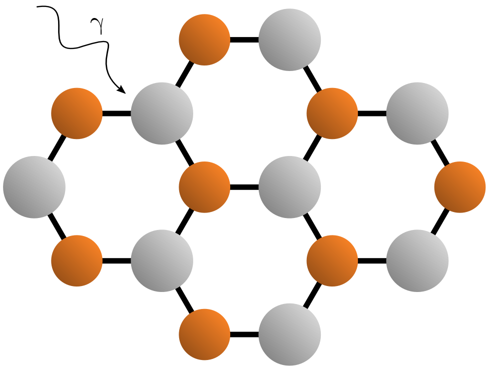
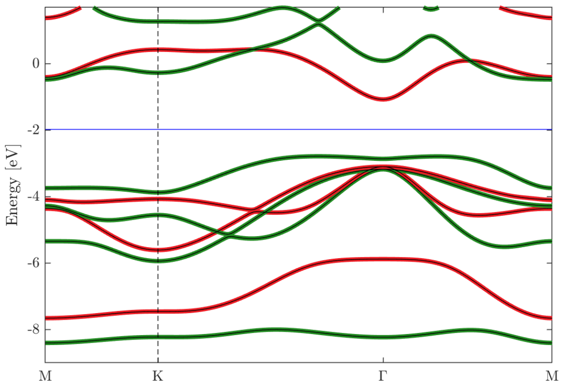
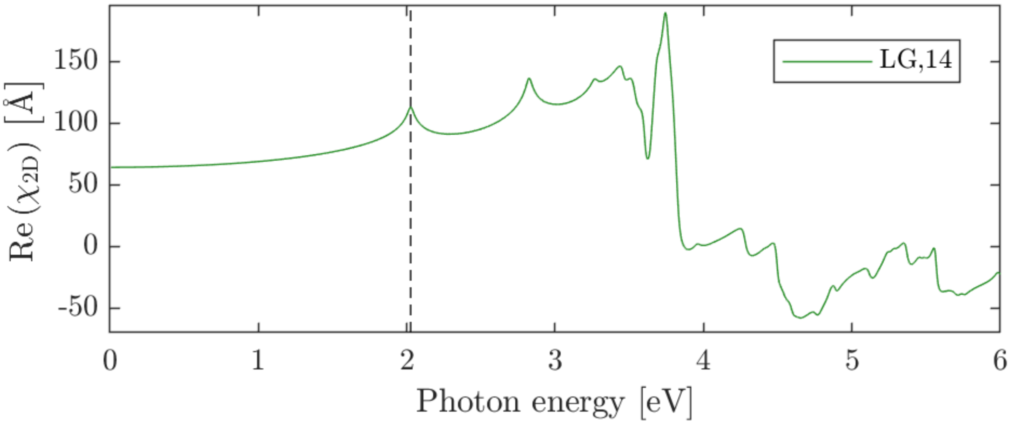
Linear and Nonlinear Optical Properties of Two-Dimensional Crystals: Length and Velocity Gauge Analysis

School of Engineering and Science
Aalborg University
GROUP 5.330
MASTER'S THESIS, 50 ECTS
2020







Aalborg Universitet
FACULTY OF ENGINEERING AND SCIENCE

Title:

Linear and Nonlinear Optical Properties of Two-Dimensional Crystals:
Length and Velocity Gauge Analysis

Project Period:

Master's Thesis
2 September 2019 - June 3 2020

ECTS:

50

Project Group:

5.330

Members of the Group:

Aleksander Bach Lorentzen

Martin Ovesen

Supervisor:

Thomas Garm Pedersen,
Professor, PhD

Number of Pages: 99

Abstract:

In this thesis the theory needed to calculate linear and nonlinear optical properties of crystals are presented with applications to atomically thin materials in mind. This covers (among other things) ab initio tight-binding parametrisation, crystal symmetry considerations, a rigorous introduction of linear and nonlinear single-particle response functions, the Bethe-Salpeter equation, the Rytova-Keldysh potential and expressions for linear and nonlinear excitonic response functions. This enables the description of nonlinear phenomena such as second-harmonic generation and optical rectification.

This theory is applied on mainly monolayer In_2Se_2 and MoS_2. First-order responses are successfully calculated in both the single-particle and excitonic cases. In one of the two approaches to the second-order response, a computational problem is identified. A solution for this is proposed for the the single-particle case, but invalidates one half of the calculated excitonic second-order response.

Comparisons of theory with four experiments found in literature is presented and limited agreement is found using the non-faulty half of the excitonic calculations.

An automatisation procedure to enable large scale calculations on other materials is proposed and implemented with limited success. Automatic symmetry detection and nonzero tensor element calculation is successfully implemented.

Preface

This 50 ECTS master's thesis has been written by two MSc physics students at Aalborg University with the aim of understanding the linear and non-linear optical properties of two-dimensional solids with and without the inclusion of excitonic effects. The thesis is composed as such: We start out in chapter 1 by introducing the subjects which will be studied in the thesis and give motivations for why one these subjects are interesting. Moving on to chapter 2, the necessary theory to model two-dimensional solids is presented. This includes a brief summary of density functional theory, Wannier interpolation and tensorial properties of crystals, but the main body of the theory chapter is concerned with rigorously establishing the single-particle perturbation theory needed to calculate the response of a solid, while the excitonic counterparts are simply introduced. In chapter 3 and chapter 4 the calculations done on two indium selenide compounds and MoS_2 are presented, utilising the theory of chapter 2. In chapter 5 comparisons between the theory and experiment can be found. In chapter 6 we discuss the results and the difficulties encountered during the calculations. Finally we conclude the thesis in chapter 7. After the conclusion is the bibliography. Statements from source material will be referred to as [n] where n refers to the labelled material in the bibliography. Last are the appendices which contains further calculations and other material which was found to be less central to the flow of the thesis.

In order to get the most out of this writing, the reader needs to have an understanding of quantum mechanics, solid state physics, electromagnetism and linear algebra.

.

Aleksander Bach Lorentzen	Martin Ovesen

Contents

Appendices

1 | Introduction

The last 15 years has seen the rise of planar materials, commonly referred to as two-dimensional (2D) materials, with graphene being the first material to be extensively studied because of its outstanding electric and thermal properties.[1] [2] Since then many more materials have come under research with transition metal dichalcogenides (TMDCs) such as $\mathrm{MoS_2}$ and $\mathrm{WSe_2}$ being examples. Many 2D materials have been predicted to exist, and C2DB, a recent high-throughput 2D-material effort, has over 2000 stable 2D materials in its database.[3] Several of these 2D materials have also been fabricated in a laboratory.[4] [5] [6]

With the advent of density functional theory (DFT) around 1990, it has since become much easier to study arbitrary systems without any experimental input, known as *ab initio* calculations.[7] This makes it possible to study these flat structures with an arbitrary number of layers, and the effects of trimming away layers of atoms in a bulk crystal can be predicted. Commonly, the resulting 2D crystal has significantly altered properties, as is the case with the $\mathrm{MoS_2}$ crystal which transitions from an indirect to a direct band gap semiconductor when going from a bulk crystal or even bilayer crystal to a single atomic layer. The 2D geometry also affects how the electrons interact as the usual presence of a polarisable dielectric medium in all directions is replaced by only a thin slab of atoms meaning the electric field lines can circumvent the dielectric medium, and the electrons start to interact with each other much more strongly.[8] Because of the altered electronic properties, these materials also have interesting responses when it interacts with light.

A simple way to categorise the response of a material to an external pertubation is analagous to a Taylor-series, where the response can be thought of as a sum of a set of coefficients multiplied onto the different powers of the size of the perturbation:

$$\text{Response } = a_1 \cdot \text{Perturbation}^1 + a_2 \cdot \text{Perturbation}^2 + \dots \tag{1.1}$$

In the limit of a small perturbations, the only term needed to describe the material will be the linear term which describes a simple interaction where a single photon is absorbed by the crystal.[9] When the electric field which perturbs the crystal is larger than the characteristic electric field $E_{\text{char}} \approx 500$ V/nm, second-order effects become relevant.[9] These are more generally known as nonlinear effects. The two second-order effects which are investigated in this report are second-harmonic generation (SHG), where two photons with the same frequency are absorbed simultaneously such that a frequency-doubled photon can then be emitted, and optical rectification (OR) where a static polarisation is created in the crystal. In effect, this means we will be calculating the two parameters a_1 and a_2 from first principles which in the full analysis with a perturbation on vector-form turn out to be tensor objects acting on the perturbation.

The spatial shape and the symmetries intrinsic to the crystal is also of fundamental importance to its linear and nonlinear responses. Large differences in the nonlinear response can for example be observed when comparing monolayers with bilayers of the same material as the two crystal structures may have a crucial difference in their symmetries. An application of this fact in material characterisation is for gathering information about stacking angles of heterostructures and using the nonlinear response to categorise the number of layers of an atomically thin film into even or odd number of layers.[2] The class of 2D materials has a wide range of band gaps, ranging from graphene's zero band gap to the ~ 7 eV of monolayer hexagonal boron-nitride, meaning light in and around the visible spectrum can potentially be manipulated with these materials.[3] [10]

Numerical frameworks to solve Maxwell's equations with nonlinear terms in it has also been developed.[10] This means that if nonlinear response functions can be calculated ab initio, one can also start modelling nonlinear devices with a wide variety of different materials, using computational power to weed out non-suitable configurations. This means much effort of physically producing and testing devices can be circumvented, but necessitates that accurate modelling of the nonlinear response can be done.

Our calculations will be done on the basis of a tight-binding (TB) model which can be obtained from a DFT-calculation by using what is known as Wannier interpolation.[11] This will allow us to reduce the computational burden related to a DFT-calculation, but it will not come without disadvantages which we will also discuss.

2 | Theory

2.1 Density Functional Theory

The place to start when doing calculations on crystals is to determine the energy levels of the solid. The standard way to do so is to employ density functional theory (DFT) which is a way to reduce the many-body calculation of the ground state energy to a fictitious independent-particle problem, solvable by standard numerical methods. It is possible to do so within the Born-Oppenheimer approximation as the electronic wave function Ψ has the defining property

$$\Psi(..., \mathbf{x}_i, ..., \mathbf{x}_j, ...) = -\Psi(..., \mathbf{x}_j, ..., \mathbf{x}_i, ...)\,. \tag{2.1}$$

Here, the variable $\mathbf{x}_i$ contain both the position $\mathbf{r}_i$ and the spin ς_i of an electron. In order to determine the ground state energy of this wave function, the evaluation of energy

$$E = \left\langle \Psi \middle| \hat{\mathcal{H}}^{\mathrm{MB}} \middle| \Psi \right\rangle = \int \Psi^*(\mathbf{x}_1, ..., \mathbf{x}_N)\, \hat{\mathcal{H}}^{\mathrm{MB}} \Psi(\mathbf{x}_1, ..., \mathbf{x}_2)\, \mathrm{d}^4\mathbf{x}_1....\mathrm{d}^4\mathbf{x}_N \tag{2.2}$$

is needed. The electronic many-body Hamiltonian $\hat{\mathcal{H}}^{\mathrm{MB}}$ is given as[12]

$$\hat{\mathcal{H}}^{\mathrm{MB}} = -\frac{\hbar^2}{2m}\sum_i^N \nabla_i^2 + \sum_i^N v_{\mathrm{lat}}(\mathbf{r}_i) + \frac{e^2}{8\pi\epsilon_0}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \tag{2.3}$$

The three terms can be identified as the kinetic energy, lattice potential and electron repulsion terms, respectively. Because the electrons are indistinguishable, these individual energy terms can be evaluated using only what are known as the one- and two-electron density matrices, defined as[12]

$$\varrho(\mathbf{r}, \mathbf{r}') = N \int \Psi^*(\mathbf{r}', \varsigma_1, \mathbf{x}_2, ...\mathbf{r}_N)\, \Psi(\mathbf{r}, \varsigma_1, \mathbf{x}_2, ..., \mathbf{x}_N)\, \mathrm{d}\varsigma_1 \mathrm{d}^4\mathbf{x}_2...\mathrm{d}^4\mathbf{x}_N, \tag{2.4}$$

$$\varrho_2(\mathbf{r}_1, \mathbf{r}_2) = \frac{N(N-1)}{2} \int \Psi^*(\mathbf{x}_1, \mathbf{x}_2, ...\mathbf{x}_N)\, \Psi(\mathbf{x}_1, \mathbf{x}_2, ...\mathbf{r}_N)\, \mathrm{d}\varsigma_1 \mathrm{d}\varsigma_2 \mathrm{d}^4\mathbf{x}_3...\mathrm{d}^4\mathbf{x}_N. \tag{2.5}$$

From these simpler quantities in comparison to the wave function, the energy (2.2) can be evaluated. Now owing to the Hohenberg-Kohn theorem,[12] a variational formulation of the ground state also exists in terms of the diagonal elements of (2.4). This fact makes it possible to construct a set of independent-particle equations of which the solution reproduces the correct density. Here, the fictitious independent particles move in a effective potential resulting from the presence of the other particles. The independent-particle framework is encapsulated in the Kohn-Sham equations, given as

$$\hat{\mathcal{H}}^{\mathrm{KS}}\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}) \quad \text{with} \quad \hat{\mathcal{H}}^{\mathrm{KS}} = -\frac{\hbar^2}{2m}\nabla^2 + v(\mathbf{r}) + \frac{e^2}{4\pi\epsilon_0}\int \frac{\varrho(\mathbf{r_2})}{|\mathbf{r} - \mathbf{r}_2|}\mathrm{d}\mathbf{r}_2 + v_{\mathrm{XC}}\left[\varrho(\mathbf{r})\right], \tag{2.6}$$

where $\hat{\mathcal{H}}^{\mathrm{KS}}$ is known as the Kohn-Sham Hamiltonian, the ψ_i's are known as the Kohn-Sham orbitals or pseudo-orbitals and v_{XC} is a potential which depends on the diagonal elements of the one-electron density matrix (2.4). It can be thought of as the mean field of all the other particles in the solid. This equation is the main problem to solve in order to get the ground state of the solid. Since it depends on the density itself, which is unknown, it should be solved iteratively until a self-consistent solution is found.

The pseudo wave functions obtained from the Kohn-Sham equations can serve as the single-particle orbitals of the solid. DFT does however have a tendency to underestimate the band gaps of semiconductors and solids.[13] The simplest way to remedy this problem inherent to DFT is to introduce a scissors operator which simply adds a constant energy term between the valence and conduction states. This constant energy term is obtained from C2DB and is the result of a G0W0 calculation.[3]

2.2 Wannier Functions: From DFT Calculation to a TB Model

Doing a DFT calculation on a crystal involves the use of a plane wave basis as this is very well suited for the periodic nature of the electronic density in the crystal. It is, however, relatively computationally intensive to diagonalise the Kohn-Sham Hamiltonian in this basis, meaning it is only feasible for a fairly modest number of $\mathbf{k}$-points. Another option is to use a TB Hamiltonian where the wave function is constructed from a set of orbitals $\omega_m(\mathbf{r}-\mathbf{R})$ which are then repeated in period with the lattice. Such a wave function is on the form[14]

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \sum_m c_m^{n\mathbf{k}} \chi_m(\mathbf{r}) \qquad \text{with} \qquad \chi_n(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{R}} \omega_n(\mathbf{r}-\mathbf{R}), \tag{2.7}$$

where N is the number of lattice cells under consideration, and the sum over $\mathbf{R}$ is considered as a sum over lattice vectors. The advantage of such a representation of the wave function is that the matrix elements of the Hamiltonian in the $\{\chi_n\}$ basis becomes a Bloch sum over the matrix elements of the localised orbitals

$$\mathcal{H}_{nm}^{\mathrm{TB}}(\mathbf{k}) = \sum_{\mathbf{R}} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{R}} \left\langle \omega_n(\mathbf{r}) \middle| \hat{\mathcal{H}}^{\mathrm{KS}} \middle| \omega_m(\mathbf{r}-\mathbf{R}) \right\rangle, \tag{2.8}$$

where the sum can be truncated at a relatively modest amount of lattice vectors $\mathbf{R}$. Converting the wave-functions of N_b bands in a plane-wave basis to a TB representation can be done by applying a well-chosen unitary transformation. The unitary transformation will transform the $\{\psi_{n\mathbf{k}}\}$ basis to some $\{\xi_{n\mathbf{R}}\}$ basis through[11]

$$\xi_{n\mathbf{R}}(\mathbf{r}) = \xi_n(\mathbf{r}-\mathbf{R}) = \frac{A_{\mathrm{UC}}}{(2\pi)^2} \int_{\mathrm{BZ}} \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{R}} \sum_{m=1}^{N_b} U_{mn}^{\mathbf{k}} \psi_{n\mathbf{k}} \mathrm{d}^2\mathbf{k}, \tag{2.9}$$

where A_{UC} is the primitive cell Area and $U^{\mathbf{k}}$ is a unitary matrix, resulting in both mixing between the wave functions and the addition of a $\mathbf{k}$-dependent phase factor. $U^{\mathbf{k}}$ can therefore be chosen in many ways. The argument $\mathbf{r}-\mathbf{R}$ in the left-hand side of (2.9) is justified by the fact that the phase $\mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{R}}$ gives a spacial translation per the shift theorem for the Fourier transform. If the ξ_n of (2.9) were to replace ω_n in (2.7), the fact that the $\mathbf{R}$-sum and the $\mathbf{k}$-integral with respective phase factors are each others' inverse transformations[11] results in $\chi_n(\mathbf{r})$ being a mix of the original $\psi_{n\mathbf{k}}$'s. As the transformations are only unitary, the span of $\{\xi_n\}$ in the resulting basis is the same as the $\{\psi_{n\mathbf{k}}\}$, and the resulting eigenvalues obtained using (2.8) are the same as those obtained from the DFT calculation. The problem has, however, been reduced to a matter of finding matrix elements of the

Hamiltonian in the $\{\xi_n\}$ basis. The significance of this is the aforementioned truncation of the Bloch sum, and the resulting relatively small computing power needed to diagonalise the Hamiltonian in this basis.

The freedom to mix the initial wave functions through $U^{\mathbf{k}}$ is a major advantage as the real space orbitals $\xi_n(\mathbf{r}-\mathbf{R})$ can be chosen such that they are maximally localised in terms of their spread. The spread functional Λ is expressed through[11]

$$\Lambda = \sum_n \left[\left\langle \xi_n(\mathbf{r}) \middle| r^2 \middle| \xi_n(\mathbf{r}) \right\rangle - \left\langle \xi_n(\mathbf{r}) \middle| \mathbf{r} \middle| \xi_n(\mathbf{r}) \right\rangle^2 \right], \tag{2.10}$$

and the ξ_n's are said to form a basis of maximally localised Wannier functions (MLWFs) once $U^{\mathbf{k}}$ has been chosen in an optimal way that minimises Λ. When the localised orbitals have been found, the matrix elements of the Kohn-Sham Hamiltonian can be calculated. The program Wannier90[15] and its interface to Quantum ESPRESSO[16] [17] has been used for this purpose during this project. The initial form of the orbitals the program minimises can be chosen to have s, p, d, and so on. If these are sufficiently "close" to the actual MLWFs, this symmetry will ease the incorporation of a spin-orbit coupling term in the Hamiltonian. Alternatively, the orbitals can be found using the so-called SCDM-**k** method.[18] An automated procedure for finding a TB model using Quantum ESPRESSO and Wannier90 is described in Appendix E.

A final note on the Bloch function χ_n of (2.7) is that one can also change $\mathbf{R} \to \mathbf{R} + \mathbf{r}_i$ where $\mathbf{r}_i$ is the centre of i'th Wannier function $\omega_i(\mathbf{r})$. This is merely equivalent to another unitary transformation, but it has the effect that the expansion coefficients instead represents the lattice-periodic part in $\psi_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r})\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}$.[19]

A TB model can thus describe the electronic band structure of a solid which is a central part of determining the properties of the material in the optical energy regime. It is furthermore also possible to get matrix elements of e.g. momentum by using the $\hat{p}_i \to \frac{m_e}{\hbar}\frac{\partial}{\partial k_i}[\hat{\mathcal{H}}]^{\mathrm{TB}}$ substitution.[20] These matrix elements are of great use when we want to calculate how a system responds to an external perturbation, as we will see in the coming section. Within the scissors-approximation mentioned in the previous section, there is a correction to be made to the matrix-elements of the transitions between the conduction and valence bands in order to keep the oscillator strength consistent with the DFT calculation.[21] In effect this means one can correct the momentum matrix elements as [21]

$$p_{cv}(\mathbf{k}) = p_{cv}^{\mathrm{TB}}(\mathbf{k}) \frac{E_{cv}^{\mathrm{Scissors}}(\mathbf{k})}{E_{cv}^{\mathrm{TB}}(\mathbf{k})}, \tag{2.11}$$

to account for the larger band gap.

2.3 Density Matrix Formalism of Quantum Mechanical Perturbation Theory

This section seeks to establish a density matrix formalism of quantum mechanics which can handle the time-evolution of perturbed electronic states in crystals. This will in the end allow for the calculation of nonlinear response spectra.

2.3.1 The Quantum Liouville Equation

First, consider a quantum system which is known to be in a certain pure state s. The physical properties of this system are determined by a single wave function $\Psi_s(\mathbf{r}, t)$ which must of course obey the time-dependent Schrödinger equation[9]

$$\mathrm{i}\hbar \frac{\partial}{\partial t} \Big|\Psi_s\Big\rangle = \hat{\mathcal{H}} \Big|\Psi_s\Big\rangle . \tag{2.12}$$

For the purposes of the independent-particle approximation (IPA), the total Hamiltonian $\hat{\mathcal{H}}$ is written on the form

$$\hat{\mathcal{H}}(\mathbf{r}, t) = \hat{\mathcal{H}}_0(\mathbf{r}) + \epsilon\hat{\mathcal{V}}(\mathbf{r}, t) , \tag{2.13}$$

i.e. it consists of an unperturbed, time-independent, single-particle Hamiltonian $\hat{\mathcal{H}}_0$ and a time-dependent perturbation energy operator $\hat{\mathcal{V}}$ modulated by a unitless perturbation parameter ϵ which can take values between 0 (no perturbation) and 1 (full perturbation). The eigenstates of $\hat{\mathcal{H}}_0$ are denoted

$$\Big|\psi_m\Big\rangle , \quad \text{for} \quad m = 1, 2, \dots , \tag{2.14}$$

and must fulfil

$$\hat{\mathcal{H}}_0 \Big|\psi_m\Big\rangle = E_m \Big|\psi_m\Big\rangle , \tag{2.15}$$

where E_m is the energy of the m'th state. These are stationary states which form a complete orthonormal basis, hence Ψ_s can be written as a linear combination of ψ_m's where the time-dependence must be carried by the weights (probability amplitudes) $C_{s,m}$, so

$$\Big|\Psi_s\Big\rangle = \sum_m C_{s,m} \Big|\psi_m\Big\rangle \quad \text{with} \quad \Big\langle\psi_n\Big|\psi_m\Big\rangle = \int \psi_n^*(\mathbf{r})\, \psi_m(\mathbf{r})\, \mathrm{d}^3\mathbf{r} = \delta_{nm}. \tag{2.16}$$

Inserting this expression for Ψ into (2.12) then yields

$$\mathrm{i}\hbar \sum_m \frac{\partial C_{s,m}}{\partial t} \Big|\psi_m\Big\rangle = \sum_m C_{s,m}\hat{\mathcal{H}} \Big|\psi_m\Big\rangle , \tag{2.17}$$

which can be multiplied from the left by $\langle\psi_n|$. Orthonormality of the eigenfunctions then ensures that every term in the left-hand sum is eliminated except for the n'th term, so

$$\frac{\partial C_{s,n}}{\partial t} = \frac{1}{\mathrm{i}\hbar} \sum_m C_{s,m}\mathcal{H}_{nm}, \tag{2.18}$$

or alternatively, by taking the complex conjugate

$$\left(\frac{\partial C_{s,n}}{\partial t}\right)^* = \frac{1}{-\mathrm{i}\hbar} \sum_m C_{s,m}^* \mathcal{H}_{nm}^* \implies \frac{\partial C_{s,n}^*}{\partial t} = -\frac{1}{\mathrm{i}\hbar} \sum_m C_{s,m}^* \mathcal{H}_{mn}, \tag{2.19}$$

where the matrix elements of the Hamiltonian have been introduced as

$$\mathcal{H}_{nm} = \Big\langle\psi_n\Big| \hat{\mathcal{H}} \Big|\psi_m\Big\rangle = \int \psi_n^*(\mathbf{r})\, \hat{\mathcal{H}}(\mathbf{r}, t)\, \psi_m(\mathbf{r})\, \mathrm{d}^3\mathbf{r}. \tag{2.20}$$

The matrix elements of any quantum mechanical operator $\hat{\mathcal{O}}$ are defined equivalently such that their expectation value $\langle\mathcal{O}\rangle$ in the state s is given by

$$\langle\mathcal{O}\rangle_s = \left\langle\Psi_s\middle|\,\hat{\mathcal{O}}\,\middle|\Psi_s\right\rangle = \sum_{nm} C_{s,m} C^*_{s,n} \mathcal{O}_{nm}. \tag{2.21}$$

Electrons are fermions which obey the Pauli exclusion principle. They will thus fill up the electronic states starting from the lowest-energy states, and at absolute zero temperature conditions, the occupancy of states will abruptly stop at the Fermi energy. However, in real crystals the atoms will be vibrating due to their thermal energy which may excite electrons to higher energy states, and electrons in higher energy states will naturally relax to vacant states with lower energy. The thermal vibrations of atoms and the (de-)excitations of electrons will induce uncertainties in regards to the shapes and occupancies of quantum states. If the Born-Oppenheimer approximation still holds then the shape of the energy-basis wave functions remain unchanged, and there is only an uncertainty in regards to the occupancy of states. This uncertainty is not quantum mechanical; it is classical in the sense that it reflects a lack of knowledge about the state of the system.

Thus, in the case of nonzero temperature, electrons no longer occupy the previously described pure states; instead, they occupy "mixed states" which are statistical ensembles of all pure states.[9] The usual quantum mechanical expectation value of operators (2.21) is no longer sufficient as well; instead, the ensemble average of $\langle\mathcal{O}\rangle$ must be determined. The ensemble average of a quantity shall be denoted by an overline, and mathematically for the case of expectation values, it is

$$\overline{\langle\mathcal{O}\rangle} = \sum_s p_s(t)\,\langle\mathcal{O}\rangle_s = \sum_s p_s(t) \sum_{nm} C_{s,m} C^*_{s,n} \mathcal{O}_{nm}, \tag{2.22}$$

i.e. it is simply a weighted sum of the $\langle\mathcal{O}\rangle_s$'s over all states s where the weights are the probabilities $p_s(t)$ that the system is in state s.[9]

It is now helpful to define the time-dependent density operator $\hat{\rho}$ and its matrix elements ρ_{nm} by[9]

$$\hat{\rho}(t) = \sum_{mn} \rho_{mn}(t) \left|\psi_m\right\rangle \left\langle\psi_n\right| \quad \text{with} \quad \rho_{mn}(t) = \sum_s p_s(t)\, C_{s,m}(t)\, C^*_{s,n}(t)\,. \tag{2.23}$$

Inserting (2.23) into (2.22) then gives

$$\overline{\langle\mathcal{O}\rangle} = \sum_{mn} \rho_{mn} \mathcal{O}_{nm} = \sum_m \left[\hat{\rho}\hat{\mathcal{O}}\right]_{mm} = \mathrm{tr}\Big(\hat{\rho}\hat{\mathcal{O}}\Big)\,, \tag{2.24}$$

where $[\,\cdot\,]_{mm}$ refers to the m'th diagonal element of the operator inside the brackets, and tr $(\,\cdot\,)$ is the trace of an operator which is a shorthand notation for the sum of all of its diagonal elements. Equation (2.24) shows that if $\hat{\rho}$ is known then the expectation value of any observable can be determined. Additionally, in order to describe the time-evolution of an operator's expectation value, the time-evolution of the density matrix must be determined. The time-derivative of the density matrix' elements are on the form

$$\frac{\mathrm{d}\rho_{mn}}{\mathrm{d}t} = \sum_s \frac{\mathrm{d}p_s(t)}{\mathrm{d}t} C_{s,m} C^*_{s,n} + \sum_s p_s(t) \left(\frac{\mathrm{d}C_{s,m}}{\mathrm{d}t} C^*_{s,n} + C_{s,m} \frac{\mathrm{d}C^*_{s,n}}{\mathrm{d}t} \right). \tag{2.25}$$

For now, the term with the derivatives of p_s shall be disregarded for all s. On the other hand, the derivatives of the probability amplitudes C have forms given by equations (2.18) and (2.19). Inserting these equations and using l as the summation index yields

$$\begin{aligned}\frac{\mathrm{d}\rho_{mn}}{\mathrm{d}t} &= \sum_s \frac{p_s(t)}{\mathrm{i}\hbar} \sum_l \left(C_{s,l} C^*_{s,n} \mathcal{H}_{ml} - C_{s,m} C^*_{s,l} \mathcal{H}_{ln} \right) \\ &= \frac{1}{\mathrm{i}\hbar} \sum_l \left(\mathcal{H}_{ml} \rho_{ln} - \rho_{ml} \mathcal{H}_{ln} \right) = \frac{1}{\mathrm{i}\hbar} \left[\hat{\mathcal{H}}\hat{\rho} - \hat{\rho}\hat{\mathcal{H}} \right]_{mn}, \end{aligned} \tag{2.26}$$

where the second equality stems from the definition of the density matrix elements, and the third equality simply performs the summation over all l. By using the definition of the commutator between operators, we then derive

$$\frac{\mathrm{d}\rho_{mn}}{\mathrm{d}t} = \frac{1}{\mathrm{i}\hbar}\left[\hat{\mathcal{H}}, \hat{\rho}\right]_{mn} \implies \frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = \frac{1}{\mathrm{i}\hbar}\left[\hat{\mathcal{H}}, \hat{\rho}\right], \tag{2.27}$$

which is known as the quantum Liouville equation. [9] This shows that if the density matrix has been determined at some point in time, and if the Hamiltonian is known, the evolution of the system can be mapped.

The effects of nonvanishing $\mathrm{d}p_s(t)/\mathrm{d}t$ can be included heuristically by subtracting a phenomenological damping term from each entrance of the quantum Liouville equation. The damping term can be written as $\eta_{mn}\left(\rho_{mn} - \rho_{mn}^{(\mathrm{eq})}\right)$ where η_{mn} is a tunable damping factor. This term ensures that the density matrix will approach the unperturbed equilibrium density matrix $\hat{\rho}^{(\mathrm{eq})}$ as the time t approaches $-\infty$. For the sake of simplicity, the η_{mn}'s are assumed to remain the same for every combination of m and n, so $\eta_{mn} = \eta$. The quantum Liouville equation shall then be

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \frac{1}{\mathrm{i}\hbar}\left[\hat{\mathcal{H}}, \hat{\rho}\right] - \eta\left(\hat{\rho} - \hat{\rho}^{(\mathrm{eq})}\right). \tag{2.28}$$

The Hamiltonian from (2.13) shall now be inserted into (2.28) to give the elements

$$\begin{aligned}\frac{\mathrm{d}\rho_{mn}}{\mathrm{d}t} =& \frac{1}{\mathrm{i}\hbar}\left[\hat{\mathcal{H}}_0 + \epsilon\hat{\mathcal{V}}, \hat{\rho}\right]_{mn} - \eta\left(\rho_{mn} - \rho_{mn}^{(\mathrm{eq})}\right) \\ =& \frac{1}{\mathrm{i}\hbar}\left[\hat{\mathcal{H}}_0, \hat{\rho}\right]_{mn} + \frac{\epsilon}{\mathrm{i}\hbar}\left[\hat{\mathcal{V}}, \hat{\rho}\right]_{mn} - \eta\left(\rho_{mn} - \rho_{mn}^{(\mathrm{eq})}\right).\end{aligned} \tag{2.29}$$

The commutator for the unperturbed Hamiltonian can be reduced greatly by using the fact that the ψ_m's are orthonormal eigenfunctions of $\hat{\mathcal{H}}_0$, so its matrix elements are

$$\mathcal{H}_{0,nm} = \left\langle \psi_n \middle| \hat{\mathcal{H}}_0 \middle| \psi_m \right\rangle = E_m \left\langle \psi_n \middle| \psi_m \right\rangle = E_m \delta_{nm}, \tag{2.30}$$

such that

$$\left[\hat{\mathcal{H}}_0, \hat{\rho}\right]_{mn} = \sum_l \left(\mathcal{H}_{0,ml}\rho_{ln} - \rho_{ml}\mathcal{H}_{0,ln}\right) = \left(E_m - E_n\right)\rho_{mn} = E_{mn}\rho_{mn}, \tag{2.31}$$

where the notation $E_{mn} = E_m - E_n$ has been adopted. This identity then simplifies the quantum Liouville equation to

$$\frac{\mathrm{d}\rho_{mn}}{\mathrm{d}t} = \frac{E_{mn}}{\mathrm{i}\hbar}\rho_{mn} + \frac{\epsilon}{\mathrm{i}\hbar}\left[\hat{\mathcal{V}}, \hat{\rho}\right]_{mn} - \eta\left(\rho_{mn} - \rho_{mn}^{(\mathrm{eq})}\right). \tag{2.32}$$

For the purposes of this thesis, finding analytic expressions for ρ_{mn} which solve the quantum Liouville equation is unrealistic, if not impossible, so certain approximations must be made.

2.3.2 Perturbing the System

Each matrix element of the density operator shall now be written as a power series in the perturbation parameter ϵ, so

$$\rho_{mn} = \rho_{mn}^{(0)} + \epsilon\rho_{mn}^{(1)} + \epsilon^2\rho_{mn}^{(2)} + \ldots, \tag{2.33}$$

where the steady-state density matrix $\rho_{mn}^{(0)} = \rho_{mn}^{(\mathrm{eq})}$. Using such an expression for ρ_{mn} in (2.32) produces an equation that must be satisfied for every ϵ. Equivalently, this means that each of the coefficients in front of a particular order of ϵ in this equation must separately satisfy the equation. The equations established can quickly be seen from (2.32) to be on the form

$$\begin{aligned} \frac{\mathrm{d}\rho_{mn}^{(0)}}{\mathrm{d}t} &= \frac{E_{mn}}{\mathrm{i}\hbar}\rho_{mn}^{(0)}, \\ \frac{\mathrm{d}\rho_{mn}^{(M)}}{\mathrm{d}t} &= \left(\frac{E_{mn}}{\mathrm{i}\hbar} - \eta\right)\rho_{mn}^{(M)} + \frac{1}{\mathrm{i}\hbar}\left[\hat{\mathcal{V}}, \hat{\rho}^{(M-1)}\right]_{mn} \quad \text{for} \quad M = 1, 2, \dots \end{aligned} \tag{2.34}$$

Going forward, the notation $E'_{mn} = E_{mn} - \mathrm{i}\hbar\eta$ shall be adopted to reduce the size of the equations. In the case of no perturbation with $\epsilon = 0$, the system will be in thermal equilibrium with $\rho_{mn} = \rho_{mn}^{(0)}$. As previously mentioned, some electrons are thermally excited to states above the Fermi level E_F; however, it can be assumed that thermal excitations cannot create a coherent superposition of states, so we can write $\rho_{mn}^{(0)} = 0$ for $m \neq n$.[9] The fermionic electrons will occupy states in accordance with the Fermi-Dirac distribution defined as

$$f(E_m) = f_m = \left(\exp\left(\frac{E_m - E_F}{k_B T}\right) + 1\right)^{-1}, \tag{2.35}$$

so $\rho_{mm}^{(0)} = f_m$.[22] This means that the zero-order equation is constantly 0 which is also expected since it should describe stationary states in thermal equilibrium. On the other hand, the M'th-order equations can be solved by guessing that the solution should be on the form

$$\rho_{mn}^{(M)}(t) = S_{mn}^{(M)}(t)\exp\left(\frac{E'_{mn}t}{\mathrm{i}\hbar}\right), \tag{2.36}$$

with an unknown function $S_{mn}^{(M)}$.[9] Differentiating this function with respect to time results in

$$\frac{\mathrm{d}\rho_{mn}^{(M)}}{\mathrm{d}t} = \frac{E'_{mn}}{\mathrm{i}\hbar}\rho_{mn}^{(M)} + \frac{\mathrm{d}S_{mn}^{(M)}}{\mathrm{d}t}\exp\left(\frac{E'_{mn}t}{\mathrm{i}\hbar}\right). \tag{2.37}$$

Combining (2.34) and (2.37) and isolating the derivative of S gives

$$\begin{gathered} \frac{\mathrm{d}S_{mn}^{(M)}}{\mathrm{d}t} = \frac{1}{\mathrm{i}\hbar}\left[\hat{\mathcal{V}}, \hat{\rho}^{(M-1)}\right]_{mn}\exp\left(\frac{\mathrm{i}E'_{mn}t}{\hbar}\right) \\ \Downarrow \\ S_{mn}^{(M)}(t) = \int_{-\infty}^{t}\frac{1}{\mathrm{i}\hbar}\left[\hat{\mathcal{V}}(t'), \hat{\rho}^{(M-1)}(t')\right]_{mn}\exp\left(\frac{\mathrm{i}E'_{mn}t'}{\hbar}\right)\mathrm{d}t'. \end{gathered} \tag{2.38}$$

It then follows that the first-order equation has a solution on the form

$$\rho_{mn}^{(M)}(t) = \frac{1}{\mathrm{i}\hbar}\int_{-\infty}^{t}\left[\hat{\mathcal{V}}(t'), \hat{\rho}^{(M-1)}(t')\right]_{mn}\exp\left(\frac{\mathrm{i}E'_{mn}}{\hbar}(t'-t)\right)\mathrm{d}t' \quad \text{for} \quad M = 1, 2, \dots, \tag{2.39}$$

i.e. every solution can be determined iteratively from the previous order.

2.3.3 First- and Second-Order Solutions with Time-Harmonic Perturbation

This section shall seek to further reduce the expressions in (2.39) in the case of time-harmonic perturbations with perturbation energy operators on the form

$$\hat{\mathcal{V}}(t) = \frac{1}{2}\sum_{\mathrm{p}}\hat{\mathcal{V}}(\omega_{\mathrm{p}})\exp(-\mathrm{i}\omega_{\mathrm{p}}t), \tag{2.40}$$

i.e. as a sum over time-harmonic terms with frequencies ω_{p}. For the case of $M = 1$, the commutator in (2.39) is expanded, and the newly defined expression for $\hat{\mathcal{V}}$ is inserted, so

$$\left[\hat{\mathcal{V}}(t'), \hat{\rho}^{(0)}(t')\right]_{mn} = \frac{1}{2}\sum_{\mathrm{p}}\left[\hat{\mathcal{V}}(\omega_{\mathrm{p}})\,\hat{\rho}^{(0)}(t') - \hat{\rho}^{(0)}(t')\,\hat{\mathcal{V}}(\omega_{\mathrm{p}})\right]_{mn}\exp(-\mathrm{i}\omega_{\mathrm{p}}t'). \tag{2.41}$$

Expanding the matrix element and adopting the notation $\langle\psi_m|\,\hat{\mathcal{V}}(\omega)\,|\psi_n\rangle = \mathcal{V}_{mn}(\omega)$ results in

$$\left[\hat{\mathcal{V}}(t'),\hat{\rho}^{(0)}(t')\right]_{mn} = \frac{1}{2}\sum_{\mathrm{p}}\sum_{l}\Big(\mathcal{V}_{ml}(\omega_{\mathrm{p}})\,\rho^{(0)}_{ln}(t') - \rho^{(0)}_{ml}(t')\,\mathcal{V}_{ln}(\omega_{\mathrm{p}})\Big)\exp(-\mathrm{i}\omega_{\mathrm{p}}t'). \tag{2.42}$$

Now as was discussed in subsection 2.3.2, the zero-order density matrix elements are time-independent and on the form $\rho^{(0)}_{mn} = \delta_{mn}f_m$. This simplifies the commutator to

$$\left[\hat{\mathcal{V}}(t'),\hat{\rho}^{(0)}(t')\right]_{mn} = \frac{f_{nm}}{2}\sum_{\mathrm{p}}\mathcal{V}_{mn}(\omega_{\mathrm{p}})\exp(-\mathrm{i}\omega_{\mathrm{p}}t'), \tag{2.43}$$

where $f_{nm} = f_n - f_m$. This expression for the commutator can now be inserted into (2.39) to give

$$\rho^{(1)}_{mn}(t) = \frac{\exp(-\mathrm{i}E'_{mn}t/\hbar)}{2\mathrm{i}\hbar}f_{nm}\sum_{\mathrm{p}}\mathcal{V}_{mn}(\omega_{\mathrm{p}})\int_{-\infty}^{t}\exp\left(\mathrm{i}\left(\frac{E'_{mn}}{\hbar}-\omega_{\mathrm{p}}\right)t'\right)\mathrm{d}t'. \tag{2.44}$$

This integral converges solely due to the presence of the damping factor inside E'_{mn}. It will effectively multiply a factor $\mathrm{e}^{\eta t'}$ onto the integrand, forcing it to 0 as $t' \to -\infty$. Evaluating the integral then allows us to write

$$\begin{aligned}\rho^{(1)}_{mn}(t) &= \frac{\exp(-\mathrm{i}E'_{mn}t/\hbar)}{2\mathrm{i}\hbar}f_{nm}\sum_{\mathrm{p}}\mathcal{V}_{mn}(\omega_{\mathrm{p}})\,\frac{\exp(\mathrm{i}(E'_{mn}/\hbar-\omega_{\mathrm{p}})\,t)}{\mathrm{i}(E'_{mn}/\hbar-\omega_{\mathrm{p}})}\\ &= \frac{f_{nm}}{2}\sum_{\mathrm{p}}\frac{\mathcal{V}_{mn}(\omega_{\mathrm{p}})}{\hbar\omega_{\mathrm{p}}-E'_{mn}}\exp(-\mathrm{i}\omega_{\mathrm{p}}t).\end{aligned} \tag{2.45}$$

For the second-order density matrix elements, the commutator in (2.39) with $M = 2$ can be similarly expanded with the time-harmonic perturbation to give

$$\left[\hat{\mathcal{V}}(t'),\hat{\rho}^{(1)}(t')\right]_{mn} = \frac{1}{2}\sum_{\mathrm{q}}\sum_{l}\Big(\mathcal{V}_{ml}(\omega_{\mathrm{q}})\,\rho^{(1)}_{ln}(t') - \rho^{(1)}_{ml}(t')\,\mathcal{V}_{ln}(\omega_{\mathrm{q}})\Big)\exp(-\mathrm{i}\omega_{\mathrm{q}}t'). \tag{2.46}$$

The expression for the first-order density matrix can then be inserted to derive

$$\begin{aligned}\left[\hat{\mathcal{V}}(t'),\hat{\rho}^{(1)}(t')\right]_{mn} = \frac{1}{4}\sum_{\mathrm{p,q}}\sum_{l}\Bigg(&\frac{f_{nl}\mathcal{V}_{ln}(\omega_{\mathrm{p}})\,\mathcal{V}_{ml}(\omega_{\mathrm{q}})}{\hbar\omega_{\mathrm{p}}-E'_{ln}}\\ &-\frac{f_{lm}\mathcal{V}_{ml}(\omega_{\mathrm{p}})\,\mathcal{V}_{ln}(\omega_{\mathrm{q}})}{\hbar\omega_{\mathrm{p}}-E'_{ml}}\Bigg)\exp(-\mathrm{i}(\omega_{\mathrm{p}}+\omega_{\mathrm{q}})\,t').\end{aligned} \tag{2.47}$$

It can readily be seen that the commutator's time-dependence lies entirely in the complex exponential. Using this expression for the commutator in (2.39) with $M = 2$ produces an integral which again converges only due to the damping term in the E'_{mn}'s. The integral is then solved, and the equation rearranged, to give the following expression for the second-order density matrix elements:

$$\rho^{(2)}_{mn}(t) = \frac{1}{4}\sum_{\mathrm{p,q}}\frac{\exp(-\mathrm{i}(\omega_{\mathrm{p}}+\omega_{\mathrm{q}})\,t)}{\hbar\omega_{\mathrm{p}}+\hbar\omega_{\mathrm{q}}-E'_{mn}}\sum_{l}\left(\frac{f_{nl}\mathcal{V}_{ln}(\omega_{\mathrm{p}})\,\mathcal{V}_{ml}(\omega_{\mathrm{q}})}{\hbar\omega_{\mathrm{p}}-E'_{ln}}-\frac{f_{lm}\mathcal{V}_{ml}(\omega_{\mathrm{p}})\,\mathcal{V}_{ln}(\omega_{\mathrm{q}})}{\hbar\omega_{\mathrm{p}}-E'_{ml}}\right). \tag{2.48}$$

Hamiltonians with Second-Order Perturbation Terms

For this derivation, it was assumed that the Hamiltonian was on the form (2.13); however, it is also of interest to consider Hamiltonians on the form

$$\hat{\mathcal{H}}(\mathbf{r},t) = \hat{\mathcal{H}}_0(\mathbf{r}) + \epsilon\hat{\mathcal{V}}^{(1)}(\mathbf{r},t) + \epsilon^2\hat{\mathcal{V}}^{(2)}(\mathbf{r},t)\,. \tag{2.49}$$

For such Hamiltonians, the quantum Liouville equation will instead yield

$$\begin{aligned}\frac{\mathrm{d}\rho^{(0)}_{mn}}{\mathrm{d}t} &= \frac{E_{mn}}{\mathrm{i}\hbar}\rho^{(0)}_{mn},\\ \frac{\mathrm{d}\rho^{(1)}_{mn}}{\mathrm{d}t} &= \frac{E'_{mn}}{\mathrm{i}\hbar}\rho^{(1)}_{mn} + \frac{1}{\mathrm{i}\hbar}\left[\hat{\mathcal{V}}^{(1)},\hat{\rho}^{(0)}\right]_{mn},\\ \frac{\mathrm{d}\rho^{(M)}_{mn}}{\mathrm{d}t} &= \frac{E'_{mn}}{\mathrm{i}\hbar}\rho^{(M)}_{mn} + \frac{1}{\mathrm{i}\hbar}\left[\hat{\mathcal{V}}^{(1)},\hat{\rho}^{(M-1)}\right]_{mn} + \frac{1}{\mathrm{i}\hbar}\left[\hat{\mathcal{V}}^{(2)},\hat{\rho}^{(M-2)}\right]_{mn} \quad \text{for} \quad M = 2,3,\dots,\end{aligned} \tag{2.50}$$

and by following the same procedure as previously, the solutions are found to be

$$\begin{aligned}\rho^{(0)}_{mn}(t) &= 0,\\ \rho^{(1)}_{mn}(t) &= \frac{1}{\mathrm{i}\hbar}\int_{-\infty}^{t}\left[\hat{\mathcal{V}}^{(1)}(t'),\hat{\rho}^{(0)}(t')\right]_{mn}\exp\!\left(\frac{\mathrm{i}E'_{mn}}{\hbar}(t'-t)\right)\mathrm{d}t',\\ \rho^{(M)}_{mn}(t) &= \frac{1}{\mathrm{i}\hbar}\int_{-\infty}^{t}\left(\left[\hat{\mathcal{V}}^{(1)}(t'),\hat{\rho}^{(M-1)}(t')\right]_{mn}\right.\\ &\qquad\left.+\left[\hat{\mathcal{V}}^{(2)}(t'),\hat{\rho}^{(M-2)}(t')\right]_{mn}\right)\exp\!\left(\frac{\mathrm{i}E'_{mn}}{\hbar}(t'-t)\right)\mathrm{d}t' \quad \text{for} \quad M = 2,3,\dots\end{aligned} \tag{2.51}$$

These solutions are very similar to those found for Hamiltonians of the previous type, although more care needs to be taken in regards to the order.

Expressions for the first- and second-order density matrix elements shall be written for perturbation energy operators on the form

$$\hat{\mathcal{V}}^{(1)}(t) = \frac{1}{2}\sum_{\mathrm{p}}\hat{\mathcal{V}}^{(1)}(\omega_{\mathrm{p}})\exp(-\mathrm{i}\omega_{\mathrm{p}}t) \tag{2.52}$$

and

$$\hat{\mathcal{V}}^{(2)}(t) = \frac{1}{4}\sum_{\mathrm{p,q}}\hat{\mathcal{V}}^{(2)}(\omega_{\mathrm{p}},\omega_{\mathrm{q}})\exp(-\mathrm{i}(\omega_{\mathrm{p}}+\omega_{\mathrm{q}})\,t). \tag{2.53}$$

The commutators involving $\hat{\mathcal{V}}^{(1)}$ have essentially already been determined previously, and since $\rho^{(1)}_{mn}$ only needs $\hat{\mathcal{V}}^{(1)}$, the first-order density matrix elements are on the exact same form

$$\rho^{(1)}_{mn}(t) = \frac{f_{nm}}{2}\sum_{\mathrm{p}}\frac{\mathcal{V}^{(1)}_{mn}(\omega_{\mathrm{p}})}{\hbar\omega_{\mathrm{p}} - E'_{mn}}\exp(-\mathrm{i}\omega_{\mathrm{p}}t). \tag{2.54}$$

To establish an expression for $\rho^{(2)}_{mn}$, the last thing we need is the commutator between $\hat{\mathcal{V}}^{(2)}$ and $\rho^{(0)}_{mn}$. This can quickly be seen to be

$$\left[\hat{\mathcal{V}}^{(2)}(t'),\hat{\rho}^{(0)}(t')\right]_{mn} = \frac{f_{nm}}{4}\sum_{\mathrm{p,q}}\mathcal{V}^{(2)}_{mn}(\omega_{\mathrm{p}},\omega_{\mathrm{q}})\exp(-\mathrm{i}(\omega_{\mathrm{p}}+\omega_{\mathrm{q}})\,t'), \tag{2.55}$$

and as such the second-order density matrix elements can be determined through

$$\begin{aligned}\rho^{(2)}_{mn}(t) = \frac{1}{4}\sum_{\mathrm{p,q}}\frac{\exp(-\mathrm{i}(\omega_{\mathrm{p}}+\omega_{\mathrm{q}})\,t)}{\hbar\omega_{\mathrm{p}}+\hbar\omega_{\mathrm{q}}-E'_{mn}}&\left\{f_{nm}\mathcal{V}^{(2)}_{mn}(\omega_{\mathrm{p}},\omega_{\mathrm{q}})\right.\\ &\left.+\sum_{l}\left(\frac{f_{nl}\mathcal{V}^{(1)}_{ln}(\omega_{\mathrm{p}})\,\mathcal{V}^{(1)}_{ml}(\omega_{\mathrm{q}})}{\hbar\omega_{\mathrm{p}}-E'_{ln}} - \frac{f_{lm}\mathcal{V}^{(1)}_{ml}(\omega_{\mathrm{p}})\,\mathcal{V}^{(1)}_{ln}(\omega_{\mathrm{q}})}{\hbar\omega_{\mathrm{p}}-E'_{ml}}\right)\right\}.\end{aligned} \tag{2.56}$$

2.4 Tensors in Crystals

This section shall explain the role and properties of tensors in crystals.

