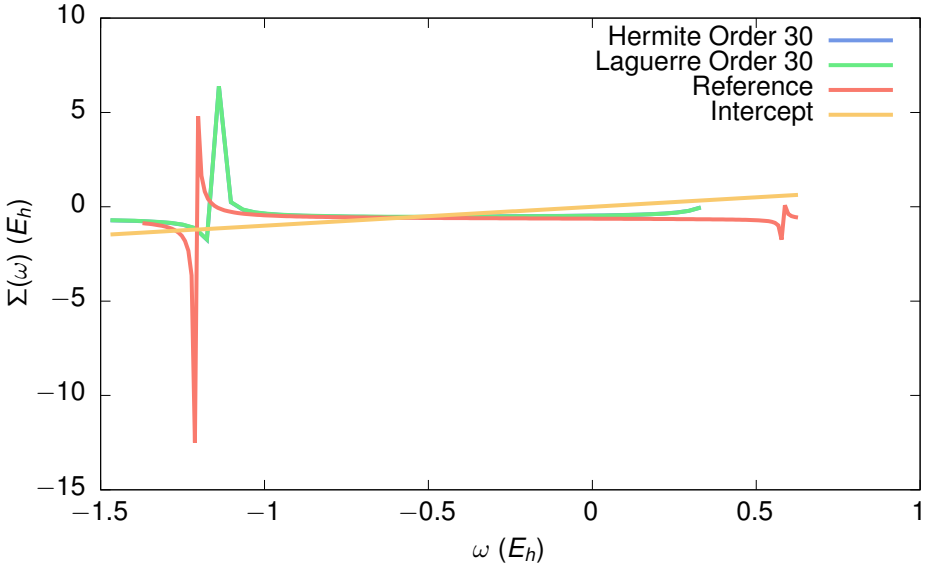

Figure 7.7: The plot shows the real part of the self-energy diagonal matrix element of the HOMO level of a hydrogen molecule. The results were computed by combining the COHSEX and the vertical integration approach. That way, the reference data can be reproduced very well. Especially the intercept is approximated very well as can also be seen from table 1.

Table 7.1: The table shows the quasi particle energy of the HOMO level of a Hydrogen molecule computed with different integration method.

Method used	QP energy (E_h)
Reference	-0.5823
Shift approach	-0.2171
Vertical approach	-0.6313
COHSEX approach	-0.4902
COHSEX vertical	-0.5510

accurate computation of the quasi particle energy.

Table 7.1 shows the quasi particle energies computed from the different approaches analysed in this chapter. The reference data was produced using votca:xtp's G0W0 method, that works along the lines of the approach outlined in chapter 2. It can be seen that the combined approach of COHSEX and vertical integration yields the best result. The purely vertical approach and the COHSEX approach are also close the the reference value but not as good as their combination. The shifted approach performs worst as it only approximates the frequency integration and a shift of at least one eV is necessary to achieve reliable and stable results.

After good results were achieved with the vertical integration COHSEX approach, the different spatial integration methods were tested again to see how the improved frequency integration affects the spatial integration schemes. In this test only the spatial integration of the COH part was altered because for the SEX part no frequency integration needs to be performed, and thus the time consumption of the spatial integration for that part is not as critical.

From figure 7.8 it can be observed that the regular cubic grids do not reproduce the results from the spherical grid. Independent of the number of grid points used the COH part produced from the cubic grid is too small. The results are very close to the SEX part of the self-energy. A possible reason for this is that the cubic grid does not take the structure of the molecule into account and thus wastes grid points in areas where the screened Coulomb potential is very small and, consequently, the grid points are arranged to coarse in areas where a finer grid would be needed. This suggests that the spherical grid is the preferred grid to use for the spatial integration performed for the Sternheimer GW method.

As a last test, the Sternheimer GW method was used to calculate the self energy of water. The quasi particle energy could not be reproduced as well as for hydrogen as the reference data gives $-0.4142\ E_h$, while the GW Sternheimer method yields $0.67\ E_h$. This can be explained by the spatial numerical integration. As figure 7.8 suggests the structure of the graph is mainly determined by the SEX part of the self-energy. The intercept also depends on the COH part, that in turn depends on the spatial integration grid. Consequently, the more complex the molecule is the more crucial a fine grid is. Choosing finer girds however is very costly.