For the purposes of this text, tensors shall represent arrays which contain weights that can relate a perturbation $\mathbf{A}$ to a certain response $\mathbf{B}$. For example a perturbation could be an electric field incident on a crystal and the response would be an induced current. This section shall cover the cases where the perturbations and responses are three-dimensional Cartesian vectors, so

$$\mathbf{A} = \begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix} \qquad \text{and} \qquad \mathbf{B} = \begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix} . \tag{2.57}$$

Generally, $\mathbf{B}$ can be written as a power series of a dimensionless perturbation parameter ϵ as was done with the density matrix in section 2.3, so

$$\mathbf{B} = \mathbf{B}^{(0)} + \epsilon \mathbf{B}^{(1)} + \epsilon^2 \mathbf{B}^{(2)} + \dots, \tag{2.58}$$

where $\mathbf{B}^{(0)}$ is the zero-order response interpreted as a vector with entrances that are independent of the perturbation $\mathbf{A}$. On the other hand, the first-order response $\mathbf{B}^{(1)}$ is a weighted sum over all the terms which contain *one* instance of an entrance of $\mathbf{A}$. The entrances of $\mathbf{B}^{(1)}$ can thus be written as

$$B_\lambda^{(1)} = \sum_{\alpha=x,y,z} T_{\lambda\alpha} A_\alpha \qquad \text{for} \qquad \lambda = x, y, z. \tag{2.59}$$

There are 9 combinations of λ and α and thus 9 weights are necessary. These weights can be collected in a 3×3 matrix as

$$\overleftrightarrow{T}^{[2]} = \begin{bmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{yx} & T_{yy} & T_{yz} \\ T_{zx} & T_{zy} & T_{zz} \end{bmatrix} , \tag{2.60}$$

where the superscripted [2] refers to $\overleftrightarrow{T}^{[2]}$ being a two-dimensional array or a tensor of rank 2.

Next, consider the second-order effect $\mathbf{B}^{(2)}$. It contains a weighted sum over all the terms which have *two* instances of entrances of $\mathbf{A}$, so we can write

$$B_\lambda^{(2)} = \sum_{\alpha,\beta=x,y,z} T_{\lambda\alpha\beta} A_\alpha A_\beta \qquad \text{for} \qquad \lambda = x, y, z. \tag{2.61}$$

There are 27 combinations of λ, α and β for the second-order effect, and as such there are 27 weights which can be collected in a $3 \times 3 \times 3$ array as

$$\overleftrightarrow{T}^{[3]} = \left[\begin{array}{ccc|ccc|ccc} T_{xxx} & T_{xxy} & T_{xxz} & T_{yxx} & T_{yxy} & T_{yxz} & T_{zxx} & T_{zxy} & T_{zxz} \\ T_{xyx} & T_{xyy} & T_{xyz} & T_{yyx} & T_{yyy} & T_{yyz} & T_{zyx} & T_{zyy} & T_{zyz} \\ T_{xzx} & T_{xzy} & T_{xzz} & T_{yzx} & T_{yzy} & T_{yzz} & T_{zzx} & T_{zzy} & T_{zzz} \end{array} \right] , \tag{2.62}$$

where $\overleftrightarrow{T}^{[3]}$ is a tensor of rank 3. This pattern continues, so the third-order response is determined from a tensor of rank 4 and so on.

Now, without specifying which perturbation and response that is considered, some general statements can be made for the elements of their tensors. Since the tensor relates two inherently physical quantities, any choice of a coordinate system should be valid. Because the representation of these physical quantities may differ, the tensor $\overleftrightarrow{T}$ has to transform accordingly. For the purposes of the crystals considered in this thesis, it is sufficient to consider only linear transformations between coordinates on the form $\mathbf{r}' = \hat{S}\mathbf{r}$ in three dimensional space. Here, $\mathbf{r}$ and $\mathbf{r}'$ are Cartesian position vectors while $\hat{S}$ is an operator relating the two coordinate systems. The general prescription for how the components of a tensor transforms under any coordinate transformation is given as[23]

$$T'_{\lambda\mu\dots\nu} = \sum_{\alpha,\beta,\dots,\gamma} \frac{\partial r_\alpha}{\partial r'_\lambda}\frac{\partial r_\beta}{\partial r'_\mu}\cdots\frac{\partial r_\gamma}{\partial r'_\nu} T_{\alpha\beta\dots\gamma} \qquad \text{for} \qquad \lambda, \mu, \dots, \nu = x, y, z. \tag{2.63}$$

Here, the left-hand side has the tensor in the primed coordinates while the right-hand side sums over the dummy indices $\alpha, \beta, ..., \gamma$. The formula above reduces to a linear system when only linear transformations of coordinates are considered. In such a case, the elements of $\hat{S}$ can be written as $S_{\lambda\alpha} = \partial r_\alpha / \partial r'_\lambda$, so[24]

$$T'_{\lambda\mu\dots\nu} = \sum_{\alpha,\beta,\dots,\gamma} S_{\lambda\alpha}S_{\mu\beta}\cdots S_{\nu\gamma}T_{\alpha\beta\dots\gamma} \qquad \text{for} \qquad \lambda, \mu, \dots, \nu = x, y, z. \tag{2.64}$$

Furthermore, the environment in which the physics takes place may have some symmetries, e.g. a crystal. If $\hat{S}$ is symmetry operator of the crystal, i.e. if the crystal is symmetric in regards to the transformation represented by $\hat{S}$, then the elements of the tensor should also be symmetric under $\hat{S}$, otherwise the relation between perturbation and response would be different when observing physically equivalent systems. Lastly, this means that $T'_{\lambda\mu\dots\nu} = T_{\lambda\mu\dots\nu}$, and the elements of the tensor must then obey

$$T_{\lambda\mu\dots\nu} = \sum_{\alpha,\beta,\dots,\gamma} S_{\lambda\alpha}S_{\mu\beta}\cdots S_{\nu\gamma}T_{\alpha\beta\dots\gamma} \qquad \text{for} \qquad \lambda, \mu, \dots, \nu = x, y, z, \tag{2.65}$$

whenever $\hat{S}$ represents a symmetry transformation in regards to the crystal. This is known as Neumann's principle.

The equations from (2.65) form nothing more than a linear system, and while this system cannot be used to yield specific values for every tensor element, it can be used to relate the tensor elements to each other which drastically reduces the amount of elements to be calculated.

As an example, consider a tensor of rank 2 in a crystal where flipping the sign of the z-coordinate is a symmetric operation, i.e. it has a reflection symmetry along the xy-plane. The z-inversion can be represented by the matrix

$$\hat{S}_z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}. \tag{2.66}$$

Choosing the tensor elements T_{xz} and T_{zz} and the symmetry operator $\hat{S}_z$ for equation (2.65) yields

$$T_{xz} = -T_{xz} \implies T_{xz} = 0, \quad \text{and} \quad T_{zz} = (-1)(-1)T_{zz} = T_{zz}. \tag{2.67}$$

For both cases, the only term in the sum that doesn't vanish is the term with the tensor element that has the same indices. For T_{xz}, the z-inversion yields that the tensor element must be equal to the negative of itself, so, of course, this element must be equal to zero. On the other hand, T_{zz} has two z-indices, and the z-inversion cancels itself out.

Generally if a crystal is symmetric under z-inversion, every tensor element that has an odd amount of z-indices will vanish. Equivalent rules also exist for crystals with x- or y-inversion symmetry.

Obtaining the symmetry operations are evidently of great use to find out which tensor elements one should calculate. These operations can be obtained from Quantum Espresso, which uses Spglib to obtain these symmetry operations.[25] Having the symmetry operations of a crystal, the system (2.65) is readily solved for a tensor of arbitrary order.

2.5 Current Density and the Conductivity Tensors

This section shall consider the induced current when an electric field is applied to a two-dimensional crystal. The main focus shall be on developing general equations for the conductivity tensors which relate the electric field to the current density.

The electric field that perturbs the crystal is assumed to be on time-harmonic form

$$\boldsymbol{\mathcal{E}}(t) = \frac{1}{2}\sum_{\mathrm{p}}\sum_{\alpha} \overline{\mathbf{e}}_{\alpha}\mathcal{E}_{\alpha}(\omega_{\mathrm{p}})\exp(-\mathrm{i}\omega_{\mathrm{p}}t), \tag{2.68}$$

where the sum over q imply a sum over time-harmonic oscillations with both negative and positive frequencies, the sum over α is a sum over the Cartesian x-, y-, and z-directions, and $\overline{\mathbf{e}}_{\alpha}$ is a unit vector in the α'th direction.

As explained in section 2.4, the current density response can generally be written as a power series in the unitless perturbation parameter

$$\mathbf{J}(t) = \mathbf{J}^{(0)} + \epsilon\mathbf{J}^{(1)}(t) + \epsilon^2\mathbf{J}^{(2)}(t) + \ldots, \tag{2.69}$$

where the entrances of each term shall be considered on the forms

$$J_{\lambda}^{(1)}(t) = \sum_{\mathrm{p}}\sum_{\alpha}\sigma_{\lambda\alpha}(\omega_{\mathrm{p}})\,\mathcal{E}_{\alpha}(\omega_{\mathrm{p}})\exp(-\mathrm{i}\omega_{\mathrm{p}}t), \tag{2.70}$$

$$J_{\lambda}^{(2)}(t) = \sum_{\mathrm{p,q}}\sum_{\alpha,\beta}\sigma_{\lambda\alpha\beta}(\omega_{\mathrm{p}},\omega_{\mathrm{q}})\,\mathcal{E}_{\alpha}(\omega_{\mathrm{p}})\,\mathcal{E}_{\beta}(\omega_{\mathrm{q}})\exp(-\mathrm{i}(\omega_{\mathrm{p}}+\omega_{\mathrm{q}})\,t) \tag{2.71}$$

and so on for $\lambda = x, y, z$.

Alternatively, the current density could be considered an observable of the current density operator $\hat{\mathbf{J}}$, and it is then given as

$$\mathbf{J}(t) = \mathrm{tr}\Big(\hat{\rho}(t)\,\hat{\mathbf{J}}(t)\Big) = \mathrm{tr}\Big(\hat{\rho}^{(0)}\hat{\mathbf{J}}(t)\Big) + \epsilon\mathrm{tr}\Big(\hat{\rho}^{(1)}(t)\,\hat{\mathbf{J}}(t)\Big) + \epsilon^2\mathrm{tr}\Big(\hat{\rho}^{(2)}(t)\,\hat{\mathbf{J}}(t)\Big) + \ldots \tag{2.72}$$

The current density operator can be expressed through the momentum operator $\hat{\mathbf{p}}$ and the vector potential $\boldsymbol{\mathcal{A}}$ by[22]

$$\hat{\mathbf{J}}(t) = -\frac{eg(\hat{\mathbf{p}} + \epsilon e\boldsymbol{\mathcal{A}}(t))}{m_e A}, \tag{2.73}$$

where e is the elementary charge, m_e is the electron mass, A is the area of the two-dimensional crystal and g accounts for the spin-degeneracy where $g = 2$ is used for a tight-binding model without spin, and $g = 1$ is used for a tight-binding model with spin included. Note that for the purposes of this thesis, $\boldsymbol{\mathcal{A}}$ shall be either 0 or a sum of terms that each are proportional to the electric field magnitudes $\mathcal{E}_{\alpha}$ (See Appendix A). This suggests that $\boldsymbol{\mathcal{A}}$ must have a perturbation factor ϵ multiplied onto it. Using the above expression for the current density operator in the trace equation (2.72) yields

$$\begin{aligned}\mathbf{J}(t) = -\frac{ge}{m_e A}\Big\{&\mathrm{tr}\Big(\hat{\rho}^{(0)}\hat{\mathbf{p}}\Big) + \epsilon\mathrm{tr}\Big(\hat{\rho}^{(1)}(t)\,\hat{\mathbf{p}} + e\hat{\rho}^{(0)}\boldsymbol{\mathcal{A}}(t)\Big)\\ &+\epsilon^2\mathrm{tr}\Big(\hat{\rho}^{(2)}(t)\,\hat{\mathbf{p}} + e\hat{\rho}^{(1)}(t)\,\boldsymbol{\mathcal{A}}(t)\Big) + \ldots\Big\}.\end{aligned} \tag{2.74}$$

Comparing the coefficients in front of the orders of ϵ in (2.74) and (2.69) then shows that

$$\mathbf{J}^{(0)} = -\frac{ge}{m_e A}\mathrm{tr}\Big(\hat{\rho}^{(0)}\hat{\mathbf{p}}\Big), \tag{2.75}$$

where the equilibrium current density $\mathbf{J}^{(0)}$ is expected to be zero since no current should run in the crystal in the absence of an electric field, and

$$\mathbf{J}^{(M)}(t) = -\frac{ge}{m_e A}\mathrm{tr}\Big(\hat{\rho}^{(M)}(t)\,\hat{\mathbf{p}} + e\hat{\rho}^{(M-1)}(t)\,\boldsymbol{\mathcal{A}}(t)\Big) \qquad \text{for} \qquad M = 1, 2, \ldots \tag{2.76}$$

Expressions for $\hat{\rho}$ and $\boldsymbol{\mathcal{A}}$ differ depending on the choice of gauge. This project shall consider the length gauge (LG) and velocity gauge (VG) as outlined in Appendix A.

2.5.1 Length Gauge

Appendix A shows that using the LG sets $\boldsymbol{\mathcal{A}} = 0$ and yields a Hamiltonian on the form

$$\hat{\mathcal{H}}(\mathbf{r}, t) = \hat{\mathcal{H}}_0(\mathbf{r}) + \epsilon\hat{\mathcal{V}}(\mathbf{r}, t) \quad \text{with} \quad \hat{\mathcal{V}}(\mathbf{r}, t) = \frac{1}{2}\sum_{\mathrm{p}}\sum_{\alpha} e\hat{r}_\alpha \mathcal{E}_\alpha(\omega_\mathrm{p})\exp(-\mathrm{i}\omega_\mathrm{p} t), \tag{2.77}$$

and using the notation from subsection 2.3.3 yields

$$\hat{\mathcal{V}}(\omega_\mathrm{p}) = \sum_{\alpha} e\hat{r}_\alpha \mathcal{E}_\alpha(\omega_\mathrm{p}) \implies \mathcal{V}_{mn}(\omega_\mathrm{p}) = \sum_{\alpha} e r^{\alpha}_{mn}\mathcal{E}_\alpha(\omega_\mathrm{p})\,, \tag{2.78}$$

where the additional short-hand notation $\langle\psi_m|\,\hat{r}_\alpha\,|\psi_n\rangle = r^{\alpha}_{mn}$ has been introduced for matrix elements of vector operators such as the position and momentum operators. Now, consider (2.76) with $M = 1$. An entrance of the first-order current density vector is given by

$$J^{(1)}_\lambda = -\frac{ge}{m_e A}\mathrm{tr}\Big(\hat{\rho}^{(1)}(t)\,\hat{p}_\lambda\Big)\,. \tag{2.79}$$

Expanding the trace in (2.75) will introduce a sum over the diagonal matrix elements of the product of the operators $\hat{\rho}^{(1)}\hat{p}_\lambda$. In a crystal, the states are defined not just by an index, but also by their crystal momentum represented through the wave vector $\mathbf{k}$. This means that for every $\mathbf{k}$, the matrices of the operators are expanded with new blocks of elements, and as such the trace operator must sum over the band indices m and n and the reciprocal vectors $\mathbf{k}$ and $\mathbf{k}'$, so

$$J^{(1)}_\lambda = -\frac{ge}{m_e A}\sum_{\mathbf{k},\mathbf{k}',m,n}\rho^{(1)}_{mn}(\mathbf{k},\mathbf{k}')\,p^{\lambda}_{nm}(\mathbf{k}',\mathbf{k})\,, \tag{2.80}$$

where for example $\langle\psi_m(\mathbf{k}')|\,\hat{p}_\lambda\,|\psi_n(\mathbf{k})\rangle = p^{\lambda}_{mn}(\mathbf{k}',\mathbf{k})$. The momentum operator is diagonal in $\mathbf{k}$,[26] so

$$p^{\lambda}_{nm}(\mathbf{k}',\mathbf{k}) = p^{\lambda}_{nm}(\mathbf{k},\mathbf{k})\,\delta_{\mathbf{k}\mathbf{k}'} = p^{\lambda}_{nm}(\mathbf{k})\,\delta_{\mathbf{k}\mathbf{k}'}\,. \tag{2.81}$$

The density operator is more troublesome; however, throughout the chapter it shall be shown that it as well is diagonal in $\mathbf{k}$. For now, the sum over $\mathbf{k}$ and $\mathbf{k}'$ is assumed to only produce only non-zero terms for $\mathbf{k} = \mathbf{k}'$, and the entrances of the first-order current density is then obtained through

$$J^{(1)}_\lambda = -\frac{ge}{m_e A}\sum_{\mathbf{k},m,n}\rho^{(1)}_{mn}(\mathbf{k})\,p^{\lambda}_{nm}(\mathbf{k})\,. \tag{2.82}$$

The $\mathbf{k}$-dependence of the energies and the matrix elements shall be considered implicit from here on to further declutter the expressions. Equation (2.45) shows the developed expression for the first-order density matrix elements. Using this expression yields

$$J^{(1)}_\lambda = -\frac{ge}{2m_e A}\sum_{q}\sum_{\mathbf{k},m,n}\frac{f_{nm}p^{\lambda}_{nm}\mathcal{V}_{mn}(\omega_\mathrm{p})}{\hbar\omega_\mathrm{p} - E'_{mn}}\exp(-\mathrm{i}\omega_\mathrm{p} t), \tag{2.83}$$

and from here, the matrix elements from (2.78) are inserted, so

$$J^{(1)}_\lambda = -\frac{ge^2}{2m_e A}\sum_{\mathrm{p}}\sum_{\alpha}\sum_{\mathbf{k},m,n}\frac{f_{nm}p^{\lambda}_{nm}r^{\alpha}_{mn}}{\hbar\omega_\mathrm{p} - E'_{mn}}\mathcal{E}_\alpha(\omega_\mathrm{p})\exp(-\mathrm{i}\omega_\mathrm{p} t). \tag{2.84}$$

Comparing the general expression for the elements of the first-order current density in (2.70) with (2.84) reveals an expression for the elements of the rank-2 conductivity tensor

$$\sigma_{\lambda\alpha}(\omega_{\mathrm{p}}) = -\frac{ge^2}{2m_e A}\sum_{\mathbf{k},m,n}\frac{f_{nm}p^{\lambda}_{nm}r^{\alpha}_{mn}}{\hbar\omega_{\mathrm{p}} - E'_{mn}}. \tag{2.85}$$

The position operator is troublesome in crystals due to the translational invariance by a lattice vector. It is therefore beneficial to switch to momentum elements which can be done by [27]

$$r^{\alpha}_{mn} = \frac{-\mathrm{i}\hbar p^{\alpha}_{mn}}{m_e E_{mn}}, \tag{2.86}$$

so

$$\sigma_{\lambda\alpha}(\omega_{\mathrm{p}}) = \frac{\mathrm{i}ge^2\hbar}{2m_e^2 A}\sum_{\mathbf{k},m,n}\frac{f_{nm}p^{\lambda}_{nm}p^{\alpha}_{mn}}{E_{mn}(\hbar\omega_{\mathrm{p}} - E'_{mn})}. \tag{2.87}$$

Since only momentum matrix elements remain, which are diagonal in $\mathbf{k}$, it can readily be seen that the density matrix elements are also diagonal in $\mathbf{k}$ as was assumed. Due to the factor $1/E_{mn}$, the terms with $m = n$ are singular; however, the singularity is countered by the f_{nm}'s which are also zero for $m = n$. The singularity can thus be evaluated by considering the $\mathbf{k}$-dependence of the energies and Fermi functions, and by letting the $\mathbf{k}$ of the n'th state approach the $\mathbf{k}$ of the m'th state, so

$$\lim_{\mathbf{k}'\to\mathbf{k}}\frac{f_n(\mathbf{k}') - f_m(\mathbf{k})}{E_m(\mathbf{k}) - E_n(\mathbf{k}')} = -\frac{\mathrm{d}f_m}{\mathrm{d}E} \qquad \text{for} \qquad m = n, \tag{2.88}$$

or to avoid introducing the singularity, the identity

$$f_{nm}r^{\alpha}_{mn} = \frac{\mathrm{i}\hbar p^{\alpha}_{mn}}{m_e}\frac{\mathrm{d}f_m}{\mathrm{d}E} \qquad \text{for} \qquad m = n, \tag{2.89}$$

can be established. Evaluating the limit instead of evaluating at $\mathbf{k} = \mathbf{k}'$ can also be argued for by considering the minuscule amount of momentum which photons carry. As such, the singularity-free expression for the elements of the rank-2 conductivity tensor is

$$\sigma_{\lambda\alpha}(\omega_{\mathrm{p}}) = \frac{\mathrm{i}ge^2\hbar}{2m_e^2 A}\sum_{\substack{\mathbf{k},m,n\\ m\neq n}}\frac{f_{nm}p^{\lambda}_{nm}p^{\alpha}_{mn}}{E_{mn}(\hbar\omega_{\mathrm{p}} - E'_{mn})} - \frac{\mathrm{i}ge^2\hbar}{2m_e^2 A}\sum_{\mathbf{k},m}\frac{p^{\lambda}_{mm}p^{\alpha}_{mm}}{\hbar\omega_{\mathrm{p}} + \mathrm{i}\hbar\eta}\frac{\mathrm{d}f_m}{\mathrm{d}E}, \tag{2.90}$$

where we've also used that $E'_{mm} = -\mathrm{i}\hbar\eta$.

For the elements of the rank-3 conductivity tensor, we start from equation (2.76) with $M = 2$. An entrance of the second-order current density vector is then

$$J^{(2)}_{\lambda}(t) = -\frac{ge}{m_e A}\mathrm{tr}\Big(\hat{\rho}^{(2)}(t)\,\hat{p}_{\lambda}\Big) = -\frac{ge}{m_e A}\sum_{\mathbf{k},m,n}\rho^{(2)}_{mn}(t)\,p^{\lambda}_{nm}. \tag{2.91}$$

Similarly, equation (2.48) shows the developed expression for $\rho^{(2)}_{mn}$. Using this expression with the matrix elements from (2.78) gives

$$\begin{aligned} J^{(2)}_{\lambda}(t) = -\frac{ge}{4m_e A}\sum_{\mathrm{p,q}}\sum_{\mathbf{k},m,n,l}&\frac{p^{\lambda}_{nm}}{\hbar\omega_{\mathrm{p}} + \hbar\omega_{\mathrm{q}} - E'_{mn}}\\ &\times\left(\frac{f_{nl}\mathcal{V}_{ln}(\omega_{\mathrm{p}})\,\mathcal{V}_{ml}(\omega_{\mathrm{q}})}{\hbar\omega_{\mathrm{p}} - E'_{ln}} - \frac{f_{lm}\mathcal{V}_{ml}(\omega_{\mathrm{p}})\,\mathcal{V}_{ln}(\omega_{\mathrm{q}})}{\hbar\omega_{\mathrm{p}} - E'_{ml}}\right)\exp(-\mathrm{i}(\omega_{\mathrm{p}} + \omega_{\mathrm{q}})\,t), \end{aligned} \tag{2.92}$$

and then

$$\begin{aligned} J^{(2)}_{\lambda}(t) = -\frac{ge^3}{4m_e A}\sum_{\mathrm{p,q}}\sum_{\alpha,\beta}\sum_{\mathbf{k},m,n,l}&\frac{p^{\lambda}_{nm}}{\hbar\omega_{\mathrm{p}} + \hbar\omega_{\mathrm{q}} - E'_{mn}}\\ &\times\left(\frac{f_{nl}r^{\alpha}_{ln}r^{\beta}_{ml}}{\hbar\omega_{\mathrm{p}} - E'_{ln}} - \frac{f_{lm}r^{\alpha}_{ml}r^{\beta}_{ln}}{\hbar\omega_{\mathrm{p}} - E'_{ml}}\right)\mathcal{E}_{\alpha}(\omega_{\mathrm{p}})\,\mathcal{E}_{\beta}(\omega_{\mathrm{q}})\exp(-\mathrm{i}(\omega_{\mathrm{p}} + \omega_{\mathrm{q}})\,t). \end{aligned} \tag{2.93}$$

From here, the conductivity tensor elements can be seen from (2.71) to be

$$\sigma_{\lambda\alpha\beta}(\omega_\mathrm{p},\omega_\mathrm{q}) = -\frac{ge^3}{4m_eA}\sum_{\mathbf{k},m,n,l}\frac{p^\lambda_{nm}}{\hbar\omega_\mathrm{p}+\hbar\omega_\mathrm{q}-E'_{mn}}\left(\frac{f_{nl}r^\alpha_{ln}r^\beta_{ml}}{\hbar\omega_\mathrm{p}-E'_{ln}} - \frac{f_{lm}r^\alpha_{ml}r^\beta_{ln}}{\hbar\omega_\mathrm{p}-E'_{ml}}\right). \tag{2.94}$$

Again, the position matrix elements will yield singularities whenever their indices are equal. The sum shall therefore be split up into different cases for the band indices m, n and l.

$m \neq n \neq l \neq m$ and $m = n \neq l$

For the cases where $m \neq n \neq l \neq m$ and $m = n \neq l$ no singularities occur, and all the position operators can safely be replaced with the momentum operators through (2.86) which yields

$$\frac{ge^3\hbar^2}{4m_e^3A}\sum_{\substack{\mathbf{k},m,n,l\\ m\neq l\neq n}}\frac{p^\lambda_{nm}}{E_{ml}E_{ln}(\hbar\omega_\mathrm{p}+\hbar\omega_\mathrm{q}-E'_{mn})}\left(\frac{f_{nl}p^\alpha_{ln}p^\beta_{ml}}{\hbar\omega_\mathrm{p}-E'_{ln}} - \frac{f_{lm}p^\alpha_{ml}p^\beta_{ln}}{\hbar\omega_\mathrm{p}-E'_{ml}}\right). \tag{2.95}$$

$m \neq n = l$ and $m = l \neq n$

For the cases with $m \neq l = n$, the l-index is replaced with n, and the sum shall be over the indices m and n with the condition that $m \neq n$. The position matrix elements with different indices can again be safely replaced with momentum matrix elements, and in the left-hand term inside the parenthesis of (2.94), the singularity is dealt with through the identity (2.88), so

$$-\frac{ge^3\hbar}{4m_e^2A}\sum_{\substack{\mathbf{k},m,n\\ m\neq n}}\frac{p^\lambda_{nm}}{E_{mn}(\hbar\omega_\mathrm{p}+\hbar\omega_\mathrm{q}-E'_{mn})}\left(\frac{\hbar}{m_e}\frac{p^\alpha_{nn}p^\beta_{mn}}{\hbar\omega_\mathrm{p}+\mathrm{i}\hbar\eta}\frac{\mathrm{d}f_n}{\mathrm{d}E} + \frac{\mathrm{i}f_{nm}p^\alpha_{mn}r^\beta_{nn}}{\hbar\omega_\mathrm{p}-E'_{mn}}\right). \tag{2.96}$$

Note that the singularity in r^β_{nn} cannot be dealt with just yet. For the cases with $m = l \neq n$, the procedure is similar except that the identity (2.88) can be used in the right-hand term inside the parenthesis of (2.94), so

$$\frac{ge^3\hbar}{4m_e^2A}\sum_{\substack{\mathbf{k},m,n\\ m\neq n}}\frac{p^\lambda_{nm}}{E_{mn}(\hbar\omega_\mathrm{p}+\hbar\omega_\mathrm{q}-E'_{mn})}\left(\frac{\mathrm{i}f_{nm}p^\alpha_{mn}r^\beta_{mm}}{\hbar\omega_\mathrm{p}-E'_{mn}} + \frac{\hbar}{m_e}\frac{p^\alpha_{mm}p^\beta_{mn}}{\hbar\omega_\mathrm{p}+\mathrm{i}\hbar\eta}\frac{\mathrm{d}f_m}{\mathrm{d}E}\right). \tag{2.97}$$

Yet a singularity remains. To resolve these singular terms, (2.96) and (2.97) are added together which gives

$$\begin{aligned}\frac{ge^3\hbar}{4m_e^2A}\sum_{\substack{\mathbf{k},m,n\\ m\neq n}}\frac{p^\lambda_{nm}}{E_{mn}(\hbar\omega_\mathrm{p}+\hbar\omega_\mathrm{q}-E'_{mn})}&\left\{\frac{\hbar}{m_e}\frac{-p^\beta_{mn}}{\hbar\omega_\mathrm{p}+\mathrm{i}\hbar\eta}\left(p^\alpha_{nn}\frac{\mathrm{d}f_n}{\mathrm{d}E} - p^\alpha_{mm}\frac{\mathrm{d}f_m}{\mathrm{d}E}\right)\right.\\ &\left.+\frac{\mathrm{i}f_{nm}p^\alpha_{mn}\left(r^\beta_{mm}-r^\beta_{nn}\right)}{\hbar\omega_\mathrm{p}-E'_{mn}}\right\}.\end{aligned} \tag{2.98}$$

From here, the singularities can be removed since the difference between the momentum matrix elements will yield what's known as the generalised derivative of some of the factors; however, this must be done carefully. Essentially, every factor which previously had an index l must be included in the generalised derivative.[22] As such, we can write for (2.98) specifically

$$\left(r^\beta_{mm}-r^\beta_{nn}\right)\frac{f_{nm}p^\alpha_{mn}}{E_{mn}(\hbar\omega_\mathrm{p}-E'_{mn})} = \mathrm{i}\left(\frac{f_{nm}p^\alpha_{mn}}{E_{mn}(\hbar\omega_\mathrm{p}-E'_{mn})}\right)_{;k_\beta}, \tag{2.99}$$

where $(\,\cdot\,)_{;k_\beta}$ is the generalised derivative with respect to k_β. The generalised derivative of the matrix elements of an operator $\hat{\mathcal{O}}$ in the eigenbasis is elaborated on in Appendix B and is defined as[27]

$$(\mathcal{O}_{mn})_{;\mathbf{k}} = \frac{\partial}{\partial\mathbf{k}}\mathcal{O}_{mn} - \mathrm{i}(\boldsymbol{\Omega}_{mm}-\boldsymbol{\Omega}_{nn})\,\mathcal{O}_{mn}, \tag{2.100}$$

with the $\mathbf{\Omega}_{mm}$'s as the Berry connection diagonal matrix elements. Thus, the terms in the sum for the two cases $m \neq l = n$ and $m = l \neq n$ added together will be on the form

$$\begin{aligned}
&-\frac{ge^3\hbar^2}{4m_e^3A}\sum_{\substack{\mathbf{k},m,n\\ m\neq n}}\frac{p^\lambda_{nm}p^\beta_{mn}}{E_{mn}(\hbar\omega_\mathrm{p}+\mathrm{i}\hbar\eta)(\hbar\omega_\mathrm{p}+\hbar\omega_\mathrm{q}-E'_{mn})}\left(p^\alpha_{nn}\frac{\mathrm{d}f_n}{\mathrm{d}E}-p^\alpha_{mm}\frac{\mathrm{d}f_m}{\mathrm{d}E}\right)\\
&-\frac{ge^3\hbar}{4m_e^2A}\sum_{\substack{\mathbf{k},m,n\\ m\neq n}}\frac{p^\lambda_{nm}}{\hbar\omega_\mathrm{p}+\hbar\omega_\mathrm{q}-E'_{mn}}\left(\frac{f_{nm}p^\alpha_{mn}}{E_{mn}(\hbar\omega_\mathrm{p}-E'_{mn})}\right)_{;k_\beta}.
\end{aligned} \tag{2.101}$$

$m = n = l$

The last case that needs to be covered is for $m = n = l$. Carefully, we start by only setting $l = n$ in the left-hand term and $l = m$ in right-hand term in the brackets of (2.94) which yields

$$\frac{ge^3\hbar^2}{4m_e^3A}\sum_{\substack{\mathbf{k},m,n\\ m=n}}\frac{p^\lambda_{nm}p^\beta_{mn}}{(\hbar\omega_\mathrm{p}+\mathrm{i}\hbar\eta)(\hbar\omega_\mathrm{p}+\hbar\omega_\mathrm{q}-E'_{mn})}\left(\frac{p^\alpha_{mm}}{E_{mn}}\frac{\mathrm{d}f_m}{\mathrm{d}E}-\frac{p^\alpha_{nn}}{E_{mn}}\frac{\mathrm{d}f_n}{\mathrm{d}E}\right). \tag{2.102}$$

We can quickly identify this as a derivative with respect to energy through

$$\lim_{\mathbf{k}'\to\mathbf{k}}\frac{p^\alpha_{mm}(\mathbf{k})\frac{\mathrm{d}f_m}{\mathrm{d}E}(\mathbf{k})-p^\alpha_{nn}(\mathbf{k}')\frac{\mathrm{d}f_n}{\mathrm{d}E}(\mathbf{k}')}{E_m(\mathbf{k})-E_n(\mathbf{k}')}=\frac{\mathrm{d}}{\mathrm{d}E}\left(p^\alpha_{mm}\frac{\mathrm{d}f_m}{\mathrm{d}E}\right),\qquad\text{for}\qquad m=n. \tag{2.103}$$

Typically, the momentum matrix elements are approximated as

$$p^\alpha_{mn}\approx\frac{m_e}{\hbar}\left\langle\psi_m\right|\frac{\partial[\hat{\mathcal{H}}]^\mathrm{TB}}{\partial k_\alpha}\left|\psi_n\right\rangle, \tag{2.104}$$

from which the Hellmann-Feynman theorem yields

$$p^\alpha_{mm}\approx\frac{m_e}{\hbar}\left\langle\psi_m\right|\frac{\partial[\hat{\mathcal{H}}]^\mathrm{TB}}{\partial k_\alpha}\left|\psi_m\right\rangle=\frac{m_e}{\hbar}\frac{\partial E_m}{\partial k_\alpha}. \tag{2.105}$$

Thus (2.102) becomes

$$\frac{ge^3}{4m_eA}\sum_{\mathbf{k},m}\frac{p^\lambda_{mm}}{(\hbar\omega_\mathrm{p}+\mathrm{i}\hbar\eta)(\hbar\omega_\mathrm{p}+\hbar\omega_\mathrm{q}+\mathrm{i}\hbar\eta)}\frac{\mathrm{d}^2f_m}{\partial k_\alpha\partial k_\beta}. \tag{2.106}$$

So finally by adding (2.95), (2.101) and (2.106), a singularity-free expression for the elements of the rank-3 conductivity tensor has been derived. It is

$$\begin{aligned}
\sigma_{\lambda\alpha\beta}(\omega_\mathrm{p},\omega_\mathrm{q})=&\frac{ge^3\hbar^2}{4m_e^3A}\sum_{\substack{\mathbf{k},m,n,l\\ m\neq l\neq n}}\frac{p^\lambda_{nm}}{E_{ml}E_{ln}(\hbar\omega_\mathrm{p}+\hbar\omega_\mathrm{q}-E'_{mn})}\left(\frac{f_{nl}p^\alpha_{ln}p^\beta_{ml}}{\hbar\omega_\mathrm{p}-E'_{ln}}-\frac{f_{lm}p^\alpha_{ml}p^\beta_{ln}}{\hbar\omega_\mathrm{p}-E'_{ml}}\right)\\
&-\frac{ge^3\hbar}{4m_e^2A}\sum_{\substack{\mathbf{k},m,n\\ m\neq n}}\frac{p^\lambda_{nm}p^\beta_{mn}}{E_{mn}(\hbar\omega_\mathrm{p}+\mathrm{i}\hbar\eta)(\hbar\omega_\mathrm{p}+\hbar\omega_\mathrm{q}-E'_{mn})}\frac{\partial f_{nm}}{\partial k_\alpha}\\
&-\frac{ge^3\hbar}{4m_e^2A}\sum_{\substack{\mathbf{k},m,n\\ m\neq n}}\frac{p^\lambda_{nm}}{\hbar\omega_\mathrm{p}+\hbar\omega_\mathrm{q}-E'_{mn}}\left(\frac{f_{nm}p^\alpha_{mn}}{E_{mn}(\hbar\omega_\mathrm{p}-E'_{mn})}\right)_{;k_\beta}\\
&+\frac{ge^3}{4m_eA}\sum_{\mathbf{k},m}\frac{p^\lambda_{mm}}{(\hbar\omega_\mathrm{p}+\mathrm{i}\hbar\eta)(\hbar\omega_\mathrm{p}+\hbar\omega_\mathrm{q}+\mathrm{i}\hbar\eta)}\frac{\partial^2f_m}{\partial k_\alpha\partial k_\beta}.
\end{aligned} \tag{2.107}$$

Lastly before we move on to the VG, the calculation of the generalised derivative shall be discussed. It can be determined if the Berry connection is known; however, a different expression which foregoes the Berry connection can also be derived. First, the k_β-derivative is performed

$$\begin{aligned}\frac{\partial}{\partial k_\beta}\left(\frac{f_{nm}p^\alpha_{mn}}{E_{mn}(\hbar\omega_\mathrm{p}-E'_{mn})}\right) &= \left(p^\alpha_{mn}\frac{\partial f_{nm}}{\partial k_\beta}+f_{nm}\frac{\partial p^\alpha_{mn}}{\partial k_\beta}\right)\bigg/\Big(E_{mn}(\hbar\omega_\mathrm{p}-E'_{mn})\Big)\\ &\quad-\frac{\hbar}{m_e}\frac{f_{nm}p^\alpha_{mn}(\hbar\omega_\mathrm{p}+\mathrm{i}\hbar\eta-2E_{mn})\big(p^\beta_{mm}-p^\beta_{nn}\big)}{E^2_{mn}(\hbar\omega_\mathrm{p}-E'_{mn})^2},\end{aligned} \tag{2.108}$$

where the Hellmann-Feynman theorem has again been used. From here, it can quickly be seen that the generalised derivative must be

$$\begin{aligned}\left(\frac{f_{nm}p^\alpha_{mn}}{E_{mn}(\hbar\omega_\mathrm{p}-E'_{mn})}\right)_{;k_\beta} &= \left(p^\alpha_{mn}\frac{\partial f_{nm}}{\partial k_\beta}+f_{nm}(p^\alpha_{mn})_{;k_\beta}\right)\bigg/\Big(E_{mn}(\hbar\omega_\mathrm{p}-E'_{mn})\Big)\\ &\quad-\frac{\hbar}{m_e}\frac{f_{nm}p^\alpha_{mn}(\hbar\omega_\mathrm{p}+\mathrm{i}\hbar\eta-2E_{mn})\big(p^\beta_{mm}-p^\beta_{nn}\big)}{E^2_{mn}(\hbar\omega_\mathrm{p}-E'_{mn})^2},\end{aligned} \tag{2.109}$$

i.e. the generalised derivative of only the momentum matrix elements is necessary. These can be found through the sum rule[22]

$$(p^\alpha_{mn})_{;k_\beta}=\hbar\delta_{\alpha\beta}\delta_{mn}+\frac{\hbar}{m_e}\left[\sum_{\substack{l\\ l\neq m}}\frac{p^\alpha_{ml}p^\beta_{ln}}{E_{ml}}-\sum_{\substack{l\\ l\neq n}}\frac{p^\beta_{ml}p^\alpha_{ln}}{E_{ln}}\right]. \tag{2.110}$$

Thus using these expressions avoids the Berry connection as well as numerically calculating the derivatives of the momentum matrix elements in the eigenbasis.

2.5.2 Velocity Gauge

Appendix A shows that using the velocity gauge sets

$$\boldsymbol{\mathcal{A}}(t) = \frac{1}{2}\sum_{\mathrm{p}}\sum_{\alpha}\frac{\overline{\mathbf{e}}_\alpha \mathcal{E}_\alpha(\omega_\mathrm{p})}{\mathrm{i}\omega_\mathrm{p}}\exp(-\mathrm{i}\omega_\mathrm{p} t), \tag{2.111}$$

and yields a Hamiltonian on the form

$$\begin{aligned} \hat{\mathcal{H}}(\mathbf{r},t) &= \hat{\mathcal{H}}_0(\mathbf{r}) + \epsilon\hat{\mathcal{V}}^{(1)}(t) + \epsilon^2\hat{\mathcal{V}}^{(2)}(t)\,, \qquad \text{with} \\ \hat{\mathcal{V}}^{(1)}(t) &= \frac{e}{2m_e}\sum_{\mathrm{p}}\sum_{\alpha}\frac{\hat{p}_\alpha \mathcal{E}_\alpha(\omega_\mathrm{p})}{\mathrm{i}\omega_\mathrm{p}}\exp(-\mathrm{i}\omega_\mathrm{p} t), \qquad \text{and} \\ \hat{\mathcal{V}}^{(2)}(t) &= \frac{-e^2}{8m_e}\sum_{\mathrm{p,q}}\sum_{\alpha}\frac{\mathcal{E}_\alpha(\omega_\mathrm{p})\,\mathcal{E}_\alpha(\omega_\mathrm{q})}{\omega_\mathrm{p}\omega_\mathrm{q}}\exp(-\mathrm{i}(\omega_\mathrm{p}+\omega_\mathrm{q})\,t). \end{aligned} \tag{2.112}$$

Again, considering the notation developed in subsection 2.3.3, we have

$$\hat{\mathcal{V}}^{(1)}(\omega_\mathrm{p}) = \frac{e}{m_e}\sum_{\alpha}\frac{\hat{p}_\alpha \mathcal{E}_\alpha(\omega_\mathrm{p})}{\mathrm{i}\omega_\mathrm{p}} \implies \mathcal{V}^{(1)}_{mn}(\omega_\mathrm{p}) = \frac{e}{m_e}\sum_{\alpha}\frac{p^\alpha_{mn}\mathcal{E}_\alpha(\omega_\mathrm{p})}{\mathrm{i}\omega_\mathrm{p}}, \tag{2.113}$$

and

$$\hat{\mathcal{V}}^{(2)}(\omega_\mathrm{p},\omega_\mathrm{q}) = -\frac{e^2}{2m_e}\sum_{\alpha}\frac{\mathcal{E}_\alpha(\omega_\mathrm{p})\,\mathcal{E}_\alpha(\omega_\mathrm{q})}{\omega_\mathrm{p}\omega_\mathrm{q}} \implies \mathcal{V}^{(2)}_{mn}(\omega_\mathrm{p},\omega_\mathrm{q}) = \delta_{mn}\hat{\mathcal{V}}^{(2)}(\omega_\mathrm{p},\omega_\mathrm{q})\,. \tag{2.114}$$

The first goal is to determine the elements of the rank-2 conductivity tensor which can be done through (2.76) with $M=1$. An entrance of the first-order current density vector is then

$$\begin{aligned} J^{(1)}_\lambda(t) &= -\frac{ge}{m_eA}\mathrm{tr}\Big(\hat{\rho}^{(1)}(t)\,\hat{p}_\lambda + e\hat{\rho}^{(0)}\mathcal{A}_\lambda(t)\Big) \\ &= -\frac{ge}{m_eA}\left(\sum_{\mathbf{k},m,n}\rho^{(1)}_{mn}(t)\,p^\lambda_{nm} + e\sum_{\mathbf{k},m}\rho^{(0)}_{mm}\mathcal{A}_\lambda(t)\right). \end{aligned} \tag{2.115}$$

Equations (2.54) and (2.111) show the developed expressions for $\rho^{(1)}_{mn}$ and $\boldsymbol{\mathcal{A}}$, respectively. Using these expressions along with $\rho^{(0)}_{mm} = f_m$ then gives

$$J^{(1)}_\lambda(t) = -\frac{ge}{2m_eA}\sum_{\mathrm{p}}\left(\sum_{\mathbf{k},m,n}\frac{f_{nm}p^\lambda_{nm}\mathcal{V}^{(1)}_{mn}(\omega_\mathrm{p})}{\hbar\omega_\mathrm{p} - E'_{mn}} + e\sum_{\mathbf{k},m}\frac{f_m\mathcal{E}_\lambda(\omega_\mathrm{p})}{\mathrm{i}\omega_\mathrm{p}}\right)\exp(-\mathrm{i}\omega_\mathrm{p} t). \tag{2.116}$$

Once again, matrix elements are inserted

$$J^{(1)}_\lambda(t) = \frac{\mathrm{i}ge\hbar}{2m_eA}\sum_{\mathrm{p}}\sum_{\alpha}\left(\frac{e}{m_e}\sum_{\mathbf{k},m,n}\frac{f_{nm}p^\lambda_{nm}p^\alpha_{mn}}{\hbar\omega_\mathrm{p}(\hbar\omega_\mathrm{p} - E'_{mn})} + e\delta_{\lambda\alpha}\sum_{\mathbf{k},m}\frac{f_m}{\hbar\omega_\mathrm{p}}\right)\mathcal{E}_\alpha(\omega_\mathrm{p})\exp(-\mathrm{i}\omega_\mathrm{p} t), \tag{2.117}$$

and from here, the elements of the conductivity tensor can be identified as

$$\sigma_{\lambda\alpha}(\omega_\mathrm{p}) = \frac{\mathrm{i}ge^2\hbar}{2m_e^2A}\sum_{\mathbf{k},m,n}\frac{f_{nm}p^\lambda_{nm}p^\alpha_{mn}}{\hbar\omega_\mathrm{p}(\hbar\omega_\mathrm{p} - E'_{mn})} + \delta_{\lambda\alpha}\frac{\mathrm{i}ge^2\hbar}{2m_eA}\sum_{\mathbf{k},m}\frac{f_m}{\hbar\omega_\mathrm{p}}. \tag{2.118}$$

Lastly, the rank-3 conductivity tensor can be found from the second-order current density vector.

$$\begin{aligned} J^{(2)}_\lambda(t) &= -\frac{ge}{m_eA}\mathrm{tr}\Big(\hat{\rho}^{(2)}(t)\,\hat{p}_\lambda + e\hat{\rho}^{(1)}\mathcal{A}_\lambda(t)\Big) \\ &= -\frac{ge}{m_eA}\left(\sum_{\mathbf{k},m,n}\rho^{(2)}_{mn}(t)\,p^\lambda_{nm} + e\sum_{\mathbf{k},m}\rho^{(1)}_{mm}\mathcal{A}_\lambda(t)\right). \end{aligned} \tag{2.119}$$

It can quickly be seen that $\rho^{(1)}_{mm} = 0$ due to the presence of $f_{mm} = 0$, so this sum vanishes. Now, insert equation (2.56) which shows the relevant expression for the $\rho^{(2)}_{mn}$'s to get

$$\begin{aligned} J^{(2)}_\lambda(t) = &-\frac{ge}{4m_e A}\sum_{\mathrm{p,q}}\sum_{\mathbf{k},m,n} p^\lambda_{nm}\frac{\exp(-\mathrm{i}(\omega_\mathrm{p}+\omega_\mathrm{q})\,t)}{\hbar\omega_\mathrm{p}+\hbar\omega_\mathrm{q}-E'_{mn}}\Bigg\{ f_{nm}\mathcal{V}^{(2)}_{mn}(\omega_\mathrm{p},\omega_\mathrm{q}) \\ &+\sum_l\left(\frac{f_{nl}\mathcal{V}^{(1)}_{ln}(\omega_\mathrm{p})\,\mathcal{V}^{(1)}_{ml}(\omega_\mathrm{q})}{\hbar\omega_\mathrm{p}-E'_{ln}} - \frac{f_{lm}\mathcal{V}^{(1)}_{ml}(\omega_\mathrm{p})\,\mathcal{V}^{(1)}_{ln}(\omega_\mathrm{q})}{\hbar\omega_\mathrm{p}-E'_{ml}}\right)\Bigg\}. \end{aligned} \tag{2.120}$$

Expressions for the matrix elements shall again be inserted; however, since $\mathcal{V}^{(2)}_{mn} = 0$ for $m \neq n$ and $f_{nm} = 0$ for $m = n$, it can clearly be seen that $\mathcal{V}^{(2)}_{mn}(\omega_\mathrm{p},\omega_\mathrm{q})\,f_{nm} = 0$ for all combinations of m and n. So,

$$\begin{aligned} J^{(2)}_\lambda(t) = &\frac{ge^3\hbar^2}{4m_e^3 A}\sum_{\mathrm{p,q}}\sum_{\alpha,\beta}\sum_{\mathbf{k},m,n,l}\frac{p^\lambda_{nm}}{\hbar\omega_\mathrm{p}\hbar\omega_\mathrm{q}(\hbar\omega_\mathrm{p}+\hbar\omega_\mathrm{q}-E'_{mn})} \\ &\times\left(\frac{f_{nl}p^\alpha_{ln}p^\beta_{ml}}{\hbar\omega_\mathrm{p}-E'_{ln}} - \frac{f_{lm}p^\alpha_{ml}p^\beta_{ln}}{\hbar\omega_\mathrm{p}-E'_{ml}}\right)\mathcal{E}_\alpha(\omega_\mathrm{p})\,\mathcal{E}_\beta(\omega_\mathrm{q})\exp(-\mathrm{i}(\omega_\mathrm{p}+\omega_\mathrm{q})\,t). \end{aligned} \tag{2.121}$$

Finally, the expression for rank-3 conductivity tensor in the velocity gauge can be identified through

$$\sigma_{\lambda\alpha\beta}(\omega_\mathrm{p},\omega_\mathrm{q}) = \frac{ge^3\hbar^2}{4m_e^3 A}\sum_{\mathbf{k},m,n,l}\frac{p^\lambda_{nm}}{\hbar\omega_\mathrm{p}\hbar\omega_\mathrm{q}(\hbar\omega_\mathrm{p}+\hbar\omega_\mathrm{q}-E'_{mn})}\left(\frac{f_{nl}p^\alpha_{ln}p^\beta_{ml}}{\hbar\omega_\mathrm{p}-E'_{ln}} - \frac{f_{lm}p^\alpha_{ml}p^\beta_{ln}}{\hbar\omega_\mathrm{p}-E'_{ml}}\right). \tag{2.122}$$

As a closing fact for this section, it shall be mentioned that the conductivity tensor elements obey the equations

$$\sigma_{\alpha\alpha}(\omega_\mathrm{p}) = \sigma^*_{\alpha\alpha}(-\omega_\mathrm{p}) \qquad \text{and} \qquad \sigma_{\alpha\alpha\alpha}(\omega_\mathrm{p},\omega_\mathrm{q}) = \sigma^*_{\alpha\alpha\alpha}(-\omega_\mathrm{p},-\omega_\mathrm{q})\,.$$

This shall only be proven explicitly for the rank-2 conductivity tensor in the velocity gauge. To start, the fraction in (2.118) is extended with the complex conjugate of its denominator

$$\sigma_{\alpha\alpha}(\omega_\mathrm{p}) = \frac{ge^2\hbar}{2m_e^2 A}\sum_{\mathbf{k},m,n}\frac{f_{nm}\,|p^\alpha_{mn}|^2(\hbar\eta+\mathrm{i}(\hbar\omega_\mathrm{p}-E_{mn}))}{\hbar\omega_\mathrm{p}\Big((\hbar\omega_\mathrm{p}-E_{mn})^2+(\hbar\eta)^2\Big)} + \delta_{\lambda\alpha}\frac{\mathrm{i}ge^2\hbar}{2m_e A}\sum_{\mathbf{k},m}\frac{f_m}{\hbar\omega_\mathrm{p}}. \tag{2.123}$$

From here, the real and imaginary parts can quickly be identified as

$$\mathrm{Re}(\sigma_{\alpha\alpha}(\omega_\mathrm{p})) = \frac{ge^2\hbar}{2m_e^2 A}\sum_{\mathbf{k},m,n}\frac{f_{nm}\,|p^\alpha_{mn}|^2\,\hbar\eta}{\hbar\omega_\mathrm{p}\Big((\hbar\omega_\mathrm{p}-E_{mn})^2+(\hbar\eta)^2\Big)}, \tag{2.124}$$

and

$$\mathrm{Im}(\sigma_{\alpha\alpha}(\omega_\mathrm{p})) = \frac{ge^2\hbar}{2m_e^2 A}\sum_{\mathbf{k},m,n}\frac{f_{nm}\,|p^\alpha_{mn}|^2(\hbar\omega_\mathrm{p}-E_{mn})}{\hbar\omega_\mathrm{p}\Big((\hbar\omega_\mathrm{p}-E_{mn})^2+(\hbar\eta)^2\Big)} + \delta_{\lambda\alpha}\frac{ge^2\hbar}{2m_e A}\sum_{\mathbf{k},m}\frac{f_m}{\hbar\omega_\mathrm{p}}. \tag{2.125}$$

Now, the frequency can be replaced by $-\omega_\mathrm{p}$, the identities $E_{mn} = -E_{nm}$ and $f_{nm} = -f_{mn}$ can be used and lastly the notation for the band indices can be switched. This quickly shows that

$$\mathrm{Re}(\sigma_{\alpha\alpha}(\omega_\mathrm{p})) = \mathrm{Re}(\sigma_{\alpha\alpha}(-\omega_\mathrm{p})) \qquad \text{and} \qquad \mathrm{Im}(\sigma_{\alpha\alpha}(\omega_\mathrm{p})) = -\mathrm{Im}(\sigma_{\alpha\alpha}(-\omega_\mathrm{p})) \tag{2.126}$$

such that $\sigma_{\alpha\alpha}(\omega_\mathrm{p}) = \sigma^*_{\alpha\alpha}(-\omega_\mathrm{p})$.