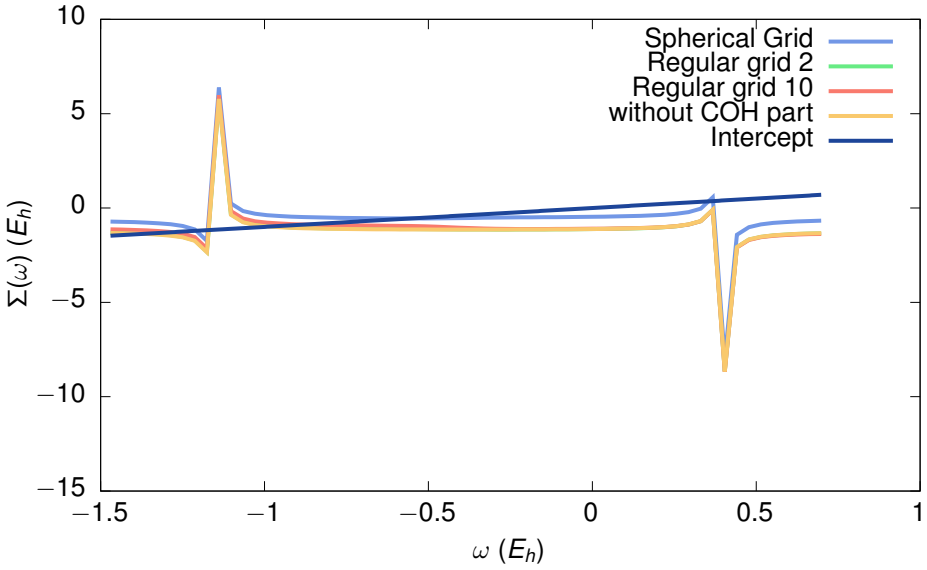

Figure 7.8: The plot shows the real part of the self-energy diagonal matrix element of the HOMO level of a hydrogen molecule. As the spherical grid results are very close to the reference data, no additional reference plot was included. A plot of the SEX part of the self-energy is included. It can be observed that the results obtained using the regular cubic grids are very close to the plot that does not contain the COH part at all suggesting the cubic grids are not well suited for this integration. Also increasing the number of grid points does not compensate for that as the results for a 10x10x10 are very similar to a 2x2x2 grid.

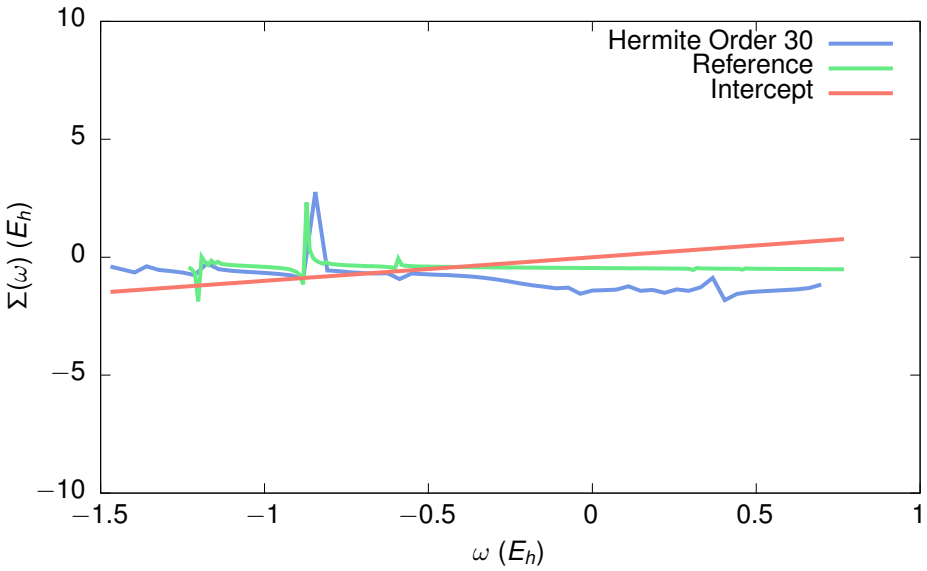

Figure 7.9: The plot shows the real part of the self-energy diagonal matrix element of the HOMO level of a water molecule. The result obtained with the *GW* Sternheimer is compared to the data produced using votca xtp. While the reference data could not be approximated as well as for Hydrogen, the structure of the plot is reproduced well. The intercept at $-0.4142\ E_h$ is only approximated as $0.67\ E_h$.