2.6 The Bethe-Salpeter Equation: Excitonic States

The previous section has been developing the framework of how to calculate the linear and nonlinear response of a crystal in the independent particle regime. This is the situation we have when we represent the many-body Hamiltonian (2.3) through only single-particle operators. However as mentioned in the introduction, 2D materials have a geometry which makes an advantageous setting for the Coulomb interaction between the electrons. Therefore we move on to how to deal with the many-body term beyond the ground state.

Having done a ground state calculation and obtained the tight parametrisation of the band structure, one has the needed ingredients for a many-body calculation. Firstly, we can take the independent particle orbitals and combine them in a Slater determinant where one valence orbital is excited into a conduction state to represent the excitation of an electron into a conduction state of same spin at the same **k**-point.[12] We write such a state as $\left|\overset{\pm}{v} \to \overset{\pm}{c}, \mathbf{k}\right\rangle$. A singlet state can be formed by taking the linear combination $|v \to c, \mathbf{k}\rangle = \frac{1}{\sqrt{2}}\left(\left|\overset{+}{v} \to \overset{+}{c}, \mathbf{k}\right\rangle + \left|\overset{-}{v} \to \overset{-}{c}, \mathbf{k}\right\rangle\right)$.[14] These singlets then form the basis on which the many-body calculation takes place, meaning we write the many-body wave function Ψ as[14]

$$\Psi = \sum_{vc\mathbf{k}} A_{vc\mathbf{k}} \left|v \to c, \mathbf{k}\right\rangle . \tag{2.127}$$

The spin-degenerate many-body Schrödinger matrix problem in terms of the expansion coefficients $A_{cv\mathbf{k}}$ then reads[14]

$$\sum_{v'c'\mathbf{k}'} \mathcal{H}^{\mathrm{MB}}_{vc\mathbf{k},v'c'\mathbf{k}'} A_{c'v'\mathbf{k}'} = E_{\mathrm{exc}} A_{cv\mathbf{k}}, \tag{2.128}$$

where E_{exc} is the excitation energy relative to the ground state $|0\rangle$. This is known as the Bethe-Salpeter equation. The matrix elements of the many-body Hamiltonian in (2.3) in the basis of singlets are given as[14]

$$\hat{\mathcal{H}}^{\mathrm{MB}}_{vc\mathbf{k},v'c'\mathbf{k}'} = \tilde{E}_{cv}\delta_{\mathbf{k}\mathbf{k}'}\delta_{vv'}\delta_{cc'} - \left\langle vc\mathbf{k}\right| W_C \left|v'c'\mathbf{k}'\right\rangle + 2\left\langle vc\mathbf{k}\right| V_X \left|v'c'\mathbf{k}'\right\rangle . \tag{2.129}$$

Here $\tilde{E}_{cv} = \tilde{E}_c(\mathbf{k}) - \tilde{E}_v(\mathbf{k})$ are the quasiparticle corrected energy differences between the valence and conduction bands and the matrix elements of the screened Coulomb interaction and the unscreened exchange interaction are given as[14]

$$\left\langle vc\mathbf{k}\right| W_C \left|v'c'\mathbf{k}'\right\rangle = \int\int \psi^*_{c\mathbf{k}}(\mathbf{r}_1)\psi_{c'\mathbf{k}'}(\mathbf{r}_1)W(\mathbf{r}_1-\mathbf{r}_2)\psi_{v\mathbf{k}}(\mathbf{r}_2)\psi^*_{v'\mathbf{k}'}(\mathbf{r}_2)\mathrm{d}^3\mathbf{r}_1\mathrm{d}^3\mathbf{r}_2 \tag{2.130}$$

$$\left\langle vc\mathbf{k}\right| V_X \left|v'c'\mathbf{k}'\right\rangle = \int\int \psi^*_{c\mathbf{k}}(\mathbf{r}_1)\psi_{c'\mathbf{k}'}(\mathbf{r}_2)V(\mathbf{r}_1-\mathbf{r}_2)\psi_{v\mathbf{k}}(\mathbf{r}_1)\psi^*_{v'\mathbf{k}'}(\mathbf{r}_2)\mathrm{d}^3\mathbf{r}_1\mathrm{d}^3\mathbf{r}_2 \tag{2.131}$$

It should be noted that the integration over $\mathbf{r}_2$ in the above integrals are convolutions between the potential and the wave functions which have $\mathbf{r}_2$ as argument. We can define the parity of the products of single-particle wave-functions depending only on $\mathbf{r}_1$ or $\mathbf{r}_2$ as $P_{ij} = P_i P_j$. Furthermore if the the ψ's have definite z-parity, then because the potentials do not mix different parity states, the condition for nonzero matrix elements in both cases

$$P_{cc'} = P_{vv'} \Leftrightarrow P_{cv} = P_{c'v'} . \tag{2.132}$$

are equivalent. This means certain combinations of states decouple and therefore they can considered separately when doing calculations.

The result of solving the Bethe-Salpeter equation with the matrix elements (2.129) are the excitonic states with zero total momentum. This is because we excited the valence band electron to a conduction band with the same $\mathbf{k}$ and the vacant hole in the valence band can be thought of as carrying a momentum of $-\mathbf{k}$, summing to zero in total. These are quasiparticle states responsible for modifying the conductivity of the solid to varying extent. The screened potential also has to be chosen with care in order to capture the screening effects present in the 2D material. A way to do this (See Appendix D) is to consider the Poisson equation for the potential from a point charge in the presence of planar slab with relative dielectric constant $\varepsilon_{\mathrm{ML}}$ and thickness d, sandwiched between two media with dielectric constant ε_{t} and ε_{b}. Doing so and taking the 2D limit yields the Rytova-Keldysh potential in $\mathbf{k}$-space as[8]

$$W_C(k_\rho) = \frac{e^2}{2\varepsilon_0(\kappa + r_0 k_\rho)\, k_\rho}, \tag{2.133}$$

where $k_\rho = \sqrt{k_x^2 + k_y^2}$, $\kappa = (\varepsilon_{\mathrm{t}} + \varepsilon_{\mathrm{b}})/2$ and $r_0 \approx d(\varepsilon_{\mathrm{ML}} - 1)/2$. The two terms κ and r_0 are screening terms related to the surrounding dielectric and screening from the planar slab. r_0 is also called the screening length.

2.6.1 Calculation of the Coulomb and Exchange Matrix Elements

The slow-rapid approximation in the tight-binding scheme

In the slow-rapid approximation, the evaluation of the Coulomb and exchange matrix elements of (2.130) and (2.131) hinges on writing the wave functions as on the Bloch form $\psi_{n\mathbf{k}} = \frac{1}{\sqrt{A}} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r})$ where $u_{n\mathbf{k}}$ is a lattice-periodic function, normed to $\int_{\mathrm{UC}} |u_{n\mathbf{k}}(\mathbf{r})|^2 \, \mathrm{d}^3\mathbf{r} = A_{\mathrm{UC}}$.[14] In order to evaluate (2.130) one can say that the product $\psi_{c\mathbf{k}}(\mathbf{r})\psi^*_{c'\mathbf{k}'}(\mathbf{r})$ contains a slowly varying factor $\mathrm{e}^{\mathrm{i}(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}}$ when the distance $|\mathbf{k} - \mathbf{k}'|$ is small and a rapidly varying factor $u_{c\mathbf{k}} u^*_{c\mathbf{k}'}$. Another comparatively slowly varying factor of (2.130) is the potential $W_C(\mathbf{r} - \mathbf{r}')$. Such an integral with an integrand with slowly and rapidly varying factors can be approximated as an integral over the slowly and rapidly varying factors individually.[14] Using this approximation twice for both the $\mathbf{r}$ and $\mathbf{r}'$ integration in the Coulomb matrix element yields[14]

$$\left\langle vc\mathbf{k} \middle| W_C \middle| v'c'\mathbf{k}' \right\rangle \approx \frac{1}{A} I_{v\mathbf{k},v'\mathbf{k}'} I^*_{c\mathbf{k},c'\mathbf{k}'} W_C(\mathbf{k} - \mathbf{k}'), \tag{2.134}$$

$$I_{n\mathbf{k},n'\mathbf{k}'} = \frac{1}{A_{\mathrm{UC}}} \int u_{n\mathbf{k}}(\mathbf{r}) u^*_{n'\mathbf{k}'}(\mathbf{r}) \mathrm{d}^3\mathbf{r}. \tag{2.135}$$

Here the phase factors has given the Fourier transform of the potential, notably in 2D instead of 3D as in the previous section. For the exchange integral, the phase factors instead cancel, yielding[14]

$$\left\langle vc\mathbf{k} \middle| V_X \middle| v'c'\mathbf{k}' \right\rangle \approx \frac{1}{A} I_{v\mathbf{k},c\mathbf{k}} I^*_{v'\mathbf{k}',c'\mathbf{k}'} \int V_X(\mathbf{r}) \mathrm{d}^3\mathbf{r}. \tag{2.136}$$

Following the approaches taken in Trolle et al.[28] and Ridolfi et al.[29] (and to be touched briefly in the discussion) we write $I_{n\mathbf{k},n'\mathbf{k}'}$ as

$$I_{n\mathbf{k},n'\mathbf{k}'} = \sum_i c_i^{n\mathbf{k}} c_i^{*n'\mathbf{k}'} \tag{2.137}$$

This is simply the dot-product between the two eigenvectors in the orthogonal Bloch-basis, meaning the exchange term will be zero in this approximation. Furthermore, $I_{n\mathbf{k},n'\mathbf{k}'}$ is zero if the states $\psi_{n\mathbf{k}}$ and $\psi_{n'\mathbf{k}'}$ have different parity. This is the special case of $P_{cc'} = 1 = P_{vv'}$ from (2.132) and results in the decoupling of the Hamiltonian into four blocks. The block overlaps $I_{n\mathbf{k},n'\mathbf{k}'}$ are furthermore not gauge-invariant.[30]

Using the relation $\frac{1}{A} = \frac{\mathrm{d}^2\mathbf{k}}{4\pi^2}$ we have the Bethe-Salpeter equation in this approximation reads

$$\tilde{E}_{vc\mathbf{k}} A_{vc\mathbf{k}} - \frac{1}{4\pi^2} \sum_{v',c'} \int_{BZ} W_C^{vc\mathbf{k},v'c'\mathbf{k}'} A_{v'c'\mathbf{k}'} \mathrm{d}^2\mathbf{k}' = E_{\mathrm{exc}} A_{vc\mathbf{k}}. \tag{2.138}$$

This equation is in principle an infinite dimensional eigenvalue problem, but can of course be discretised on a fine **k**-mesh. To include spin into this framework, W_C can be taken to be diagonal in spin,[31] meaning if two well-defined spin states exist, they will not couple through W_C and the BSE can then be solved seperately for each spin state.
It is straight-forward to discretise (2.138), of course being careful with the diagonal elements. In the case of isotropic screening, one encounters the integral over a small volume $\mathrm{d}A_k$

$$\int_{\mathrm{d}A_k} \frac{1}{k(\kappa + r_0 k)} \mathrm{d}^2 k = 2\pi \int_0^{R_k} \frac{1}{\kappa + r_0 k} \mathrm{d}k = 2\pi \frac{\log(1 + r_0 R_k/\kappa)}{r_0}. \tag{2.139}$$

In the case of a nonisotropic screening (e.g phosphorene, Figure E.5), introduced in r_0, one can instead do the integral numerically on the same small volume element $\mathrm{d}A_k$.

Equation (2.134) is actually an approximation since, strictly speaking, **k** is not restricted to the first Brillouin zone. A more elaborate approach to the Coulomb and exchange terms therefore involves summing over all reciprocal lattice vectors and incorporating a factor $\mathrm{e}^{\mathrm{i}\mathbf{G}\cdot\mathbf{r}}$ into the integral (2.135).[32] However, for localised excitons, the effect of summing over all reciprocal lattice vectors becomes negligible.[29] Equation (2.138) does of course also apply to the case of nonzero exchange as is the case when more **G**'s are involved where the exchange term is just put under the **k**-integral in the same fashion as the Coulomb term.

The Wannier Model

In the case of two isolated bands with the well-known effective mass dispersion relation $E_{vc} = E_g + \frac{\hbar^2 k^2}{2m_e^*} + \frac{\hbar^2 k^2}{2m_h^*} \equiv E_g + \frac{\hbar^2 k^2}{2\mu}$ the slow-rapid approximation in its roughest approximation yields the equation[14]

$$\left[E_g - \frac{\hbar^2}{2\mu}\nabla^2 - W(\mathbf{r})\right]\psi(\mathbf{r}) = E_{\mathrm{exc}}\psi(\mathbf{r}), \tag{2.140}$$

meaning the excitons can be thought of as bound states of hydrogenic atom, with the modified potential W and zero potential at E_g. In such a generalised hydrogen model using the Rytova-Keldish potential from (2.133), the binding energy of the lowest s-exciton relative to E_g reads[33]

$$E_B^{\mathrm{2D}} = \frac{8\mu}{\left(1 + \sqrt{1 + \frac{16 r_0 \mu}{3}}\right)^2} \quad , \tag{2.141}$$

in units of Hartree. Within the Wannier model we will have excitons with regular s-, p- d-symmetries and so fourth. However, when we solve the Bethe-Salpeter equation (2.138) we do so in **k**-space. Within the Wannier model one writes the wave function in **k**-space as simply the Fourier transform of $\psi(\mathbf{r})$. In particular, this means if $\psi(\mathbf{r}) = R(r)Y(\theta)$, where in 2D $Y(\theta) = \mathrm{e}^{\mathrm{i}m_l\theta}$,[34] then $\psi(\mathbf{k})$ will be given as[35]

$$\begin{aligned} \psi(k, \theta_{\mathbf{k}}) &= \int_0^\infty R(r) r \int_0^{2\pi} \mathrm{e}^{-\mathrm{i}(kr\cos(\theta_k - \theta) - m_l\theta)} \mathrm{d}\theta \mathrm{d}r \\ &= -2\pi \mathrm{i}^{-m_l} \mathrm{e}^{\mathrm{i}m_l\theta_{\mathbf{k}}} \int_0^\infty R(r) r J_{m_l}(kr) \mathrm{d}r, \end{aligned} \tag{2.142}$$

resulting in $\psi(\mathbf{k})$ also having the same angular symmetry with an additional dependence encapsulated in the integral over r of R and the Bessel function J_{m_l}. This means the symmetry in $\mathbf{k}$-space also implies a symmetry in $\mathbf{r}$-space, and within the Wannier model one can therefore tell the angular symmetry of the exciton from its $\mathbf{k}$-space shape.

2.7 The Many-Body Response

The ability to find the many-body response is as mentioned in the introduction important in 2D materials because of the modified screening and 2D geometry of the materials considered in this thesis. This subsection will establish how to calculate this linear and non-linear response in a variety of ways. We simply state many of the results needed for the many-body responses.

2.7.1 First-Order Response Through the Green's Function of the System

The theory of section 2.6 allows for the calculation of the excitonic states and thereby allows the response hereof to be calculated. To first-order, the many-body susceptibility reads[14]

$$\chi = \frac{2\mathrm{e}^2\hbar^2}{\epsilon_0 m_e^2 A}\sum_i \frac{\left|\langle 0|\,\hat{P}\,|i\rangle\right|^2}{E_i(E_i^2-\hbar^2(\omega+\mathrm{i}\Gamma)^2)}, \tag{2.143}$$

where $|i\rangle$ is an eigenstate of the many-body Hamiltonian (2.129) and A the crystal area. The sum includes spin-summation. The many-body momentum operator is given as $\hat{P} = \sum_{i=1}^{N} \hat{p}_i$. Furthermore, in the limit as $\Gamma \to 0$, each term in the sum, behaves as a Dirac δ-function, which is evident from the calculation[12]

$$\begin{aligned}\lim_{\eta\to 0}\mathrm{Im}\frac{1}{E_i(E_i^2-\hbar^2(\omega+\mathrm{i}\eta)^2)} &= \lim_{\eta\to 0}\frac{2\hbar^2\omega\eta E_i}{(E_i^3-\hbar^2(\omega^2-\eta^2)E_i)^2+4\hbar^4\omega^2\eta^2E_i^2}\\ &= \lim_{\eta\to 0}\frac{2\hbar^2\omega\eta E_i}{E_i^2(E_i^2-\hbar^2(\omega^2-\eta^2))^2+(2\hbar^2\omega\eta E_i)^2}\\ &= \pi\delta(E_i^2(E_i-\hbar\omega)) = \pi\frac{\delta(E_i-\hbar\omega)}{E_i^2},\end{aligned} \tag{2.144}$$

meaning the imaginary part of the susceptibility can be written

$$\mathrm{Im}\chi = \frac{2\pi\mathrm{e}^2}{\epsilon_0 m_e^2\omega^2 A}\sum_i \left|\left\langle 0\right|\hat{P}\left|i\right\rangle\right|^2\delta(E_i-\hbar\omega) \tag{2.145}$$

when using $E_i = \hbar\omega$ for all non-zero terms in the sum. This sum actually turns out to be a diagonal element of what is known as the Green's function of the system.[14] If we define the Green's function as the operator[14]

$$\hat{G}(z) = \sum_i \frac{|i\rangle\langle i|}{z-E_i}, \tag{2.146}$$

The most interesting property of $\hat{G}$ is seen when computing imaginary part of the diagonal matrix element with the state $|P\rangle = \hat{P}\,|0\rangle$:

$$\begin{aligned}\lim_{\eta\to 0}\mathrm{Im}\left\langle P\right|\hat{G}(\hbar\omega-\mathrm{i}\eta)\left|P\right\rangle &= \lim_{\eta\to 0}\mathrm{Im}\sum_i\frac{\langle 0|\,\hat{P}\,|i\rangle\langle i|\,\hat{P}\,|0\rangle}{\hbar\omega-\mathrm{i}\eta-E_i}\\ &= \sum_i\left|\left\langle 0\right|\hat{P}\left|i\right\rangle\right|^2\delta(E_i-\hbar\omega)\\ &\propto \mathrm{Im}\chi.\end{aligned} \tag{2.147}$$

Only a factor is missing in order for this diagonal element to be equal to the imaginary part of the susceptibility. This factor is obtained by comparison with (2.145) and taking into account whenever the Green's function under consideration is spin-degenerate or not. The iη-term added in the argument of $\hat{G}$ also works as a broadening term, equivalent to the iη term of (2.143). While the relation (2.147) requires $\eta = 0^+$, for actual calculations η is just taken as a small quantity. We will furthermore in chapter 3 give the simple relation between the susceptibility χ and the conductivity σ.

The starting point of this subsection was (2.143), which involves squares of momentum matrix-elements, meaning the χ considered here is a diagonal element of the susceptibility tensor of the material. Off-diagonal components of the first-order susceptibility tensor would involve off-diagonal elements of $\hat{G}$ on the form $\langle P_i | \hat{G} | P_j \rangle$.[36] The point-group symmetries of $\mathrm{In_2Se_2}$ and $\mathrm{MoS_2}$ ensures we need not worry about these components in those cases, but were we to consider more exotic situations, the relations[36]

$$\left\langle P_i + \mathrm{i}P_j \middle| \hat{G} \middle| P_i + \mathrm{i}P_j \right\rangle = \left\langle P_i \middle| \hat{G} \middle| P_i \right\rangle + \left\langle P_j \middle| \hat{G} \middle| P_j \right\rangle + 2\mathrm{iIm} \left\langle P_i \middle| \hat{G} \middle| P_j \right\rangle, \tag{2.148}$$

$$\left\langle P_i + P_j \middle| \hat{G} \middle| P_i + P_j \right\rangle = \left\langle P_i \middle| \hat{G} \middle| P_i \right\rangle + \left\langle P_j \middle| \hat{G} \middle| P_j \right\rangle + 2\mathrm{Re} \left\langle P_i \middle| \hat{G} \middle| P_j \right\rangle, \tag{2.149}$$

would enable the calculation of such susceptibility elements when all the diagonal elements are known. $\hat{G}$ is not a hermitian operator with $z \in \mathcal{C}$ and therefore its diagonal elements can contain a nonzero imaginary part. The above equations show that $\hat{G}$ is instead a skew-hermitian operator.

The Lanczos-Haydock Routine

The above calculations show that the diagonal matrix element $\langle P | \hat{G}(\hbar\omega - \mathrm{i}\eta) | P \rangle$ is the only quantity needed to calculate the susceptibility to first-order, therefore how to calculate this quantity is an important question. The way to do this is through the Lanczos-Haydock routine, which turns out to be able to compute this matrix-element knowing only $|P\rangle$ and the Hamiltonian $\hat{\mathcal{H}}^{\mathrm{MB}}$.[14] We already know $\hat{\mathcal{H}}^{\mathrm{MB}}$ from section 2.6 and therefore need $|P\rangle$. The basis in which the Hamiltonian was written in section 2.6 was a basis of singly excited states $|v_i \to c_j\rangle = \frac{1}{\sqrt{2}} \left(\left| \overset{+}{v}_i \to \overset{+}{c}_j \right\rangle + \left| \overset{-}{v}_i \to \overset{-}{c}_j \right\rangle \right)$. This basis does not contain $|0\rangle$ since Brillouin's theorem states that $\hat{\mathcal{H}}^{\mathrm{MB}}$ does not couple it to singly excited states and therefore it is redundant in the basis. However, using basic linear algebra, the state $|P\rangle$ can be expanded in this basis as[14]

$$\left| P \right\rangle = \sum_{i,j} \left\langle v_i \to c_j \middle| \hat{P} \middle| 0 \right\rangle \left| v_i \to c_j \right\rangle = \sqrt{2} \sum_{i,j} \left\langle v_i \middle| \hat{p} \middle| c_j \right\rangle^* \left| v_i \to c_j \right\rangle, \tag{2.150}$$

where the general relation $\langle \Psi | \hat{F} | \Psi' \rangle = \langle \psi_k | \hat{f} | \psi'_k \rangle$ where $\hat{F}$ is a sum of single-electron operators like $\hat{P}$, and Ψ and Ψ' are Slater-determinants differing only by the k'th spin-orbital like $\left| \overset{\pm}{v}_i \to \overset{\pm}{c}_j \right\rangle$ and $|0\rangle$.[12] The single-particle momentum matrix elements can be obtained from the TB model, giving $|P\rangle$.

Turning now to the Lanczos-Haydock routine, it starts by setting[37] $b_1 = 0, |\psi_0\rangle = 0$ and $|\psi_1\rangle = |P\rangle / ||P\rangle|$ and then calculating the next step as[38] [37] [14]

$$\begin{aligned} a_i &= \left\langle \psi_i \middle| \hat{\mathcal{H}}^{\mathrm{MB}} \middle| \psi_i \right\rangle \\ b_{i+1} &= \left| \left| \hat{\mathcal{H}}^{\mathrm{MB}} \left| \psi_i \right\rangle - a_i \left| \psi_i \right\rangle - b_i \left| \psi_{i-1} \right\rangle \right| \right| \\ \left| \psi_{i+1} \right\rangle &= \frac{\hat{\mathcal{H}}^{\mathrm{MB}} |\psi_i\rangle - a_i |\psi_i\rangle - b_i |\psi_{i-1}\rangle}{b_{i+1}}. \end{aligned} \tag{2.151}$$

The definition of ψ_{i+1}, a_i and b_i are so cleverly chosen that $|\psi_{i+1}\rangle$ is orthogonal on the previous states, ensuring an othonormal basis and that $\hat{\mathcal{H}}^{\mathrm{MB}}$ is a tridiagonal matrix in this basis. From here the matrix element $\langle P|\,\hat{G}\,|P\rangle$ is calculated as[14]

$$\left\langle P\middle|\,\hat{G}(z)\,\middle|P\right\rangle = \left|\middle|P\right\rangle\Big|^2 \frac{1}{z - a_1 - \frac{b_2^2}{z-a_2-\frac{b_3^2}{\cdots}}}. \tag{2.152}$$

This technique circumvents the need to know the eigenstates $|i\rangle$ from (2.143), which is especially convenient in the case of large matrices, which cannot be stored in the RAM of a computer. Instead, only matrix vector products need to be calculated. These are more easily handled, as blocks of the matrix can be loaded into the RAM individually and the matrix-vector product calculated. To obtain a converted result from the recursive fraction (2.152), only a small number of all the b_i's and a_i's need to be calculated. Therefore, instead of the computational cost scaling like N^3 like a normal eigenvalue problem, this problem scales as kN^2, where $k << N$.[37]

For numerical stability, $|\psi_{i+1}\rangle$ has to be orthogonalised to the previous states, or else degenerate eigenvalues will appear in the truncated Hamiltonian. This orthogonalisation can be done using the Gram-Schmidt procedure. Also, for states with different parity, since the Hamiltonian is block-diagonal in different parity states, the response can be calculated separately for each block and added in the end. Only the upper half of the matrix of the Hamiltonian needs calculation since (2.129) is hermitian, and the matrix-vector products involving $[\hat{\mathcal{H}}^{\mathrm{MB}}]$ is straight-forwardly calculated from this.

2.7.2 The Response Through Diagonalising the Two-Body Hamiltonian

The density-matrix framework described in subsection 2.3.3 can also be generalised to include many-body effects, as is done in Taghizadeh and Pedersen[32]. Here the conductivity tensors equivalent to (2.90), (2.118), (2.107) and (2.122) are derived in terms of many-body momentum and position matrix elements, namely the normalised position operator $\mathbf{X} = \frac{m_e}{\hbar}\mathbf{R}$, the Heisenberg momentum operator $\mathbf{\Pi} \equiv \mathrm{i}\left[\hat{\mathcal{H}}^{\mathrm{MB}}, \mathbf{X}\right]$, and the regular momentum operator $\mathbf{P}$.[32] Because of the general commutator relation $\langle n|\left[\hat{\mathcal{H}}, \hat{o}\right]|m\rangle = E_{nm}\langle n|\,\hat{o}\,|m\rangle$, the matrix elements of $\mathbf{\Pi}$ can be found as $\mathbf{\Pi}_{nm} = \mathrm{i}(E_n - E_m)\mathbf{X}_{nm}$, while the matrix elements of the usual position and momentum operators. The matrix elements needed for calculation of first and second order conductivities are transitions between the ground state $|0\rangle$ and an excited state $|n\rangle$, together with the matrix elements between excited states $|n\rangle$ and $|m\rangle$. These read[32]

$$\mathbf{P}_{0n} = \sum_{vc\mathbf{k}} A^n_{vc\mathbf{k}}\mathbf{p}_{vc\mathbf{k}} \quad , \quad \mathbf{P}_{nm} = \sum_{vc\mathbf{k}} A^{n*}_{vc\mathbf{k}}\left[\sum_{c'} A^m_{vc'\mathbf{k}}\mathbf{p}_{cc'} - \sum_{v'} A^m_{v'c\mathbf{k}}\mathbf{p}_{v'v}\right], \tag{2.153}$$

$$\mathbf{X}_{0n} = \sum_{vc\mathbf{k}} A^n_{vc\mathbf{k}}\mathbf{x}_{vc\mathbf{k}} \quad , \quad \mathbf{X}_{nm} = \sum_{vc\mathbf{k}} A^{n*}_{vc\mathbf{k}}\left[\sum_{c'} A^m_{vc'\mathbf{k}}\mathbf{x}_{cc'} - \sum_{v'} A^m_{v'c\mathbf{k}}\mathbf{x}_{v'v}\right]. \tag{2.154}$$

These matrix elements originate from the density matrix framework and elements between excited states are similar, but not equal, to the result of taking matrix elements of Slater determinants. Again, the intraband part of (2.154) can be handled by identifying the commutator between the expansion coefficients $A^m_{vc\mathbf{k}}$ and $\mathbf{x}^{(i)}$, as done in (B.2) and splitting position matrix elements into $\mathbf{X}_{nm} = \mathbf{Y}_{nm} + \frac{m_e}{\hbar}\mathbf{Q}_{nm}$, where $\mathbf{Y}_{nm}$ and $\mathbf{Q}_{nm}$ contains the interband and intraband contributions respectively.[32] $\mathbf{Q}_{nm}$ is given as[32]

$$\mathbf{Q}_{nm} = \mathrm{i}\sum_{vc\mathbf{k}} A^{n*}_{vc\mathbf{k}}(A^m_{vc\mathbf{k}})_{;\mathbf{k}} \tag{2.155}$$

Having calculated these matrix elements from the solution of the BSE, the current response to first-order with the length gauge can be calculated as[32]

$$\boldsymbol{\sigma}^{[1]}(\omega) = -\frac{ge^2\hbar}{m_e^2 A}\sum_n \left[\frac{\mathbf{\Pi}_{0n}\mathbf{X}_{0n}^*}{\hbar\omega' - E_n} - \frac{\mathbf{\Pi}_{0n}^*\mathbf{X}_{0n}}{\hbar\omega' + E_n}\right], \tag{2.156}$$

where $\omega' = \omega + \mathrm{i}\eta$.

The second order current response can be calculated in the length gauge as[32]

$$\begin{aligned}\boldsymbol{\sigma}^{[2]} = -\frac{ge^3\hbar^2}{4m_e^3 A}\sum_{nm}\Bigg[&\frac{\mathbf{\Pi}_{0n}\mathbf{X}_{nm}\mathbf{X}_{0m}^*}{(\hbar\omega_2' - E_n)(\hbar\omega_\mathrm{q}' - E_m)} + \frac{\mathbf{\Pi}_{0n}^*\mathbf{X}_{nm}^*\mathbf{X}_{0m}}{(\hbar\omega_2' - E_n)(\hbar\omega_\mathrm{q}' - E_m)} \\ &- \frac{\mathbf{X}_{0n}\mathbf{\Pi}_{nm}\mathbf{X}_{0m}^*}{(\omega_\mathrm{q}' + E_n)(\hbar\omega_\mathrm{p}' - E_m)}\Bigg],\end{aligned} \tag{2.157}$$

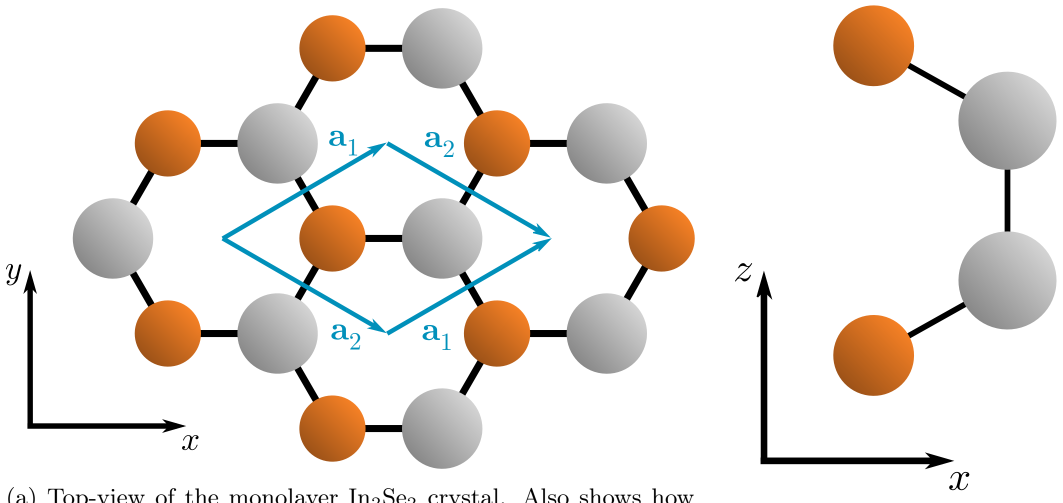
and within the conventional velocity gauge as[32]

$$\begin{aligned}\boldsymbol{\sigma}^{[2]} = \frac{ge^3}{4m_e^3 A\omega_\mathrm{p}\omega_\mathrm{q}}\sum_{nm}\Bigg[&\frac{\mathbf{P}_{0n}\mathbf{P}_{nm}\mathbf{P}_{0m}^*}{(\hbar\omega_2' - E_n)(\hbar\omega_\mathrm{q}' - E_m)} + \frac{\mathbf{P}_{0n}^*\mathbf{P}_{nm}^*\mathbf{P}_{0m}}{(\hbar\omega_2' - E_n)(\hbar\omega_\mathrm{q}' - E_m)} \\ &- \frac{\mathbf{P}_{0n}\mathbf{P}_{nm}\mathbf{P}_{0m}^*}{(\hbar\omega_\mathrm{q}' + E_n)(\hbar\omega_\mathrm{p}' - E_m)}\Bigg]\end{aligned} \tag{2.158}$$

In these expressions we set $\omega_2' = \omega_\mathrm{p} + \omega_\mathrm{q} + \mathrm{i}\eta$. As the expressions are given above, they discriminate between ω_q and ω_p, but by adding similar terms with ω_q and ω_p interchanged fixes this. These equations require the knowledge of the eigenstates $|n\rangle$, with means the BSE Hamiltonian needs to be diagonalised. However, a hybrid approach using the Lanczos-Haydock routine to truncate the matrix of the Hamiltonian to a subspace of much lower dimension compared to the original is also possible, as the lowest energy states converges reasonably fast.

3 | Indium(II) Selenide Monolayer

The goal of this chapter is to perform optical response calculations on a monolayer indium(II) selenide (In_2Se_2) crystal through the use of a TB model with a basis of maximally localised Wannier functions. First, the structure of the crystal shall be presented along with the details of the DFT calculations. Afterwards, the details of the Wannierisation procedures shall be covered, and the DFT and TB band structures are compared. The focus shall then turn to the current density response to a specific electric field and to the calculation of the crystal's conductivity tensor; the complexity of which is reduced greatly by taking into account the symmetries of the crystal. These calculations shall be carried out including a varied number of bands. Furthermore, two different gauges are employed, the length gauge and the velocity, and the results from the two different methods shall be compared. Lastly, excitonic effects are included in the conductivity calculations in order to increase precision.



(a) Top-view of the monolayer In_2Se_2 crystal. Also shows how the lattice vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ encapsulate the unit cell.

(b) Side-view of the unit cell.

Figure 3.1: The hexagonal lattice structure of a monolayer In_2Se_2 crystal. Grey spheres represent indium atoms and orange spheres selenium atoms.

3.1 Crystal Structure and Band Diagram

The crystal structure of the In_2Se_2 monolayer, shown in Figure 3.1, is almost identical to that of hexagonal transition metal dichalcogenides monolayers with the exception that the centre transition metal is replaced with a dimer of the post-transition metal indium.[39] As such, the primitive lattice vectors $\mathbf{a}$ and reciprocal lattice vectors $\mathbf{G}$ are on the form

$$\mathbf{a}_1 = \frac{a}{2}\begin{bmatrix}\sqrt{3}\\1\end{bmatrix} \quad \text{and} \quad \mathbf{a}_2 = \frac{a}{2}\begin{bmatrix}\sqrt{3}\\-1\end{bmatrix}, \qquad \mathbf{G}_1 = \frac{2\pi}{\sqrt{3}a}\begin{bmatrix}1\\\sqrt{3}\end{bmatrix} \quad \text{and} \quad \mathbf{G}_2 = \frac{2\pi}{\sqrt{3}a}\begin{bmatrix}1\\-\sqrt{3}\end{bmatrix}, \tag{3.1}$$

with a lattice constant a. Furthermore, with the coordinate system applied in Figure 3.1 it is clear that the crystal structure has inversion symmetries along the y- and z-axes with an additional 120°-rotational symmetry about the z-axis. Therefore, only three distances are necessary in order to construct the crystal: the lattice constant a and the bond lengths of the dimers d_{In} and d_{Se}. Initially, these distances are[39]

$$a = 4.05 \text{ Å}, \qquad d_{\text{In}} = 2.79 \text{ Å} \qquad \text{and} \qquad d_{\text{Se}} = 5.41 \text{ Å}.$$

The primitive reciprocal lattice vectors $\mathbf{G}$ establish the first Brillouin zone as shown in Figure 3.2. Additionally, the crystal's point group symmetries and time-reversal symmetry can be applied to the first Brillouin zone to yield the irreducible Brillouin zone. Figure 3.2 also shows how the high symmetry points Γ, M and K form a triangle which encapsulates the irreducible Brillouin zone. In a coordinate system spanned by $\mathbf{G}_1$ and $\mathbf{G}_2$, these points have the coordinates $(0,0)$, $\left(\frac{1}{2},\frac{1}{2}\right)$ and $\left(\frac{2}{3},\frac{1}{3}\right)$, respectively.

For the DFT calculations, Quantum ESPRESSO version 6.4.1 is used[16] [17] along with PAW-type pseudopotentials[40] from pslibrary 1.0.0.[41] The chosen pseudopotentials are scalar relativistic (no spin-orbit coupling) and have a nonlinear core correction. For the exchange correlation functionals, the choice is between PBE[42] and PBEsol.[43] PBEsol is chosen since this functional has been shown to increase accuracy of the crystal structure parameters when relaxing the structure.[44]

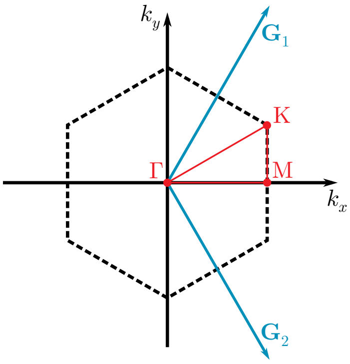
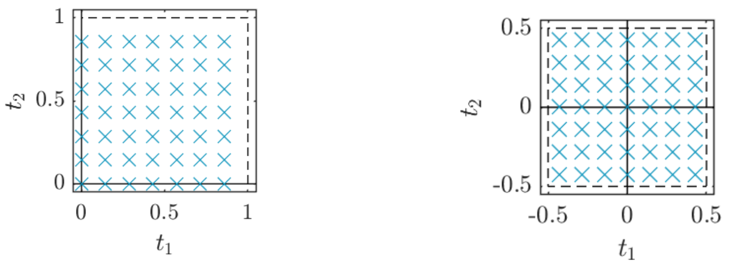


Figure 3.2: The hexagonal first Brillouin zone of a monolayer In_2Se_2 crystal. The red triangle encapsulates the irreducible Brillouin zone.

The pseudopotential for each indium atom has 13 valence electrons, and by comparing with its electron configuration, 10 of these valence electrons can be attributed to d-orbitals, 2 to s-orbitals and 1 to p-orbitals. For the selenium atoms, the pseudopotential has 6 valence electrons; 2 of which can be attributed to s-orbitals and 4 to p-orbitals.

Aside from the pseudopotentials, Quantum ESPRESSO needs two cutoff energies, which serve to fix the number of plane waves in the basis sets, and a grid of $\mathbf{k}$-points from which to sample. For the DFT calculations in this chapter, the cutoff for the wave functions and for the potential are chosen to be 70 Ry and 700 Ry, respectively. There are two common choices for the $\mathbf{k}$-point grid: The Monkhorst-Pack grid and a Γ-centred grid. Examples of both grid types are shown in Figure 3.3, and for self-consistent field calculations in this chapter, a Γ-centred 25 × 25 $\mathbf{k}$-grid is used.



(a) Monkhorst-Pack.

(b) Γ-centred.

Figure 3.3: Two common sampling grids with 7 × 7 $\mathbf{k}$-points. t_1 and t_2 form a wavevector through $\mathbf{k} = t_1\mathbf{G}_1 + t_2\mathbf{G}_2$.

Now with the established parameters, the In_2Se_2 crystal structure is relaxed which gives the distance parameters

$$a = 3.98\ \text{Å}, \qquad d_{\text{In}} = 2.77\ \text{Å} \qquad \text{and} \qquad d_{\text{Se}} = 5.36\ \text{Å}.$$

These distances are always used in this chapter from here on. After the relaxation procedure, we run an SCF calculation from which the band diagram can be determined. The result is shown in Figure 3.4 where the general shape of the bands is in good agreement with literature.[45] An indirect band gap $E_g = 1.71$ is formed between the lowest conduction state at the Γ-point and the highest valence state between the K- and Γ-point which is an improvement over the 1.55 eV band gap which you can get from PBE.

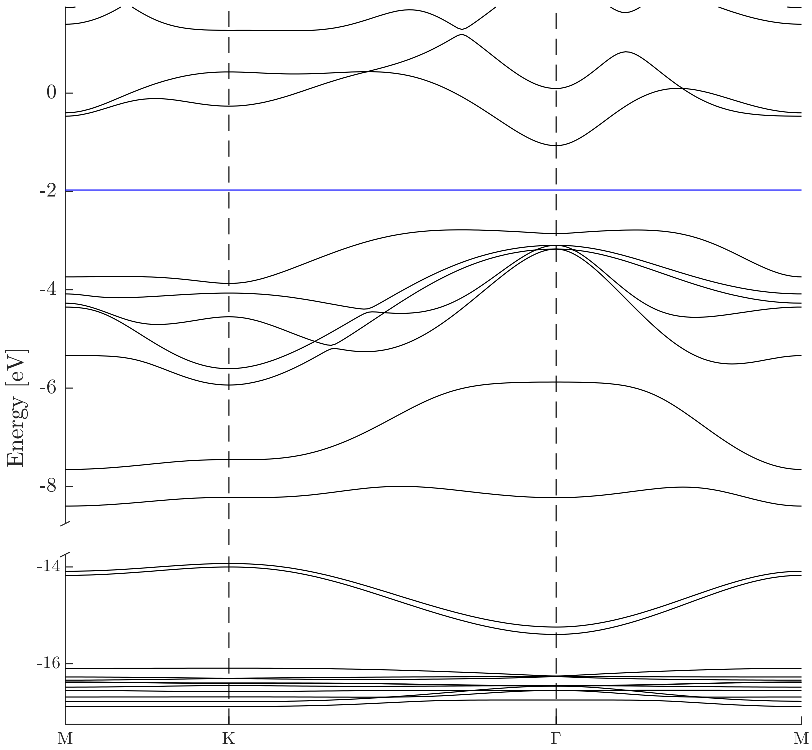


Figure 3.4: Spin-degenerate DFT band structure for the relaxed In_2Se_2 crystal. $\mathbf{k}$-points are scanned on straight lines between the high-symmetry points. The Fermi level $E_F = -1.97$ eV is marked as a blue line.

The DFT band structure has 19 valence bands; 12 of which lie at the bottom separated from the upper valence states by at least 5.5 eV. These bottom valence bands can house 24 electrons, and we postulate that they can be attributed to the d-orbitals of indium and s-orbitals of selenium. This suggest that pseudopotentials which count these electrons as part of the frozen core may have been beneficial for computation time, alas such pseudopotentials were unavailable at the time of writing.

The next step is the Wannierisation of the Bloch orbitals from the DFT calculations. For this, the program Wannier90 version 3.0.0[46] is used. This program mainly needs a set of bands to include and a set of initial projectors from which to start the Wannierisation procedure. Based on the brief discussion of the DFT band structure, the bottom 12 valence bands can be excluded and the initial projectors can be chosen as the s- and p-orbitals of indium and the p-orbitals of selenium. As such, the TB band structure will consist of 7 valence bands and 7 conduction bands. Additional conduction bands can be included in the Wannierisation procedure without changing the number of bands in the final TB model. This engages the disentanglement procedure[47] which increases the precision of the TB model during our calculations. Now, a 15 × 15 Γ-centred $\mathbf{k}$-grid is used, and Wannier90 then produces a set of parameters $\mathcal{H}_{mn,\mathbf{R}}$ from the DFT data which will yield the TB Hamiltonian matrix elements through a Bloch sum

$$\mathcal{H}_{mn}(\mathbf{k}) = \sum_{\mathbf{R}} \frac{\mathcal{H}_{mn,\mathbf{R}}}{d_{\mathbf{R}}} \exp(i\mathbf{k}\cdot(\mathbf{R}+\mathbf{r}_n-\mathbf{r}_m)), \tag{3.2}$$

where $\mathbf{r}_m$ is the centre of the m'th Wannier function and $d_{\mathbf{R}}$ is the degeneracy of the Wigner-Seitz cell associated with $\mathbf{R}$. The TB Hamiltonian can then be constructed and its eigenvectors and energy eigenvalues found for various values of $\mathbf{k}$.

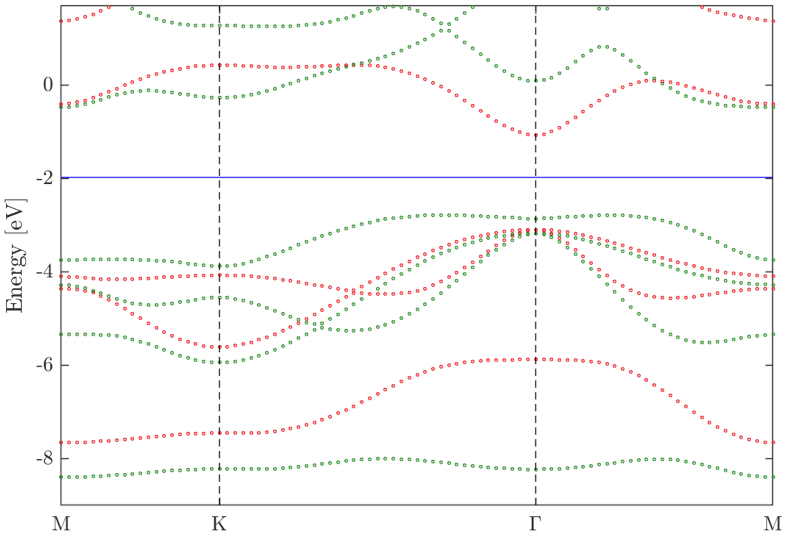


Figure 3.5: TB band structure obtained through Wannier90. Green-coloured states consist of at least 99.9% positive parity states and red-coloured states 99.9 % negative parity states.

Monolayer In_2Se_2 has z-inversion symmetry which means the Hamiltonian commutes with the z-inversion operator, implying that the eigenstates have either positive or negative z-parity. It should therefore be possible to construct new basis functions with either positive or negative z-parity from the previous basis through multiplication with a unitary matrix U_p. The entrances of this unitary matrix can be deduced from the entrances of a single eigenvector and the choice of initial projectors. For example for each orbital on one selenium atom, there must be an equivalent orbital on the opposing selenium atom. The basis functions with these two orbitals can be added and subtracted from each other and renormalised by a factor $1/\sqrt{2}$ to form two new basis functions with definite z-parity.

Each energy eigenvalue can now be associated with either positive or negative z-parity through the eigenvectors in the new basis. The TB band diagram can be seen in Figure 3.5 where green points are energy states which consist of at least 99.9 % positive parity basis functions, and red points consist of 99.9 % negative parity basis functions. Numerically, the positive and negative parity states still couple; however, this small coupling shall be disregarded. As such, the transformed TB Hamiltonian $U_p\mathcal{H}^{\mathrm{TB}}U_p^\dagger$ can be separated into four blocks

$$U_p\mathcal{H}^{\mathrm{TB}}U_p^\dagger = \begin{bmatrix} \mathcal{H}^{++} & \mathcal{H}^{+-} \\ \mathcal{H}^{-+} & \mathcal{H}^{--} \end{bmatrix} \qquad \text{where} \qquad \mathcal{H}^{-+} \approx \mathcal{H}^{+-} \approx 0, \tag{3.3}$$

reducing the eigenvalue problem from one 14×14 matrix to two 7×7 matrices, halving the number of elements. Finally in Figure 3.6, the DFT band structure is compared with the parity-decoupled TB band structure, showing good agreement. Interestingly, the parity decoupling shows how the indirect band gap from the DFT structure splits into a direct band gap for the negative parity states and an indirect band gap for the positive parity states. The direct band gap is then $E_g^- = 2.026$ eV.

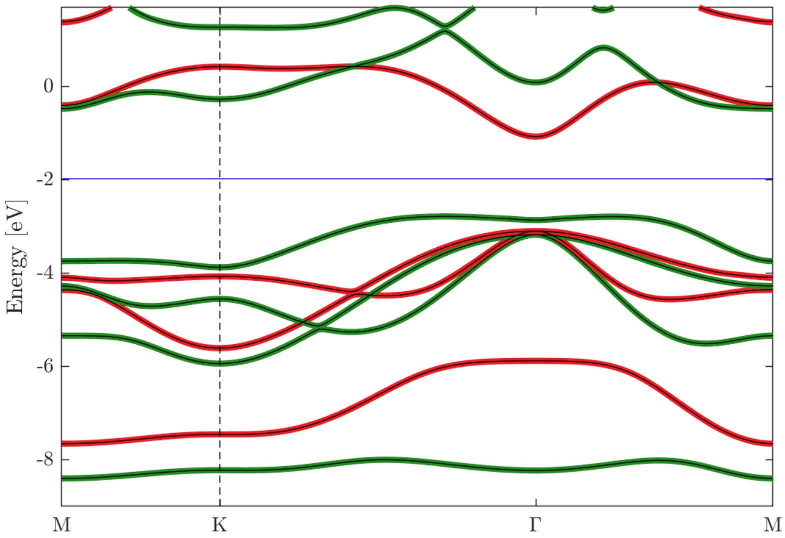


Figure 3.6: Thin black lines representing the DFT band structure obtained through Quantum ESPRESSO superimposed on top of coloured lines representing the parity-decoupled TB band structure obtained through Wannier90 and a unitary transformation.

3.2 Response Spectra in the Independent-Particle Approximation

The focus shall now turn to calculating the first- and second-order independent-particle response spectra for indium selenide crystals. First, it shall be shown how the crystal symmetries reduce the conductivity tensor by using the theory developed in section 2.4. Respectively, the matrices for the y- and z-inversion symmetries and the 120°-rotational symmetry about the z-axis are

$$S_y = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad S_z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \quad \text{and} \quad S_{\text{rot}} = \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{3.4}$$

Now, consider the conductivity tensors of rank 2 and 3 which yield the first- and second-order current density responses

$$\begin{gathered} \overleftrightarrow{\sigma}^{[2]} = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{bmatrix} \quad \text{and} \\ \overleftrightarrow{\sigma}^{[3]} = \left[\begin{array}{ccc|ccc|ccc} \sigma_{xxx} & \sigma_{xxy} & \sigma_{xxz} & \sigma_{yxx} & \sigma_{yxy} & \sigma_{yxz} & \sigma_{zxx} & \sigma_{zxy} & \sigma_{zxz} \\ \sigma_{xyx} & \sigma_{xyy} & \sigma_{xyz} & \sigma_{yyx} & \sigma_{yyy} & \sigma_{yyz} & \sigma_{zyx} & \sigma_{zyy} & \sigma_{zyz} \\ \sigma_{xzx} & \sigma_{xzy} & \sigma_{xzz} & \sigma_{yzx} & \sigma_{yzy} & \sigma_{yzz} & \sigma_{zzx} & \sigma_{zzy} & \sigma_{zzz} \end{array}\right]. \end{gathered} \tag{3.5}$$

The easiest symmetries to apply are S_y and S_z which imply that every tensor element with an odd amount of either y- or z-indices are zero, so

$$\begin{gathered} \overleftrightarrow{\sigma}^{[2]} = \begin{bmatrix} \sigma_{xx} & 0 & 0 \\ 0 & \sigma_{yy} & 0 \\ 0 & 0 & \sigma_{zz} \end{bmatrix} \quad \text{and} \\ \overleftrightarrow{\sigma}^{[3]} = \left[\begin{array}{ccc|ccc|ccc} \sigma_{xxx} & 0 & 0 & 0 & \sigma_{yxy} & 0 & 0 & 0 & \sigma_{zxz} \\ 0 & \sigma_{xyy} & 0 & \sigma_{yyx} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{xzz} & 0 & 0 & 0 & \sigma_{zzx} & 0 & 0 \end{array}\right]. \end{gathered} \tag{3.6}$$

Next the effects of S_{rot} are considered, and since the inversion symmetries eliminated a great number of elements, this process is simplified. For the elements of $\overleftrightarrow{\sigma}^{[2]}$, rotation symmetry quickly yields that σ_{zz} is independent, and that $\sigma_{xx} = \sigma_{yy}$. For the elements of $\overleftrightarrow{\sigma}^{[3]}$, the rotation matrix instead reveals a system of equations

$$\begin{aligned} 3\sigma_{xxx} &= -\sigma_{xyy} - \sigma_{yxy} - \sigma_{yyx} \\ 3\sigma_{xyy} &= -\sigma_{xxx} + \sigma_{yxy} + \sigma_{yyx} \\ 3\sigma_{yxy} &= -\sigma_{xxx} + \sigma_{xyy} + \sigma_{yyx} \\ 3\sigma_{yyx} &= -\sigma_{xxx} + \sigma_{xyy} + \sigma_{yxy} \end{aligned} \quad \text{and} \quad \begin{aligned} \sigma_{xzz} &= -2\sigma_{xzz} \\ \sigma_{zxz} &= -2\sigma_{zxz} \\ \sigma_{zzx} &= -2\sigma_{zzx} \end{aligned}, \tag{3.7}$$

which is solved by $\sigma_{xxx} = -\sigma_{xyy} = -\sigma_{yxy} = -\sigma_{yyx}$ and $\sigma_{xzz} = \sigma_{zxz} = \sigma_{zzx} = 0$. So, the final form of the conductivity tensors of rank 2 and 3 in the indium selenide crystal is

$$\begin{gathered} \overleftrightarrow{\sigma}^{[2]} = \begin{bmatrix} \sigma_{xx} & 0 & 0 \\ 0 & \sigma_{xx} & 0 \\ 0 & 0 & \sigma_{zz} \end{bmatrix} \quad \text{and} \\ \overleftrightarrow{\sigma}^{[3]} = \left[\begin{array}{ccc|ccc|ccc} \sigma_{xxx} & 0 & 0 & 0 & -\sigma_{xxx} & 0 & 0 & 0 & 0 \\ 0 & -\sigma_{xxx} & 0 & -\sigma_{xxx} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]. \end{gathered} \tag{3.8}$$

Thus by using Neumann's principle in a system with y-inversion symmetry, z-inversion symmetry and 120°-rotation symmetry about the z-axis, the conductivity tensors of rank 2 and 3 have been reduced from containing 36 elements in total to only 3 distinct elements.

3.2.1 First-Order Response

Now, the crystal shall be perturbed by a monochromatic laser beam with frequency ω and amplitude $\mathcal{E}_0$. The electric field is chosen to have an in-plane polarisation

$$\boldsymbol{\mathcal{E}}(t) = \frac{\mathcal{E}_0}{\sqrt{2}} \cos(\omega t)(\overline{\mathbf{e}}_x + \overline{\mathbf{e}}_y) = \frac{\mathcal{E}_0}{2\sqrt{2}} \left(\mathrm{e}^{-\mathrm{i}\omega t} + \mathrm{e}^{\mathrm{i}\omega t}\right)(\overline{\mathbf{e}}_x + \overline{\mathbf{e}}_y) \,. \tag{3.9}$$

The first-order response corresponds to the situation where an electron absorbs a single photon whenever the energy of the photon matches a transition energy, at which point they can resonate. From section 2.5, it can be seen that an electric field on this form gives rise to a first-order current response

$$\begin{aligned} J_x^{(1)}(t) &= \frac{\mathcal{E}_0}{\sqrt{2}} \Big[(\sigma_{xx}(\omega) + \sigma_{xy}(\omega)) \exp(-\mathrm{i}\omega t) + (\sigma_{xx}(-\omega) + \sigma_{xy}(-\omega)) \exp(\mathrm{i}\omega t) \Big] \,, \\ J_y^{(1)}(t) &= \frac{\mathcal{E}_0}{\sqrt{2}} \Big[(\sigma_{yx}(\omega) + \sigma_{yy}(\omega)) \exp(-\mathrm{i}\omega t) + (\sigma_{yx}(-\omega) + \sigma_{yy}(-\omega)) \exp(\mathrm{i}\omega t) \Big] \,, \\ J_z^{(1)}(t) &= \frac{\mathcal{E}_0}{\sqrt{2}} \Big[(\sigma_{zx}(\omega) + \sigma_{zy}(\omega)) \exp(-\mathrm{i}\omega t) + (\sigma_{zx}(-\omega) + \sigma_{zy}(-\omega)) \exp(\mathrm{i}\omega t) \Big] \,. \end{aligned} \tag{3.10}$$

By using the rank 2 conductivity tensor elements from (3.7), the expressions reduce to

$$\mathbf{J}^{(1)}(t) = \frac{\mathcal{E}_0}{\sqrt{2}} \Big[\sigma_{xx}(\omega) \exp(-\mathrm{i}\omega t) + \sigma_{xx}(-\omega) \exp(\mathrm{i}\omega t) \Big] (\overline{\mathbf{e}}_x + \overline{\mathbf{e}}_y) \,. \tag{3.11}$$

$\mathbf{J}^{(1)}$ and $\boldsymbol{\mathcal{E}}$ point in the same direction, and if an electric field with a different in-plane polarisation was chosen, this would still hold, showing that the first-order in-plane response is isotropic. Now by using that $\sigma_{xx}(\omega_\mathrm{p}) = \sigma_{xx}^*(-\omega_\mathrm{p})$, we get

$$\mathbf{J}^{(1)}(t) = \frac{2\mathcal{E}_0}{\sqrt{2}} \Big[\mathrm{Re}(\sigma_{xx}(\omega)) \cos(\omega t) + \mathrm{Im}(\sigma_{xx}(\omega)) \sin(\omega t) \Big] (\overline{\mathbf{e}}_x + \overline{\mathbf{e}}_y) \,. \tag{3.12}$$

This shows that in order to determine the full linear current density response, only $\sigma_{xx}(\omega)$ has to be calculated. Moreover, the current will be an alternating current oscillating between $\mathrm{Re}(\sigma_{xx}(\omega))$, $\mathrm{Im}(\sigma_{xx}(\omega))$, $-\mathrm{Re}(\sigma_{xx}(\omega))$, $-\mathrm{Im}(\sigma_{xx}(\omega))$ and then back to $\mathrm{Re}(\sigma_{xx}(\omega))$ where a single cycle of course takes ω^{-1} units of time.

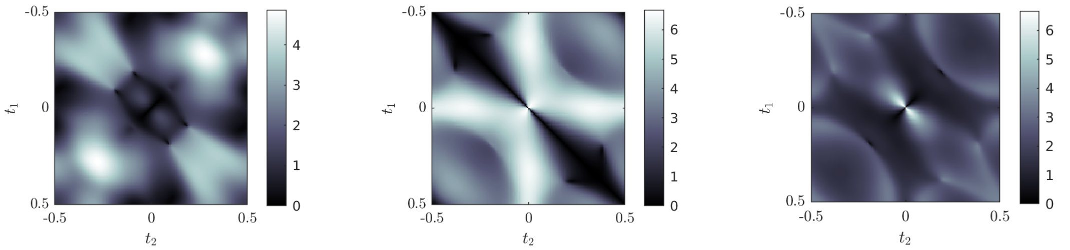
In the LG and VG, the general formula for the rank 2 tensor elements are (2.90) and (2.118). Specifically for $\sigma_{xx}(\omega)$ using a spin-degenerate TB model, they are

$$\sigma_{xx}^{(\mathrm{LG})}(\omega) = \frac{\mathrm{i}e^2\hbar}{m_e^2 A} \sum_{\substack{\mathbf{k},m,n \\ m \neq n}} \frac{f_{nm} \left| p_{mn}^x \right|^2}{E_{mn}(\hbar\omega - E'_{mn})} - \frac{\mathrm{i}e^2\hbar}{m_e^2 A} \sum_{\mathbf{k},m} \frac{(p_{mm}^x)^2}{\hbar\omega + \mathrm{i}\hbar\eta} \frac{\mathrm{d}f_m}{\mathrm{d}E} \,, \tag{3.13}$$

and

$$\sigma_{xx}^{(\mathrm{VG})}(\omega) = \frac{\mathrm{i}e^2\hbar}{m_e^2 A} \sum_{\mathbf{k},m,n} \frac{f_{nm} \left| p_{mn}^x \right|^2}{\hbar\omega(\hbar\omega - E'_{mn})} + \frac{\mathrm{i}e^2\hbar}{m_e A} \sum_{\mathbf{k},m} \frac{f_m}{\hbar\omega} \,. \tag{3.14}$$

where the $\mathbf{k}$-sum implies an integral over the first Brillouin zone through $4\pi^2 \sum_{\mathbf{k}} \to A \int_{\mathrm{BZ}} \mathrm{d}^2\mathbf{k}$,[22] and the momentum matrix elements are calculated through (2.104) which can be taken analytically since the Hamiltonian is built from a Bloch sum. Plots of some momentum matrix elements in $\mathbf{k}$-space is shown in Figure 3.7



(a) The highest valence state and lowest conduction state of $\mathcal{H}^{++}$.

(b) The highest valence state and lowest conduction state of $\mathcal{H}^{--}$.

(c) The highest valence state and second-lowest conduction state of $\mathcal{H}^{--}$.

Figure 3.7: Plots of the absolute value of $\frac{\hbar}{m_e}p^x_{mn}(\mathbf{k})$ in units of eV · Å where t_1 and t_2 form a wavevector through $\mathbf{k} = t_1\mathbf{G}_1 + t_2\mathbf{G}_2$.

$\sigma_{xx}(\omega)$ shall now be calculated using the parity-sorted TB model while varying the number of bands in the sum. We use the two highest valence bands and two lowest conduction bands, then the four highest valence bands and four lowest conduction bands, and lastly all states are summed over. The sorting order of the bands is defined right at the Γ-point. Additionally, we always choose $\hbar\eta = 15$ meV, $k_BT = 25$ meV, a unit of measurement $\sigma_0 = e^2/4\hbar = 6.0853 \cdot 10^{-5}$ S, and the integral over the Brillouin zone is performed using a Γ-centred grid with a little over 200,000 **k**-points. The real parts of the conductivity summing over the previously mentioned bands are then plotted in Figure 3.8 where the vertical line represents E_g^-. This plot shows that the real parts of σ_{xx} coincide very well in the two different gauges and both spectra step up right at the direct band gap energy. Of course, the inclusion of more bands displays more peaks and features of the spectra.

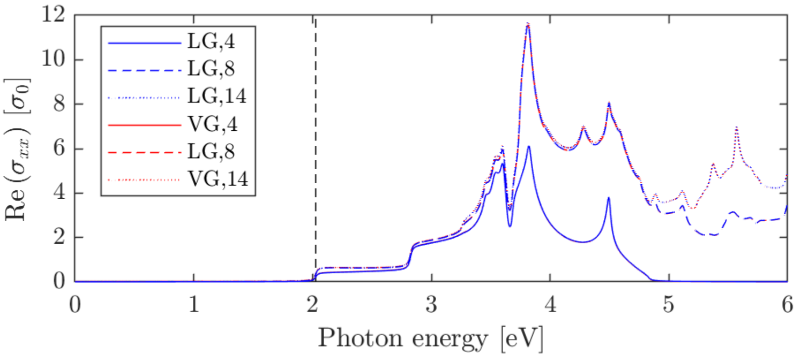
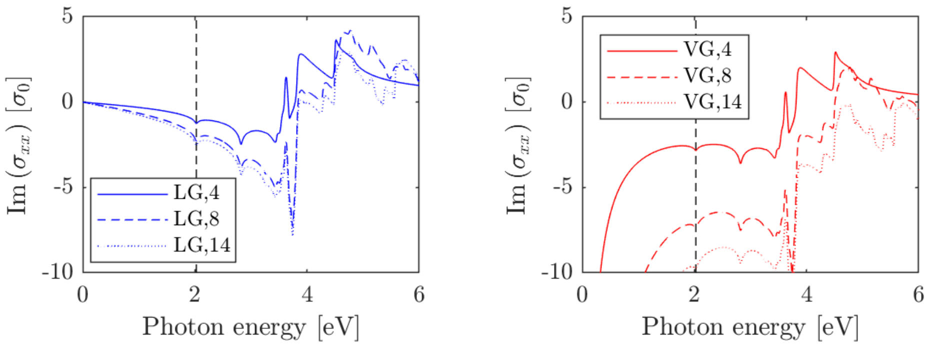
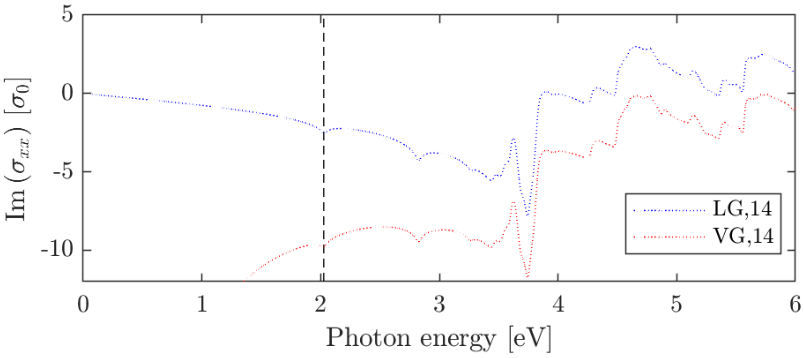


Figure 3.8: Real parts of the linear response conductivity using 4, 8 and 14 bands in the LG and VG.

On the other hand, the imaginary parts shown in Figure 3.9 are not nearly as clean. Both results in the LG and VG change significantly from using 4 to 8 bands; however, when going from 8 to 14 bands the LG spectrum's change is small as it seems to reach convergence while the divergence of the VG spectrum at $\hbar\omega = 0$ worsens. In spite of this, it should still be mentioned that the first peak appears correctly at E_g^- in all spectra, and using 14 bands with both gauges reveal spectra with the same features although the spectra do not coincide.



(a) LG using 4, 8 and 14 bands. (b) VG using 4, 8 and 14 bands.



(c) LG and VG with 14 bands.

Figure 3.9: Comparison between the imaginary parts of the linear response conductivity.

To explain and remedy these differences, first note that while the conductivity in the LG and VG should theoretically be equivalent, this is only true when summing over infinitely many bands in both cases. Since, In_2Se_2 is a semiconductor as is clear from the band diagram, and the calculations were done at relatively low thermal energy, the large conductivity at low photon energy must be an incorrect quirk of the VG. Indeed, the divergence at $\hbar\omega = 0$ is clearly visible in equation (3.14), and it has also been shown that the LG converges faster than the VG with respect to the number of basis functions.[22] As such, the LG spectra is considered as the target, and a correction to the VG spectra must be formulated. Due to the divergence towards $-\infty$, it is clear that a positive imaginary part is missing from the VG spectra. At least some of these missing imaginary parts can be found in the latter sum of (3.14). It is essentially a sum over valence bands, since $f_m \approx 1$ for a valence band and $f_m \approx 0$ for a conduction band in a cold semiconductor. The integral over $\mathbf{k}$ will in this case just yield the area of the first Brillouin zone $4\pi^2/A_{\mathrm{UC}}$. So, the latter sum could instead be approximated as

$$\frac{\mathrm{i}e^2}{m_e\omega A_{\mathrm{UC}}}\sum_m 1 = \frac{\mathrm{i}N_{\mathrm{val}}e^2}{m_e\omega A_{\mathrm{UC}}}.$$

The TB model only includes 7 valence bands; the rest are assumed to lie in the frozen core. Additional valence states can easily be included by increasing N_{val}. Each In_2Se_2 unit cell has 166 electrons which translates to 83 bands in the spin-degenerate case. This leaves 76 frozen valence bands unaccounted for in the TB model, and the correction which must be added to (3.14) is

$$\sigma_{\mathrm{corr}}^{(\mathrm{VG})} = \frac{\mathrm{i}76e^2}{m_e\omega A_{\mathrm{UC}}} = \frac{\mathrm{i}\hbar^2}{m_e}\frac{304}{\hbar\omega A_{\mathrm{UC}}}\,[\sigma_0]\,.$$

Adding $\sigma_{\mathrm{corr}}^{(\mathrm{VG})}$ to the VG spectra in Figure 3.9 and comparing these new spectra to the 14-band LG spectra yields Figure 3.10. While the correction does help, especially for the 8-band spectra, it doesn't quite overcome the divergence induced by basis truncation when all 14 bands are included.

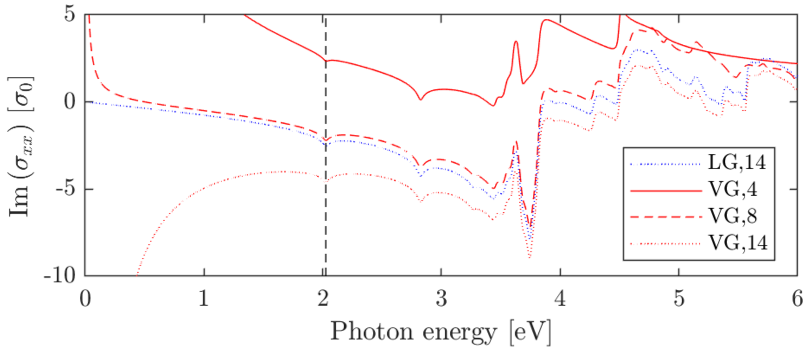


Figure 3.10: Imaginary parts of the linear response conductivity using 4, 8 and 14 bands in the VG with the frozen core correction.

Lastly, the screening length r_0, which is needed in order to determine the excitonic spectra in section 3.3, can be determined through $r_0 \approx (\chi_{2D}(\omega = 0))/2$ as shown in Appendix D, and the 2D susceptibility must then be calculated which is usually done through[48]

$$\chi_{2D}(\omega) = \frac{i\sigma(\omega)}{\varepsilon_0\omega}.$$

Since In_2Se_2 is a semiconductor, we must require that there is little to no absorption close to $\hbar\omega = 0$, so we're only interested in the real part of χ_{2D}. In our case, this can be found through

$$\text{Re}(\chi_{2D}(\omega)) = \frac{i\sigma_{xx}(\omega)}{\varepsilon_0\omega} - \frac{i\sigma_{xx}(-\omega)}{\varepsilon_0\omega} = -\frac{2\text{Im}(\sigma_{xx}(\omega))}{\varepsilon_0\omega},$$

The result using the most precise spectrum, the 14-band LG spectrum, is shown in Figure 3.11 which gives $r_0 = 32$ Å.

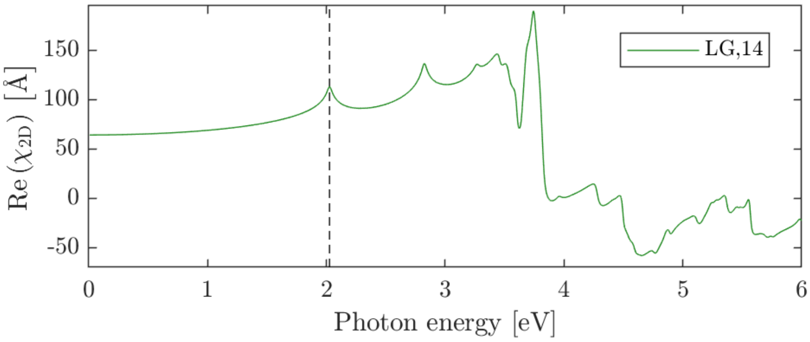


Figure 3.11: Real part of the 2D susceptibility calculated using the linear response conductivity with 14 bands in the LG.

3.2.2 Second-Order Response

The focus shall now turn to the second-order response with the same electric field. The x-entrance to the second-order current density will initially be on the complicated form

$$\begin{aligned}J_x^{(2)}(t) = \frac{\mathcal{E}_0^2}{2}\Big[&(\sigma_{xxx}(\omega,\omega)+\sigma_{xxy}(\omega,\omega)+\sigma_{xyx}(\omega,\omega)+\sigma_{xyy}(\omega,\omega))\,\mathrm{e}^{-2\mathrm{i}\omega t}\\ &+(\sigma_{xxx}(-\omega,-\omega)+\sigma_{xxy}(-\omega,-\omega)+\sigma_{xyx}(-\omega,-\omega)+\sigma_{xyy}(-\omega,-\omega))\,\mathrm{e}^{2\mathrm{i}\omega t}\\ &+\sigma_{xxx}(\omega,-\omega)+\sigma_{xxy}(\omega,-\omega)+\sigma_{xyx}(\omega,-\omega)+\sigma_{xyy}(\omega,-\omega)\\ &+\sigma_{xxx}(-\omega,\omega)+\sigma_{xxy}(-\omega,\omega)+\sigma_{xyx}(-\omega,\omega)+\sigma_{xyy}(-\omega,\omega)\Big]\end{aligned} \tag{3.15}$$

Again the point group symmetry of the crystal can be used to reduce these expressions. We find that the x- and z-entrances vanishes such that the entire current density vector becomes

$$\begin{aligned}\mathbf{J}^{(2)}(t) &= -\,\mathcal{E}_0^2\Big[\sigma_{xxx}(\omega,\omega)\,\mathrm{e}^{-2\mathrm{i}\omega t}+\sigma_{xxx}(-\omega,-\omega)\,\mathrm{e}^{2\mathrm{i}\omega t}+\sigma_{xxx}(\omega,-\omega)+\sigma_{xxx}(-\omega,\omega)\Big]\,\bar{\mathbf{e}}_y\\ &= -\,2\mathcal{E}_0^2\Big[\mathrm{Re}(\sigma_{xxx}(\omega,\omega))\cos(2\omega t)+\mathrm{Im}(\sigma_{xxx}(\omega,\omega))\sin(2\omega t)+\mathrm{Re}(\sigma_{xxx}(\omega,-\omega))\Big]\,\bar{\mathbf{e}}_y\end{aligned} \tag{3.16}$$

from $\sigma_{xxx}(\omega_\mathrm{p},\omega_\mathrm{q}) = \sigma^*_{xxx}(-\omega_\mathrm{p},-\omega_\mathrm{q})$. This shows that in order to calculate the entire second-order current density vector, only the real and imaginary parts of $\sigma_{xxx}(\omega,\omega)$ and the real part of $\sigma_{xxx}(\omega,-\omega)$ are necessary.

The first two terms represent oscillating alternating currents as with linear response, although the oscillation period is now $\omega^{-1}/2$, and $\sigma_{xxx}(\omega,\omega)$ itself represents the situation where two photons with the same frequency combine to create a photon with double the frequency. The spectra should therefore show peaks whenever this new photon has an energy that matches the energy difference between a valence and conduction band. $\sigma_{xxx}(\omega,\omega)$ is therefore commonly called the frequency doubling spectrum or the second-harmonic generation (SHG). On the other hand, the peculiar last term is constant with time; it represents a direct current. This term is known as optical rectification (OR) since the oscillating electric field is "rectified" to yield a constant response similar to how a diode may rectify AC to DC.

In reality, an oscillating electric field with a constant intensity cannot induce a direct current; however, it can induce a constant polarisation density.[49] Changing the intensity with time will then change the polarisation density which yields a current. The constant polarisation density can be understood through the fact that the optical rectification terms vanish in a crystal with inversion symmetry which is clear through the use of Neumann's principle. Electrons in crystals like $\mathrm{In_2Se_2}$, which lack inversion symmetry, are in asymmetric, anharmonic potentials. As such, they will on average be pushed further to one side when affected by an electric field which is what can be attributed to the constant part of the polarisation density.

Contrary to the first-order response, if the light is polarised in another direction with a perturbation on the form $\boldsymbol{\mathcal{E}}(t) = \mathcal{E}_0\cos(\omega t)\bar{\mathbf{e}}_x$ (light linearly polarised in the horizontal direction in Figure 3.1) then this yields the second-order current response from (2.71) as

$$\mathbf{J}^{(2)}(t) = \mathcal{E}_0^2\Big[\mathrm{Re}(\sigma_{xxx}(\omega,\omega))\cos(2\omega t)+\mathrm{Im}(\sigma_{xxx}(\omega,\omega))\sin(2\omega t)+\mathrm{Re}(\sigma_{xxx}(\omega,-\omega))\Big]\,\bar{\mathbf{e}}_x. \tag{3.17}$$

The angle between the electric field direction and the response direction is different depending on the polarisation, and the amplitude is different as well. The second-order response is therefore anisotropic. Similarly, for y-polarised light, the negative of the above result will be obtained, because of the tensor form (3.8).

The SHG spectrum shall now be plotted in the LG and VG. They can be calculated through (2.107) and (2.122) which reduce to

$$\begin{aligned}
\sigma^{(\mathrm{LG})}_{xxx}(\omega,\omega) = {} & \frac{e^3\hbar^2}{2m_e^3 A}\sum_{\substack{\mathbf{k},m,n,l\\ m\neq l\neq n}} \frac{p^x_{nm}p^x_{ml}p^x_{ln}}{E_{ml}E_{ln}(2\hbar\omega - E'_{mn})}\left(\frac{f_{nl}}{\hbar\omega - E'_{ln}} - \frac{f_{lm}}{\hbar\omega - E'_{ml}}\right) \\
& - \frac{e^3\hbar}{2m_e^2 A}\sum_{\substack{\mathbf{k},m,n\\ m\neq n}} \frac{|p^x_{mn}|^2}{E_{mn}(\hbar\omega + \mathrm{i}\hbar\eta)(2\hbar\omega - E'_{mn})}\frac{\partial f_{nm}}{\partial k_x} \\
& - \frac{e^3\hbar}{2m_e^2 A}\sum_{\substack{\mathbf{k},m,n\\ m\neq n}} \frac{p^x_{nm}}{2\hbar\omega - E'_{mn}}\left(\frac{f_{nm}p^x_{mn}}{E_{mn}(\hbar\omega - E'_{mn})}\right)_{;k_x} \\
& + \frac{e^3}{2m_e A}\sum_{\mathbf{k},m}\frac{p^x_{mm}}{(\hbar\omega + \mathrm{i}\hbar\eta)(2\hbar\omega + \mathrm{i}\hbar\eta)}\frac{\partial^2 f_m}{\partial k_x^2},
\end{aligned} \tag{3.18}$$

and

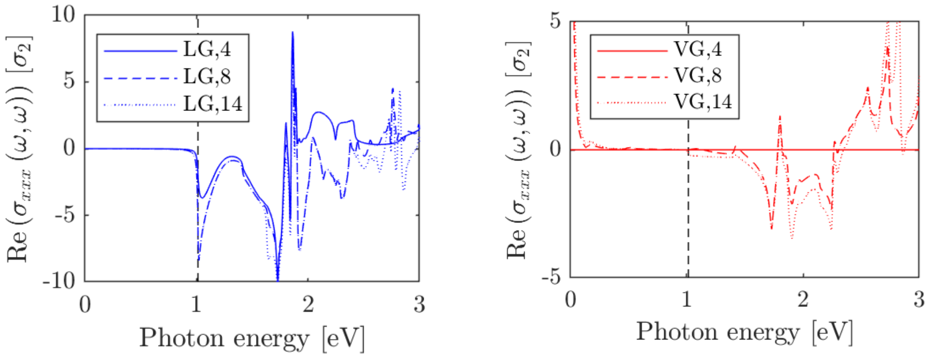
$$\sigma^{(\mathrm{VG})}_{xxx}(\omega,\omega) = \frac{e^3\hbar^2}{2m_e^3 A}\sum_{\mathbf{k},m,n,l}\frac{p^x_{nm}p^x_{ml}p^x_{ln}}{(\hbar\omega)^2(2\hbar\omega - E'_{mn})}\left(\frac{f_{nl}}{\hbar\omega - E'_{ln}} - \frac{f_{lm}}{\hbar\omega - E'_{ml}}\right). \tag{3.19}$$

The momentum matrix elements can still be calculated analytically; however, the rest of the k_x derivatives must be taken numerically. Furthermore, (2.109) shows that the generalised derivative of specifically the momentum matrix element is needed. This can be calculated through (B.2) where we need the k_x derivative of the momentum matrix elements in the eigenbasis which are found numerically. When performing the numerical derivatives, a phase convention for the eigenvectors must be chosen, so we choose a simple convention where the first entrance of each eigenvector must be real since there is no band crossing for the bottom bands for each parity. Additionally, the Berry connection is needed for the generalised derivative. It is calculated through (B.4). We note that the Berry connections from Wannier90 are not Hermitian in regards to band switching, although they theoretically should be.[50] In spite of this, by calculating the Chern numbers through (B.5), it does still yield an integer (0, specifically) down to the fifth decimal.

Now for the actual calculation of the SHG spectra, we again choose $\hbar\eta = 15$ meV and $k_B T = 25$ meV. The unit of measurement shall in this case be $\sigma_2 = e^3/8\hbar \cdot 1\ \text{Å}/1\ \text{eV} = 3.0427 \cdot 10^{-15}\ \mathrm{S \cdot m \cdot V^{-1}}$, and the numerical integration shall be performed by using a Γ-centred $\mathbf{k}$-grid with a little over 300,000 $\mathbf{k}$-points. The results for the real parts are shown in Figure 3.12, and the imaginary parts in Figure 3.13. The vertical line in the plots represent $E_g^-/2$.

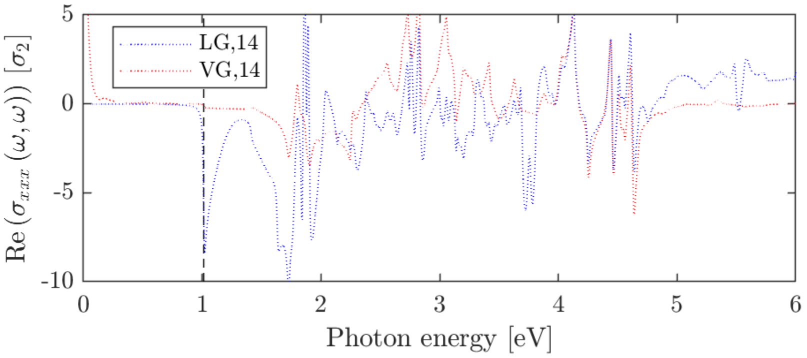
First, it can be seen that the LG and VG spectra are generally very different except around $\hbar\omega = 4.1$ eV where the spectra match up briefly.

It takes more than 4 bands in order for the LG spectra to converge at the first peak around $E_g^-/2$. Meanwhile, the 4-band VG spectra is simply flat as it does not have contain enough information. The 14-band VG spectra shows a small step-up at $E_g^-/2$ while the LG spectra shows a large peak. We're overall more inclined to believe the LG spectrum result as the LG spectrum was shown to be less dependent on the number of bands in the case of linear response. It is therefore possible that 14 bands is not enough for convergence in the VG. This is discussed further in section 6.1



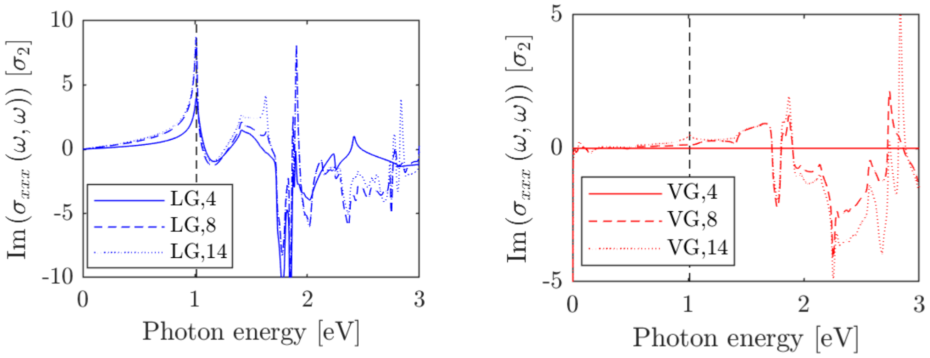
(a) LG using 4, 8 and 14 bands.

(b) VG using 4, 8 and 14 bands.



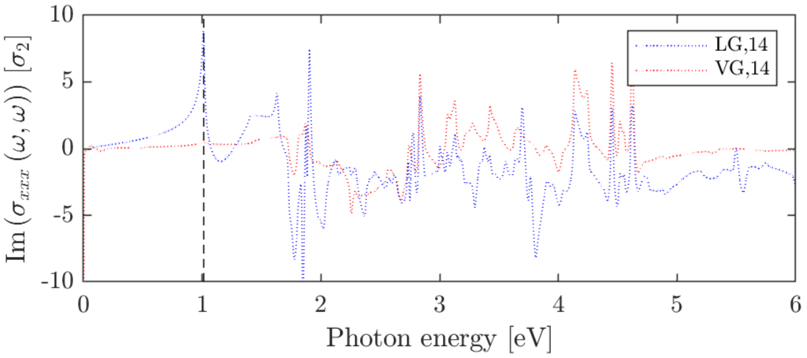
(c) LG and VG with 14 bands.

Figure 3.12: Comparison between the real parts of the SHG spectra.



(a) LG using 4, 8 and 14 bands.

(b) VG using 4, 8 and 14 bands.



(c) LG and VG with 14 bands.

Figure 3.13: Comparison between the imaginary parts of the SHG spectra.

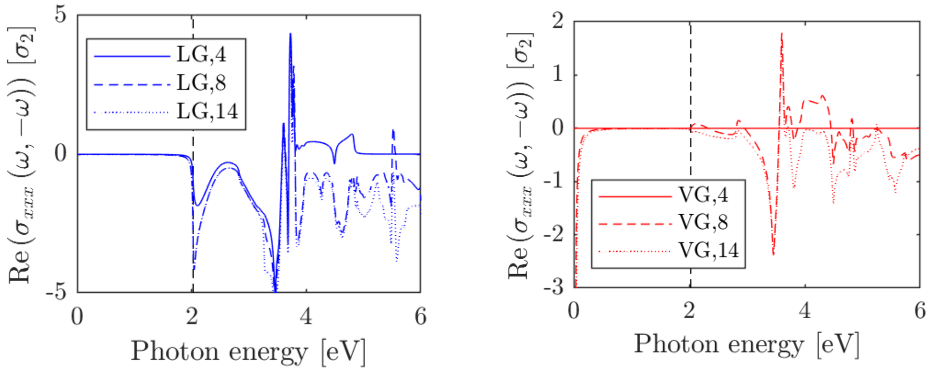
Lastly before considering excitonic effects, the OR spectra shall be calculated. In the case of optical rectification, (2.107) and (2.122) reduce to

$$\begin{aligned}\sigma_{xxx}^{(\mathrm{LG})}(\omega,-\omega) = {}& \frac{e^3\hbar^2}{2m_e^3 A}\sum_{\substack{\mathbf{k},m,n,l\\ m\neq l\neq n}} \frac{-p_{nm}^x p_{ml}^x p_{ln}^x}{E'_{mn}E_{ml}E_{ln}}\left(\frac{f_{nl}}{\hbar\omega - E'_{ln}} - \frac{f_{lm}}{\hbar\omega - E'_{ml}}\right)\\ &+ \frac{e^3\hbar}{2m_e^2 A}\sum_{\substack{\mathbf{k},m,n\\ m\neq n}}\left(\frac{|p_{mn}^x|^2}{E'_{mn}E_{mn}(\hbar\omega + \mathrm{i}\hbar\eta)}\frac{\partial f_{nm}}{\partial k_x} + \frac{p_{nm}^x}{E'_{mn}}\left(\frac{f_{nm}p_{mn}^x}{E_{mn}(\hbar\omega - E'_{mn})}\right)_{;k_x}\right) \qquad (3.20)\\ &+ \frac{e^3}{2m_e A}\sum_{\mathbf{k},m}\frac{p_{mm}^x}{\mathrm{i}\hbar\eta(\hbar\omega+\mathrm{i}\hbar\eta)}\frac{\partial^2 f_m}{\partial k_x^2},\end{aligned}$$

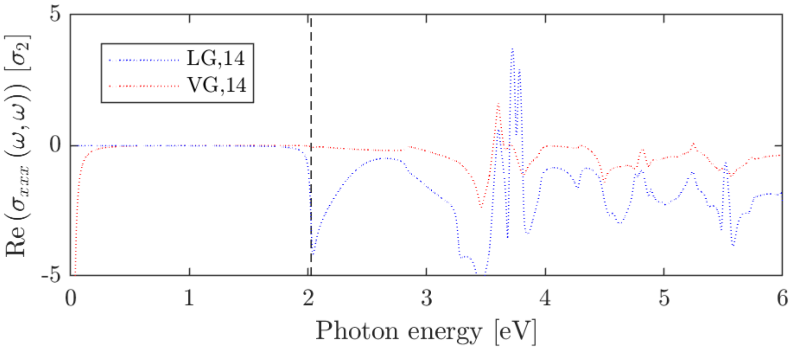
and

$$\sigma_{xxx}^{(\mathrm{VG})}(\omega,-\omega) = \frac{e^3\hbar^2}{2m_e^3 A}\sum_{\mathbf{k},m,n,l}\frac{p_{nm}^x p_{ml}^x p_{ln}^x}{(\hbar\omega)^2 E'_{mn}}\left(\frac{f_{nl}}{\hbar\omega - E'_{ln}} - \frac{f_{lm}}{\hbar\omega - E'_{ml}}\right). \qquad (3.21)$$

The OR spectra shall be calculated using the same settings as the SHG spectra, and the results are shown in Figure 3.14 with the vertical line representing E_g^-. Generally, the same comments in regards to convergence can be made with the OR spectra as with the SHG spectra. An interesting observation is that the OR spectrum looks similar to the SHG spectrum when the axes are scaled appropriately.



(a) LG using 4, 8 and 14 bands. (b) VG using 4, 8 and 14 bands.



(c) LG and VG with 14 bands.

Figure 3.14: Comparison between the real parts of the OR spectra.

3.3 Excitonic States

In the IPA, electrons only interact via a mean field as prescribed by the Kohn-Sham equations. In particular, bound states known as excitons are not accounted for within the IPA. The BSE (2.128) does however account for this, at the expense of having to solve a problem where every **k**-point couples to ever other **k**-point.

To remedy the small DFT band gap, a scissors correction is chosen such that the electronic band gap is 2.74 eV,[3] and because $\mathrm{In_2Se_2}$ has its (slightly indirect) band gap almost at the Γ-point, we generate a Monkhorst-Pack **k**-grid, but also center this **k**-grid at Γ. The screening length calculated previously is $r_0 = 32$ Å. Spin-orbit coupling is not included, even though it does change the band diagram at Γ somewhat.[3]

Without spin-orbit coupling, we have four blocks within the slow-rapid approximation: $[\hat{\mathcal{H}}^{\mathrm{MB}}]^{PP}$, $[\hat{\mathcal{H}}^{\mathrm{MB}}]^{PN}$, $[\hat{\mathcal{H}}^{\mathrm{MB}}]^{NP}$ and $[\hat{\mathcal{H}}^{\mathrm{MB}}]^{NN}$ of which only the PP and NN blocks will contribute to the in-plane response. Considering the band diagram from Figure 3.5, the negative parity bands in red have the smallest direct band gap and are thus the bands to expect bound states to be present. The positive parity bands in green in Figure 3.5 have instead an indirect band gap.

For the response spectra to follow in the coming sections, the momentum matrix elements has not been corrected with (2.11).

The NN Block

We focus on the bound states of the negative parity bands for the optical response and note the scissors-corrected electronic band gap between these bands is actually $E_g^{NN} = E_g^{-} + \Delta + 0.227$ eV $= 2.96$ eV. We also note that depending on the exchange correlation functional used, the difference in energy between the top two valence bands of opposing parity will vary.
Solving the BSE using one negative parity conduction band and two negative parity valence bands within the slow-rapid approximation from (2.138) yields the lowest porton of the exciton spectrum, of which the lowest 20 energies are seen in Figure 3.15.

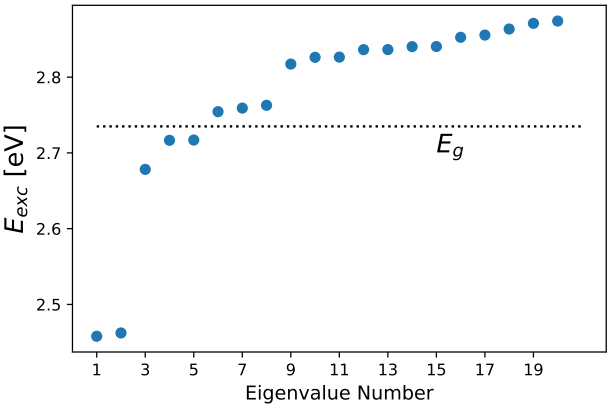
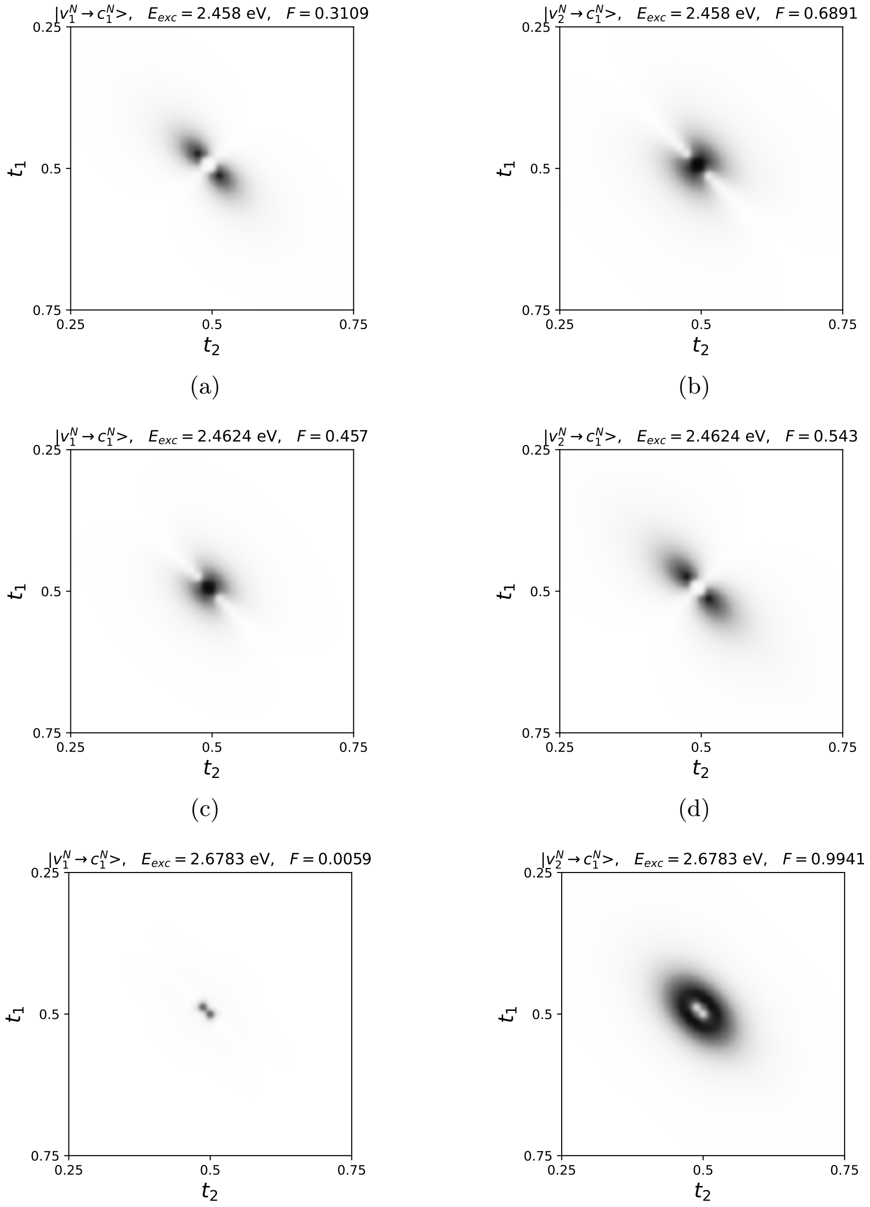


Figure 3.15: The first 20 eigenvalues of the $[H]^{NN}$ blocks with 80×80 **k**-points. $E_g = 2.74$ eV. Calculated by diagonalising the truncated Hamiltonian obtained by the Lanczos-Haydock routine with 800 iterations.

The energies of the states are not equal to enough decimals to say they are degenerate, even though it may seem so from the figure. The grid chosen is symmetrically distributed around the Γ-point and therefore this lack of degeneracy does not seem to be attributable to a skewed **k**-grid.

Some of the eigenvectors of $[\hat{\mathcal{H}}]^{NN}$ in **k**-space can be seen in Figure 3.16. Classification into definite angular momentum states is not obvious as firstly the eigenstates consist of Slater determinants of several bands, and secondly, even looking at the individual coefficients of the Slater determinants, they are not all perfectly s-, p- or d-shaped. While e.g. Figure 3.16b does have some s-characteristic around the Γ-point, one might also argue it looks like a p-state. The two lowest excitons does also seem similar in terms of their expansion coefficients but on opposite valence bands. The state of Figure 3.16f also bears great resemblance to an $2s$-state, but also seems to have two minima near the Γ-point. It is such a state one will suspect of being optically active.



(a) (b)

(c) (d)

(e) (f)

Figure 3.16: The absolute value of three lowest eigenvectors of $[\mathcal{H}]^{NN}$ with $n = 1, 2, 3$ plotted in the entire Brillouin zone. F is the sum of absolute squares of the this subset of entries in the eigenvector. $\mathbf{k}$-points are given as $\mathbf{k} = t_1\mathbf{b}_1 + t_2\mathbf{b}_2 - (\mathbf{b}_1 + \mathbf{b}_2)/2$ meaning the Γ-point is located at $(1/2, 1/2)$ which is where the excitons are localised. Calculated with 80×80 $\mathbf{k}$-points. Smeared using a gaussian profile.

Without further speculation into this, we examine the matrix elements of the position operator X^x_{0n}, seen in Figure 3.17. This quantity, by (2.156), reveal the optically active states. The $2s$ state can among other be seen not to have a significant matrix element. Instead, among the lowest

20 eigenstates, it is the n=1, 4, 7, 14 and 16 states one should expect to see contribute to the conductivity spectrum.

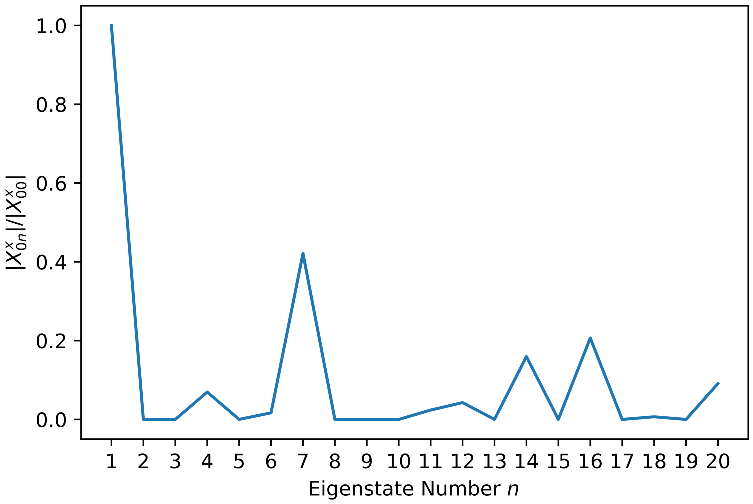
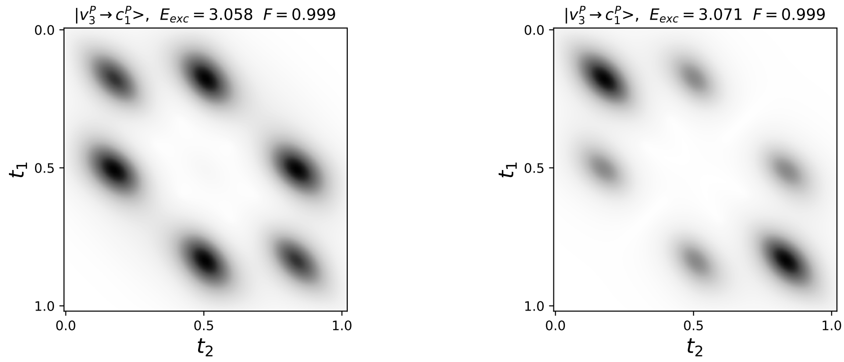


Figure 3.17: Absolute value of the matrix elements of the normalised positon operator in units of the matrix element of the lowest exciton.

The PP Block

A very similar calculation can of course be done on the PP-block of the Hamiltonian, but as is seen from Figure 3.5 the positive parity bands do not have a direct band gap and there turns out to be several minima in E^+_{cv}. The latter means the **k**-grid cannot centred on the region where one suspects the excitons to be located. This is examplified in Figure 3.18. Contrary to the well-localised excitons of the NN block, these are spread out over much of the first Brillouin zone, and more importantly they come close to the edges of the **k**-grid. The sum over the reciprocal lattice vectors **G**, mentioned in subsection 2.6.1, is not present, meaning the interactions with points outside the first Brillouin zone has been neglected. Aside from this fact, the excitons seems reasonably well localised, albeit not to a single region in the first Brillouin zone. We therefore carry on the calculation.



(a) (b)

Figure 3.18: The lowest energy Exciton with $n = 1$ together with the $n = 3$ exciton of the PP-block, calculated on a 80×80 **k**-grid with one conduction band and three valence bands. Smoothed with Gaussian smearing.

The absolute value of the position matrix elements can be seen in Figure 3.19. The lowest energy exciton is in this case not the one with the largest position matrix element, it is instead the one seen in Figure 3.18b. This state does seem s-like and is located a bit further out on the x-axis (along the diagonal from top left corner to bottom right corner) than the M-point. All eigenstates of this block seem to be as poorly localised as the one shown in Figure 3.18, meaning we should be careful to draw conclusions from the calculations done on this block of the Hamiltonian which will follow.

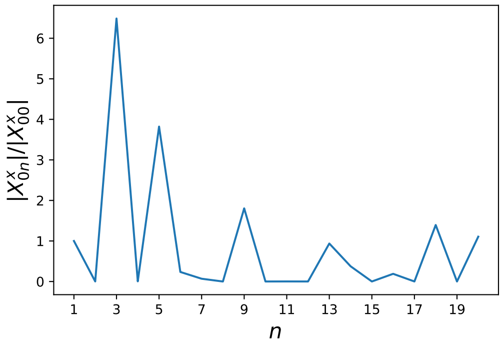


Figure 3.19: The absolute values of the normalised position matrix element in units of the $|X^x_{00}|$.

3.3.1 First-Order Excitonic Response

We calculate the conductivity of In_2Se_2 to first-order, now including the presence of excitons. We have already seen the matrix elements of the position operator in Figures 3.17 and 3.19. Together with the excitonic energies seen in Figure 3.15, one can calculate the conductivity with (2.156) of the $\hat{\mathcal{H}}^{NN}$ and $\hat{\mathcal{H}}^{PP}$ blocks. We also have the other option of using the Green's function approach from (2.147) and (2.151) to calculate the optical conductivity, needing only the momentum matrix elements and Hamiltonian to calculate the optical conductivity. Figure 3.20 shows the conductivity of $\hat{\mathcal{H}}^{NN}$, while Figure 3.21 shows the conductivity of $\hat{\mathcal{H}}^{PP}$, calculated with both methods.

In these spectra of the two blocks of the Hamiltonian, we see some clearly defined peaks on the discrete photon frequencies matching the exciton energy. Peaks which cannot be attributed to a single exciton are also seen, e.g. the peak just above 4 eV of Figure 3.20. In regards to the response, the peak conductivity for the Hamiltonian consisting of the bands we picked out will have a large peak just above 4 eV as both the NN and PP blocks have a large peak here.

In regards to recovering the features of the IPA spectrum Figure 3.8 we should look at the peaks of the IPA spectrum, keep the scissors shift $\Delta = 0.95$ eV in mind, and look for peaks in Figures 3.20 and 3.21 which could match. However, no convincing parallels between the IPA result and the excitonic spectra of the above figures can be found.