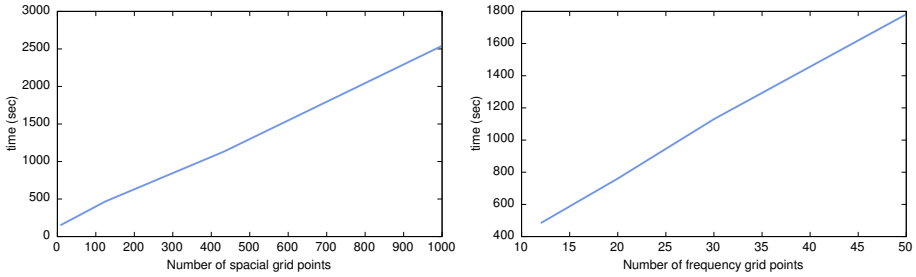

Figure 7.10: The plot shows the time needed to evaluate the *GW* Sternheimer method for Hydrogen for 10 frequencies, depending on the number of spatial grid points needed and the number of frequency grid points. The computation time depends linearly on the the number of grid points. A comparable computation using a standard *GW* approach used to obtain the reference results only took a few seconds.

The plots in figure 7.10 show how the computation time increases with the number of grid points. As expected the time needed to evaluate the *GW* Sternheimer methods increases linearly with the number of frequency and spatial grid points. The frequency grid points are less crucial because they don't need to be increased for larger molecules. The number of spatial grid points however increase with the molecule size simply because the integration needs to be performed over a larger space. The main problem is that even for the smallest number of grid points tested the computation time is greater than for other *GW* methods, that need only a few seconds to compute the self energy and quasi particle energy of hydrogen. Thus the *GW* Sternheimer method is not able to provide a speed up compared to other *GW* methods when using localized basis sets.

In conclusion it was possible to improve the performance and accuracy of the numerical integrations from the initial approaches, but they are still major issues of the *GW* Sternheimer method. As suspected in the Introduction, they form a major bottleneck from a computational workload and from an accuracy perspective. While it was possible to stabilize the frequency integration by using the COHSEX approach combined with transforming the integration interval leading to reliable and accurate results in the test cases, the main problem remains the spatial integration as it needs a large amount of grid points, particularly for larger and more complex molecules. So while it was possible to reproduce the quasi particle energy of a hydrogen molecule, to accurately approximate the quasi particle energy of water, a finer grid is needed. In practice this is not feasible, especially since the desired speed-up could not be achieved. Finer spatial grids would lead to unreasonable computing times as the timing of the code in figure 7.10 suggests.

Chapter 8

Summary and Outlook

The goal of this thesis was to implement the Sternheimer GW method for localized basis sets. Five topics where presented that would need special attention throughout this work because they could be problematic when using localized basis sets. To summarize the results of this thesis, all the crucial topics explained will be revisited and it will be evaluated how well they were handled.

1. **Convergence:** The convergence of the self-consistent Sternheimer cycle is problematic when no mixing methods are used. However, using the Anderson mixing scheme introduced in chapter 4, the self-consistent Sternheimer cycle converged quickly for all test cases.

2. **Integration:** The numerical integration poses the biggest issue of this implementation of the GW Sternheimer method. The numerical integrations are very time consuming. It was possible to improve the computation time and accuracy of the integration from the initial naive approaches. However, while reference results could be reproduced, the computation time needed is still not on par with other GW methods for localized basis sets.

3. **Parameters:** In chapter 5 several parameter sets were compared and it was found that the results were not affected much by changing the mentioned parameters. Although the parameters always need to be chosen careful, the method is robust to minor changes of parameters like the convergence tolerance.

4. **Analytic Continuation:** The Padé-approximation is a useful tool in polarizability tensor calculations, see chapter 5, as one can avoid evaluations on the real axis. For the GW calculations however, in chapter 7 it was shown that analytic continuation of the self-energy evaluated on the imaginary axis onto the real axis did not produce reliable results.

5. **Basis expansion:** Equations to express the self-energy in terms of Gaussian basis function were derived in chapter 6.