As a check of the implementation, one can take the limit of infinite screening i.e $\kappa \to \infty$ in order to eliminate the coupling between states of different $\mathbf{k}$-points. This should reproduce the IPA spectrum of Figure 3.8, with an additional scissors shift. Setting $\kappa = 100$ produces the spectrum seen in Figure 3.22, matching in the lower part of the spectrum, when accounting for the fact that a larger broadening term of $\hbar\eta = 50$ meV has been applied contrary to the $\hbar\eta = 15$ meV in Figure 3.8.

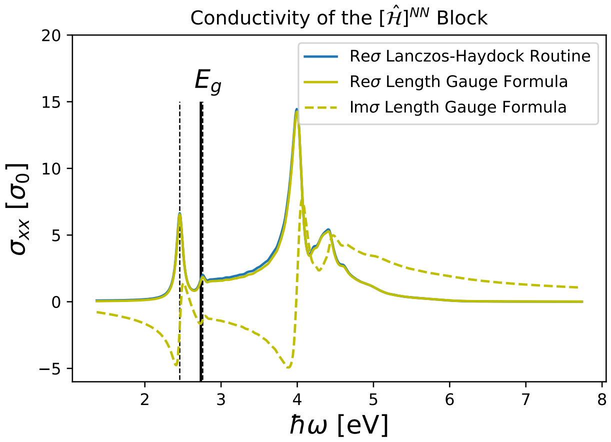


Figure 3.20: Calculated using 80 × 80 **k**-points, two valence bands and one conduction band. Dotted lines are the exciton energies E^1_{exc} and E^7_{exc}. $\hbar\eta = 50$ meV.

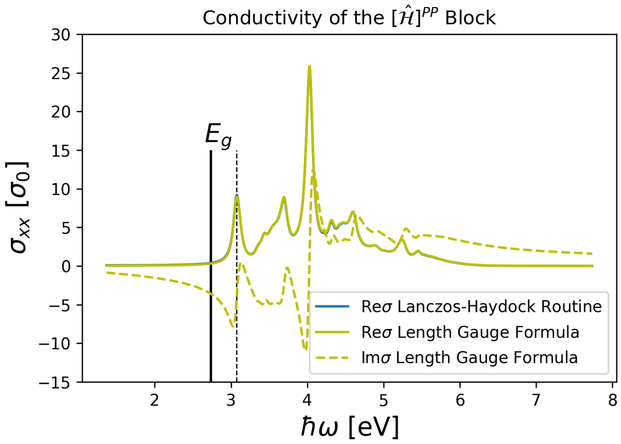


Figure 3.21: Calculated with the Lanczos-Haydock method using 80 × 80 **k**-points, three positive parity valence bands and one positive parity conduction band. Dotted line is the exciton energy E^2_{exc}. Broadening term $\hbar\eta = 50$ meV.

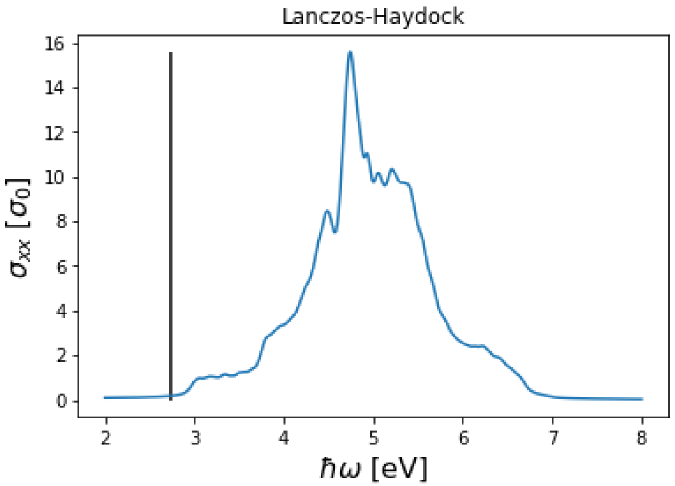
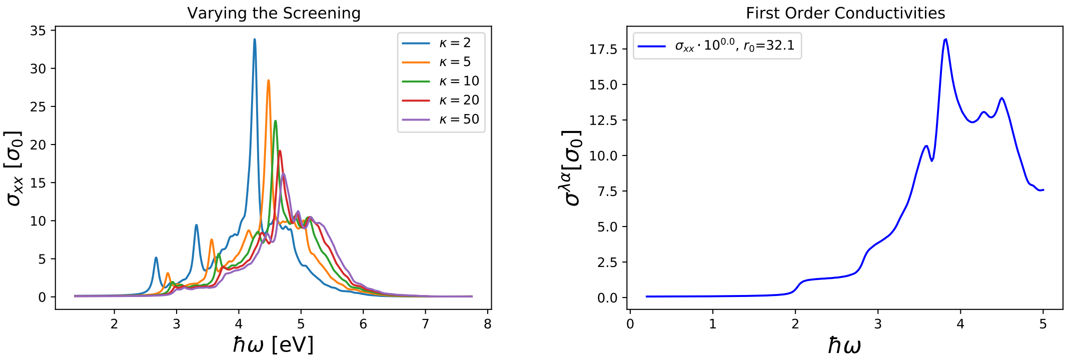


Figure 3.22: BSE first-order $\sigma_{xx}(\omega)$ with $\kappa = 100$ calculated with 200 iterations of the Lanczos-Haydock routine, 60 × 60 **k**-points, one conduction band of each parity, two negative valence bands and three positive parity valence bands. $\hbar\eta = 50$ meV.

Because of the problematic nonlocalised exciton states of the PP block, claiming we have modelled $\hat{\mathcal{H}}^{PP}$ accurately is not something we will do. However the situation with infinite screening is a situation where we model $\hat{\mathcal{H}}^{PP}$ accurately as we get agreement between the IPA approximation and the BSE calculation. One can then also think that the modelling of $\hat{\mathcal{H}}^{PP}$ is accurate for screenings smaller than the large $\kappa = 100$ chosen in Figure 3.22, but still larger than $\kappa = 1$. We therefore plot the total conductivity for a variety of values of κ in Figure 3.23a. The act of "turning on" the Coulomb interaction redshifts and enlarges the peaks present in Figure 3.8, while the appearance of the bound states also become more apparent. Since κ is the average of the dielectric constants of the materials encapsulating the In$_2$Se$_2$ monolayer, one can think of materials with large dielectric constants, where the our calculated $\hat{\mathcal{H}}^{PP}$ could capture the physics of what is happening.



(a) κ varied in the BSE calculation. The real part of the total conductivity of both the NN and PP blocks. Same valence and conduction bands included as the previous calculations.

(b) IPA spectrum. $\hbar\omega$ in eV.

Figure 3.23: $\hbar\eta = 50$ meV. No scissorshift in IPA spectrum. $\kappa = 50$ resemble the IPA spectrum on the right

3.3.2 Second-Order Excitonic Response

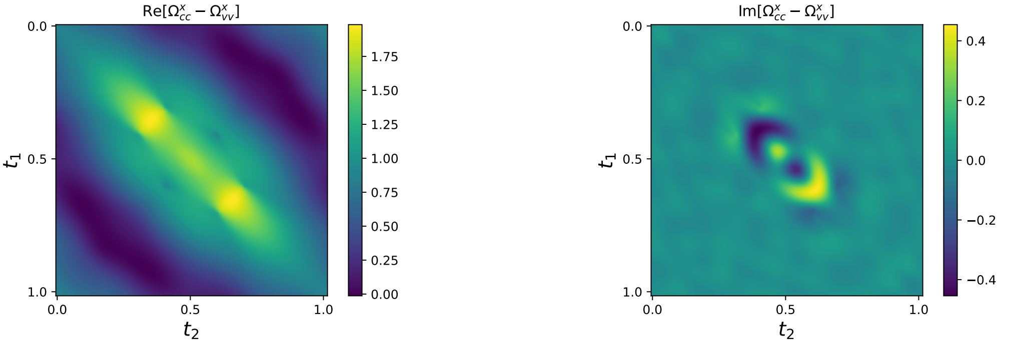
We now investigate the second-order conductivity which is, as discussed in the IPA section, associated with the phenomena of optical rectification and second-harmonic generation, and furthermore include excitonic effects. Our starting point is the framework discussed in subsection 2.7.2. While the VG expression (2.158) has matrix elements that are straight forward to compute, it has been shown to need more conduction bands to converge.[32] The LG expression Equation 2.157 does however have the feature of having to compute the generalised derivative which we will elaborate on in a bit.

We also again treat the PP and NN blocks of the Hamiltonian separately and keep in mind we should add the results in the end. This is still allowed by the decoupling of different parity states for the inplane response. While (2.158) is relatively simple from a computational standpoint, (2.157) on the other hand needs more sophisticated input in the form of a generalised derivative of the exciton expansion coefficients.

The $\mathbf{k}$-gradient acting on $A^n_{vc\mathbf{k}}$ is calculated using FD to find the directional derivatives along $\mathbf{b}_1$ and $\mathbf{b}_2$ from which the components of the gradient can be obtained from using the chain rule

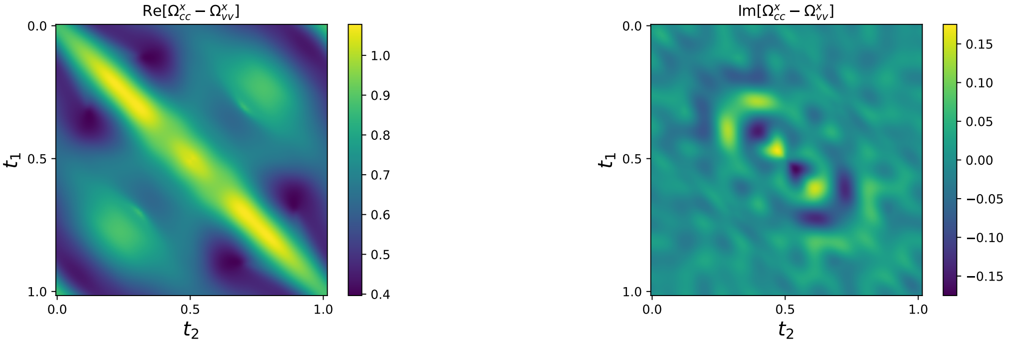
$$\begin{bmatrix} \frac{\partial f}{\partial t_1} \\ \frac{\partial f}{\partial t_2} \end{bmatrix} = \begin{bmatrix} \frac{\partial k_x}{\partial t_1} & \frac{\partial k_y}{\partial t_1} \\ \frac{\partial k_x}{\partial t_2} & \frac{\partial k_y}{\partial t_2} \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial k_x} \\ \frac{\partial f}{\partial k_y} \end{bmatrix} \qquad \text{where} \quad \mathbf{k} = t_1\mathbf{b}_1 + t_2\mathbf{b}_2, \tag{3.22}$$

and lastly doing a matrix inversion. This step is needed as we use a Γ-centered Monkhorst-Pack grid which does not have its axes aligned with the x and y-directions. We will be calculating $\sigma^{(2)}_{xxx}$, meaning it is $\Omega^x_{c_1c_1} - \Omega^x_{v_2v_2}$ seen in Figure 3.24 that will contribute to the generalised derivative. This will take care of the diagonal elements of the normalised position operator required for the calculation of X^x_{mn} from (2.154) while the nondiagonal elements are found using the formula $x_{mn} = -\mathrm{i}\frac{p^x_{mn}}{E_{mn}}$. We return to the issue of the Berry connection not being Hermitian as it should be[50] in the discussion.



(a) Re[$\Omega^x_{cc} - \Omega^x_{vv}$] for the top positive parity valence band and bottom positive parity conduction band.

(b) Im[$\Omega^x_{cc} - \Omega^x_{vv}$] for the top positive parity valence band and bottom positive parity conduction band.



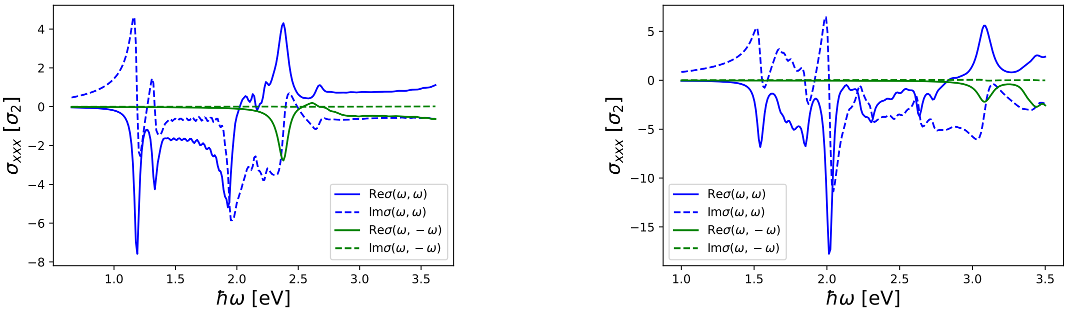
(c) Re[$\Omega^x_{cc} - \Omega^x_{vv}$] for the top negative parity valence band and bottom negative parity conduction band.

(d) Im[$\Omega^x_{cc} - \Omega^x_{vv}$] for the top negative parity valence band and bottom negative parity conduction band.

Figure 3.24: The difference in Berry connection between valence and conduction bands throughout the first Brillouin zone. Chern numbers from (B.5) zero to 5'th decimal.

All the ingredients to calculate the second-order conductivity tensor of the two blocks is now in order. The LG second harmonic and optical rectification parts for the NN-block are seen in Figure 3.25a while the same parts for the PP-block are seen in Figure 3.25b. Clear peaks related to the lowest energy excitons are seen while regions where the attribution of the conductivity to a single state is more troublesome. In Figure 3.25a the region between 1.5 - 2 eV is oscillating related to the coarseness of the **k**-grid. In this region the NN part of the conductivity is almost constant over a region starting in the top infrared going into the red part of the visible spectrum.

The optical rectification part of the NN conductivity has a well-defined peak related to the lowest exciton and otherwise being reasonably constant when $\hbar\omega > 3$ eV. As the optical rectification term does not seem to be affected by excitonic states matching double the frequency and only the directly matching transitions, one has a frequency region up to $\hbar\omega \sim 7$ eV, well into the ultraviolent, before the transitions related to the next negative parity bands should become significant.



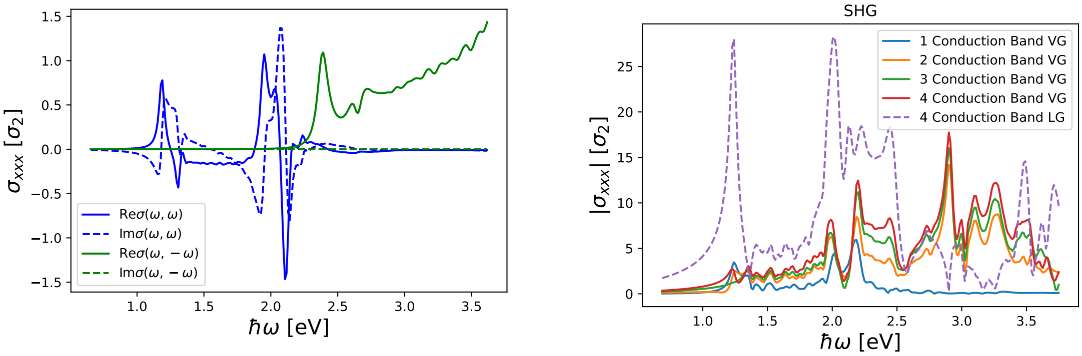
(a) The second-order LG conductivity of the In_2Se_2 NN-block, calculated with 90 × 90 **k**-points with two valence bands, one conduction band, and the 400 eigenstates of the truncated BSE Hamiltonian. Two first peaks matches half the energies of E^1_{exc}, E^6_{exc}. $\hbar\eta = 50$ meV

(b) The second-order LG conductivity of the In_2Se_2 PP-block, calculated with 80 × 80 **k**-points with three valence bands, one conduction band, and the 400 eigenstates of the truncated BSE Hamiltonian. First peak matches half the energy of E^1_{exc}. $\hbar\eta = 50$ meV

Figure 3.25

The more troublesome PP-block also has a significant response and has a peak clearly matching half the energy of the lowest states of Figure 3.18, while the rest does not seem to be attributable to a single state. The magnitude of response is of comparable size as the NN-block.

We also have the VG expression from (2.158). The TB model from Figure 3.6 we have three available conduction bands, and given the VG requires more bands, we add these one by one in Figures 3.26a and 3.26b. This calculation, because of the inclusion of more bands needs fewer **k**-points to be computationally feasible. We therefore do not think further of the jagged lines caused by a to rough **k**-grid.

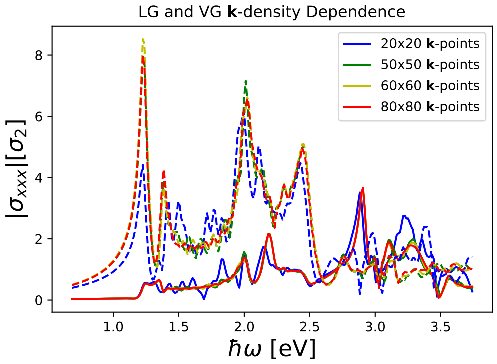


(a) VG conductivity of NN block using 2 valence bands and one conduction band. 90 × 90 **k**-points.

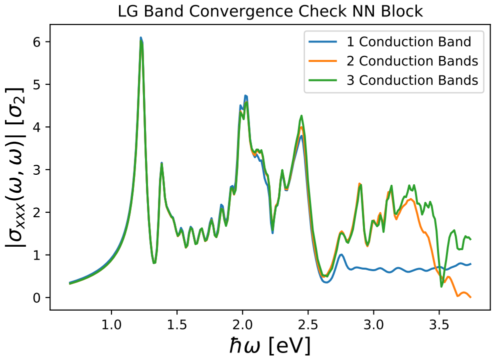
(b) Calculated with 40×40 **k**-points for the NN block.

Figure 3.26: The second-order conductivity of the In_2Se_2 NN-block. $\hbar\eta = 50$ meV in a) and 40meV in b).

As can be seen in Figures 3.26a and 3.26b, the VG expression is not converged to the LG result with respect to the number of conduction bands included, even in the lower part of the spectrum. This is contrary to the LG expression which is clearly recognised when comparing with Figure 3.25a. The spectra do however disagree wildly on the second-order response both of the bound states and the higher frequency of the spectrum. The VG does however tend to vary less with the conduction bands in the lower part of the spectrum.



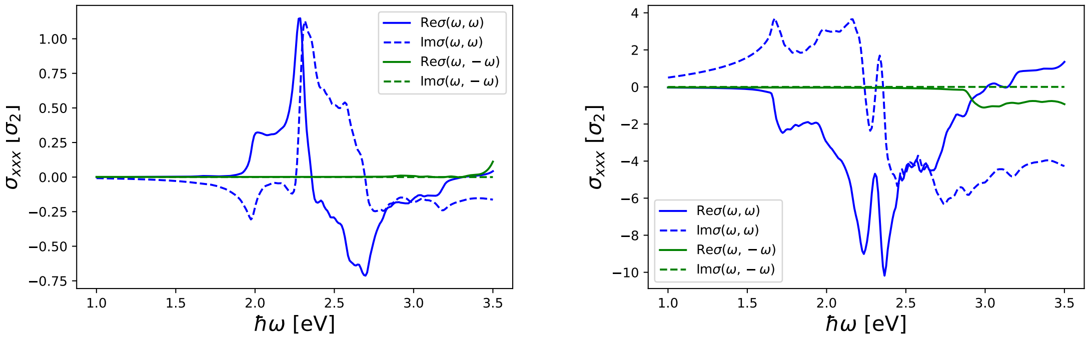
(a) Convergence check with respect to **k**-points using two valence and conduction bands of the NN block. Dotted lines are LG result, while fully drawn are VG result. 300 LH iterations used.



(b) Convergence check with respect to conduction using two valence bands of the NN block. 300 LH iterations and 60×60 **k**-points used.

The convergence with respect to **k**-grid density is seen in Figure 3.27a and shows reasonably well convergence even for 50×50 grids. The conduction band convergence check is seen in Figure 3.27b and shows the LG converges very rapidly with the number of included conduction bands.

We can again examine the limit of large screening of the sum of the second-order conductivities of both blocks, which is seen in Figure 3.28. We should obtain the spectrum in Figure 3.12, which is roughly the case shapewise, especially at the start. The bands used are not exactly the same, so for larger $\hbar\omega$, it is expected the results of Figure 3.12 and Figure 3.28 should deviate. The VG expression is zero in Figure 3.12, so not much comparison can be done.



(a) VG (b) LG

Figure 3.28: The second-order conductivity of the In_2Se_2 with $\kappa = 100$, calculated with 70×70 **k**-points with all valence bands except the two lowest, one conduction band of each parity. 800 eigenstates of the truncated BSE Hamiltonian used. $\hbar\eta = 50$ meV

3.4 α-In$_2$Se$_3$ Monolayer

One can add an extra selenium atom to the In$_2$Se$_2$ monolayer and still obtain a stable material. With the labelling going from bottom to top in Figure 3.29, we obtain the relaxed parameters

$$\begin{aligned} a &= 4.018\ \text{Å}, \\ z_{\text{Se}_1,\text{In}_1} &= 1.254\ \text{Å}, \quad z_{\text{In}_1,\text{Se}_2} = 2.518\ \text{Å}, \\ z_{\text{Se}_2,\text{In}_2} &= 1.659\ \text{Å}, \quad z_{\text{In}_2,\text{Se}_3} = 1.325\ \text{Å}, \end{aligned} \tag{3.23}$$

by using the "bfgs"-mode for ion- and cell-dynamics in the Quantum Espresso code. Looking at the yz-plane of the relaxed structure in Figure 3.29, it is evident the z-inversion symmetry of In$_2$Se$_2$ is gone. The point-group now reads P3m1, which is a point group still associated with three-fold rotational symmetry and y-inversion like In$_2$Se$_2$, but with the lack of z-inversion. The outcome of this are seen in the tensor elements T_{zzz} and the ones containing one z-index $T_{...z...}$ are no longer necessarily zero from the crystal symmetry in addition to the tensor elements $T_{xxx} = -T_{yxy} = -T_{yyx} = -T_{xyy}$. The TB model can be obtained in much the same fashion as In$_2$Se$_2$ and can be seen in Figure 3.30. A new valence band has appeared, touching the other top valence band at Γ. One can still recognise some features reminiscent of the In$_2$Se$_2$ monolayer band structure.

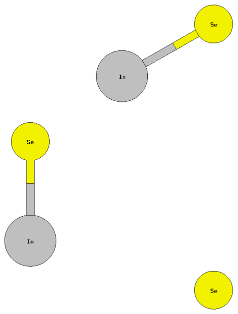


Figure 3.29: yz-plane of In$_2$Se$_3$ monolayer unit cell. Selenium in yellow, indium in grey. Figure made using XCrySDen.[51]

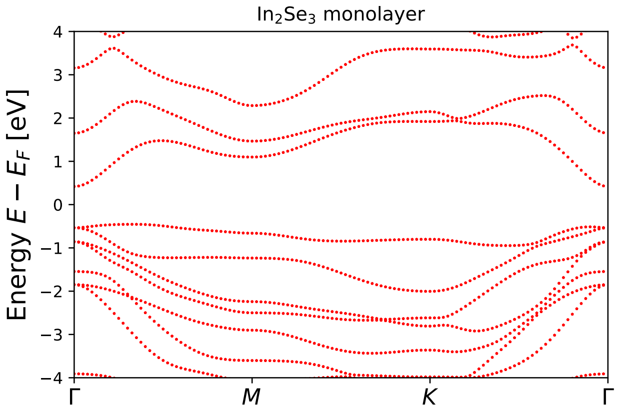
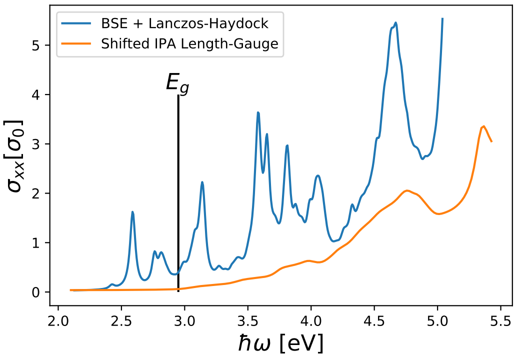
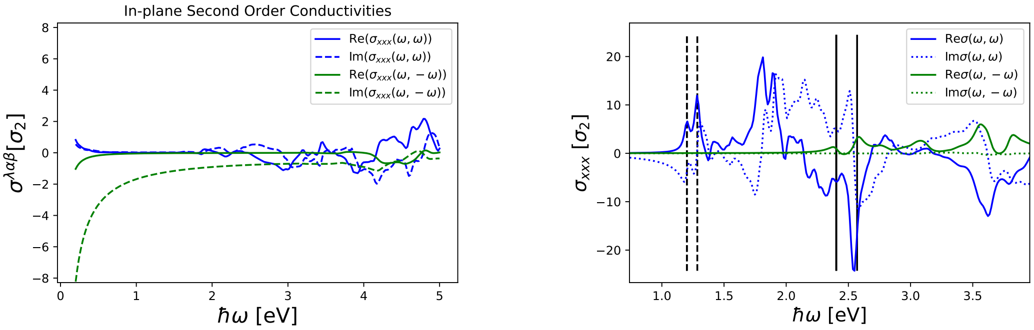


Figure 3.30: Band diagram for α-In$_2$Se$_3$.E_g = 0.95 eV. In accordance with Debbichi et al.[52].

The same can be said for the two lowest conduction bands, which now also do not cross from Γ to M or Γ to K. The calculation of the response runs along the same vein as presented before, but since the states are not of definite parity, the BSE, calculation does not decouple as nicely as in the previous subsection. We choose a scissor $\Delta = 2$ eV.[52] This results in the band gap being $E_g = 2.9$ eV, not far off the one in In$_2$Se$_2$. The two bottom valence conduction bands and four highest valence bands are found to be responsible for the first-order response in the region $\hbar\omega$ = 0-4 eV, and are therefore the ones included in the BSE calculation, seen in Figure 3.31a. The second-order VG IPA spectrum is seen in Figures 3.31b and 3.31c shows the BSE VG conductivity.



(a) First-order BSE and IPA conductivity. Four valence bands and two conduction bands included. 60×60 **k**-points for the BSE calculation. $\hbar\eta = 50$ meV.



(b) IPA second-order VG conductivity. Calculated using ~ 65000 **k**-points. $\hbar\eta = 50$ meV.

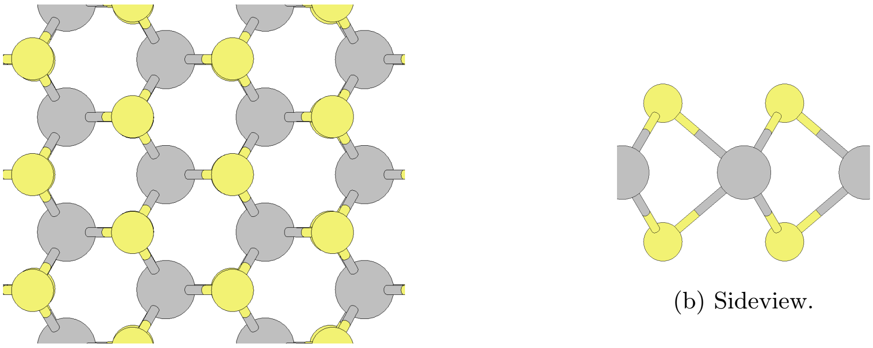
(c) BSE second-order LG conductivity. Calculated using 50×50 **k**-points. Dotted line are half E^1_{exc} and E^2_{exc}, while the full drawn are just these same energies. $\hbar\eta = 50$ meV

Figure 3.31

4 | MoS$_2$ Monolayer

Monolayer molybdenum disulfide or MoS$_2$, which crystal structure is seen in Figure 4.1, has received attention the past few years because of its direct band gap which is unique for the monolayer.[53] The spin-orbit coupling is furthermore significant in this material, meaning a spin-orbit term has to be incorporated into the TB Hamiltonian. By using orbitals with s, p and d symmetry in the TB model, this is can be done with relative ease.[53] [14] MoS$_2$ is a hexagonal TMDC with point-group P-6m2, meaning the general tensor-properties of monolayer In$_2$Se$_2$ also applies to the current density tensor in this material. MoS$_2$ is a crystal without inversion symmetry which means the energies in the crystal do not need to be spin-degenerate.[54]

Even though the calculations done on MoS$_2$ are interesting in themselves, good TB models for other TMDCs can be obtained using almost the same settings as in MoS$_2$, meaning these calculations are fairly easily replicable on other TMDCs. Beyond the calculations presented in this chapter and in chapter 3, a handful calculations on other materials can be found in Appendix E.



(a) Top-down view.

(b) Sideview.

Figure 4.1: Two different perspectives on the hexagonal MoS$_2$ crystal structure. Molybdenum atoms are in grey while the sulphur atoms are in yellow. The lattice constant is 3.184Å.[3] Figures made using XCrySDen.[51]

The band diagram of MoS$_2$ without spin-orbit coupling (SOC) is seen in Figure 4.2 while the band diagram with SOC included is seen in Figure 4.3. From these figures, it is evident that SOC splits the band structure of MoS$_2$ significantly around the K-point, and there are a few regions where the wave function is significantly hybridised between spin-up and spin-down states. However, for transitions near the band gap, the wave function is not is not significantly spin-hybridised, meaning it is possible to exclude the matrix-elements of $\hat{\mathcal{H}}_{\mathrm{SO}}$ that are nondiagonal entries wrt. spin is a decent approximation, thus decoupling the spin-up and down states in the BSE-calculation. One might expect some effects related to this spin-hybridisation, which we henceforth will ignore.

When we ignore SOC, the Hamiltonian will have four blocks that do not couple, like in In$_2$Se$_2$, and only $[\hat{\mathcal{H}}^{\mathrm{MB}}]^{PP}$ and $[\hat{\mathcal{H}}^{\mathrm{MB}}]^{NN}$ will contribute to the in-plane conductivity, as the momentum matrix elements of different parity states are zero, except for the z-direction. As is evident from

Figure 4.2 the lowest conduction band and highest valence band has the same parity, therefore contributing to the in-plane absorption spectrum. In many of the plots coming sections, we will be plotting a quantity throughout the first Brillouin zone. This will be done on a rectangular grid with coordinates (t_1, t_2), which corresponds to a $\mathbf{k}$-point through the relation $\mathbf{k} = \mathbf{G}_1 t_1 + \mathbf{G}_2 t_2$.

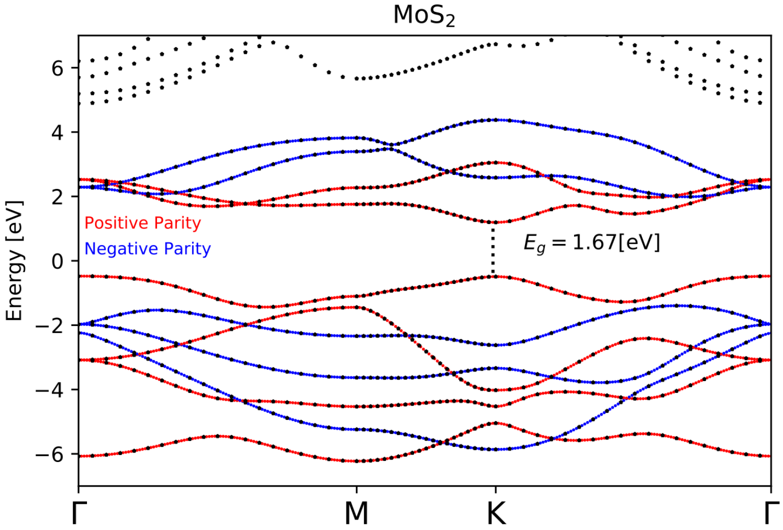


Figure 4.2: The band diagram of $\mathrm{MoS_2}$ with no SOC. Black dots are DFT-calculated bands using a PBE-exchange-correlation functional[55] while the coloured are a Mo-d S-p TB model. The parity is well-defined because the Hamiltonian w/o SOC commutes with the parity operator. $E_F = 0.53$[eV], $m_h^* = 0.567 m_e$ and $m_e^* = 0.454 m_e$.

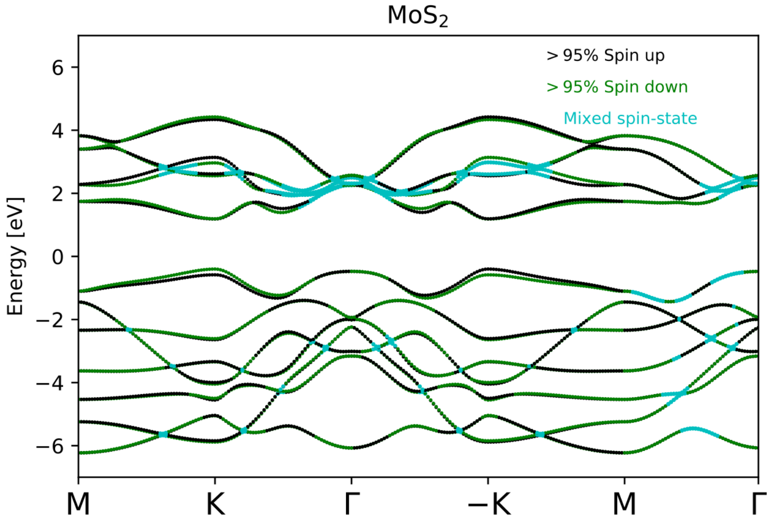
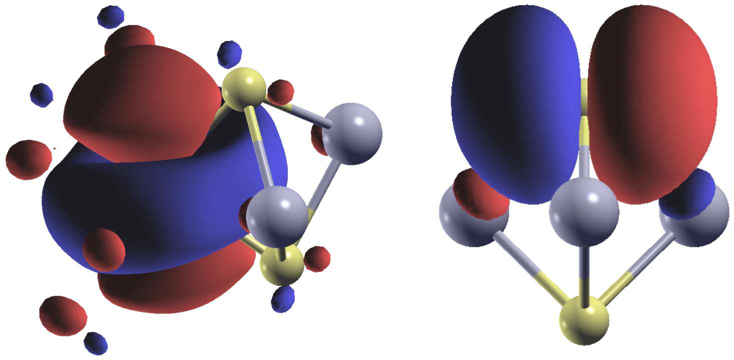


Figure 4.3: The band diagram of $\mathrm{MoS_2}$ with SOC. Top valence band split and lowest conduction band split at K-point is 183 meV and 3 meV. The spin-state is approximately well defined where the lines are green and black. Percentage calculated from absolute square of the TB-eigenvectors.

To correct the faulty DFT band gap, a scissors shift Δ can make the the band gap match a G0W0 calculation with $E_g^{\mathrm{G0W0}} \approx 2.53\,\mathrm{eV}$.[3] The out-of-plane momentum matrix elements (matrix elements of $-\mathrm{i}\hbar\frac{\partial}{\partial z}$) in the MLWF-basis can of course also can calculated by means of a finite-difference (FD) approach, given that the explicit form of the MLWFs are known. The MLWFs from Wannier90 can be extracted, and for $\mathrm{MoS_2}$ these functions show very close resemblance to the atomic d and p orbitals. Two MLWFs can be seen in Figure 4.4. Inspecting the form of the p-like orbital in

Figure 4.4b, it can be seen that a unitary transformation is needed in order to make the decoupling of different parity states obvious in the spin-diagonal TB Hamiltonian.



(a) A d_{z^2}-like MLWF centered on the Mo-atom.

(b) A p_y-like MLWF centered on a S-atom.

Figure 4.4: 3D contour-plot of two MLWFs of MoS_2. Figures made using XCrySDen.[51]

The symmetry of the orbital in question will of course dictate if it contributes to negative or positive parity bands. The d_{z^2}-orbital of Figure 4.4a as an example, will contribute to the positive parity bands of Figure 4.2. For all the calculations the phase is fixed by choosing the first entry of the eigenvectors to be real and if this element should happen to have a magnitude smaller than a threshold-value due to parity, the next entry is chosen real. For the exciton calculations, the parity is made explicit between the basis Wannier functions by finding an appropriate unitary transformation. In practice, this means combining the p-orbitals on sulphur as $\tilde{\omega}^{+/-} = \frac{1}{\sqrt{2}}(\omega_{p,S_1} \pm \omega_{p,S_2})$. Wannier90 turns out not to be consistent with the sign on the p_z orbitals one can plot and the sign with which the output Hamiltonian has to be transformed.

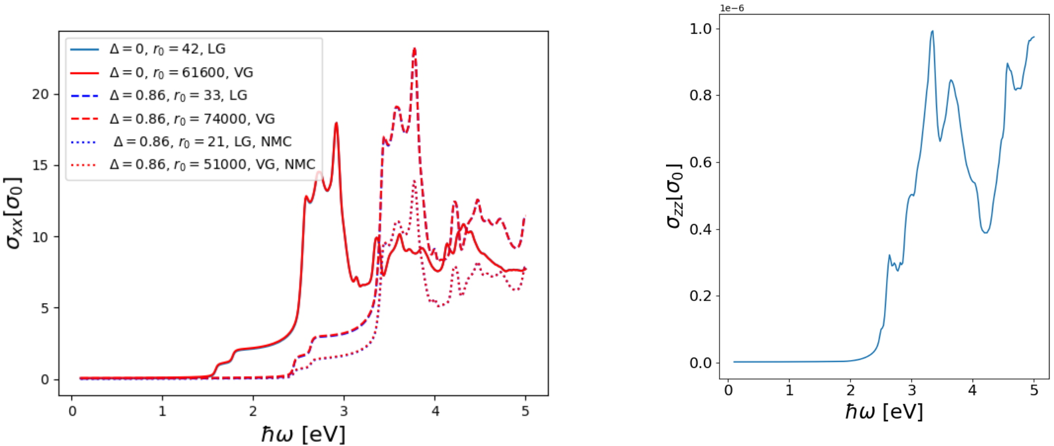
4.1 Response Spectra in the Independent-Particle-Approximation

4.1.1 First-Order Response

The first-order conductivity within the IPA is given in section 2.5 and is calculated without the before mentioned scissors shift as the static susceptibility is underestimated when applying the scissors shift. The primary purpose of this calculation is to determine the screening length r_0 and to show the shortcomings of the IPA for TMDC-materials.

The response can be calculated using both the LG expression (2.107) and VG expression (2.118). The conductivity of monolayer MoS_2 is plotted in Figure 4.5a, with and without a scissors shift, with and without the correction of the momentum matrix elements from (2.11). The VG and LG expressions are in agreement for the first-order conductivity. The imaginary part of the VG does however again diverge as $\omega \to 0$ leading to a much too large r_0. The scissors shift with (2.11) enlarges the shape of the conductivity somewhat, but still underestimates the screening length significantly compared to the MoS_2 screening length $r_0 = 44.3$Å from Pedersen[14] which the non-scissors shifted LG conductivity matches most closely. For this reason we chose to calculate r_0 from the non scissors shifted band structure.

σ_{zz} first becomes noticeably different from zero when the photon-energy matches the transition energy between two bands of opposite parity, see Figure 4.2. There is approximately seven magnitudes of difference of the in-plane and out-of-plane response when comparing the responses in Figure 4.5.



(a) NMC means without the correction of (2.11). r_0 is in Å and Δ in eV.

(b) $\sigma^{(1)}_{zz}$ calculated using the LG and no scissors shift. 5×5 unit cells used when evaluating matrix elements of $\frac{\partial}{\partial z}$

Figure 4.5: The real part of the IPA conductivities of a MoS$_2$ monolayer in units of $\sigma_0 = \frac{e^2}{4\hbar}$. Cutoff at $\hbar\omega = 0.2$ eV. A broadening term of $\hbar\eta = 20$ meV has been used. SOC-term included in Hamiltonian. $\sim 65500\mathbf{k}$-points

4.1.2 Second-Order Response

The second-order conductivity can be calculated using the two different expressions of section 2.5 and will be done using a scissors shift $\Delta = 0.86$ eV and the momentum correction (2.11). In Figures 4.6 and 4.7 these conductivity tensors are calculated in the LG and VG respectively. During these calculations, the phase was fixed by making the first nonzero element of the eigenvector real and positive. This derivative of the momentum matrix elements of the eigenstates was evaluated using a finite-difference approach directly on the momentum matrix element $\langle n\mathbf{k}|\,\hat{p}\,|m\mathbf{k}\rangle$. Using the unit $\sigma_2 = \frac{e^3\text{Å}}{8\text{eV}\hbar}$, one ends up multiplying a factor $\frac{g\mathrm{d}^2\mathbf{k}}{2\pi^2}$ onto the sums of (2.107) and (2.122). The results using the LG is seen in Figure 4.6 while the VG result is seen in Figure 4.7 for two different TB models.

The VG expression evidently diverges when $\omega \to 0$, contrary to the LG calculation. The two curves agree in terms of where $E_g/2$ is as the SHG terms kicks in here. The LG does however have a more well-defined peak at this position. Both gauges agree on the peak at $\hbar\omega \approx 1.9$ eV, but differ in size. Similarly to the In$_2$Se$_2$ calculation, the SHG spectrum is very similar to the OR spectrum halved in frequency.

The LG expression (2.107) has two intra-interband transition terms contributing to σ_{ijk}. The latter of these terms furthermore depends on the Berry connection Ω^x_{ii} of the material. In Figure 4.8, the x-component is seen for the top valence and bottom conduction band respectively, the difference of with enters the second-order conductivity expression (2.107). This difference is evidently significant, meaning the generalised derivative contributes significantly.

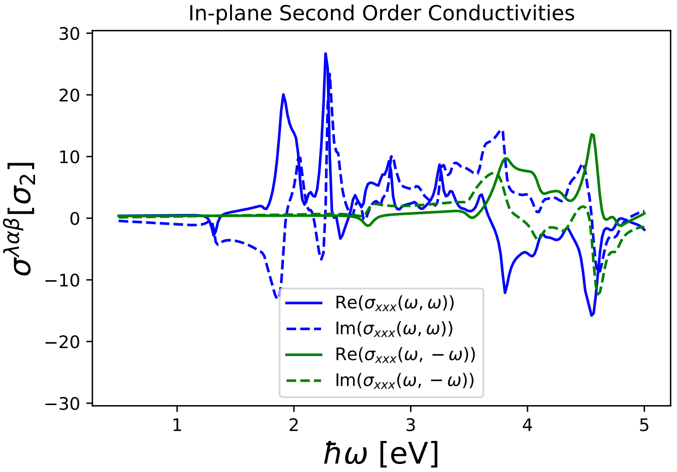
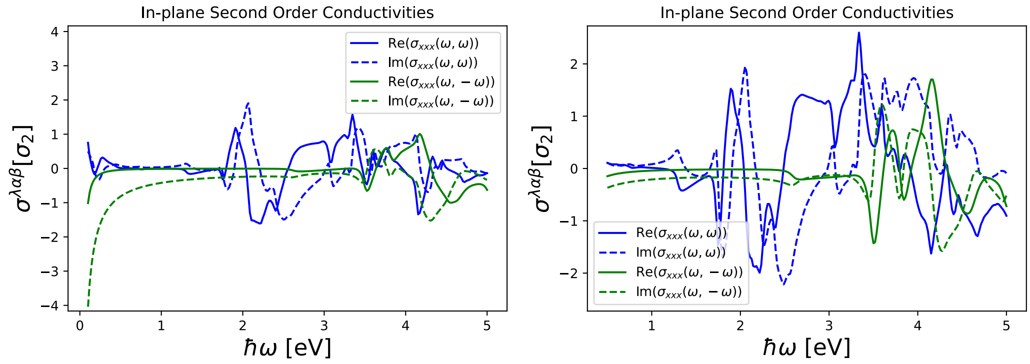


Figure 4.6: σ_{xxx} in units of $\sigma_2 = \frac{e^3 \text{Å}}{8\text{eV}\hbar}$ calculated in the LG without SOC. Broadening $\hbar\eta = 50$ meV and ~ 65500**k**-points-points.



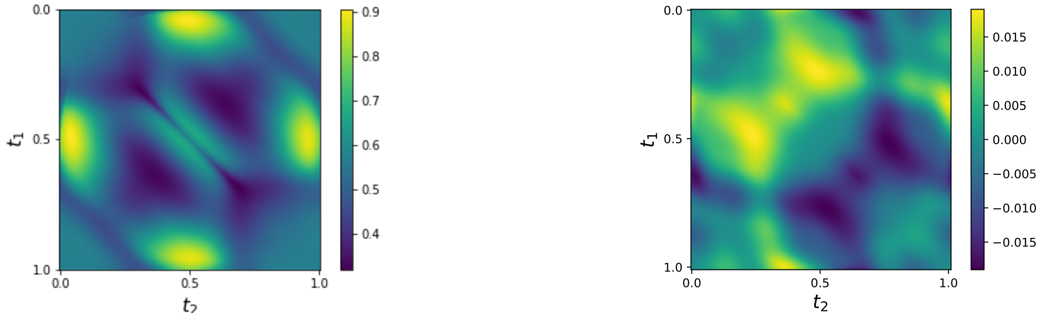
(a) 11 band model.

(b) 27 band model.

Figure 4.7: VG second order conductivity without SOC. $\hbar\eta = 50$meV and ~ 65500 **k**-points.

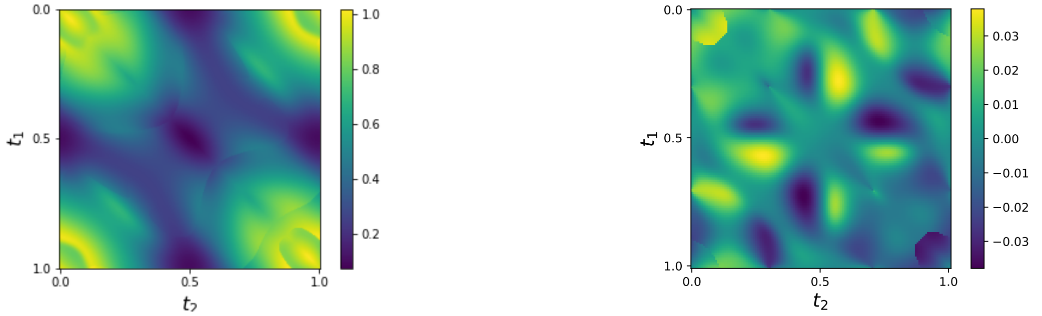
The generalised derivative of the momentum matrix element between the lowest conduction and top valence band can be seen in Figure 4.9. White points in these figures means the absolute value is very large and the point would otherwise skew the colorscheme of the plot is not excluded. Figure 4.9 shows the generalised derivative is significant at the K-points. The third term in (2.107) does depend on other things than the generalised derivative, but a nonzero $(p^x)_{;\mathbf{k}}$ is a prequisite for this term to contribute and it is nonzero even at $T = 0$K. Figure 4.9 shows the generalised derivative of the momentum matrix element is significant, in particular around the K-points in the first Brillouin zone, where it is large and negative. This makes this region around the K-point fit with the first negative peak in Figure 4.6. x-direction is along line from the top left to the bottom right corner in Figure 4.9.

The problem of nonzero imaginary part on the diagonal entries of $\mathbf{\Omega}_{mn}$ persists in MoS$_2$, however to a lesser degree as is seen in Figure 4.8, where the magnitudes on the imaginary part is now about $\sim \frac{1}{50}$ the magnitude of the real part. With SOC introduced and using the (B.5), one gets the Chern numbers integrate to zero to third decimal and without SOC, this is instead to seventh decimal.



(a) ReΩ^x_{ii} with i the top valence band.

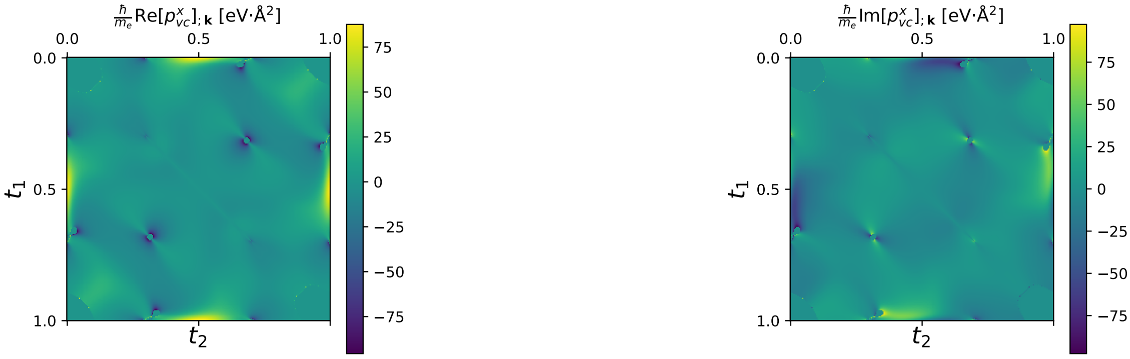
(b) ImΩ^x_{ii} with i the top valence band.



(c) ReΩ^x_{ii} with i the lowest conduction band.

(d) ImΩ^x_{ii} with i the lowest conduction band.

Figure 4.8: The Berry connection throughout the Brillouin zone in units of Å.



(a)

(b)

Figure 4.9: $(p^x)_{;\mathbf{k}}$between the top valence and lowest conduction band without SOC in the x-direction.

4.2 Excitonic States

We now turn to the calculation of the excitons of the system. As a first estimate on the binding energy of the lowest exciton, we can use the effective masses from Figure 4.2, $r_0 = 42$Å and the binding energy from (2.141), yielding a binding energy of $E_B^{1s} = 0.42$ eV. This is compared to a binding energy of 0.55 eV reported on C2DB using a full BSE-approach.[3] We are evidently quite far off with this value so we turn to the slow-rapid version of the BSE from (2.138). The BSE calculation will follows the same vein as In_2Se_2, thus we start at (2.138) with a scissors shift $\Delta = 0.86$ eV, the previously calculated $r_0 \approx 42$Å, $\kappa = 1$, and the spin-diagonal 2D version of (2.133) on the form[31]

$$W_C^{\varsigma\varsigma'}(k) = \delta_{\varsigma\varsigma'} \frac{e^2}{\epsilon_0 k(\kappa + r_0 k)}. \tag{4.1}$$

As $W(k)$ stands here, the Hamiltonian will depend on the particular way the **k**-grid is made. For MoS_2 this is not a problem,[29] as the excitons will turn out to be localised well within the Monkhorst-Pack grid we chose previously.

The quasiparticle energy $\tilde{E}^s_{cv}(\mathbf{k})$ is now instead calculated from diagonalising a spin-dependent TB Hamiltonian on the form

$$\hat{\mathcal{H}}^{\varsigma\varsigma'} = \hat{\mathcal{H}}^{\mathrm{TB}} + \delta_{\varsigma\varsigma'}(\hat{\mathcal{H}}^{\mathrm{SO}}_{\varsigma,\varsigma'} + \hat{\Delta}), \quad \text{with} \quad \hat{\Delta} = \Delta \sum_c \Big|c\Big\rangle \Big\langle c\Big|, \tag{4.2}$$

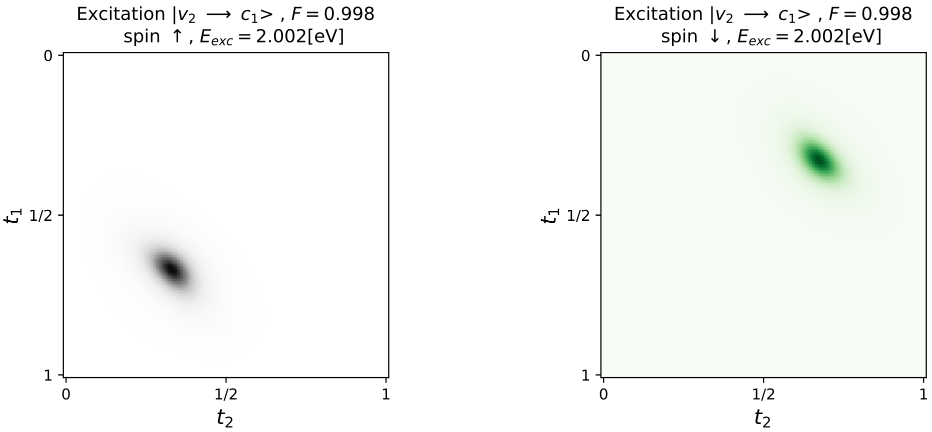
where the matrix elements of $\mathcal{H}^{\mathrm{TB}}$ are the matrix elements from (2.8). This keeps the parity well-defined and the spins decoupled. Plugging this into (2.138) with **k**-points in a $N_k \times N_k$ Monkhorst-pack grid, shifted by $\frac{\mathbf{b}_1+\mathbf{b}_2}{2N_k}$ gives the BSE Hamiltonian within the slow-rapid approximation, with a total of eight decoupled blocks. We discard the two lowest positive parity valence bands and the single lowest negative parity band from Figure 4.2.

In Figure 4.10, plots of the coefficients of the dominant Slater determinants in the wave function and their energies is seen, obtained from solving the eigenvalue-problem. Here, the lowest lying excitons of each spin-band are seen to be strongly confined to K- and $-K$-valleys while also being degenerate. Their binding energy $E_b = 0.526$ eV from the caption of Figure 4.10 is more in line with the one reported on C2DB.

The ten lowest energy of the spin up and down PP-blocks of the Hamiltonian have also been calculated and is seen in Figure 4.11. The two states of the spin-up and down blocks of $\hat{\mathcal{H}}^{\mathrm{MB}}$ depicted in Figure 4.10b and Figure 4.10a are degenerate, but not located in the same K-valley because of spin-flipping seen in Figure 4.3.

Looking at the lowest part of the calculated energy spectrum in Figure 4.11 it is evident the spectrum is not hydrogenic as the binding energy for a 2D hydrogen model is on the form $E_B = \frac{1}{2(n-\frac{1}{2})^2}$, $n = 1, 2, 3...$ with an appropriate choice of units[34] and would have the second-lowest exciton much closer to $E_{\mathrm{exc}} \approx 2.48$ eV. The nonhydrogenic spectrum in Figure 4.11 is a result of the reduced screening present in the monolayer and is encapsulated in the Rytova-Keldysh potential.

The states depicted in Figures 4.10a and 4.10b are s-excitons as we see from their (lack of) angular dependence. Because an actual s-state is indifferent to rotation, its dipole moment can be large. These are therefore the states one can expect to be clearly defined in the optical spectrum of the material.



(a) Lowest energy exciton for the spin↑ PP block.

(b) Lowest energy exciton for the spin↓ PP block.

Figure 4.10: The absolute value of the eigenvector of $[\mathcal{H}^{\uparrow/\downarrow}]^{\mathrm{PP}}$ containing the transitions from the highest valence band to the lowest conduction band, plotted in the entire Brillouin zone. Two top valence bands and two lowest conduction bands used. F is the sum of absolute squares of the this subset of entries in the eigenvector. The K-point is located at $(2/3, 1/3)$ while the $-K$-point is located at $(1/3, 2/3)$. $E_g = 2.528$ eV, giving a binding energy $E_b = 0.526$ eV. Calculated with 60×60 $\mathbf{k}$-points. The two excitons are degenerate to ninth decimal.

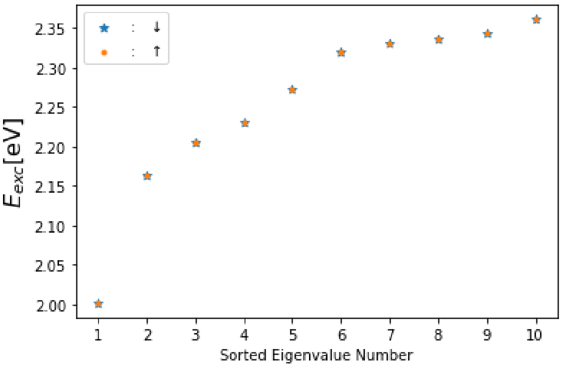
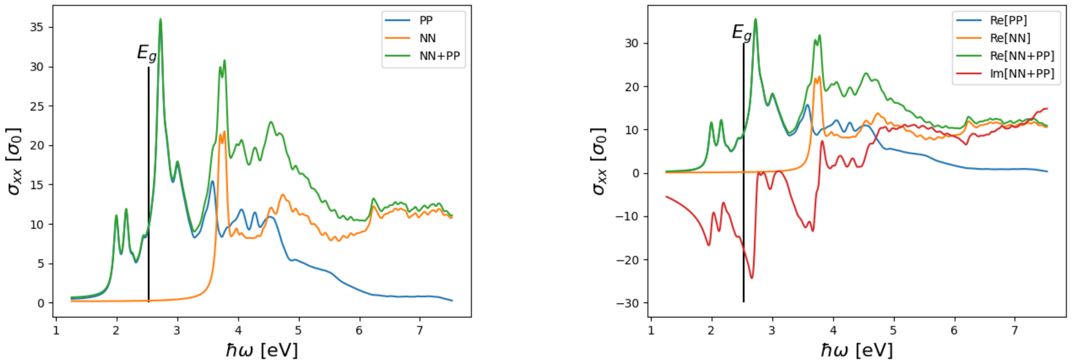


Figure 4.11: The first 10 eigenvalues of the $[H]^{PP}$ blocks with 60×60 $\mathbf{k}$-points. $E_g = 2.528$ eV. Calculated by diagonalising the truncated Hamiltonian obtained by the Lanczos-Haydock routine with 1000 iterations.

4.2.1 First-Order Excitonic Response

Turning now to the optical response with excitonic effects included, we have two approaches. We can find the diagonal element $\langle P|\, \hat{G}\, |P\rangle$ of the Green's function using (2.151) and (2.152), or we can use the excitonic LG expression (2.156) in conjunction with a diagonalisation of the truncated BSE Hamiltonian. The results of both these approaches is seen in Figures 4.12a and 4.12b. The correction (2.11) has not been employed for these spectra.



(a) Real part of the first-order conductivity calculated with the Lanczos-Haydock method using 60×60 **k**-points, two valence and conduction bands of each spin state. $\hbar\eta = 50$ meV.

(b) Calculated with (2.156) and 800 eigenstates of the 800×800 truncated BSE Hamiltonian using 60×60 **k**-points and two valence and conduction bands of each spin state. $\hbar\eta = 50$ meV.

Figure 4.12: First-order conductivity of monolayer MoS_2 including excitonic effects calculated using two different methods.

The distinct peaks at $\hbar\omega = 2.00$ eV and $\hbar\omega = 2.16$ eV are clearly seen. The splitting of 0.16 eV between these two peaks are reasonably in line with the 183 meV spin-orbit splitting of the valence band. The other bound states from Figure 4.11 are not as distinct as the two lowest excitons. The conductivity does not approach zero with the damping chosen for Figure 4.12a, but can be made arbitrarily small by making the damping smaller. The two lowest peaks are the two 1s-excitons from Figure 4.10b and their corresponding one in the other K-valley between the same spin-bands.

The two spectra using the two different methods do not have any noticeable disagreements on the real part of the conductivity. Throughout most of the visible spectrum it is the states of the PP block of the Hamiltonian which contributes to the absorption. We also note the spectrum is in good agreement with the one found on C2DB. The actual excitonic peaks are located about 0.1 eV below the calculated positions,[48] but a modified scissors shift would bring the calculated spectrum in line with the observed peaks.

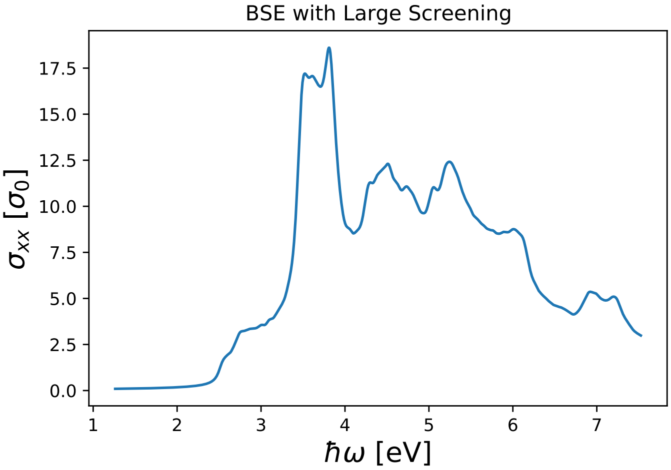
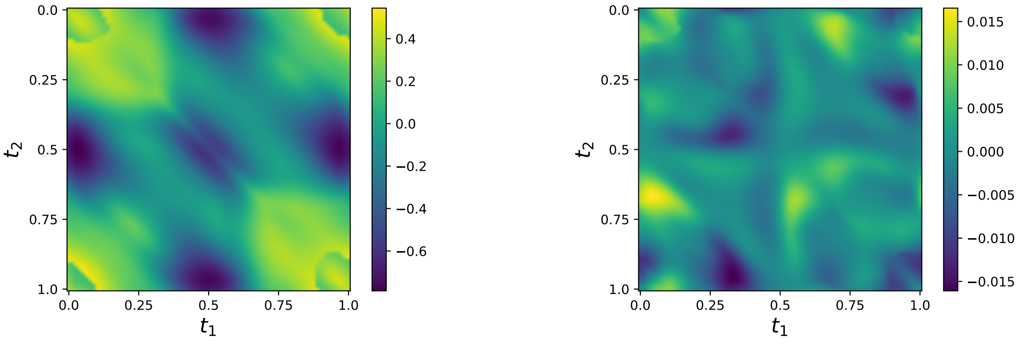


Figure 4.13: Real part of the conductivity calculated with the LG expression with large screening. 70×70 **k**-points and 2 valence and 2 conduction bands of each parity and spin. $\eta = 50$ meV and $\kappa = 100$.

It is evident from Figure 4.12 that the excitonic effects modifies the optical conductivity significantly compared to the IPA result from Figure 4.5a. The constant conductivity region from $2.5 - 3.2$ eV is absent, but the peak at 3.7 eV when transitions between negative parity states becomes significant is however present, but appears less significantly in this spectrum. Significantly higher peaks values are also obtained when excitonic effects are included. The IPA result from Figure 4.5a should however be obtained in the limit of infinite screening, modelled by letting $\kappa \to \infty$. This can also be seen in Figure 4.13 to be the case with a reasonable margin of error, albeit deviating above 4 eV.

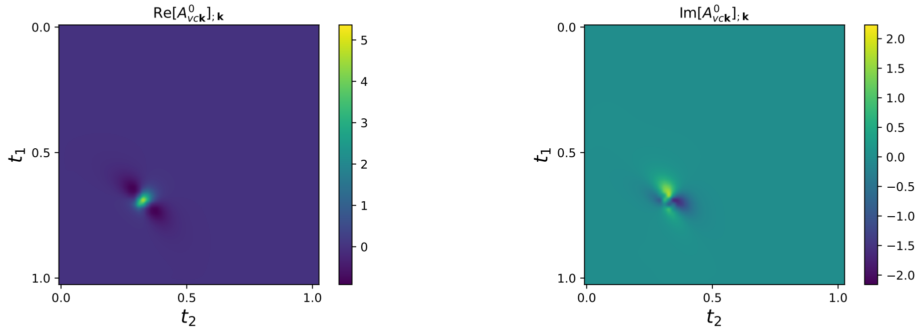
4.2.2 Second-Order Excitonic Response

We approach the calculation of the second-order conductivity in much the same way as in In$_2$Se$_2$. We consider first the difference in Berry connection Ω^x_{mn} between the top valence and lowest conduction bands, which can be seen in Figure 4.14. The generalised derivative of the lowest excitonic state can be seen in Figure 4.15.



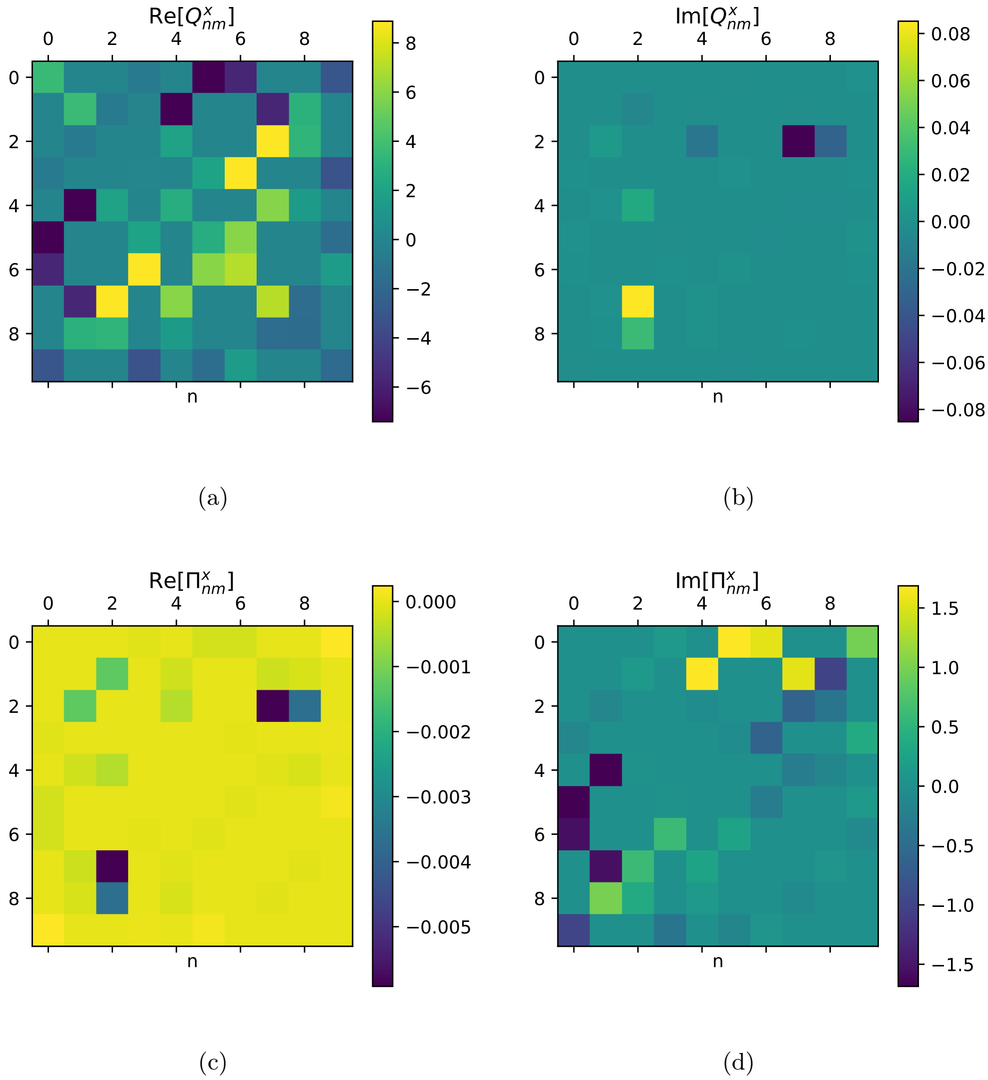
(a) Re$[\Omega^x_{c_1c_1} - \Omega^x_{v_2v_2}]$ (b) Im$[\Omega^x_{c_1c_1} - \Omega^x_{v_2v_2}]$

Figure 4.14: Difference in the Berry connection of the top valence band and bottom conduction band for the spin-up bands.



(a) Real part. (b) Imaginary part.

Figure 4.15: The generalised derivative of the expansion coefficient $A_{cv\mathbf{k}}$ in the k_x-direction (along the diagonal from the top left corner to the bottom right corner) for the excitation between the top spin-up valence band and lowest spin-up conduction band. Gaussian interpolation have been used to smooth the plot, which was calculated on using 60×60 $\mathbf{k}$-points. In units of Å.



(a) (b)

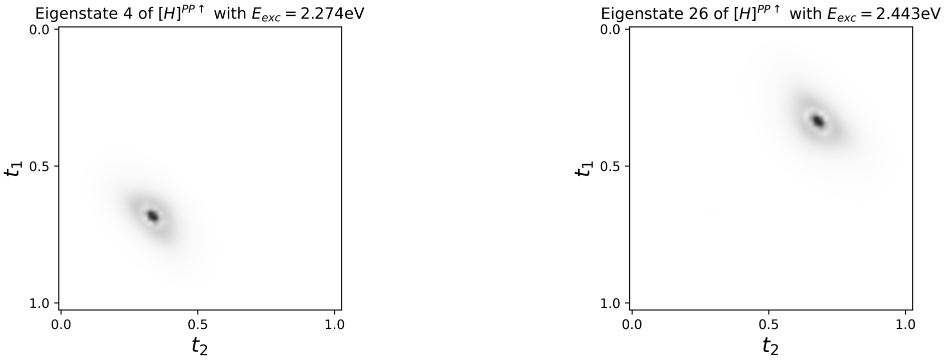
(c) (d)

Figure 4.16: Matrix elements of the excited states of $\mathrm{MoS_2}$ without SOC. Calculated on a 80×80 grid with one conduction and one valence band. Q^x in units of Å and Π^x in units of eV·Å.

The value of the matrix elements $\mathbf{X}_{nm}$ will then in part be determined by the intraband elements $\mathbf{x}_{cc'}$ and $\mathbf{x}_{v'v}$ from (2.154), but also on the generalised derivative of the matrix elements. The generalised derivative in turn is large if the $A_{cv\mathbf{k}}$ is strongly localised in $\mathbf{k}$ and the difference in the Berry connections of the bands is large. In Figure 4.16 one can see the matrix elements $X^x_{nm} = \frac{m_e}{\hbar} Q^x_{nm}$ for the 10 lowest lying eigenstates of the $[\hat{\mathcal{H}}]^{PP\uparrow}$ block in the case of one valence and one conduction band. From this figure it can be seen e.g. that the Q^x_{01} is insignificant, which is because these two states are located in different K-valleys.

Using integration by parts on a localised state, the generalised derivate is a hermitian operator, which can also be seen from Figure 4.16, as both Q^x_{nm} and Q^x_{mn} have been calculated. For the matrix elements involving the first 100 eigenstates, the maximum $\max_{nm}\left[|Q^{x*}_{nm} - Q^x_{mn}|/\mathrm{Re}[Q^x_{00}]\right] \approx 4.58 \cdot 10^{-4}$. This maximum does however increase in size as one considers higher energy states obtained from the Lanczos-Haydock procedure.

For the following plots of the second-order conductivity, we plot it in units of $\sigma_2 = \frac{e^3 \text{Å}}{8\text{eV}\hbar}$, the same unit as the IPA spectra. They are furthermore computed without the correction (2.11). The second-order conductivity for MoS$_2$ is seen in Figure 4.18 and shows the excitonic transitions has a clear influence on the second harmonic generation and optical rectification. Whereas the optical rectification was close to zero around the band gap in the IPA response in Figure 4.6, the presence of the excitonic transitions means it is now significant below the quasiparticle band gap. The spectra seen in Figures 4.18 and 4.19 are the MoS$_2$ second-order conductivity within the LG, both calculated using two valence and two conduction states, but the former has been calculated with 80 × 80 **k**-points while the latter has been calculated with only 45 × 45 **k**-points (only $\approx \frac{2}{5}$ the amount of the former). The similar appearance indicate we have a converged calculation. The distinct peaks in $\sigma^{(2)}_{xxx}(\omega, \omega)$ in the region 1-1.25 eV are related to bound excitonic states below the band gap while larger peak from 1.4-1.7 eV seems to be caused by the continuum of states above the electronic band gap. This is because the peaks below $E_g/2$ appear after only including the 30 lowest states while the larger peak only converges after including 500 states from the Lanczos-Haydock routine. The excitonic spectrum furthermore has two additional peaks stemming from higher 2s-like excitons. The states responsible for these peaks can be seen in Figure 4.17. In an hydrogen-atom, the spin-orbit coupling does not split s-states,[14] but because the underlying mechanism for the splitting is within the underlying band structure, this is to be expected.



(a) 2s state responsible for the third peak in Figure 4.18

(b) 2s state responsible for the fourth peak in Figure 4.18

Figure 4.17: 2s excitons of MoS$_2$ in each K-valley.

In Figure 4.19a the second-order conductivity obtained using the conventional VG is seen with only four bands. Within this gauge, beyond being factor 2 smaller in overall magnitude compared to the LG, the peaks related to the excitonic structure is much less well-defined in comparison to the very apparent peaks related to the s-states in the LG conductivity. They do however become more apparent in Figure 4.19b, where all the positive parity bands has been included. The spectrum in the VG is however not converged with respect to the number of bands.

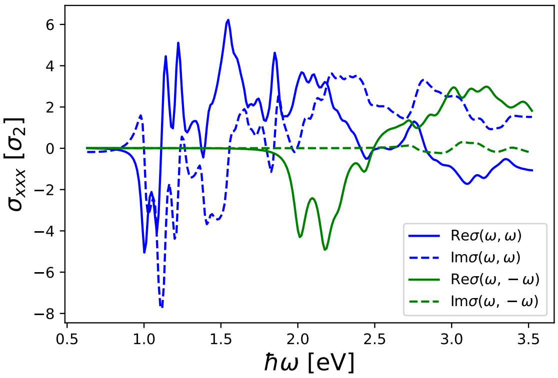
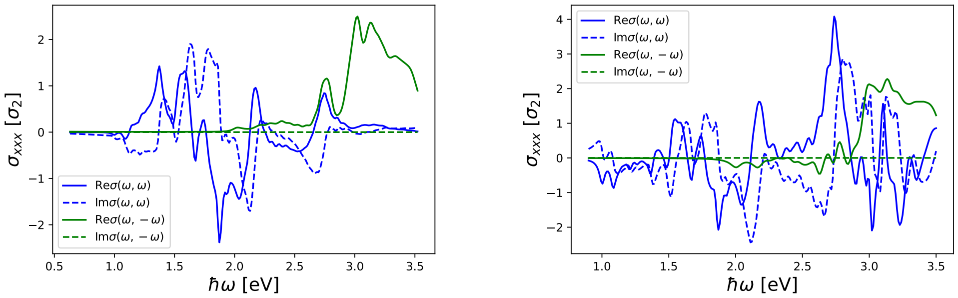


Figure 4.18: The second-order conductivity of MoS$_2$ using the LG, calculated with 70×70 **k**-points with two conduction and two valence bands of positive parity and the 500 eigenstates of the truncated BSE Hamiltonian. Four first peaks matches the energies of of the four s-excitons. $\hbar\eta = 50$ meV



(a) The second-order conductivity of MoS$_2$ using the conventional VG, calculated with 80×80 **k**-points with two conduction and two valence bands of positive parity and the 800 eigenstates of the truncated BSE Hamiltonian.

(b) 50×50 **k**-points but all positive parity bands included in 11-band model. 400 eigenstates of the truncated BSE Hamiltonian.

Figure 4.19: $\hbar\eta = 50$ meV in both figures.

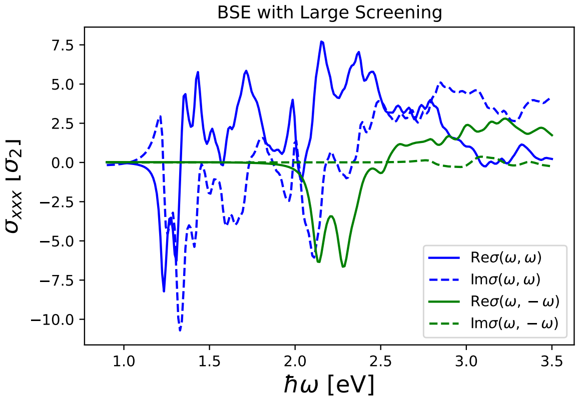


Figure 4.20: The second-order conductivity of MoS$_2$, calculated with 45×45 **k**-points with two conduction and two valence bands of each parity and spin and the 500 eigenstates of the truncated BSE Hamiltonians. $\hbar\eta = 50$ meV.

The OR in the excitonic spectrum and IPA spectrum seems to have in common that it does not start at half the excitation energy, but has to have a matching transition in order for the optical rectification term to be nonzero. The LG and VG also disagrees on the sign of the optical rectification term. Given the way the optical rectification has been calculated with the symmetrisation in ω_p and ω_q, the optical rectification should not yield an imaginary part. As seen in e.g. Figure 4.18, this ceases to be the case for higher photon energies, and seems to be caused by the eigenstates of the truncated Hamiltonian not being close enough to the actual eigenstates.

The lowest transition between the bands of $[\hat{\mathcal{H}}]^{NN}$ is $E_{\text{exc}}^{0,NN} = 3.70$ eV, meaning at half this photon energy its contribution to the second-order spectrum becomes significant.

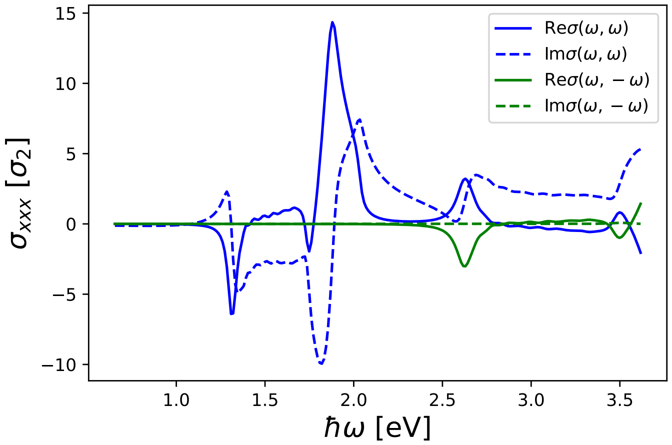
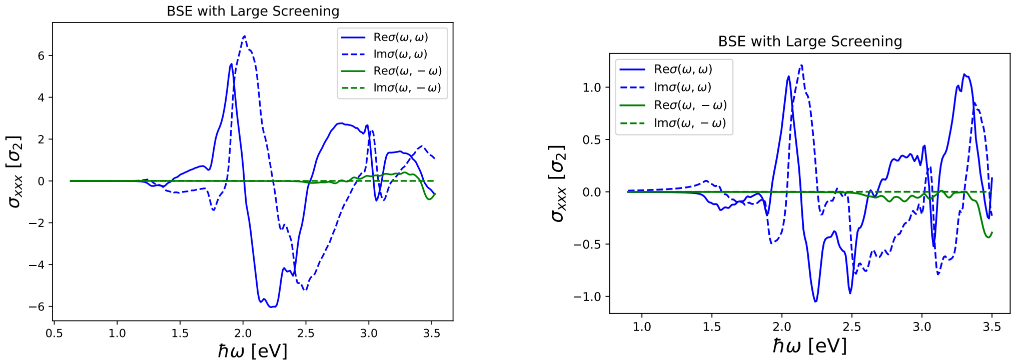


Figure 4.21: Second-order BSE conductivity with the LG. Calculated with 70×70 **k**-points and 2 valence and two conduction bands of each spin state and parity.

The spectrum with large screening $\kappa = 100$ is seen in Figure 4.21. Figure 4.21, in terms of shape, fits nicely in first half of the spectrum when comparing to the LG IPA spectrum from Figure 4.6, when accounting for the fact that Figure 4.6 was calculated without spin-orbit coupling. A difference is however the magnitude, as the LG result of Figure 4.21 is a bit larger than the LG IPA result of Figure 4.6. The large peak at 1.9 eV in the $\sigma_{xxx}(\omega,\omega)$ seems a common feature of the IPA spectra, and the presence of excitons appears to broaden and shift this peak, redshifting it to ≈ 1.6 eV in Figure 4.18. The conventional VG conductivity in Figure 4.22, apart from a factor two, is more in line with the IPA VG conductivity from Figure 4.7.

No clear-cut p-, d-, or f-states seems to be present when inspecting the eigenstates of the $[\hat{\mathcal{H}}]^{PP}$ block of the Hamiltonian. Lastly, we demonstrate the we have a converged result using the 800 states obtained from the Lanczos-Haydock procedure. Using just the two bands we obtain plot the convergence check in Figure 4.23, 200 iterations and above is enough



(a) Second-order BSE conductivity with the conventional VG. Calculated with 70 × 70 **k**-points and 2 valence and two conduction bands of each spin state and parity (16 bands).

(b) Second-order BSE conductivity with the conventional VG. Calculated with 50 × 50 **k**-points and all positive parity bands.

Figure 4.22

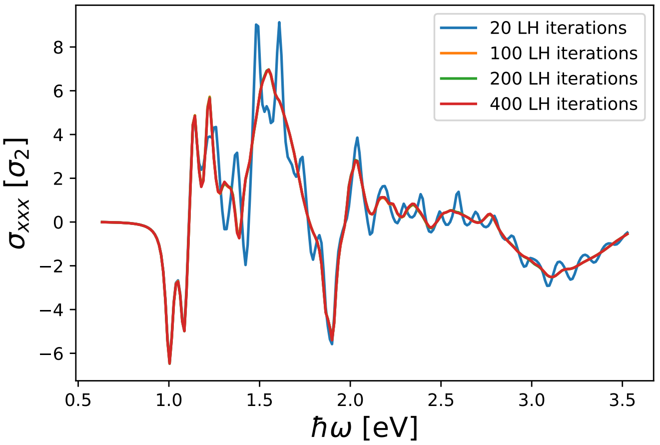


Figure 4.23: Convergence check using one valence and conduction band, 60×60 **k**-points, plotting the real part of the SHG spectrum. $\hbar\eta = 50$ meV.

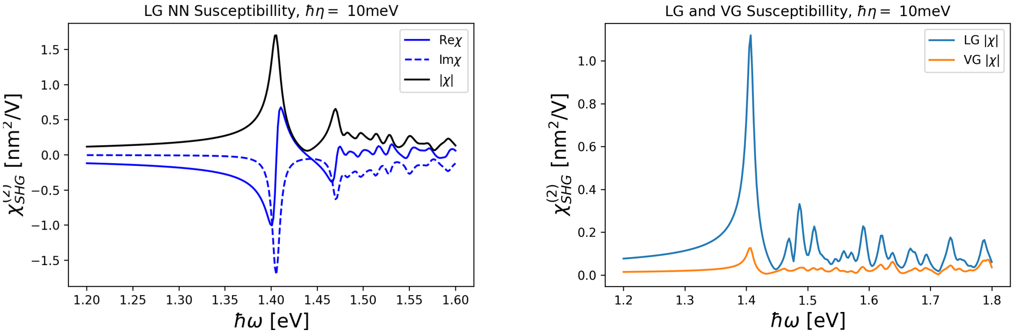
5 | Comparison with Experiments

Having been through the theory and applications, we now turn to see how the the calculations of the nonlinear response fits with a handful of experiments one can find in the literature on monolayers.

The second order response of a solid is in much of the literature given as a susceptibility, often measured in sheet susceptibility with units $\frac{\mathrm{nm}^2}{\mathrm{V}}$. As we have measured in units of σ_2 throughout the thesis, a factor of conversion is needed. The conversion happens as $\chi^{(2)}_{\mathrm{SHG}} = \frac{\sigma(\omega,\omega)}{-2i\omega\epsilon_0}$.[22] In effect, given all our second order conductivity plots are in units of σ_2, this results in a numerical factor of $\approx \frac{0.112}{(\hbar\omega)_{\mathrm{eV}}}$ to convert between the two units of measure of the response.

In_2Se_2 Monolayer

In Leisgang et al.[56] they examine few-layer In_2Se_2 crystals encapsulated in h-BN, and among these is the monolayer variant of In_2Se_2. Because the band gap of h-BN is much larger than In_2Se_2 we need not worry about resonances of BN, but we can regard it as a dielectric. At a wavelength of 405 nm$\sim$ 3.06 eV they observe a spike in the nonlinear sheet susceptibility of $(17.9 \pm 11) \cdot 10^{-2} \frac{\mathrm{nm}^2}{V}$.[56] Obtaining a susceptibility in the region $0.1 - 0.3 \frac{\mathrm{nm}^2}{V}$ would therefore be success. Looking at the vacuum-monolayer-vacuum response of Figure 3.25a, one might suspect the lowest fundamental exciton of being responsible, and calculate a peak susceptibillity of $\sim 0.65 \frac{\mathrm{nm}^2}{V}$. Setting $\kappa = \varepsilon_{BN} = 3.8$ in the Rytova-Keldysh potential and setting our ad hoc $\hbar\eta = 10$ meV to match the smaller linewidth seen in experiment, the NN block in the LG and VG gives the non-linear susceptibility seen in Figure 5.1a.



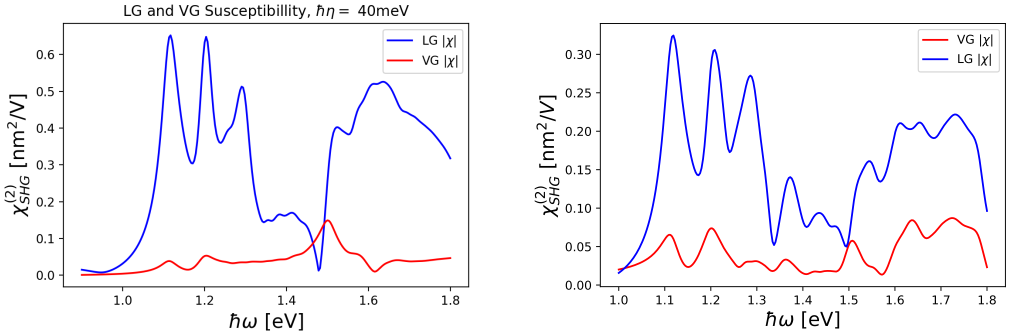
(a) Calculated two valence bands and one conduction band of the NN block using 80×80 **k**-points.

(b) Calculated using all negative parity bands but only 40×40 **k**-points.

In the LG we have two peaks while the VG only has one significant peak. In Leisgang et al.[56], they only scan a frequency region of about 0.11 eV, meaning only one of the peaks should be visible if one believes the LG result. The second peak does however not seem like something one would omit from a data set if it was present, but it is possible as their scanning window only covers a frequency difference of 0.11 eV. However, the size and the fact that the VG only has one significant peak does imply the VG result is correct here. The fact that it is 0.1 eV lower in energy, compared to the $\frac{3.06}{2}$ eV of Leisgang et al.[56] is probably down to imprecisions in the calculation of the excitonic binding energy and the scissorshift. Our modelling does also not include any effects from the interactions between BN and the In_2Se_2 monolayer, so if interactions between the BN and In_2Se_2 are strong, we are not modelling this system correctly.

MoS_2 monolayer

In Malard et al.[57] they have a similar situation to previous subsection, with MoS_2 on a quartz substrate and vacuum on top. Here they observe a peak at approximately $2\hbar\omega = 2.85$ eV (measured by eye on graph), with a second order sheet susceptibility of $6.4 \cdot 10^{-2} \frac{\text{nm}^2}{V}$ for the monolayer.[57] Setting $\kappa = \frac{\varepsilon_{\text{Quartz}}+1}{2} = 2.4$ and $\hbar\eta = 40$ meV one obtains the susceptibility in Figure 5.2.



(a) 2 valence and one conduction band, 80×80 **k**-points

(b) All positive parity bands in 11 band model, 40×40 **k**-points

Figure 5.2

It is evident from the first peaks that our exciton binding energy has gotten smaller. Inspecting the absorption spectrum found in Malard et al.[57], one does however find the lowest two exciton peaks located at ≈1.95 eV and ≈2.05 eV. the calculated lowest two exitons are however present at ≈2.2 eV and ≈2.4 eV. This means our spectra needs a shift of ≈0.1 eV in order for the transition energies to match. This means the peak at $\hbar\omega = 1.5$ eV should correspond to the one observed at a energy of 2.85 eV in Malard et al.[57], and indeed, the magnitude of the response does. The shape is however much narrower, meaning the similarities in shape this peak and the one seem in Malard et al.[57] are probably accidental.

WSe_2 monolayer

In Seyler et al.[58] they examine monolayer WSe_2 on quartz subtrate. They do a measurement of the fundamental exciton also present in WSe_2, which very similar in structure to MoS_2, both structurally and in their band diagrams. At an excitation energy of $\hbar\omega = 0.835$ eV they observe a peak value of $|\chi_{\text{SHG}}| = 0.04\frac{\text{nm}^2}{\text{V}}$ with a width of roughly 0.04 eV.[58]

The procedure is the repeated, setting $\kappa = 2.4$, $\hbar\eta = 20$ meV and do an IPA-calculation to obtain $r_0 = 36.9$Å. We scissorshift to match $E_g = 2.1$ eV from C2DB for WSe_2 and apply SOC, such that the valence bands splits by ~ 0.5 eV, roughly matching C2DB.[3] Doing so, one obtains the susceptibillity in Figure 5.3

Being roughly a factor 3 off in the VG case and a factor 60 off in the LG case is not satisfactory, but the LG result is again off by a large margin. The question of the combined accuracy of binding energy and scissors-correction can again be argued to be the cause of the mismatch in peak location. The shape of the VG result actually matches the experimental peak shape nicely to the one found in Seyler et al.[58].

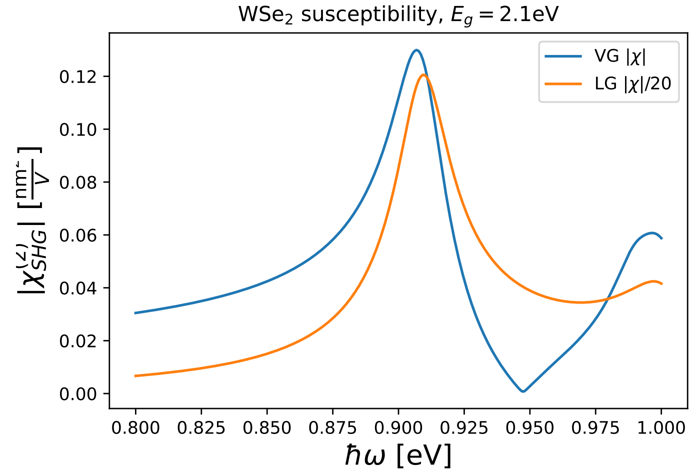


Figure 5.3: WSe$_2$ susceptibility as obtained with the previous parameters using 50 × 50 **k**-points and all positive parity bands in a 11 band model. Note the factor 1/20 on the LG result.

MoTe$_2$ monolayer

In Song et al.[59] report a strong second harmonic generation in monolayer MoTe$_2$ on top of quartz substrate at a wavelength 775 nm $\sim$ 1.60 eV corresponding to two photons of energy $\sim$ 0.80 eV combining. Song et al.[59] state a bulk susceptibility of $2.5\frac{\text{nm}}{V}$ and a monolayer thickness of 0.8 nm, yielding a sheet-susceptibility of $2.0\frac{\text{nm}^2}{V}$. The lack of frequency resolved plot of the susceptibility means we simply take $\hbar\eta = 20$ meV as in WSe$_2$. The scissors shift is done to obtain a band gap of 1.59 eV, roughly matching the band gap of 1.56 eV found on C2DB.[3] An IPA calculation reveals a $r_0 = 63.5$Å and the SOC on the molybdenum d orbitals is taken to be the same as in Mo$_2$ while the tellerium p orbitals are taken to have double the SOC as the sulfur p orbitals in MoS$_2$. While

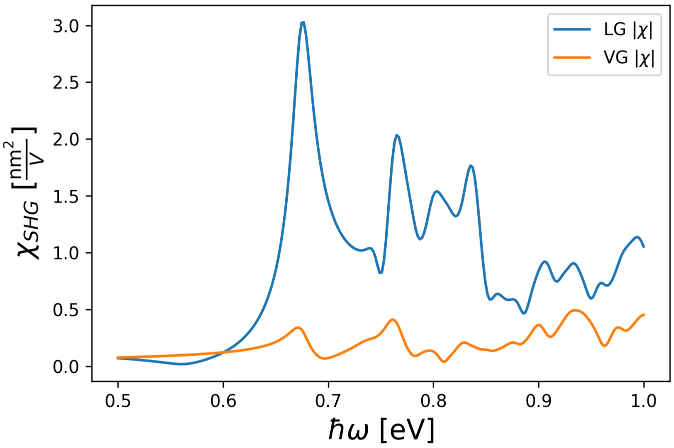


Figure 5.4

the LG result do seem the match the stated $2.0\frac{\text{nm}^2}{V}$ from Song et al.[59], but given how the LG has been far off experiment in the other cases, this seems like a coincidence. The VG result, while smaller than the $2.0\frac{\mathrm{nm}^2}{V}$, is significantly larger than in the other cases. The positions of the peaks are not a good match either, but the size of the VG result seems to also be found in experiment.

6 | Discussion

Throughout the calculations done for this thesis, various approximations, simplifications and assumptions have been employed. Here we discuss the limitations of these, some ways we could possibly have been more accurate in our calculations, the possibilities when the localised TB orbitals are known explicitly, and lastly, we also discuss various difficulties we encountered in our calculations.

6.1 Accuracy of the In_2Se_2 IPA Spectra

This section shall discuss the IPA spectra in an attempt to explain the difference between the LG and VG second-order spectra and to determine which spectra is most believable. Additionally, the effects of certain assumptions and simplifications made during the calculation of the IPA spectra shall be presented quantitatively.

6.1.1 Convergence of VG Spectra

During the calculations, it has been shown that the LG spectra generally do not match up with VG spectra when determining second-order effects. Notably, the first peak has much smaller amplitudes. This discrepancy was initially hypothesised to be caused by a lack of convergence of the VG since this gauge has been shown to be more sensitive to basis truncation. This hypothesis can be investigated by creating an additional 36-band TB model with Wannier90 for In_2Se_2. The second-order VG spectra can then be calculated and compared with the 14-band spectra as shown in Figure 6.1. While major differences are clearly visible, the first peak only changes slightly which suggests convergence of the VG spectra and casts doubt on the LG spectra. On the other hand, the divergence at $\hbar\omega = 0$ gets worse with more bands, indicating that yet more bands are necessary.

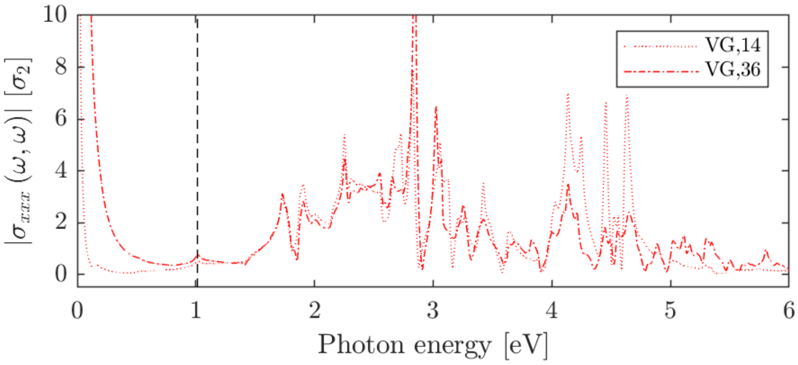
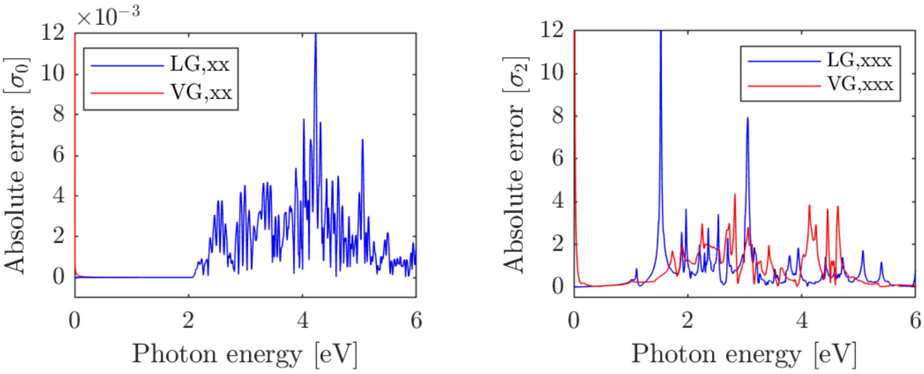


Figure 6.1: SHG spectra in the VG.

6.1.2 Effects of Parity Decoupling

Next, the effects of forcing the parity decoupling in the In_2Se_2 TB model shall be explored. We calculate the spectra in TB models with and without forced parity decoupling and then subtract the two spectra to quantitatively visualise the error generated by this decoupling. The results for the linear and SHG spectra is shown in Figure 6.2. For the linear spectra, the error is small with magnitudes roughly round 10^{-3}; however, the SHG spectra show large inconsistencies, especially in the LG. The actual LG spectra are plotted in Figure 6.3 to give a better look. Here, we see that the two spectra do follow each other somewhat closely except for a large peak in the parity coupled model at $\hbar\omega \approx 1.5$ eV. This may suggest that the unitary transformation which constructs basis functions with definite parity is on a form more complicated than the one presented in chapter 3. Alternatively, this could also suggest that the MLWFs from Wannier90 have not been constructed with high enough precision. Previously unbeknownst to us, Wannier90 has an option which allows for the construction of MLWFs with properties which correspond to the symmetry of the crystal. [60] Enabling this option may yield a better parity-decoupled TB model; however, due to time constraints this was not investigated.



(a) Linear spectra. (b) SHG spectra.

Figure 6.2: Absolute error between 14-band TB models with and without parity decoupling.

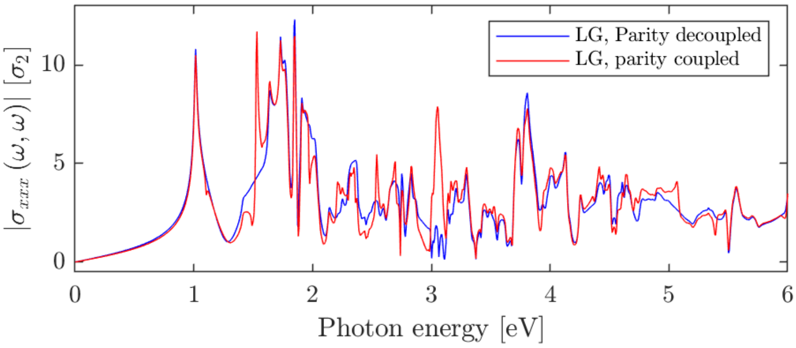
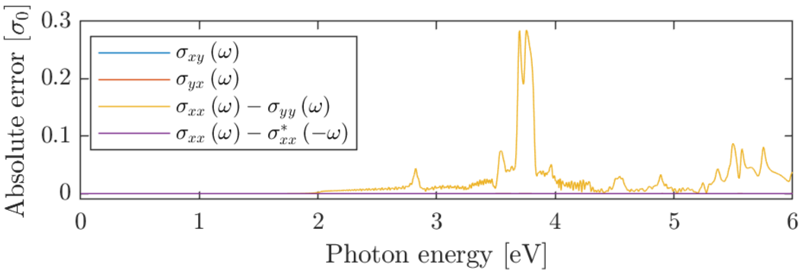


Figure 6.3: The LG spectra which yield the absolute error from Figure 6.2b.

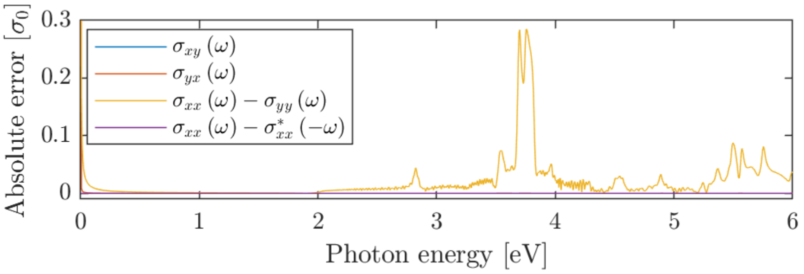
6.1.3 Equivalences of Tensor Elements

For In_2Se_2, it was argued from their point group that some tensor elements were equal and some were identically zero for all ω. First, we test the equivalences $\sigma_{xx}(\omega) = \sigma_{yy}(\omega) = \sigma^*_{xx}(-\omega)$ and $\sigma_{xy}(\omega) = \sigma_{yx}(\omega) = 0$. This is done for the In_2Se_2 LG and VG in Figure 6.4. The errors have generally the same shape in either gauge aside from the divergence in the VG. The largest error comes from the $\sigma_{xx}(\omega) = \sigma_{yy}(\omega)$ assertion while the other functions stay below at least 10^{-3}.

We also argued that $\sigma_{xxx}(\omega,\omega) = -\sigma_{xyy}(\omega,\omega) = -\sigma_{yxy}(\omega,\omega) = -\sigma_{yyx}(\omega,\omega) = \sigma^*_{xxx}(-\omega,-\omega)$ and $\sigma_{yyy}(\omega,\omega) = 0$ among others. This is tested using the VG in Figure 6.5. Here, the errors are on much larger scales except for the $\sigma_{yyy}(\omega,\omega) = 0$ and $\sigma_{xxx}(\omega,\omega) = \sigma^*_{xxx}(-\omega,-\omega)$ assertions. These have errors on the scales of 10^{-2} and 10^{-8}, respectively. The errors for the other assertions remain relatively small except for various peaks in the spectra.



(a) LG with 14 bands.



(b) VG with 14 bands.

Figure 6.4: Absolute error for the equivalences of linear response tensor elements.

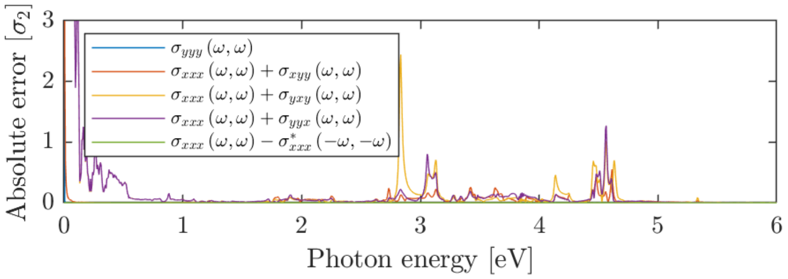


Figure 6.5: Absolute error for the equivalences of the 14-band SHG response tensor elements in the VG.

These relations do not hold as consistently in the LG. For example, when calculating $\sigma_{yyy}(\omega,\omega)$, we get Figure 6.6 which shows large discrepancies.

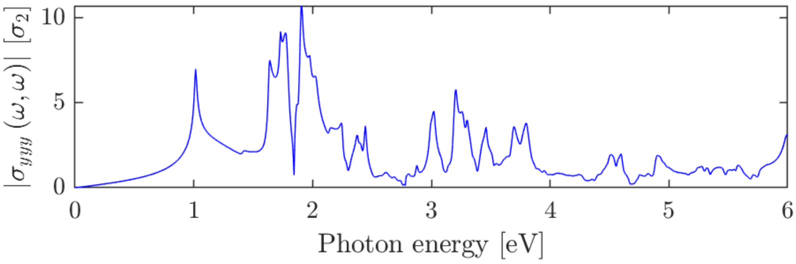


Figure 6.6: A second-order tensor element in the LG which should equal 0 for all ω.

6.1.4 Generalised Derivative Through Sum Rule

So far, the discussion has suggested that the second-order tensor elements in the LG is wrong. These large discrepancies are possibly caused either by a generally inaccurate TB model or by an error in the actual calculation. If there is an error, it most likely lies with the calculation of the generalised derivative through numerical differentiation and the Berry connection. These shall be discussed later, but for now, we consider calculating the generalised derivative through the sum rule (2.110) instead. The result is shown in Figure 6.7. Generally, this new LG spectrum still shows many differences from the VG spectrum; however, their amplitudes at the first step-up match. The disadvantage of the sum rule is that it does introduce an additional convergence issue, as the sum rule expression also requires a complete basis to be exact. However, by using the sum rule for the generalised derivative, the error associated with $\sigma_{yyy}(\omega,\omega)$ becomes much smaller which is shown in Figure 6.8

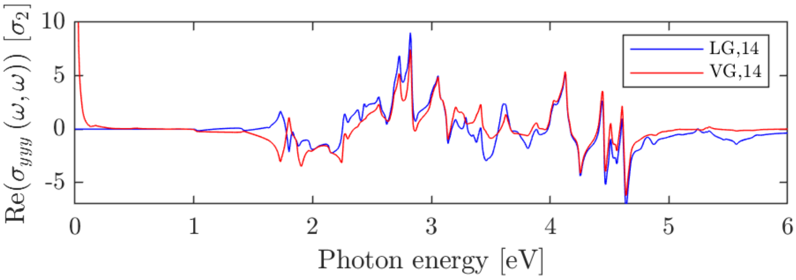


Figure 6.7: LG spectrum computed with the sum rule generalised derivative compared with the usual 14-band VG spectrum.