All in all, while all topics could be addressed in this thesis, the numerical integration is the major bottleneck of the *GW* Sternheimer method for localized basis sets. From the beginning it was clear that the Sternheimer *GW* method could not be adapted one to one from the plane waves implementation. In this thesis it was verified that the Sternheimer *GW* method is not as well suited for localized basis sets as it is for plane waves. As discussed in chapter 7 the special properties of plane waves allow for a much cheaper treatment of the numerical integrations. [1] While is was possible to reproduce quasi particle energies using the Sternheimer *GW* method for localized basis sets, the tradeoff of the Sternheimer method is not favorable. The advantage of using only occupied orbitals is negated by computational workload caused by the excessive amount of Sternheimer cycle evaluations needed for the numerical integrations over space and frequency. Although the main goal of speeding up *GW* calculation with the Sternheimer method could not be reached, there were still insights gained.

For one it was verified that the *GW* Sternheimer method provides no speed up when using localized basis sets. There are however possibilities that could improve the performance and accuracy of the *GW* Sternheimer method. For example the computations performed in the Sternheimer *GW* method could be optimized even more. It might be possible to save some parts of the integration that are used repeatedly in order to prevent performing the same operation multiple times. Particularly the screened Coulomb interaction does not depend on ω, thus, for each evaluation point the same screened Coulomb potential is calculated. It might be possible to save the screened Coulomb potential so the spatial integration has to be performed less often. Additionally it was observed that the COH part of the self energy is very close to constant with respect to the frequency. Therefore, it might be possible to reduce the number of frequencies for which the COH part is evaluated and only calculate the SEX part for the intermediate frequencies. The COH part could then be used for a set of frequencies, instead of calculating it for each individual frequency. It is questionable however, if the *GW* Sternheimer method can be optimized so far that it is on par with other approaches.

Further the self-consistent Sternheimer method explained in chapter 4 can be used to calculate various different properties. While the method does not perform well for *GW* calculation it can still be used to calculate other properties from linear perturbation theory like inter-molecular forces, electron-phonon coupling terms or the polarizability tensor. [2] The self-consistent Sternheimer method implemented in this thesis is easily adaptable to any kind of perturbation as the only thing that needs to be changed is the input perturbation. The change of the electron density Δn or the self consistent perturbation potential ΔV can then be used to obtain the desired property. Consequently, the self-consistent Sternheimer cycle is a useful tool for different perturbation calculations.

For example the frequency dependent polarizabiliy tensors calculated in chapter 5 could not be calculated previously with the votca-xtp package. Similarly to the process described in chapter 4 for the polarizability tensor, other properties can be calculated. The Sternheimer method can therefore also be

used in future project that involve perturbation calculations.

Bibliography

[1] Feliciano Giustino, Marvin L Cohen, and Steven G Louie. Gw method with the self-consistent sternheimer equation. *Physical Review B*, 81(11):115105, 2010.

[2] Stefano Baroni, Stefano De Gironcoli, Andrea Dal Corso, and Paolo Giannozzi. Phonons and related crystal properties from density-functional perturbation theory. *Reviews of Modern Physics*, 73(2):515, 2001.

[3] Hannes Hübener and Feliciano Giustino. Time-dependent density functional theory using atomic orbitals and the self-consistent sternheimer equation. *Physical Review B*, 89(8):085129, 2014.

[4] Paolo Umari, Edoardo Mosconi, and Filippo De Angelis. Relativistic gw calculations on ch 3 nh 3 pbi 3 and ch 3 nh 3 sni 3 perovskites for solar cell applications. *Scientific reports*, 4:4467, 2014.

[5] Dieter Wöhrle and Dieter Meissner. Organic solar cells. *Advanced Materials*, 3(3):129–138, 1991.

[6] H Derouiche and V Djara. Impact of the energy difference in lumo and homo of the bulk heterojunctions components on the efficiency of organic solar cells. *Solar energy materials and solar cells*, 91(13):1163–1167, 2007.

[7] Albert Messiah. *Quantum mechanics*, volume 2. Elsevier, 1981.

[8] Jean-Luc Bredas. Mind the gap! *Materials Horizons*, 1(1):17–19, 2014.

[9] Klaus Capelle. A bird's-eye view of density-functional theory. *Brazilian Journal of Physics*, 36(4A):1318–1343, 2006.

[10] Dorothea Golze, Marc Dvorak, and Patrick Rinke. The gw compendium: A practical guide to theoretical photoemission spectroscopy. *Frontiers in chemistry*, 7:377, 2019.

[11] Michael Rohlfing and Steven G Louie. Excitonic effects and the optical absorption spectrum of hydrogenated si clusters. *Physical Review Letters*, 80(15):3320, 1998.

[12] Valentin D Mihailetchi, HX Xie, Bert de Boer, LJ Anton Koster, and Paul WM Blom. Charge transport and photocurrent generation in poly (3-hexylthiophene): methanofullerene bulk-heterojunction solar cells. *Advanced Functional Materials*, 16(5):699–708, 2006.

[13] William Shockley and Hans J Queisser. Detailed balance limit of efficiency of p-n junction solar cells. *Journal of applied physics*, 32(3):510–519, 1961.

[14] EA Alsema, P Frankl, and K Kato. Energy pay-back time of photovoltaic energy systems: present status and prospects. 1998.

[15] Tracy H Schloemer, Jeffrey A Christians, Joseph M Luther, and Alan Sellinger. Doping strategies for small molecule organic hole-transport materials: impacts on perovskite solar cell performance and stability. *Chemical science*, 10(7):1904–1935, 2019.

[16] Joseph Kalowekamo and Erin Baker. Estimating the manufacturing cost of purely organic solar cells. *Solar Energy*, 83(8):1224–1231, 2009.

[17] Lars Hedin. New method for calculating the one-particle green's function with application to the electron-gas problem. *Physical Review*, 139(3A):A796, 1965.

[18] Christopher J Cramer and FM Bickelhaupt. Essentials of computational chemistry. *ANGEWANDTE CHEMIE-INTERNATIONAL EDITION IN ENGLISH-*, 42(4):381–381, 2003.

[19] Hannes Hübener, Miguel A Pérez-Osorio, Pablo Ordejón, and Feliciano Giustino. Dielectric screening in extended systems using the self-consistent sternheimer equation and localized basis sets. *Physical Review B*, 85(24):245125, 2012.

[20] Henry AR Lambert. Electronic excitations in semiconductors and insulators using the sternheimer-gw method. 2014.

[21] Daan Frenkel and Berend Smit. *Understanding Molecular Simulation*. Academic Press, San Diego, second edition, 2002.

[22] Douglas Rayner Hartree. The calculation of atomic structures. 1957.

[23] M. Born and R. Oppenheimer. Zur Quantentheorie der Molekeln. *Annalen der Physik (Berlin)*, 389:457–484, 1927.

[24] Robert L Fulton and Martin Gouterman. Vibronic coupling. i. mathematical treatment for two electronic states. *The Journal of Chemical Physics*, 35(3):1059–1071, 1961.

[25] Douglas R Hartree. The wave mechanics of an atom with a non-coulomb central field. part i. theory and methods. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 24, pages 89–110. Cambridge University Press, 1928.

[26] Christopher J Cramer. *Essentials of computational chemistry: theories and models*. John Wiley & Sons, 2013.

[27] John C Slater. The theory of complex spectra. *Physical Review*, 34(10):1293, 1929.

[28] Anthony P Scott and Leo Radom. Harmonic vibrational frequencies: an evaluation of hartree- fock, møller- plesset, quadratic configuration interaction, density functional theory, and semiempirical scale factors. *The Journal of Physical Chemistry*, 100(41):16502–16513, 1996.

[29] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Physical Review*, 136(3):864–871, 1964.

[30] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Physical Review*, 140(4A):A1133–, November 1965.

[31] R.M. Martin. *Electronic Stucture, Basic Theory and Practical Methods*. apud Cambridge, Cambridge, 2004.

[32] Miguel AL Marques, Micael JT Oliveira, and Tobias Burnus. Libxc: A library of exchange and correlation functionals for density functional theory. *Computer physics communications*, 183(10):2272–2281, 2012.

[33] Tjalling Koopmans. Über die zuordnung von wellenfunktionen und eigenwerten zu den einzelnen elektronen eines atoms. *Physica*, 1(1-6):104–113, 1934.

[34] Peter Politzer and Fakher Abu-Awwad. A comparative analysis of hartree-fock and kohn-sham orbital energies. *Theoretical Chemistry Accounts*, 99(2):83–87, 1998.