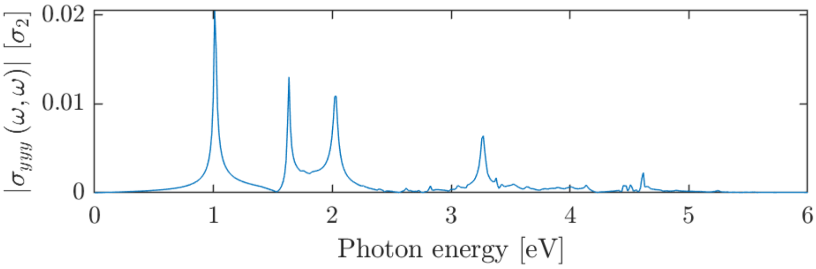


Figure 6.8: LG spectrum of $\sigma_{yyy}(\omega,\omega)$ computed with the sum rule generalised derivative.

6.2 Comparing Results For Different Crystals

So far, only the In_2Se_2 spectra have been discussed, and we have omitted the comparison of the tensor elements in the second-order length gauge results. These can be seen in Figure 6.9. In the LG result, one can still see the common features of the components, but both sign and size are off. This may be related to the troubles previously described with the non-Hermitian Berry connection.

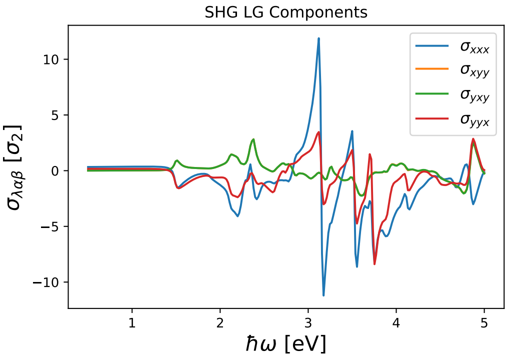


Figure 6.9: Real part of the SHG LG spectrum for the 14-band model of In_2Se_2. ~ 65000 **k**-points and $\hbar\eta = 50$ meV.

Alternatively, the least provoking result we obtained, namely the LG conductivity of hBN (same point-group as the other materials) is seen in Figure 6.10. It is not perfect, but one clearly sees the relationship between the four tensor elements as their shape are in accordance. More importantly, this shows our LG calculations are able to produce results with some amount of sense to it.

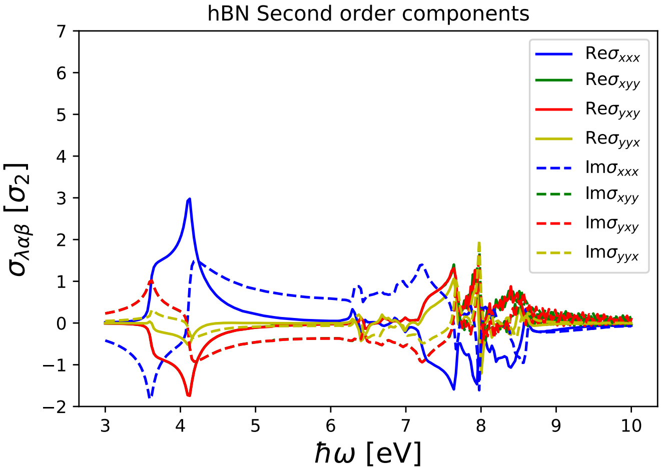
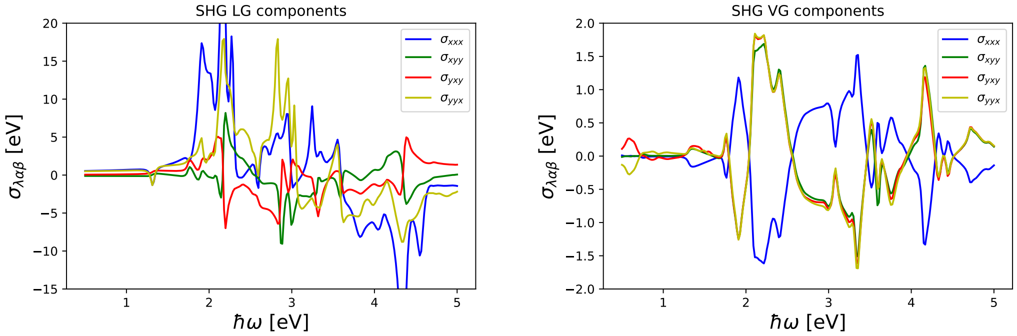


Figure 6.10: SHG spectrum of hexagonal boron-nitride with $E_g \sim 7.2$eV.[3] [32] Green line ontop of red line. ~ 65000 **k**-points

For MoS_2 the equivalent plots are seen in Figure 6.11 and the calculated LG components again violate the symmetry requirements of the system and to a worse degree than In_2Se_2.



(a) (b)

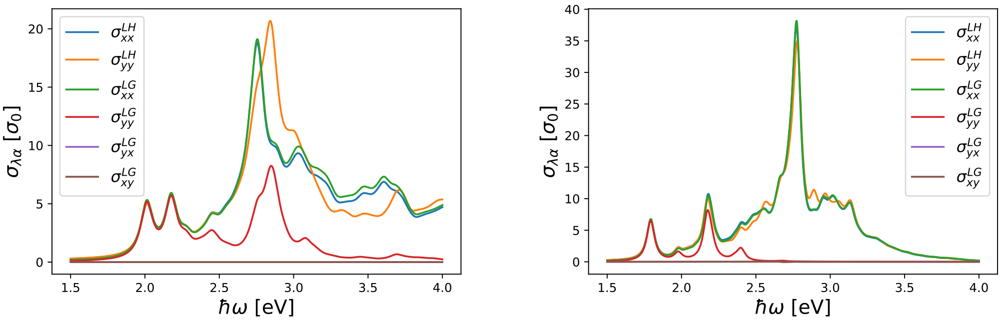
Figure 6.11: Real part of the SHG spectrum for MoS_2 using 11 band model. ~ 65000 **k**-points and $\hbar\eta = 50$ meV.

In the subsection 6.2.2, a possible cause for these disparities between the various tensor elements is given.

6.2.1 Excitonic Response

We now come to the tensorial symmetries which the conductivity tensor elements with excitonic effects should also satisfy. Starting with the first-order elements, we see in Figure 6.12a and Figure 6.12 the components of MoS_2 and In_2Se_2.

The agreement is reasonable and MoS_2 is like in the IPA case the one which fits most badly of the two. One does however note the $\sigma_{yy}^{(\mathrm{LG})}$ component, which die out much too fast, but if one instead seed the LH-procedure with $|P_y\rangle$ and use these approximate eigenstates to calculate σ_{yy}, one gets much better agreement. This is seen in Figure 6.13 for In_2Se_2 and also happens in the case of MoS_2. It does demonstrate what seems like discrimination between the x- and y-directions when using the approximate eigenstates obtained from the LH-procedure. A full diagonalisation, through not feasible in all cases, of the BSE Hamiltonian also fixes this issue.



(a) MoS_2 first-order components (b) In_2Se_2 first-order components

Figure 6.12: Real part of the first-order conductivity components

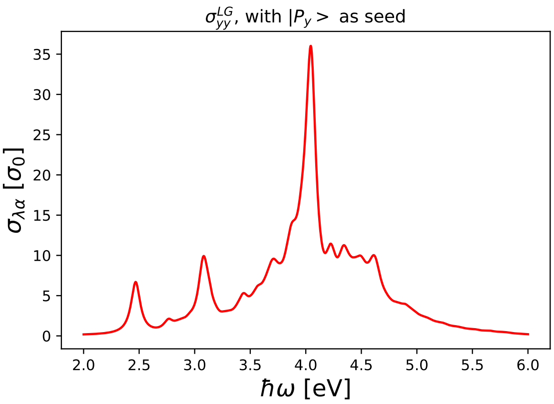
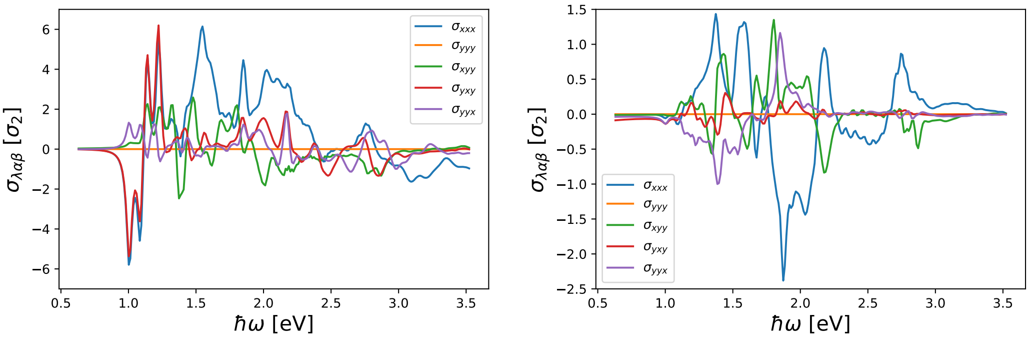
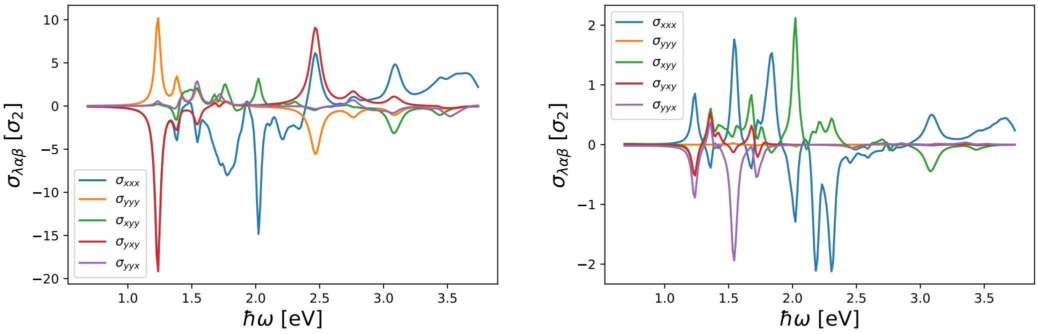


Figure 6.13: In_2Se_2 first-order yy-component on the conductivity, calculated using a truncated BSE Hamiltonian, where the LH procedure was instead seeded with the momentum in the y-direction.



(a) MoS_2 LG (b) MoS_2 VG



(c) In_2Se_2 LG (d) In_2Se_2 VG

Figure 6.14: Calculated using the regular phase-convention. 65 × 65 **k**-points and same bands as in the results section, except in MoS_2 where the negative parity bands are excluded.

We again check the second-order conductivity. The InSe and MoS_2 components in the LG and VG can be seen in Figure 6.14. The σ_{yyy} component should be zero in all cases, but the LG In_2Se_2 components do not agree with this while the MoS_2 obeys this. Strangely, in the case of In_2Se_2 the σ_{yyy} mirrors σ_{xxx} better than any of the components that actually should do this. The σ_{yxy}-component seems to equal σ_{xxx} instead of mirroring it in the first part of both LG spectra. This may be because of the generalised derivative which is dependent on the middle index and is the same for both components. Strangely while MoS_2 had its IPA LG σ_{yyy}-component be nonzero, the excitonic response surprisingly has this component be zero. Changing to the alternate phase-convention also makes the MoS_2 σ_{yyy} significantly different from zero, even through one may take from subsection 6.2.2 that this alternate phase convention is more consistent. However, in the light of problems with the LG calculation described in the coming subsection (subsection 6.2.2), this strange behaviour will come as no surprise.

One can argue the lack of other indecies than x on the σ_{xxx} element makes it simpler, and thereby less prone to error compared to the other components when considering the VG response. As the different tensor elements seem to mirror the σ_{xxx}-component in shifts and not each other, this is arguably a good sign for the VG σ_{xxx} result.

In the first-order response we also saw the strange underestimation of the yy-component when seeding with $|P_x\rangle$, which was fixed by instead seeding with $|P_y\rangle$ or a full diagonalisation. In MoS_2, using $|P_y\rangle$ as seed instead, reduces the numerical magnitude of the yyy component from $\sim 10^{-5}$ to $\sim 10^{-6}$, and for In_2Se_2 it does not produce any improvements, leaving the yyy-component close to unchanged. A full diagonalisation of the Hamiltonian does not change the outcome either. However, in the case of In_2Se_2 the spectrum changes for the xxx-element when using the $|P_y\rangle$ as seed, but also seems to converge much slower. During the first 400 iterations one will only see the two first peaks. This is seen in Figure 6.15 and even through using many more iterations than the bulk of the other SHG-spectra do not seem converged.

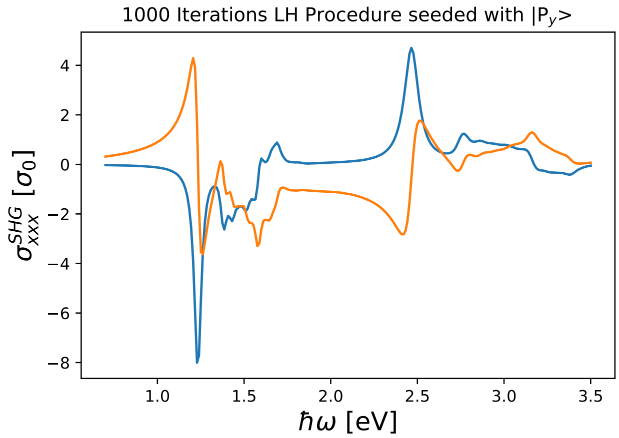


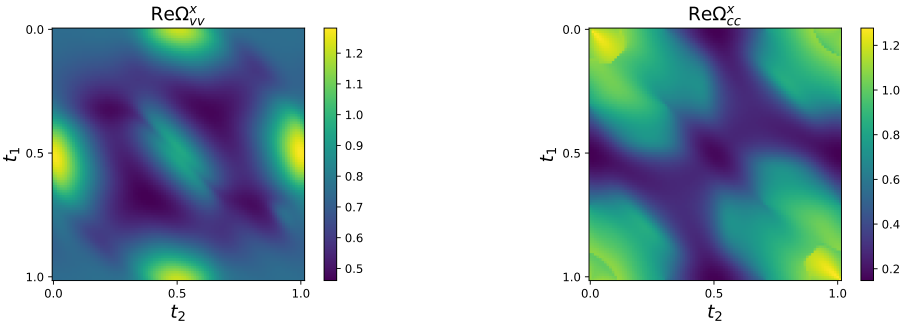
Figure 6.15: In_2Se_2 SHG xxx-component on the conductivity, calculated using a truncated BSE Hamiltonian, where the LH procedure was instead seeded with the momentum in the y-direction. 1000 Iterations used. Real part in blue, imaginary part in yellow.

For MoS_2 the spectrum does not change significantly, but does take notably more iterations to converge. The seed-vector should be arbitrary, except when calculating the first-order response using the recursive fraction, as they in the limit of iterations equal to the rank of the matrix just produces a unitary transformation. It does however seem to have an impact without full diagonalisation.

6.2.2 Difficulties in the Calculation of the Second-Order Response

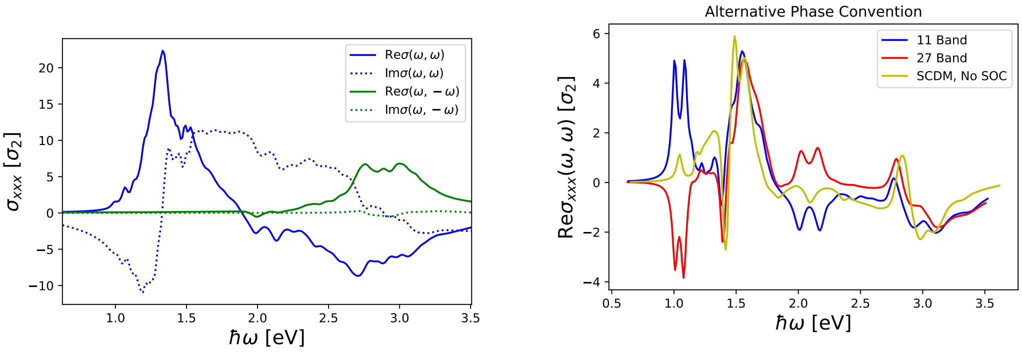
The calculated Berry connection used during the calculations of the SHG and OR had the unwanted feature of not being strictly Hermitian since many cases had nonzero imaginary part on the diagonal in the band index. It was however not dominating the magnitude of the Berry connection. Taking MoS_2 as an example, the Berry connection seen in Figure 4.8 was reproduced using a full 27-band TB model with *spd*-orbitals on all atoms. This is seen in Figure 6.16 and differs a bit in magnitude, form and is not as symmetric, but is comparable to Figure 4.8. This 27 band TB model is actually not of as good quality as the band gap at $\pm K$ varies by ~ 0.004 eV. Using this 27-band TB model, the first-order spectrum was reproduced, albeit with a somewhat larger magnitude, but most importantly with the same shape. The Berry connection was also reproduced using both norm-conserving and PAW pseudopotentials. Our confidence in the correctness hereof is therefore good.

The second-order conductivity in the LG was however troublesome to reproduce, and while the first-order spectrum does not depend on which phase convention (first entrance of eigenvector real or an alternative, elaborated on soon) we choose in the initial TB-calculation, the second-order response does however change when going from convention to convention.



(a) Top valence band Berry connection (b) Lowest conduction band Berry connection

Figure 6.16: The Berry connection of MoS_2 obtained with a 27 band TB model. Magnitude of imaginary part is again $\sim \frac{1}{50}$ the magnitude of the real part.



(a) 27 band *spd* TB model. Regular phase convention.

(b) Three different models in the alternative phase convention. See Figure E.5d for regular phase on SCDM-**k** TB model.

Figure 6.17: Top valence band and bottom conduction band used, with SOC included in 11 and 27 band models.

The BSE matrix elements containing the Bloch overlaps $I_{c\mathbf{k},c'\mathbf{k}'}$ are gauge-dependent[30] which we have also stated in subsection 2.6.1, making a choice of phase convention necessary. We also have previously evaluated a **k**-gradient of the momentum matrix elements where the phase had to be fixed from **k**-point to **k**-point. Here we fixed the phase by taking the first non-zero element in each eigenvector and multiplied a phase onto it, to make it real and positive. This phase convention was carried over to the exciton calculations. An alternate convention is also possible, by choosing to make the sum of the eigenvector entrances real.[30] These two, seemingly, equally valid ways to fix the phase does however produce different results when it comes to the excitonic SHG and OR spectra. In Figure 6.17 SHG and OR are plotted using the two phase conventions, for both TB-models using the top valence band and bottom conduction band of MoS_2. An additional TB model is also presented using the SCDM-**k** method, which is elaborated on in Appendix E.

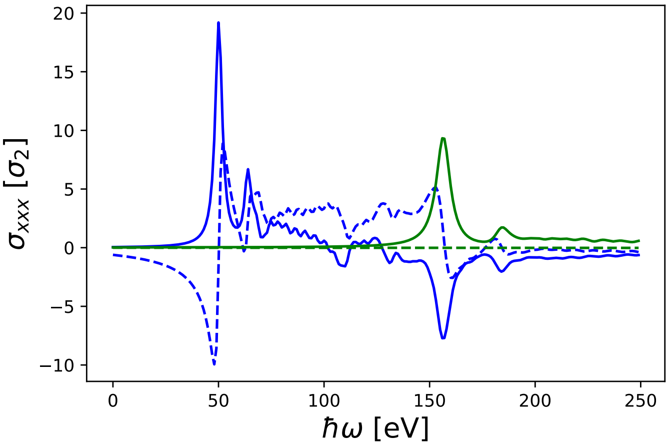


Figure 6.18: Second-order conductivity of the NN block, approximately reproducing Figure 3.25a in shape but has larger magnitude and differs by a sign. Alternate phase convention and SCDM-**k** TB model. 70 × 70 **k**-points.

We do not know why it is so, but the convention involving the sum of entrances is what makes the three models roughly agree on the SHG and OR spectra. Importantly, it also significantly diminishes the peaks corresponding to the $2s$-like states in spectra. The SCDM-**k** does however have smaller magnitude on the $1s$ exciton resonance. On the other hand using the phase convention used throughout the calculations of this thesis does not produce similar spectra from the two models. Changing the phase convention for the In_2Se_2 calculations from chapter 3 does change the spectra appreciably in size, but the shape is similar. The NN-block spectrum is reproduced in Figure 6.18 with a SCDM-**k** TB model.

The reason for this phase convention-dependence may come down to the evaluation of $\nabla_{\mathbf{k}} A_{vc\mathbf{k}}$, which is evaluated numerically using a central difference scheme and the associated problems of discontinuities in the expansion coefficients in **k**-space. The **k** derivative in the IPA second-order response had the generalised derivative evaluated using direct finite difference scheme and this turned out to be in accordance with the excitonic spectrum in the infinite screening limit. One can also take a more elaborate approach, by taking the product rule and applying it to the momentum matrix elements, yielding

$$\nabla_{\mathbf{k}}[p]^{EB} = (\nabla_{\mathbf{k}}[V]^{\dagger})[p]^{W}[V] + [V]^{\dagger}(\nabla_{\mathbf{k}}[p]^{W})[V] + [V]^{\dagger}[p]^{W}(\nabla_{\mathbf{k}}[V]), \tag{6.1}$$

where $[V]$ is the matrix containing the eigenvectors, $[p]^W$ is the momentum matrix elements in the Wannier basis and $[p]^{EB}$ is the momentum matrix elements in the basis of eigenvectors. The middle part can be evaluated analytically, and the eigenvector derivative can then be calculated by making the inner product of an eigenvector at **k** with an eigenvector at $\mathbf{k}+\delta\mathbf{k}$ real. This produces a different spectrum seen in Figure 6.19, which is not consistent with the excitonic result in the limit of large screening, but is much more in line with the VG result using 27 bands in Figure 4.7b.

The choice of phase have here been done locally so that the phase matches for evaluation of the derivative, instead of the more global prescription of a real first entry chosen at every **k**-point.

However, this notion cannot be extended to the evaluation of $\nabla_{\mathbf{k}}$ in the case of the generalised derivative of the expansion coefficients for the excitonic response. This does suggest a better way to evaluate $\nabla_{\mathbf{k}} A_{vc\mathbf{k}}$ is needed because the two phase conventions both gives the questionable result also obtained with the finite difference approach on the momentum matrix element in the limit of infinite screening. The VG on the other hand does have the significant advantage of not having to deal with a generalised derivative, and does resemble the VG 27 band IPA spectrum in the limit of infinite screening. The equivalent of Figure 6.19 for In_2Se_2 is seen in Figure 6.20 and is also much more in line with the VG result Figure 6.1,except for a sign. This does provide basis to disregard the excitonic LG calculations done throughout this report. The fact that the LG was much too large when comparing with experiment in chapter 5 also stresses this. A calculation on hexagonal boron-nitride seen in Figure E.1d does strangely enough resemble a similar calculation found in Taghizadeh and Pedersen[32] in shape but with roughly double the magnitude.

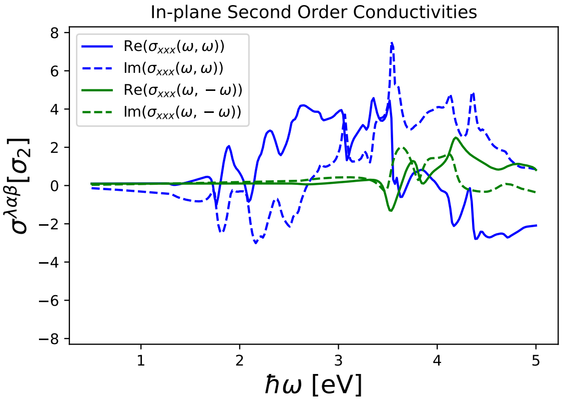


Figure 6.19: MoS_2 11 band-second-order response, evaluating $\nabla_{\mathbf{k}}[p]$ as in Equation 6.1.

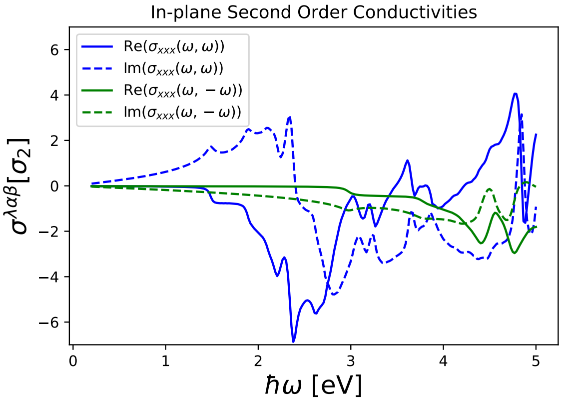


Figure 6.20: In_2Se_2 14 band second-order response, evaluating $\nabla_{\mathbf{k}}[p]$ as in Equation 6.1.

Convergence with respect to the inclusion of conduction and valence bands has, however, not been demonstrated in the case of the excitonic VG spectra. Careful parity sorting and a good TB model with many bands is a necessity for this to work.

6.3 Automatisation of Wannierisation

The DFT calculation one is able to do effortlessly when one has the structure of the material from C2DB, and Quantum ESPRESSO can choose a suitable number of bands for the ground state calculation, limiting the need for input from our side. Quantum ESPRESSO with Spglib[25] furthermore give the crystal symmetries in matrix form.

The Wannierisation was however a bit more tricky as the correct energy windows and projectors would have to be specified. One can brute force the process of choosing these projectors, but sorting the eigenstates in parity by simply looking at which entrances are zero and nonzero after applying a transformation is one of the things that become exceedingly tricky if doing so. This is because the unitary transformation becomes harder to find. Another approach for this could have been to consider the momentum matrix elements instead of the eigenvectors and from the size of these deduce which states couple to each other. This is; however, not as direct as looking at the eigenstates of the Hamiltonian and may be lead to erroneous conclusions. In order to get any significant number of **k**-points for the excitonic second-order spectra, decoupling into parity states seems a necessity. This process of determining the parity of a state would have been much easier in a plane wave basis.

A more targeted way to pick the projections is to instead look at projected density of states (PDOS) from each atomic *s*-, *p*- and *d*-orbital from each atom, which is elaborated in Appendix E. However, even with a more elaborate way of automating the projections, the success-rate for producing a good TB model is not high, at least with our implementation, and the code we wrote needed manual input for many cases.

An alternative to automating the choice of projectors was the SCDM-**k** method,[18] [61] elaborated on in Appendix E, but it has the promising feature of only needing two parameters which can be calculated from the DFT output, and the number of bands one needs. This method is smarter than the brute force method mentioned previously, but not much can be said for the unitary transformation needed to decouple the Hamiltonian into a basis of functions with definite parity as is the case when one chooses well-suiting atomic orbital projections. If the LG second-order method could be made viable, this method could be used for quick and reliable Wannierisation for high throughput calculation of second-order spectra.

6.4 Avoiding the Slow-Rapid Approximation

The matrix element with singly excited states the Coulomb and exchange matrix elements are given in (2.130) and (2.131), which are integrals involving the eigenstates. In terms of the form given in (2.7), it is shown in Appendix F that the terms can be approximated as

$$\left\langle v c\mathbf{k}\right| W_C \left|v' c' \mathbf{k}'\right\rangle = \frac{A_{\mathrm{UC}}}{4\pi^2} \mathrm{d}^2\mathbf{k} \sum_{ijkl} c_i^{c*} c_j^{c'} c_k^{v} c_l^{v'*} \sum_{\mathbf{R}} \mathrm{e}^{\mathrm{i}(\mathbf{k}-\mathbf{k}')\cdot\mathbf{R}} W_{ijkl,\mathbf{R}}, \tag{6.2}$$

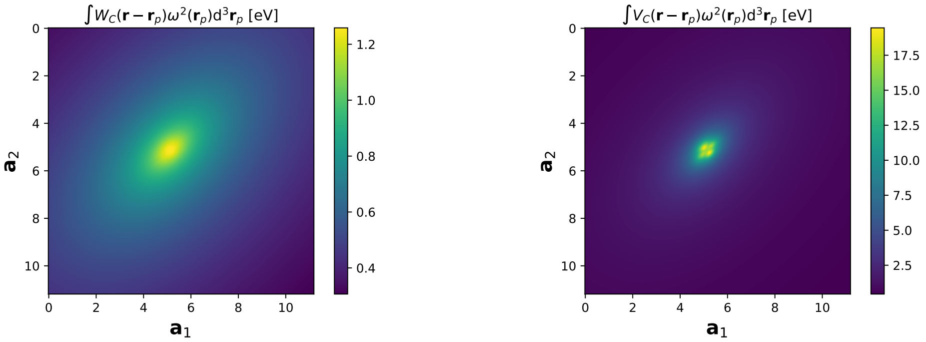
and in the case of the exchange interaction the phases instead cancel as a result of the condition that $\mathbf{R} = \mathbf{R}''$ and $\mathbf{R}' = \mathbf{R}'''$, yielding

$$\left\langle v c\mathbf{k}\right| V_X \left|v' c' \mathbf{k}'\right\rangle = \frac{A_{\mathrm{UC}}}{4\pi^2} \mathrm{d}^2\mathbf{k} \sum_{ijkl} c_i^{c*} c_j^{c'} c_k^{v} c_l^{v'*} \sum_{\mathbf{R}} V_{ikjl,\mathbf{R}}. \tag{6.3}$$

In these expressions, terms involving Wannier functions of the same coordinates on different sites have been ignored. $W_{ijkl,\mathbf{R}}$ is given as

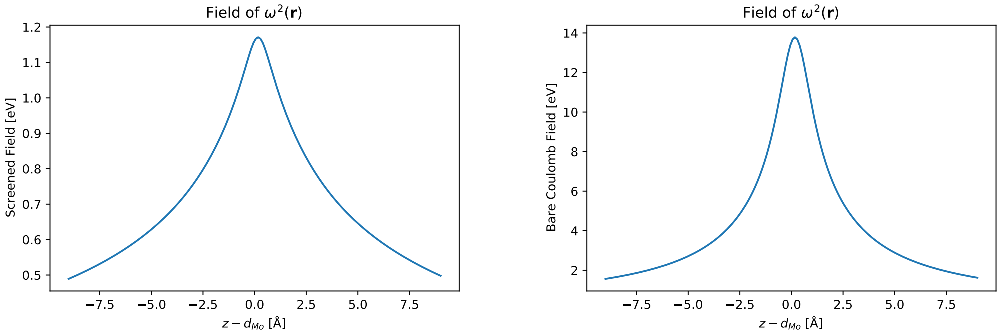
$$W_{ijkl,\mathbf{R}} = \int\int \omega_{ij}(\mathbf{r}_1 - \mathbf{R}))W(\mathbf{r}_1 - \mathbf{r}_2)\omega_{kl}(\mathbf{r}_2)\mathrm{d}^3\mathbf{r}_1\mathrm{d}^3\mathbf{r}_2 \quad \text{with} \quad \omega_{ij}(\mathbf{r}) = \omega_i(\mathbf{r})\omega_j(\mathbf{r}) \tag{6.4}$$

and similarly for $V_{ikjl,\mathbf{R}}$. These expressions for the matrix elements are actually periodic in **k**, contrary to the slow-rapid approximation employed throughout the thesis. For the evaluation of these integrals, exploiting the fact that they contain a convolution between the screened potential W and the product ω_{ij}. Using a FFT-algorithm to convolve $W(\mathbf{r})$ with ω_{kl}, with W calculated using the expression (D.22), and the product is fairly quick and only needs to be done $\frac{N(N+1)}{2}$ times. We tried to calculate the matrix elements using the above formulas and the Wannier functions from Wannier90 but the matrix elements did not seem to have been converged, even after we included 11×11 units cells, above which computer RAM becomes a problem. In Figure 6.21 the screened and bare fields of a MoS$_2$ d-orbital is seen plotted in the plane containing the Mo-atom. While this approach is more computationally demanding, it will also include the $P_{ij} = -1$ parts of the Hamiltonian from (2.132).



(a) d-orbital convolved with the screened Coulomb interaction

(b) d-orbital convolved with the bare Coulomb interaction



(c)

(d)

Figure 6.21: a) and b): Mo d-orbital convolved with the screened and bare Coulomb interaction over 11×11 unit cells. Plotted as $z = d_{\mathrm{Mo}}$ and with the real space lattice vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ on the axes. Obtained with $27 \times 27 \times 160$ points per unit cell. c) and d): Field in z-direction roughly at the location of the Mo atom. $r_0 = 42.3$ Å.

As can be seen, the screened Coulomb potential is much more slowly changing, as the slow-rapid approximation required. Contrary to the bare Coulomb integral, the screened one does not show any features of the underlying d-orbital, which the bare Coulomb potential does show features of. If one were to consider the screened potential as constant over each unit cell, one would end up with $W_{ijkl,\mathbf{R}} = \delta_{ij}\delta_{kl}W_C(\mathbf{R}, z = 0)$. Interpreting the unit cell area A_{UC} as a dA one can take the $\mathbf{R}$-sum of (6.2) as a Fourier transform yielding (2.134). This approximation also removes the exchange interaction by the same argument as in subsection 2.6.1. We thus explicitly see the reason for the success of the slow-rapid approximation here when considering the Coulomb term only.

6.5 The Dielectric Function

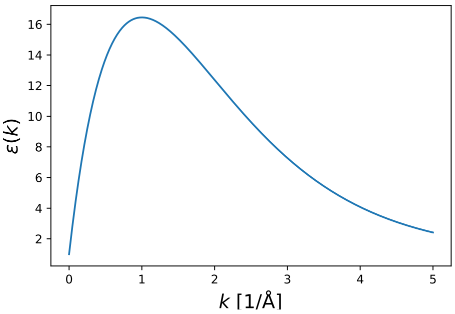


Figure 6.22: Slope close to $k = 0$ is $r_0 = 42$Å, but also has an exponential decay, only reaching its peak in $k = 1\,\frac{1}{\text{Å}}$. $a = 1$ Å.

The dielectric function of the form $\epsilon(k) = \kappa + r_0$ used throughout the calculations of the excitonic properties is actually only a first order approximation for small k, but because the coupling in the BSE is only significant for small k, this is still a good approximation. A more accurate modelling of the dielectric function of a 2D material actually has the screening become negligible when $k \to \infty$,[62] and can be written on the form[14]

$$\varepsilon(k) = 1 + e^{-ak} r_0 k \tag{6.5}$$

for a freestanding monolayer. This instead gives a Coulomb interaction that decays as $\frac{1}{k}$ at large k. Doing a calculation with the dielectric function with the limiting behaviour $\epsilon \sim \kappa + r_0 k$ as $k \to 0$, but also decaying to one with $k \to \infty$ can be seen in Figure 6.22. A comparison between the exiton energies of the usual Rytova-Keldysh dielectric function and the decaying dielectric function is seen in Figure 6.23 and are evidently close to each other. For many monolayers with appreciable band gap the the peak value of dielectric function happens roughly at $k \approx 0.5\frac{1}{\text{Å}}$. [63].

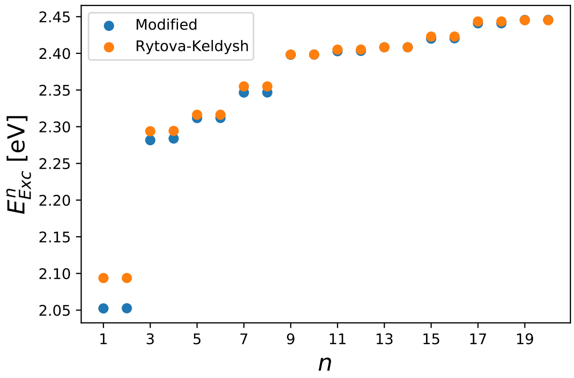


Figure 6.23: Excitonic energies of MoS_2 without SOC obtained using Rytova-Keldish dielectric function and modified decaying dielectric function from Figure 6.22. $E_g \approx$ 2.62eV.

The dielectric function in the Rytova-Keldysh form, while it is an approximation, is still very versatile as it can account for a lot of different dielectric environments through κ, not just vacuum. In any experiment on a monolayer, one will have at least a substrate on which the monolayer is located which is readily modelled. A modification to doped materials also exist.[64] [14] Anisotropic screenings as the one found in monolayer phosphorous can also be modelled.[65] A BSE calculation of the conductivity tensors can be found in Figure E.4 and clearly shows the anisotropic response. Here r_0 has been made into a diagonal matrix so that it treats k_x and k_y differently.

6.6 Adding More Reciprocal Lattice Vectors to the Coulomb and Exchange Terms

The approximation made in (2.134) neglected a sum over reciprocal lattice vectors. For materials exemplified by MoS_2, where the excitons are largely localised to the K-points, the effects of points outside the Brillouin zone can be neglegted while still retaining accuracy. hBN, as an example has much less localisation when one inspects the lowest eigenstates. A more correct of (2.134) is given as [32]

$$V^{C}_{abcd} = \sum_{\mathbf{G}} I_{a\mathbf{k},c\mathbf{k}'}(\mathbf{G}) I_{b\mathbf{k}',d\mathbf{k}}(-\mathbf{G}) W_C(\mathbf{k}-\mathbf{k}'-\mathbf{G}) \tag{6.6}$$

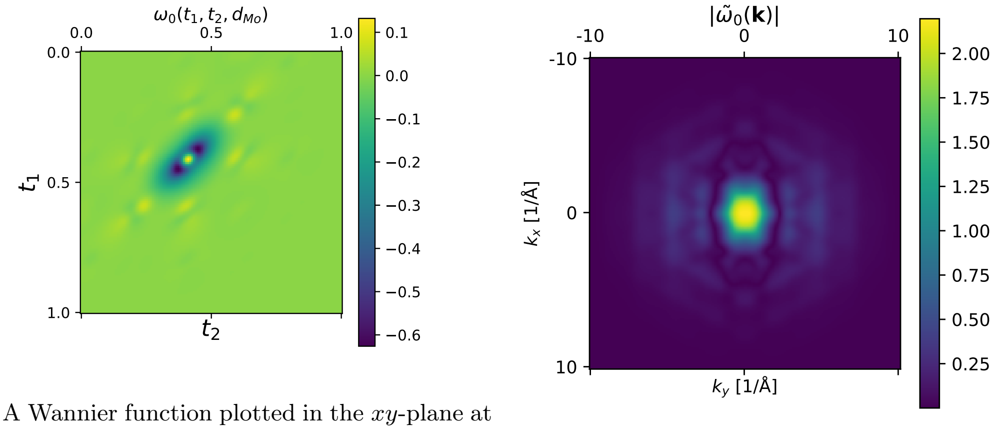
where $I_{a\mathbf{k},c\mathbf{k}'}(\mathbf{G}) = \sum_{\mathbf{G}} C^{*}_{a,\mathbf{k}}(\mathbf{G}+\mathbf{G}')C_{c\mathbf{k}'}(\mathbf{G}')$, with $C_{a\mathbf{k}}(\mathbf{G})$ being the expansion coeffiecients in an empirical pseudpotential type wave function[66] on the form $\psi_{a\mathbf{k}} = \frac{1}{\sqrt{A}}\sum_{\mathbf{G}} C_{a,\mathbf{k}}(\mathbf{G},z)\mathrm{e}^{\mathrm{i}(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}}$. Using this form and the orthogonality of the exponentials, one can obtain the expansion coefficients as

$$C_{a\mathbf{k}}(\mathbf{G},z) = \int \frac{\mathrm{e}^{-\mathrm{i}(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}}}{\sqrt{A}} \psi_{a\mathbf{k}}(\mathbf{r})\mathrm{d}^2\mathbf{r}. \tag{6.7}$$

On the other hand, using the form of $\psi_{a\mathbf{k}}$ from (2.7) yields

$$\begin{aligned} C_{a\mathbf{k}}(\mathbf{G},z) &= \frac{1}{\sqrt{AN}}\sum_i c_i^{a\mathbf{k}} \sum_{\mathbf{R}} \mathrm{e}^{\mathrm{i}\mathbf{R}\cdot\mathbf{k}} \int \omega_i(\mathbf{r}-\mathbf{R})\mathrm{e}^{-\mathrm{i}(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}}\mathrm{d}^2\mathbf{r} && (6.8)\\ &= \frac{1}{\sqrt{A_{UC}}}\sum_i c_i^{a\mathbf{k}}\tilde{\omega}_i(\mathbf{G}+\mathbf{k},z) && (6.9) \end{aligned}$$

when using the Fourier shift theorem and $\mathbf{G}\cdot\mathbf{R} = 2n\pi$. This means the coefficients $C_{a\mathbf{k}}$ can be obtained from the Wannier functions in principle, by simple Fourier transformation in the xy-plane. Using a d-orbital from MoS_2 as an example one can do the Fourier transform in the Mo-plane and obtain the Fourier spectrum seen in Figure 6.24.



(a) A Wannier function plotted in the xy-plane at the Mo-atom. 5×5 unit cells. $\omega_0(t_1, t_2, d_{Mo}) = \omega_0(5t_1\mathbf{a}_1 + 5t_2\mathbf{a}_2 + \hat{\mathbf{z}}d_{Mo})$

(b) Fourier transform of Figure 6.24a

Figure 6.24: Calculation of the Fourier transform of a Wannier function in a plane.

The downside of this is one needs to store all the z-dependency for each $\mathbf{k}$-point, giving a lot of data. One could choose a basis of Gaussians with a polynomial multiplied onto it to shrink the data needed to be stored, while also keeping the different parity states decoupled.
This approach would enable to examine materials where the excitonic states are not as localised as in the cases we have examined.

7 | Conclusion

The goal of this thesis was to study the optical properties monolayer materials such as $\mathrm{In_2Se_2}$ and $\mathrm{MoS_2}$ with two different gauges both within the IPA and with the inclusion of excitonic effects. Furthermore, the ability to apply the techniques used on these two materials to the wider class of 2D materials was also a goal.

The theory section started off by briefly introducing density function theory and the Kohn-Sham equations along with Wannier functions. Afterwards the linear and nonlinear IPA response functions were derived rigorously, starting from density matrix theory which was then applied to find the current density observable in terms of a conductivity tensor and the applied time-harmonic electric field perturbation. This was done using two approaches: the velocity gauge and the length gauge. Additionally, Neumann's principle was also derived which allows for quick simplification of tensors in crystals based on the crystal symmetries.

The inclusion of excitonic effects was then elaborated on, first by introducing the Bethe-Salpeter equation and the screened Coulomb interaction through the Rytova-Keldysh potential. We then elaborated on how to calculate the response, circumventing direct diagonalisation of the Hamiltonian through the Lanczos-Haydock procedure, and eventually reaching the nonlinear response functions which include these excitonic effects.

The application of the laid out theory was then mainly carried out on $\mathrm{In_2Se_2}$ and $\mathrm{MoS_2}$, combining DFT and the Wannierisation procedure to obtain an ab initio tight-binding model, giving access to a significantly enhanced **k**-point resolution required for the evaluation of the nonlinear conductivity tensor of the materials. The conductivity tensors were simplified greatly by Neumann's principle due to presence of y-inversion, z-inversion, and 120°-rotation symmetries. The linear response spectra went smoothly except for the velocity gauge spectra which incorrectly diverge for small frequencies. The velocity gauge is unstable and more sensitive to the effects of basis truncation which is necessary for a tight-binding model, and as such the length gauge is generally preferable. A correction was established for the velocity gauge spectra, but this correction only mitigated the divergence to a certain degree. For the second-order spectra, the velocity gauge calculations went smoothly, again aside from divergences inherent to this gauge; however, a problem was encountered in the length gauge. This problem was identified as related to the phase-convention and the **k**-gradient needed for the generalised derivative and meant our length gauge calculations were regarded as faulty.

The tight-binding models developed was also used to set up the Bethe-Salpeter equation, through the slow-rapid approximation. In $\mathrm{MoS_2}$, the excitonic binding energy was determined to good accuracy, while only the direct band gap negative parity bands of $\mathrm{In_2Se_2}$ were fit for the approximations employed in this thesis. Having obtained the excitonic states, we calculated the first-order conductivity in both cases, with limiting behaviour consistent with the IPA calculations. The second-order conductivity was also calculated in both the conventional velocity gauge and the length gauge.

The length gauge calculation for MoS_2 did in the limit of large screening approach a result which had much resemblance to the faulty length gauge calculation and was significantly different from the velocity gauge calculation with many bands. We have therefore regarded the excitonic length gauge calculations as wrong, with the evaluation of the generalised derivative being the problem. The excitonic velocity gauge spectrum was not demonstrated to be converged with respect to the number of bands in neither MoS_2 or In_2Se_2, but its limiting behaviour was consistent with the IPA velocity gauge calculations. A calculation with more bands is however needed.

Comparison with experiments in the literature was done and showed in the majority of cases that the velocity gauge calculation was matched the most with experiments. While not matching the shape of the observed frequency resolved susceptibilities, the overall size was shown to be in line with the velocity gauge. The length gauge was much too large in three out of four cases. Incidentally, a reported large susceptibility in $MoTe_2$ was also calculated in the velocity gauge, although not matching the peaks in frequency.

In the discussion, we have also suggested further ways to use the Wannier functions which can be obtained from Wannier90. This may extend the ab initio tight-binding approach beyond the slow-rapid approximation to also include a sum over reciprocal lattice vectors.

The automatisation of the generation of tight-binding models only partially successful and employed both the SCDM-**k**-method and atomic orbitals. Many tight-binding models where however obtained, ready for calculating IPA spectra. For the excitonic calculations however, only a few tight-binding models which had the wanted direct band gap was obtained. The automatic calculation of which tensor elements to calculate, given the crystal structure was successful.

Bibliography

[1] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov. Electric field effect in atomically thin carbon films. *Science*, 2004.

[2] A. V. Kolobov and J. Tominaga. *Two-dimensional transition-metal dichalcogenides.* Springer, 2016.

[3] S. Haastrup, M. Strange, M. Pandey, T. Deilmann, P. S. Schmidt, N. F. Hinsche, M. N. Gjerding, D. Torelli, P. M. Larsen, A. C. Riis-Jensen, et al. The Computational 2D Materials Database: high-throughput modeling and discovery of atomically thin crystals. *2D Materials*, 2018.

[4] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz. Atomically thin MoS2: a new direct-gap semiconductor. *Physical review letters*, 2010.

[5] L. Li, Y. Yu, G. J. Ye, Q. Ge, X. Ou, H. Wu, D. Feng, X. H. Chen, and Y. Zhang. Black phosphorus field-effect transistors. *Nature nanotechnology*, 2014.

[6] L. Ci, L. Song, C. Jin, D. Jariwala, D. Wu, Y. Li, A. Srivastava, Z. F. Wang, K. Storr, L. Balicas, et al. Atomic layers of hybridized boron nitride and graphene domains. *Nature materials*, 2010.

[7] R. O. Jones. Density functional theory: Its origins, rise to prominence, and future. *Reviews of modern physics*, 2015.

[8] B. Scharf, D. van Tuan, I. Žutić, and H. Dery. Dynamical screening of excitons in monolayer transition-metal dichalcogenides. *Journal of Physics: Condensed Matter*, 2019.

[9] R. W. Boyd. *Nonlinear optics.* Academic Press, Inc., third edition, 2008.

[10] J. W. You, S. R. Bongu, Q. Bao, and N. C. Panoiu. Nonlinear optical properties and applications of 2D materials: theoretical and experimental aspects. *Nanophotonics*, 2018.

[11] N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt. Maximally localized Wannier functions: Theory and applications. *Reviews of Modern Physics*, 2012.

[12] J. P. Dahl. *Introduction to the quantum world of atoms and molecules.* World Scientific Publishing Company, 2001.

[13] F. Hüser, T. Olsen, and K. S. Thygesen. Quasiparticle GW calculations for solids, molecules, and two-dimensional materials. *Physical Review B*, 2013.

[14] T. G. Pedersen. *Electric, optical & magnetic properties of nanostructures.* Aalborg University, 2019.

[15] A. A. Mostofi, J. R. Yates, G. Pizzi, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari. An updated version of Wannier90: A tool for obtaining maximally-localised Wannier functions. *Computer Physics Communications*, 2014.

[16] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, et al. Quantum ESPRESSO: A modular and open-source software project for quantum simulations of materials. *Journal of Physics: Condensed Matter*, 2009.

[17] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, et al. Advanced capabilities for materials modelling with Quantum ESPRESSO. *Journal of Physics: Condensed Matter*, 2017.

[18] A. Damle, L. Lin, and L. Ying. SCDM-k: Localized orbitals for solids via selected columns of the density matrix. *Journal of Computational Physics*, 2017.

[19] T. Yusufaly, D. Vanderbilt, and S. Coh. Tight-Binding Formalism in the Context of the PythTB Package. `https://www.physics.rutgers.edu/pythtb/formalism.html`, 2020. [Online; accessed 13-may-2020].

[20] T. G. Pedersen, K. Pedersen, and T. B. Kriestensen. Optical matrix elements in tight-binding calculations. *Physical Review B*, 2001.

[21] R. Del Sole and R. Girlanda. Optical properties of semiconductors within the independent-quasiparticle approximation. *Physical Review B*, 1993.

[22] A. Taghizadeh, F. Hipólito, and T. G. Pedersen. Linear and nonlinear optical response of crystals using length and velocity gauges: Effect of basis truncation. *Physical Review B*, 2017.

[23] J. H. Heinbockel. *Introduction to tensor calculus and continuum mechanics.* Trafford Victoria, Canada, 2001.

[24] R. F. Tinder. *Tensor properties of solids: Phenomenological development of the tensor properties of crystals.* Morgan & Claypool Publishers, 2008.

[25] A. Togo and I. Tanaka. Spglib: a software library for crystal symmetry search. *arXiv preprint arXiv:1808.01590*, 2018.

[26] V. Y. Irkhin, V. Y. Irkhin, and Y. P. Irkhin. *Electronic structure, correlation effects and physical properties of d-and f-metals and their compounds.* Cambridge Int Science Publishing, 2007.

[27] C. Aversa and J. E. Sipe. Nonlinear optical susceptibilities of semiconductors: Results with a length-gauge analysis. *Physical Review B*, 1995.

[28] M. L. Trolle, G. Seifert, and T. G. Pedersen. Theory of excitonic second-harmonic generation in monolayer MoS2. *Physical Review B*, 2014.

[29] E. Ridolfi, C. H. Lewenkopf, and V. M. Pereira. Excitonic structure of the optical conductivity in MoS2 monolayers. *Physical Review B*, 2018.

[30] M. Rohlfing and S. G. Louie. Electron-hole excitations and optical spectra from first principles. *Physical Review B*, 2000.

[31] J. Have, N. M. R. Peres, and T. G. Pedersen. Excitonic magneto-optics in monolayer transition metal dichalcogenides: From nanoribbons to two-dimensional response. *Physical Review B*, 2019.

[32] A. Taghizadeh and T. G. Pedersen. Gauge invariance of excitonic linear and nonlinear optical response. *Physical Review B*, 2018.

[33] T. Olsen, S. Latini, F. Rasmussen, and K. S. Thygesen. Simple screened hydrogen model of excitons in two-dimensional materials. *Physical review letters*, 2016.

[34] A. B. Lorentzen, A. Ørsted, and M. Ovesen. *Two-dimensional hydrogen in a magnetic field: Non-relativistic and relativistic cases.* Aalborg University, 2019.

[35] I. S. Gradshteyn and I. M. Ryzhik. *Table of integrals, series, and products.* Academic Press, 2014.

[36] J. Have and T. G. Pedersen. Magnetoexcitons and Faraday rotation in single-walled carbon nanotubes and graphene nanoribbons. *Physical Review B*, 2018.