[35] Ulrike Salzner and Roi Baer. Koopmans' springs to life, 2009.

[36] Eleftherios N Economou. *Green's functions in quantum physics*, volume 7. Springer Science & Business Media, 2006.

[37] F Gygi and A Baldereschi. Quasiparticle energies in semiconductors: Self-energy correction to the local-density approximation. *Physical review letters*, 62(18):2160, 1989.

[38] Stephen L Adler. Quantum theory of the dielectric constant in real solids. *Physical Review*, 126(2):413, 1962.

[39] Nathan Wiser. Dielectric constant with local field effects included. *Physical Review*, 129(1):62, 1963.

[40] RM Sternheimer. Electronic polarizabilities of ions from the hartree-fock wave functions. *Physical Review*, 96(4):951, 1954.

[41] Vladimir Goncharov. *Non-Linear Optical Response in Atoms, Molecules and Clusters: An Explicit Time Dependent Density Functional Approach.* Springer, 2014.

[42] Stefano Baroni, Paolo Giannozzi, and Andrea Testa. Green's-function approach to linear response in solids. *Physical review letters*, 58(18):1861, 1987.

[43] Robert M Pick, Morrel H Cohen, and Richard M Martin. Microscopic theory of force constants in the adiabatic approximation. *Physical Review B*, 1(2):910, 1970.

[44] John C Slater. Atomic shielding constants. *Physical Review*, 36(1):57, 1930.

[45] Simen Reine, Trygve Helgaker, and Roland Lindh. Multi-electron integrals. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2(2):290–303, 2012.

[46] Shigeru Obara and A Saika. Efficient recursive computation of molecular integrals over cartesian gaussian functions. *The Journal of chemical physics*, 84(7):3963–3974, 1986.

[47] BI Dunlap and N Rosch. The gaussian-type orbitals density-functional approach to finite systems. In *Advances in Quantum Chemistry*, volume 21, pages 317–339. Elsevier, 1990.

[48] Peter MW Gill. Molecular integrals over gaussian basis functions. In *Advances in quantum chemistry*, volume 25, pages 141–205. Elsevier, 1994.

[49] A George Jr et al. *Essentials of Padé approximants.* Elsevier, 1975.

[50] Andreas Frommer. Bicgstab () for families of shifted linear systems. *Computing*, 70(2):87–109, 2003.

[51] Duane D Johnson. Modified broyden's method for accelerating convergence in self-consistent calculations. *Physical Review B*, 38(18):12807, 1988.

[52] Russell B Thompson, K O/ Rasmussen, and Turab Lookman. Improved convergence in block copolymer self-consistent field theory by anderson mixing. *The Journal of chemical physics*, 120(1):31–34, 2004.

[53] HJ Vidberg and JW Serene. Solving the eliashberg equations by means ofn-point padé approximants. *Journal of Low Temperature Physics*, 29(3-4):179–192, 1977.

[54] Karin Eichkorn, Oliver Treutler, Holger Öhm, Marco Häser, and Reinhart Ahlrichs. Auxiliary basis sets to approximate coulomb potentials. *Chemical physics letters*, 240(4):283–290, 1995.

[55] Ferdi Aryasetiawan and Olle Gunnarsson. The gw method. *Reports on Progress in Physics*, 61(3):237, 1998.

[56] Marat Valiev, Eric J Bylaska, Niranjan Govind, Karol Kowalski, Tjerk P Straatsma, Hubertus JJ Van Dam, Dunyou Wang, Jarek Nieplocha, Edoardo Apra, Theresa L Windus, et al. Nwchem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Computer Physics Communications*, 181(9):1477–1489, 2010.

[57] Xinguo Ren, Patrick Rinke, Volker Blum, Jürgen Wieferink, Alexandre Tkatchenko, Andrea Sanfilippo, Karsten Reuter, and Matthias Scheffler. Resolution-of-identity approach to hartree–fock, hybrid density functionals, rpa, mp2 and gw with numeric atom-centered orbital basis functions. *New Journal of Physics*, 14(5):053020, 2012.

[58] MeiYue Shao, Lin Lin, Chao Yang, Fang Liu, Felipe H Da Jornada, Jack Deslippe, and Steven G Louie. Low rank approximation in g 0 w 0 calculations. *Science China Mathematics*, 59(8):1593–1612, 2016.