[37] Y. Gillet, M. Giantomassi, and X. Gonze. Efficient on-the-fly interpolation technique for Bethe–Salpeter calculations of optical spectra. *Computer Physics Communications*, 2016.

[38] Roger Haydock. The recursive solution of the Schrödinger equation. *Computer Physics Communications*, 1980.

[39] D. Ma, W. Ju, Y. Tang, and Y. Chen. First-principles study of the small molecule adsorption on the InSe monolayer. *Applied Surface Science*, 2017.

[40] P. E. Blöchl. Projector augmented-wave method. *Physical review B*, 1994.

[41] A. Dal Corso. Pseudopotentials periodic table: From H to Pu. *Computational Materials Science*, 2014.

[42] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Physical review letters*, 1996.

[43] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke. Restoring the density-gradient expansion for exchange in solids and surfaces. *Physical review letters*, 2008.

[44] G.-X. Zhang, A. M. Reilly, A. Tkatchenko, and M. Scheffler. Performance of various density-functional approximations for cohesive properties of 64 bulk solids. *New Journal of Physics*, 2018.

[45] N. T. Hung, A. R. T. Nugraha, and R. Saito. Two-dimensional InSe as a potential thermoelectric material. *Applied Physics Letters*, 2017.

[46] G. Pizzi, V. Vitale, R. Arita, S. Blügel, F. Freimuth, G. Géranton, M. Gibertini, D. Gresch, C. Johnson, T. Koretsune, et al. Wannier90 as a community code: new features and applications. *Journal of Physics: Condensed Matter*, 2020.

[47] I. Souza, N. Marzari, and D. Vanderbilt. Maximally localized Wannier functions for entangled energy bands. *Physical Review B*, 2001.

[48] A. B. Lorentzen, A. Ørsted, and M. Ovesen. *Tight-binding analysis of graphene and MoS2*. Aalborg University, 2019.

[49] M. Bass, P. A. Franken, J. F. Ward, and G. Weinreich. Optical rectification. *Physical Review Letters*, 1962.

[50] X. Wang, J. R. Yates, I. Souza, and D. Vanderbilt. Ab initio calculation of the anomalous hall conductivity by wannier interpolation. *Physical Review B*, 2006.

[51] A. Kokalj. XCrySDen—a new program for displaying crystalline structures and electron densities. *Journal of Molecular Graphics and Modelling*, 1999.

[52] L. Debbichi, O. Eriksson, and S. Lebègue. Two-dimensional indium selenides compounds: An ab initio study. *The journal of physical chemistry letters*, 2015.

[53] F. Zahid, L. Liu, Y. Zhu, J. Wang, and H. Guo. A generic tight-binding model for monolayer, bilayer and bulk MoS2. *AIP Advances*, 2013.

[54] G. Dresselhaus. Spin-orbit coupling effects in zinc blende structures. *Physical Review*, 1955.

[55] D. R. Hamann. Optimized norm-conserving vanderbilt pseudopotentials. *Physical Review B*, 2013.

[56] N. Leisgang, J. G. Roch, G. Froehlicher, M. Hamer, D. Terry, R. Gorbachev, and R. J. Warburton. Optical second harmonic generation in encapsulated single-layer inse. *AIP Advances*, 2018.

[57] L. M. Malard, T. V. Alencar, A. P. M. Barboza, K. F. Mak, and A. M. De Paula. Observation of intense second harmonic generation from MoS2 atomic crystals. *Physical Review B*, 2013.

[58] K. L. Seyler, J. R. Schaibley, P. Gong, P. Rivera, A. M. Jones, S. Wu, J. Yan, D. G. Mandrus, W. Yao, and X. Xu. Electrical control of second-harmonic generation in a WSe2 monolayer transistor. *Nature nanotechnology*, 2015.

[59] Y. Song, R. Tian, J. Yang, R. Yin, J. Zhao, and X. Gan. Second harmonic generation in atomically thin MoTe2. *Advanced Optical Materials*, 2018.

[60] R. Sakuma. Symmetry-adapted Wannier functions in the maximal localization procedure. *Physical Review B*, 2013.

[61] V. Vitale, G. Pizzi, A. Marrazzo, J. Yates, N. Marzari, and A. Mostofi. Automated high-throughput wannierisation. *arXiv preprint arXiv:1909.00433*, 2019.

[62] F. Hüser, T. Olsen, and K. S. Thygesen. How dielectric screening in two-dimensional crystals affects the convergence of excited-state calculations: Monolayer MoS2. *Physical Review B*, 2013.

[63] K. Andersen, S. Latini, and K. S. Thygesen. Dielectric genome of van der Waals heterostructures. *Nano letters*, 2015.

[64] S. Gao, Y. Liang, C. D. Spataru, and L. Yang. Dynamical excitonic effects in doped two-dimensional semiconductors. *Nano letters*, 2016.

[65] D. A. Prishchenko, V. G. Mazurenko, M. I. Katsnelson, and A. N. Rudenko. Coulomb interactions and screening effects in few-layer black phosphorus: a tight-binding consideration beyond the long-wavelength limit. *2D Materials*, 2017.

[66] T. G. Pedersen. *Chapter 36: Excitons in Gapped Graphene.* Aalborg University, To Be Published.

[67] D. Vanderbilt. *Berry Phases in Electronic Structure Theory: Electric Polarization, Orbital Magnetization and Topological Insulators.* Cambridge University Press, 2018.

[68] G. S. Adkins. Three-dimensional Fourier transforms, integrals of spherical Bessel functions, and novel delta function identities. *arXiv preprint arXiv:1302.1830*, 2013.

[69] G. Prandini, A. Marrazzo, I. E. Castelli, N. Mounet, and N. Marzari. Precision and efficiency in solid-state pseudopotential calculations. *npj Computational Materials*, 2018.

[70] K. Lejaeghere, G. Bihlmayer, T. Björkman, P. Blaha, S. Blügel, V. Blum, D. Caliste, I. E. Castelli, S. J. Clark, A. Dal Corso, et al. Reproducibility in density functional theory calculations of solids. *Science*, 2016.

[71] Y. Hinuma, G. Pizzi, Y. Kumagai, F. Oba, and I. Tanaka. Band structure diagram paths based on crystallography. *Computational Materials Science*, 2017.

A | Velocity Gauge and Length Gauge for Time-Harmonic Electric Fields

The one-electron Hamiltonian is on the form

$$\hat{\mathcal{H}}_0(\mathbf{r}) = \frac{\hat{\mathbf{p}}^2}{2m_e} + V_{\mathrm{KS}}(\mathbf{r}) \,, \tag{A.1}$$

where V_{KS} is the effective Kohn-Sham potential which ensures the correct electron density of the system.[12] In the presence of an external electromagnetic field, the one-electron Hamiltonian is altered to[12]

$$\hat{\mathcal{H}}(\mathbf{r},t) = \frac{\left(\hat{\mathbf{p}} + e\boldsymbol{\mathcal{A}}(\mathbf{r},t)\right)^2}{2m_e} - e\varphi(\mathbf{r},t) + V_{\mathrm{KS}}(\mathbf{r}) \,, \tag{A.2}$$

where φ and $\boldsymbol{\mathcal{A}}$ are the scalar and vector potentials of the field. These fields are mathematical objects which should yield the physical fields $\boldsymbol{\mathcal{B}}$ (magnetic induction) and $\boldsymbol{\mathcal{E}}$ (electric field) through[12]

$$\boldsymbol{\mathcal{B}}(\mathbf{r},t) = \nabla \times \boldsymbol{\mathcal{A}}(\mathbf{r},t) \qquad \text{and} \qquad \boldsymbol{\mathcal{E}}(\mathbf{r},t) = -\nabla\varphi(\mathbf{r},t) - \frac{\partial \boldsymbol{\mathcal{A}}(\mathbf{r},t)}{\partial t}. \tag{A.3}$$

The case of negligible magnetic induction and a uniform time-harmonic electric field on the form

$$\boldsymbol{\mathcal{E}}(t) = \frac{1}{2}\sum_{\mathrm{p}}\sum_{\alpha} \bar{\mathbf{e}}_\alpha \mathcal{E}_\alpha(\omega_{\mathrm{p}}) \exp(-\mathrm{i}\omega_{\mathrm{q}} t), \tag{A.4}$$

is investigated in this report which means that $\boldsymbol{\mathcal{A}}$ must have zero curl. By choosing a gauge where $\varphi = 0$, the electric field can be represented entirely by the vector potential. It can quickly be verified that

$$\boldsymbol{\mathcal{A}}(t) = \frac{1}{2}\sum_{\mathrm{p}}\sum_{\alpha} \frac{\bar{\mathbf{e}}_\alpha \mathcal{E}_\alpha(\omega_{\mathrm{p}})}{\mathrm{i}\omega_{\mathrm{p}}} \exp(-\mathrm{i}\omega_{\mathrm{p}} t) \tag{A.5}$$

is curl-less and satisfies

$$\boldsymbol{\mathcal{E}}(t) = -\frac{\partial \boldsymbol{\mathcal{A}}(t)}{\partial t}. \tag{A.6}$$

Inserting this uniform $\boldsymbol{\mathcal{A}}$ and $\varphi = 0$ in (A.2) yields

$$\hat{\mathcal{H}}(\mathbf{r},t) = \frac{\left(\hat{\mathbf{p}} + e\boldsymbol{\mathcal{A}}\right)^2}{2m_e} + V_{\mathrm{KS}}(\mathbf{r}) = \frac{\hat{\mathbf{p}}^2}{2m_e} + V_{\mathrm{KS}}(\mathbf{r}) + \frac{e}{m_e}\boldsymbol{\mathcal{A}}(t)\cdot\hat{\mathbf{p}} + \frac{e^2}{2m_e}\boldsymbol{\mathcal{A}}^2(t) \,. \tag{A.7}$$

In the notation of perturbation theory which was developed in subsection 2.3.3, the Hamiltonian can be written as

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_0 + \epsilon\hat{\mathcal{V}}^{(1)}(t) + \epsilon^2\hat{\mathcal{V}}^{(2)}(t) \,. \tag{A.8}$$

where the perturbation energy operators are

$$\hat{\mathcal{V}}^{(1)}(t) = \frac{e}{m_e}\boldsymbol{\mathcal{A}}(t)\cdot\hat{\mathbf{p}} = \frac{e}{2m_e}\sum_{\mathrm{p}}\sum_{\alpha} \hat{p}_\alpha \mathcal{E}_\alpha(\omega_{\mathrm{p}}) \frac{\exp(-\mathrm{i}\omega_{\mathrm{p}} t)}{\mathrm{i}\omega_{\mathrm{p}}}, \tag{A.9}$$

and

$$\hat{\mathcal{V}}^{(2)}(t) = \frac{e^2}{2m_e}\boldsymbol{\mathcal{A}}^2(t) = -\frac{e^2}{8m_e}\sum_{\mathrm{p,q}}\sum_{\alpha}\mathcal{E}_\alpha(\omega_\mathrm{p})\,\mathcal{E}_\alpha(\omega_\mathrm{q})\,\frac{\exp(-\mathrm{i}(\omega_\mathrm{p}+\omega_\mathrm{q})\,t)}{\omega_\mathrm{p}\omega_\mathrm{q}}. \tag{A.10}$$

Note that the unitless perturbation parameter ϵ is multiplied onto $\hat{\mathcal{V}}^{(2)}$ twice since it includes two instances of the electric field magnitudes in each term. This is known as the velocity gauge due to the presence of the momentum operator in the first-order perturbation energy operator.

Alternatively, the electric field could be represented entirely by the scalar potential, so $\boldsymbol{\mathcal{A}} = 0$, and it can then easily be verified that

$$\varphi(\mathbf{r}, t) = -\hat{\mathbf{r}} \cdot \boldsymbol{\mathcal{E}}(t) \tag{A.11}$$

satisfies

$$\boldsymbol{\mathcal{E}}(t) = -\nabla\varphi(\mathbf{r}, t)\,. \tag{A.12}$$

Inserting $\boldsymbol{\mathcal{A}} = 0$ and this expression for the scalar potential into (A.2) then yields

$$\hat{\mathcal{H}}(\mathbf{r}, t) = \frac{\hat{\mathbf{p}}^2}{2m_e} + V_\mathrm{KS}(\mathbf{r}) + e\hat{\mathbf{r}} \cdot \boldsymbol{\mathcal{E}}(t)\,, \tag{A.13}$$

and using the notation from subsection 2.3.3 yields

$$\hat{\mathcal{H}}(\mathbf{r}, t) = \hat{\mathcal{H}}_0(\mathbf{r}) + \epsilon\hat{\mathcal{V}}(\mathbf{r}, t)\,, \tag{A.14}$$

where

$$\hat{\mathcal{V}}(\mathbf{r}, t) = e\hat{\mathbf{r}} \cdot \boldsymbol{\mathcal{E}}(t) = \frac{e}{2}\sum_{\mathrm{p}}\sum_{\alpha}\hat{r}_\alpha\mathcal{E}_\alpha(\omega_\mathrm{p})\exp(-\mathrm{i}\omega_\mathrm{p}t). \tag{A.15}$$

In this case, the perturbation energy operator contains a position operator, so the gauge where the electric field is represented entirely by the scalar potential is known as the length gauge.

We close this appendix by reminding that when looking at electric fields from light incident on a crystal, these electric fields will of course have an oscillatory spatial dependence. The electric field can in these cases be approximated as uniform inside the crystal only if the wavelength of the light is much larger than the unit cell parameters of the crystal under inspection. This is known as the electric dipole approximation, and we note that the gauge transform from the velocity gauge to the length gauge and the length gauge itself is only valid under this approximation.

B | The Position Operator

When calculating the dipole moment in response to an electric field, a necessary quantity is the position operator $\mathbf{r}$ and the matrix elements hereof. These can be found by taking the Bloch state $\psi_{n\mathbf{k}} = \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}u_{n\mathbf{k}}(\mathbf{r})$ and letting the operator $-\mathrm{i}\nabla_{\mathbf{k}}$ act on it, yielding $\mathbf{r}\psi_{n\mathbf{k}} = \mathrm{i}\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}\nabla_{\mathbf{k}}u_{n\mathbf{k}} - \mathrm{i}\nabla_{\mathbf{k}}\psi_{n\mathbf{k}}$. From here the general matrix element becomes[14]

$$\left\langle m\mathbf{k}'\middle| \mathbf{r} \middle| n\mathbf{k}\right\rangle = \delta_{\mathbf{k}\mathbf{k}'}\mathrm{i}\int u^*_{m\mathbf{k}'}(\mathbf{r})\nabla_{\mathbf{k}}u_{n\mathbf{k}}(\mathbf{r})\mathrm{d}^3\mathbf{r} - \mathrm{i}\delta_{mn}\nabla_{\mathbf{k}}\delta_{\mathbf{k}\mathbf{k}'}. \tag{B.1}$$

The first term on the right is furthermore diagonal in $\mathbf{k}$ and is known as the Berry connection $\mathbf{\Omega}_{mn\mathbf{k}}$.[14] The use of this matrix element is most evident when considering the intraband part of the position matrix elements as the interband contribution is not problematic, it is just found using the relation $\mathbf{r}_{mn} = \frac{\hbar}{im(E_m-E_n)}\hat{\mathbf{p}}_{mn}$, which does not have any singularities, and the momentum matrix elements is easily found by the substitution $\hat{\mathbf{p}} \to \frac{m_e}{\hbar}\nabla_{\mathbf{k}}$.[20] This is contrary to the intraband case where $\mathbf{r}_{mn}$ is not obtained as easily. However, (B.1) does not contain any terms that blows up when $E_m = E_n$, making it useful for the intraband case. The treatment of the commutator of the intraband operator with a arbitrary simple operator $\hat{\mathcal{O}}$ furthermore yields the relation[27] [14]

$$\left\langle m\mathbf{k}\middle| [\mathbf{r}^i, \hat{\mathcal{O}}] \middle| \mathbf{n}, \mathbf{k}\right\rangle = \mathrm{i}\delta_{\mathbf{k}\mathbf{k}'}\left(\mathcal{O}_{mn}\right)_{;\mathbf{k}} \quad \text{with} \quad \left(\mathcal{O}_{mn}\right)_{;\mathbf{k}} = \nabla_{\mathbf{k}}\mathcal{O}_{mn} - \mathrm{i}\left[\mathbf{\Omega}_{nn\mathbf{k}} - \mathbf{\Omega}_{mm\mathbf{k}}\right]\mathcal{O}_{mn}. \tag{B.2}$$

Here the matrix elements $\mathcal{O}_{mn}$ and $\mathbf{\Omega}_{mn}$ are the matrix elements in the basis of eigenstates of the Hamiltonian.[14] The program Wannier90 gives needed matrix elements to calculate the Berry connection. The calculation of the Berry Connection is done using the above form

$$\mathbf{\Omega}_{nm}(\mathbf{k}) = \mathrm{i}\left\langle u_{n\mathbf{k}}\middle| \nabla_{\mathbf{k}} \middle| u_{m\mathbf{k}}\right\rangle \tag{B.3}$$

and writing $u_{n\mathbf{k}} = \frac{1}{\sqrt{N}}\sum_i c_i^{n\mathbf{k}}\sum_{\mathbf{R}}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\mathbf{R}-\mathbf{r})}\omega_i(\mathbf{r}-\mathbf{R}) \equiv \sum_i c_i^{n\mathbf{k}}\tilde{\chi}_i^{\mathbf{k}}(\mathbf{r})$ to yield

$$\begin{aligned}\mathbf{\Omega}_{nm}(\mathbf{k}) &= \mathrm{i}\sum_i c_i^{n\mathbf{k}}(\nabla_{\mathbf{k}}c_i^{m\mathbf{k}}) + \mathrm{i}\sum_{ij} c_i^{n\mathbf{k}*}c_j^{m\mathbf{k}}\left\langle \tilde{\chi}_i\middle| \nabla_{\mathbf{k}} \middle| \tilde{\chi}_j\right\rangle \\ &= \mathrm{i}\sum_i c_i^{*n\mathbf{k}}(\nabla_{\mathbf{k}}c_i^{m\mathbf{k}}) + \sum_{ij} c_i^{*n\mathbf{k}*}c_j^{m\mathbf{k}}\left[\sum_{\mathbf{R}}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{R}}\left\langle \omega_i\mathbf{0}\middle| \mathbf{r} \middle| \omega_j\mathbf{R}\right\rangle\right].\end{aligned} \tag{B.4}$$

According to Vanderbilt,[67] the first term can be dropped. Doing a numerical experiment, one will reach the same conclussion for the diagonal elements only. The correctness of the Berry connection can be checked using the fact that the Chern number c_n, given as[14]

$$c_n = \frac{1}{2\pi}\int_{\mathrm{BZ}}\mathcal{C}_{nn}\mathrm{d}^2\mathbf{k} \quad \text{with} \quad \mathcal{C}_{nn} = 2\mathrm{i}\sum_m \Omega^x_{nm}\Omega^y_{mn}, \tag{B.5}$$

is an integer.

C | The Spin-Orbit Interaction

Electrons also have intrinsic angular momentum, known as spin, that needs to be considered when doing calculations, especially when dealing with the heavier atoms in the periodic table. This is because the spin interacts with the regular orbital angular momentum in what is known as the spin-orbit interaction. This interaction adds a term in the Hamiltonian, which can written on the form[14] [48]

$$\hat{\mathcal{H}}^{\mathrm{SO}} = \Delta_{\mathrm{SO}}(r)\mathbf{L} \cdot \mathbf{S} \tag{C.1}$$

with $\Delta_{\mathrm{SO}}(r) = \frac{\hbar^2}{2m_e^2c^2}\frac{1}{r}\frac{\partial V}{\partial r}$, and $V(\mathbf{r}) = V(r)$. When taking the matrix elements of this interaction, the resulting integrals involving Δ_{SO} can be taken as an ad-hoc parameter fitted so that the spin-splitting in the resulting tight-binding model is correct. Another approach is to take the converged DFT-calculation (2.6) and use the effective potential from here and evaluate the resulting integrals.[14] In the case of a tight-binding model utilising orbitals with s, p, d, $\ldots$ symmetry, this interaction becomes especially simple because of the orthogonality relations between the spherical harmonics.[14] each of the p and d orbitals only couples to other p or d orbitals through $\hat{\mathcal{H}}^{\mathrm{SO}}$. The matrix elements of the spin-orbit interaction can be found in [14].

D | The Screened Coulomb Interaction in Planar Structures

This appendix shall develop an expression for the electrostatic potential energy which can approximate the interaction between two electrons in a thin plate of uniform thickness surrounded by two semi-infinite media.

First, consider two electrons located at $\mathbf{r}$ and $\mathbf{r}'$, respectively. The electrostatic potential energy V between these particles must obey a Poisson-like equation on the form

$$\nabla \cdot [\varepsilon_r(\mathbf{r})\, \nabla V(\mathbf{r}, \mathbf{r}')] = -\frac{e^2}{\varepsilon_0}\delta(\mathbf{r} - \mathbf{r}')\,, \tag{D.1}$$

where ε_0 is the vacuum permittivity and ε_r is the structure-dependent relative permittivity which contextualises the problem. For the case of two electrons in an isotropic medium, ε_r is a constant, and it stands to reason that V must only depend on the displacement vector between the two electrons, so

$$\nabla^2 V(\mathbf{r} - \mathbf{r}') = -\frac{e^2}{\varepsilon_0 \varepsilon_r}\delta(\mathbf{r} - \mathbf{r}')\,. \tag{D.2}$$

Equation (D.2) can be solved through Fourier transforms. The Fourier transform of a function $f(x)$ shall be denoted $f(q_x)$ with the convention

$$f(q_x) = \int_{-\infty}^{\infty} f(x)\, \mathrm{e}^{-\mathrm{i}xq_x} \mathrm{d}x \qquad \text{and} \qquad f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(q_x)\, \mathrm{e}^{\mathrm{i}xq_x} \mathrm{d}q_x. \tag{D.3}$$

Now by transforming $x - x' \to q_x$, $y - y' \to q_y$ and $z - z' \to q_z$, equation (D.2) becomes

$$V(q) = \frac{e^2}{\varepsilon_0 \varepsilon_r q^2}, \tag{D.4}$$

and transforming back, we get[68]

$$V_{\mathrm{C}}(|\mathbf{r} - \mathbf{r}'|) = \frac{e^2}{4\pi\varepsilon_0\varepsilon_r\, |\mathbf{r} - \mathbf{r}'|}, \tag{D.5}$$

which is the familiar Coulomb potential.

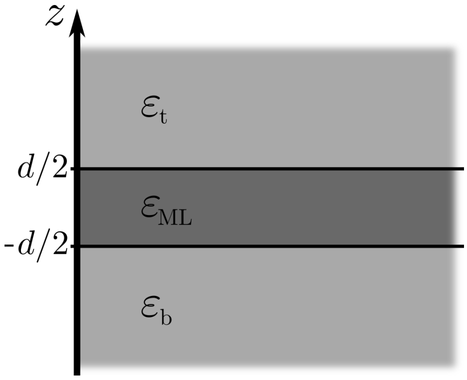


Figure D.1: A layered planar structure where a plate with uniform thickness d and dielectric constant $\varepsilon_{\mathrm{ML}}$ is surrounded by two semi-infinite media unbounded in the $\pm z$ directions and with dielectric constants ε_{t} and ε_{b}.

Next, consider the interaction between two electrons in a layered planar structure as shown in Figure D.1. For such a structure, the relative permittivity takes the form

$$\varepsilon_r(z) = \begin{cases} \varepsilon_{\mathrm{t}} & \text{for} \quad z > d/2, \\ \varepsilon_{\mathrm{ML}} & \text{for} \quad -d/2 \le z \le d/2, \\ \varepsilon_{\mathrm{b}} & \text{for} \quad z < -d/2. \end{cases}, \tag{D.6}$$

Using this relative permittivity function alters (D.1) to

$$\frac{\mathrm{d}\varepsilon_r(z)}{\mathrm{d}z}\frac{\mathrm{d}V(\boldsymbol{\rho}-\boldsymbol{\rho}',z,z')}{\mathrm{d}z} + \varepsilon_r(z)\,\nabla^2 V(\boldsymbol{\rho}-\boldsymbol{\rho}',z,z') = -\frac{e^2}{\varepsilon_0}\delta(\boldsymbol{\rho}-\boldsymbol{\rho}')\,\delta(z-z')\,, \tag{D.7}$$

where the $\boldsymbol{\rho}$'s are vectors containing only the x- and y-coordinates. It is generally assumed that at least one electron lies inside the plate, so $|z'| \le d/2$. The energy can then be found by first Fourier transforming $x - x' \to q_x$ and $y - y' \to q_y$ in (D.7) which yields

$$\frac{\mathrm{d}}{\mathrm{d}z}\left[\varepsilon_r(z)\,\frac{\mathrm{d}V(\mathbf{q}_\rho,z,z')}{\mathrm{d}z}\right] + \varepsilon_r(z)\,q_\rho^2 V(\mathbf{q}_\rho,z,z') = -\frac{e^2}{\varepsilon_0}\delta(z-z')\,, \tag{D.8}$$

and this equation can be solved by using an ansatz.[14] The two-dimensional limit of this solution, where both electrons are forced to remain in a plane at $z = z' = 0$, is[8]

$$\begin{gathered} V(q_\rho) = \frac{e^2}{2\varepsilon_0\varepsilon_{\mathrm{eff}}(q_\rho)\,q_\rho} \qquad \text{where} \\ \varepsilon_{\mathrm{eff}}(q_\rho) = \frac{\left(\cosh(q_\rho d)(\varepsilon_{\mathrm{ML}}\varepsilon_{\mathrm{t}} + \varepsilon_{\mathrm{ML}}\varepsilon_{\mathrm{b}}) + \sinh(q_\rho d)\left(\varepsilon_{\mathrm{ML}}^2 + \varepsilon_{\mathrm{t}}\varepsilon_{\mathrm{b}}\right)\right)\varepsilon_{\mathrm{ML}}}{\cosh(q_\rho d)(\varepsilon_{\mathrm{ML}}^2 + \varepsilon_{\mathrm{t}}\varepsilon_{\mathrm{b}}) + \sinh(q_\rho d)(\varepsilon_{\mathrm{ML}}\varepsilon_{\mathrm{t}} + \varepsilon_{\mathrm{ML}}\varepsilon_{\mathrm{b}}) + \varepsilon_{\mathrm{ML}}^2 - \varepsilon_{\mathrm{b}}\varepsilon_{\mathrm{t}}}. \end{gathered} \tag{D.9}$$

This effective relative permittivity can be expanded in a Taylor series at $q_\rho = 0$ which gives

$$\varepsilon_{\mathrm{eff}}(q_\rho) \approx \kappa + r_0 q_\rho \qquad \text{where} \quad \kappa = \frac{\varepsilon_{\mathrm{t}} + \varepsilon_{\mathrm{b}}}{2} \quad \text{and} \quad r_0 = d\left(\frac{\varepsilon_{\mathrm{ML}} - 1}{2} - \frac{\varepsilon_{\mathrm{t}}^2 + \varepsilon_{\mathrm{b}}^2 - 2\varepsilon_{\mathrm{ML}}}{4\varepsilon_{\mathrm{ML}}}\right). \tag{D.10}$$

r_0 is called the screening length, and the latter term in r_0 is usually small, so we approximate[14]

$$r_0 \approx d\left(\frac{\varepsilon_{\mathrm{ML}} - 1}{2}\right) = \frac{\chi_{2\mathrm{D}}(\omega = 0)}{2}.$$

Transforming (D.9) back whilst using the approximated relative permittivity function then yields[8]

$$V_{\mathrm{RK}}(|\boldsymbol{\rho}-\boldsymbol{\rho}'|) = \frac{e^2}{8\varepsilon_0 r_0}\left(\mathbf{H}_0\left(\frac{\kappa\,|\boldsymbol{\rho}-\boldsymbol{\rho}'|}{r_0}\right) - Y_0\left(\frac{\kappa\,|\boldsymbol{\rho}-\boldsymbol{\rho}'|}{r_0}\right)\right), \tag{D.11}$$

where $\mathbf{H}_0$ is a Struve function of the first kind and Y_0 is a Bessel function of the second kind. This result is the Rytova-Keldysh potential known for its use in modelling the interactions between electrons in two-dimensional materials.

In an attempt to add a z-dependence to the Rytova-Keldysh potential, the Fourier transform of the Coulomb potential is considered again except this time, the dielectric constant is replaced with the relative permittivity function from the Rytova-Keldysh potential. The Fourier transform of this modified Rytova-Keldysh potential W will thus be

$$W(q_\rho, q_z) = \frac{e^2}{\varepsilon_0 \varepsilon_r(q_\rho)\, q^2} = \frac{e^2}{\varepsilon_0 (\kappa + r_0 q_\rho)\big(q_\rho^2 + q_z^2\big)}, \tag{D.12}$$

The inverse Fourier transform of W shall now be carried out explicitly. Integrating in cylindrical coordinates gives

$$W(\mathbf{r} - \mathbf{r}') = \frac{e^2}{8\pi^3 \varepsilon_0} \int_0^\infty \frac{q_\rho}{\kappa + r_0 q_\rho} \left[\int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i} q_\rho |\boldsymbol{\rho} - \boldsymbol{\rho}'| \cos(q_\theta - \theta)} \mathrm{d}q_\theta\right] \left[\int_{-\infty}^{\infty} \frac{\mathrm{e}^{\mathrm{i} q_z (z - z')}}{q_\rho^2 + q_z^2} \mathrm{d}q_z\right] \mathrm{d}q_\rho. \tag{D.13}$$

Using the integral identities[35]

$$\int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}a\cos(t-b)} \mathrm{d}t = 2\pi J_0(a) \qquad \text{and} \qquad \int_{-\infty}^{\infty} \frac{\mathrm{e}^{\mathrm{i}at}}{b^2 + t^2} \mathrm{d}t = \frac{\pi}{b} \mathrm{e}^{-|ab|} \tag{D.14}$$

yields

$$W(|\boldsymbol{\rho} - \boldsymbol{\rho}'|, |z - z'|) = \frac{e^2}{4\pi\varepsilon_0} \int_0^\infty \frac{J_0(q_\rho |\boldsymbol{\rho} - \boldsymbol{\rho}'|)}{\kappa + r_0 q_\rho} \mathrm{e}^{-q_\rho |z - z'|} \mathrm{d}q_\rho. \tag{D.15}$$

From here, a factor $1/r_0$ is taken outside the integral, and the quantity $I(a, b, k)$ is defined as

$$I(a, b, k) \equiv \int_0^\infty \frac{J_0(at)\, \mathrm{e}^{-kt}}{b + t} \mathrm{d}t \tag{D.16}$$

such that

$$W(|\boldsymbol{\rho} - \boldsymbol{\rho}'|, |z - z'|) = \frac{e^2}{4\pi\varepsilon_0 r_0} I\Big(|\boldsymbol{\rho} - \boldsymbol{\rho}'|, \frac{\kappa}{r_0}, |z - z'|\Big). \tag{D.17}$$

The integral I shall be evaluated numerically; however, due to the oscillatory Bessel function and the improper limits, the integral shall first be rewritten to a form that is easier to evaluate through the use of Laplace transforms. By using the identity

$$\int_0^\infty f(t)\, g(t)\, \mathrm{d}t = \int_0^\infty \mathcal{L}[f](s)\, \mathcal{L}^{-1}[g](s)\, \mathrm{d}s, \tag{D.18}$$

in conjunction with the Laplace transform identities[35]

$$\mathcal{L}\left[J_0(at)\mathrm{e}^{-kt}\right](s) = \frac{1}{\sqrt{a^2 + (k+s)^2}} \qquad \text{and} \qquad \mathcal{L}^{-1}\left[\frac{1}{b + x}\right](s) = \mathrm{e}^{-bs}, \tag{D.19}$$

I becomes

$$I(a, b, k) = \int_0^\infty \frac{\mathrm{e}^{-bs}}{\sqrt{a^2 + (k+s)^2}} \mathrm{d}s = \mathrm{e}^{bk} \left\{\int_0^\infty \frac{\mathrm{e}^{-bs}}{\sqrt{a^2 + s^2}} \mathrm{d}s - \int_0^k \frac{\mathrm{e}^{-bs}}{\sqrt{a^2 + s^2}} \mathrm{d}s\right\}. \tag{D.20}$$

For the improper integral, there exists an analytic expression[35]

$$\int_0^\infty \frac{\mathrm{e}^{-bs}}{\sqrt{a^2 + s^2}} \mathrm{d}s = \frac{\pi}{2}\left[\mathbf{H}_0(ab) - Y_0(ab)\right], \tag{D.21}$$

however, the latter term must be evaluated numerically which is decidedly easier since the integrand is now strictly monotonic and the limits are finite. The modified Rytova-Keldysh potential will therefore take the form

$$\begin{aligned} W(|\boldsymbol{\rho}-\boldsymbol{\rho}'|,|z-z'|) = \exp\left(\frac{\kappa\,|z-z'|}{r_0}\right)\Bigg\{ & \frac{e^2}{8\varepsilon_0 r_0}\left[\mathbf{H}_0\left(\frac{\kappa\,|\boldsymbol{\rho}-\boldsymbol{\rho}'|}{r_0}\right) - Y_0\left(\frac{\kappa\,|\boldsymbol{\rho}-\boldsymbol{\rho}'|}{r_0}\right)\right] \\ & -\frac{e^2}{4\pi\varepsilon_0 r_0}\int_0^{|z-z'|}\exp\left(-\frac{\kappa s}{r_0}\right)\frac{1}{\sqrt{|\boldsymbol{\rho}-\boldsymbol{\rho}'|^2+s^2}}\mathrm{d}s\Bigg\}\,. \end{aligned} \tag{D.22}$$

The first term in the parenthesis can be identified as the z-independent Rytova-Keldysh potential, and as such it is clear that for $z = z'$ this modified potential reduces to the usual Rytova-Keldysh potential. For $\boldsymbol{\rho} = \boldsymbol{\rho}'$ both the Neumann function and the remaining integral diverge. Logically, these two divergences must counter each other to give a finite limit. This limit can be found by evaluating[35]

$$I(0,b,k) = \int_0^\infty \frac{\mathrm{e}^{-kt}}{b+t}\mathrm{d}t = -\mathrm{e}^{bk}\mathrm{Ei}(-bk)\,, \tag{D.23}$$

where Ei is the exponential integral. The $\boldsymbol{\rho} = \boldsymbol{\rho}'$ limit of W will thus be

$$W(0,|z-z'|) = -\frac{e^2}{4\pi\varepsilon_0 r_0}\exp\left(\frac{\kappa\,|z-z'|}{r_0}\right)\mathrm{Ei}\left(-\frac{\kappa\,|z-z'|}{r_0}\right)\,. \tag{D.24}$$

E | Automatisation Procedure

In order to function, the Quantum ESPRESSO code needs a library of pseudopotentials. For this purpose the SSSP precision library was used.[69] [70] The next step in a Quantum ESPRESSO calculation is then to specify the atoms, their positions and the lattice of the crystal in question. In most of the cases, these parameters were obtained from C2DB.[3] In the minority of cases, the structures were also relaxed by Quantum ESPRESSO. The Brillouin zone path to follow when calculating the band diagram of a crystal was obtained using SeeK-path.[71] [25]

The self-consistent-field band structure and non-self-consistent field calculations were then carried out. The symmetry operations for the crystal are obtained from the output file of the SCF calculation (with the verbose=high setting), yielding the relations between the tensor elements as described in section 2.4. Semi-core valence states can also be excluded from further processing using the information of these files. Isolated bands in terms of energy can also be found. The next step is to obtain the MLWF's using Wannier90. Two options are available for this purpose: atomic orbitals (AO) projections + disentanglement + iterative minimisation or the SCDM-**k** method[18].

The first option requires specifying suitable orbitals for the initial projections used in the Wannierisation-scheme. A way to have a reasonable guess for these is by finding the AO-resolved LDOS which is also energy-resolved using the projwfc.x code and then find a set of atomic orbitals and a fitting energy window. An easy example of this is $\mathrm{MoS_2}$ which has 7 isolated valence bands and 4 isolated conduction bands with AO projections matching Mo-d and S-p orbitals. Such cases are easy to detect. More generally, the valence states are easily projected onto localised orbitals while the conduction states gets less and less projectable the higher in energy one looks. This means that in a more versatile program, one has to scan from the lowest conduction band energy and up, and then decide a tolerance for how far the projected LDOS can deviate from the total LDOS. From here the AO projections used in the Wannierisation procedure can be chosen.

The SCDM-**k** option requires only two parameters plus the number of localised orbitals as described in Vitale et al.[61] and is somewhat simpler to implement. These two parameters called μ and σ are constants found from the **k**-resolved projectability, defined as[61]

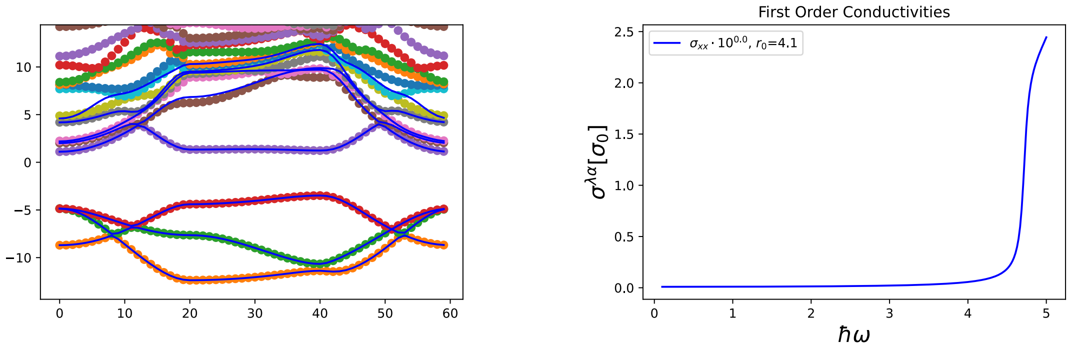
$$p_{n\mathbf{k}} = \sum_{A,n,l} \left| \left\langle \psi_{n\mathbf{k}} \middle| \phi_{A,n,l} \right\rangle \right|^2, \tag{E.1}$$

which is also obtained using the projwfc.x code from Quantum ESPRESSO. The projectability $p_{n\mathbf{k}}$ can then be plotted histogrammically as a function of the eigenvalue of the Bloch function $\psi_{n\mathbf{k}}$. To this histogram, the function $f(\epsilon, \mu, \sigma) = \frac{1}{2}\mathrm{erfc}(\frac{\epsilon-\mu}{\sigma})$ is fitted, yielding initial values for μ_{fit} and σ_{fit}. However, as described in Vitale et al.[61] these values yields bad results and they suggest using $\mu = \mu_{\mathrm{fit}} - 3\sigma_{\mathrm{fit}}$ and $\sigma = \sigma_{\mathrm{fit}}$. These values, while yielding accurate TB-band diagrams for some materials, also fail on a large portion of other materials we tested. Instead we set $\sigma = \frac{2}{3}\sigma_{\mathrm{fit}}$ and then scanned μ from the bottom of the energy window, defined $S = \{(\epsilon, p_{n\mathbf{k}}(\epsilon)) | p_{n\mathbf{k}}(\epsilon)) < 0.2\}$ and stopped when the number of points simultaneously in S and below the curve of $f(\epsilon, \mu, \sigma)$ became greater than 200.

The approach to the automatisation procedure was then to try and get an AO-TB model as these orbitals are well-localised by virtue of the iterative minimisation of the spread hereof. In most cases it is possible to get an AO-TB model. Because the SCDM-**k** method seems to yield worse band structures when minimising their spread with Wannier90, this was not done. This also results in the SCDM-**k** method being fast compared to the AO approach. Because all the many of the time-consuming steps already have been done and written to files for the TB-AO calculation, the SCDM-**k** method is rerun after all TB-AO calculations taking only a couple of extra minutes of computing time.

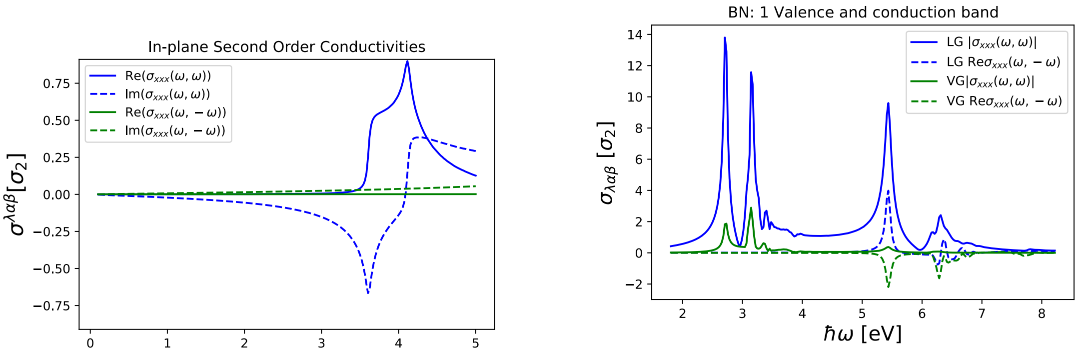
E.1 A Selection of the Generated Tight-Binding Models

The described automatisation procedure has been run on a variety of 2D materials, whose structure has primarily been obtained from C2DB.[3] A selection band-structure of the produced TB-models, the calculated $\sigma^{(2)}$, and when it has been computationally feasable, excitonic response, can be found in this part of the appendix. DFT-calculated bands are plotted with colored dots, while the Wannier-interpolated bands are plotted with blue lines. The following selection of TB-models are matches the DFT-bands nicely, and there where of course several cases where this was not the case. The Fermi-level will located above the lowest group of bands. The program also finds the non-zero first and second order tensor-elements of the material. The scissor-shifts applied makes the band gaps of the materials roughly match the G0W0 band gaps on C2DB.[3]



(a) BN Bandstructure. $\Gamma \to M \to K \to \Gamma$.

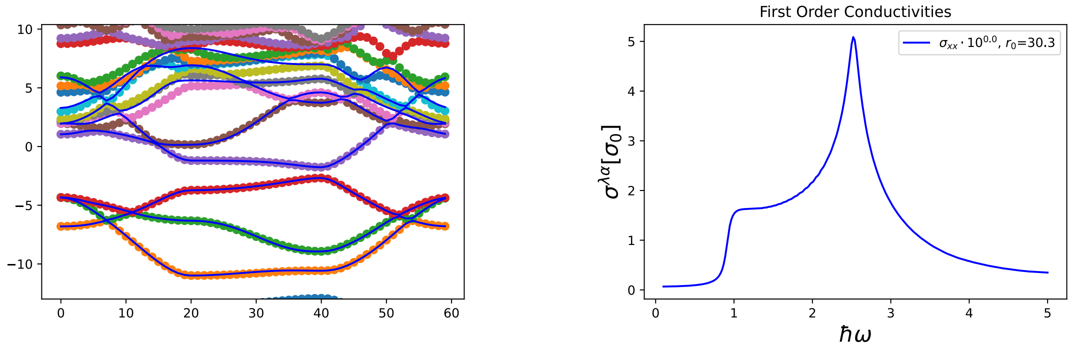
(b) Linear IPA LG conductivity.



(c) IPA second order LG conductivity.

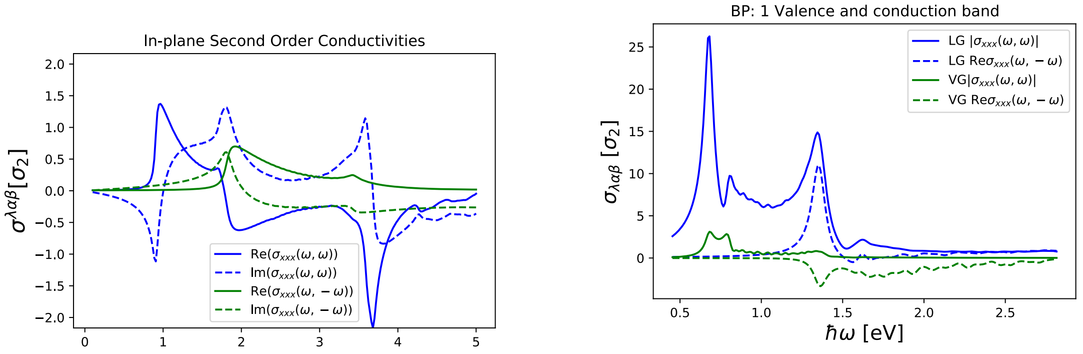
(d) BSE second order LG conductivity σ_{xxx}. 1 valence band, 1 conduction band, 70×70 **k**-points

Figure E.1: hBN TB-model using atomic orbitals. Same point-group as MoS_2. $\Delta = 2.5$eV. $r_0 = 10$Å used instead.[32] M-centered grid.



(a) BP Bandstructure. $\Gamma \to M \to K \to \Gamma$.

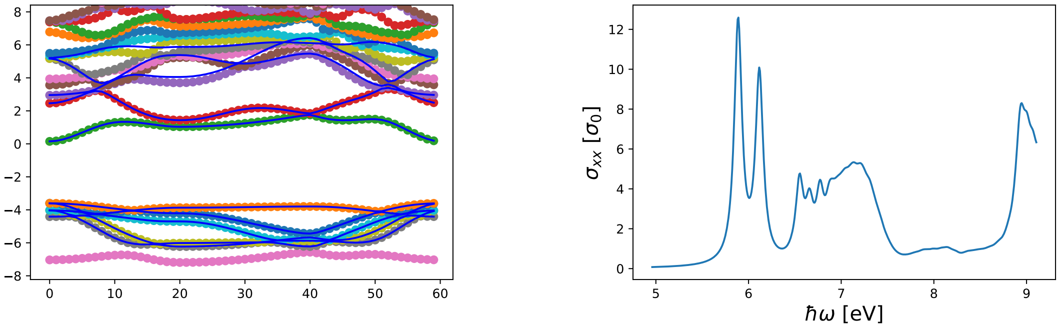
(b) IPA linear LG conductivity.



(c) IPA second order LG conductivity

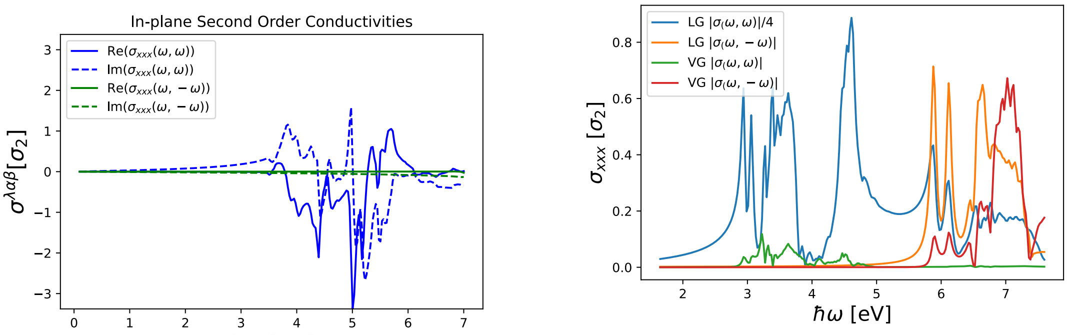
(d) BSE Second order LG conductivity σ_{xxx}. 1 valence band, 1 conduction band, 70×70 **k**-points

Figure E.2: hBP TB-model using atomic orbitals. Same point-group as MoS$_2$. $\Delta = 0.9$eV. M-centered grid.



(a) MgBr$_2$ Band Structure. $\Gamma \to M \to K \to \Gamma$.

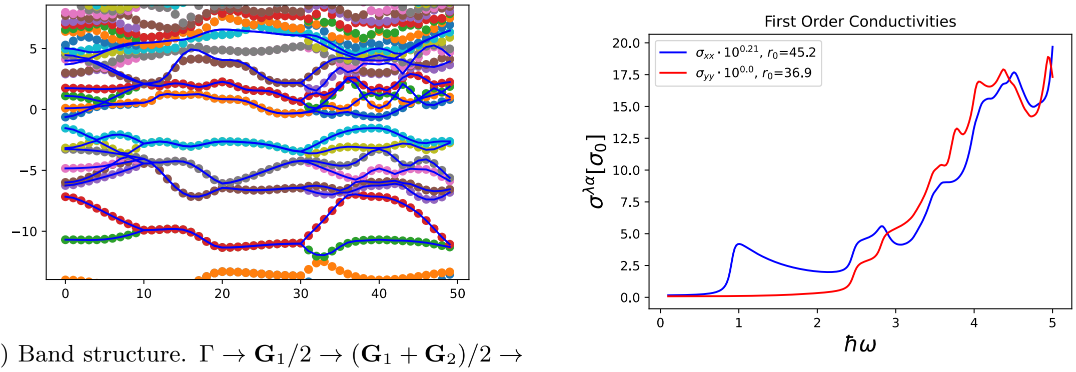
(b) BSE First order conductivity



(c) IPA second order LG conductivity

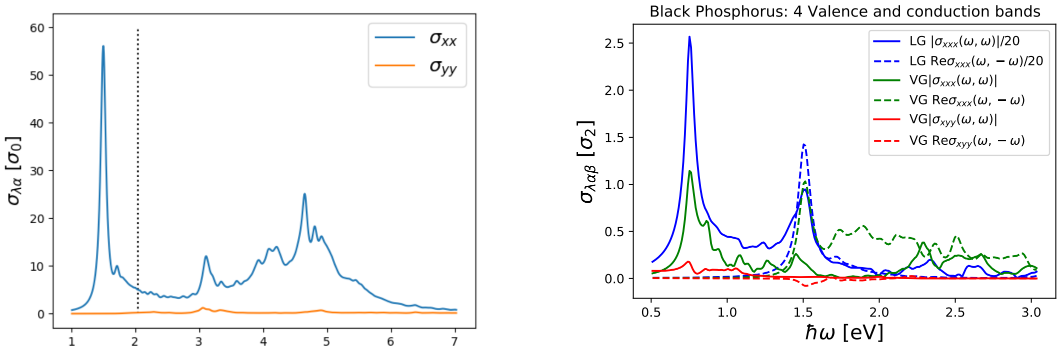
(d) BSE second order LG conductivity. LG SHG result has a factor 1/4 on it.

Figure E.3: MgBr$_2$ TB-model using atomic orbitals. Same point-group as MoS$_2$. Significant SOC[3] on Br atoms, splitting the first peak in the first order conductivity. $\Delta = 2.93$eV. Γ-centered grid with 60×60 **k**-points, one conduction band and four valence bands for BSE calculation. $r_0 \approx 11$Å from IPA calculation.



(a) Band structure. $\Gamma \to \mathbf{G}_1/2 \to (\mathbf{G}_1 + \mathbf{G}_2)/2 \to \mathbf{G}_2/2 \to \Gamma \to (\mathbf{G}_1 + \mathbf{G}_2)/2$.

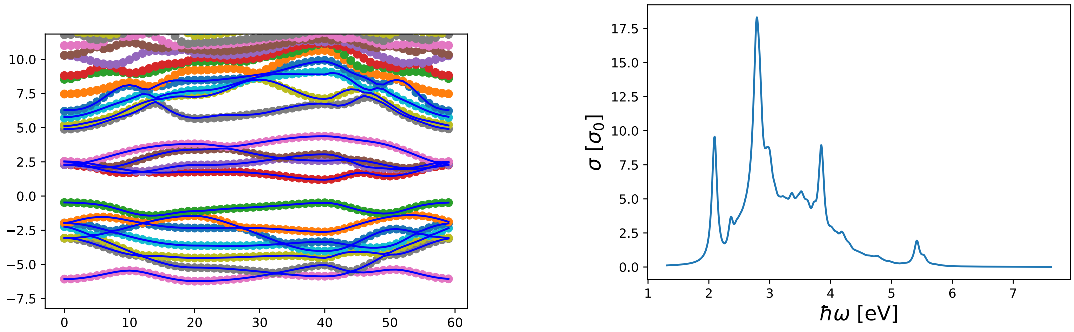
(b) IPA first order conductivity.



(c) BSE first order conductivity.

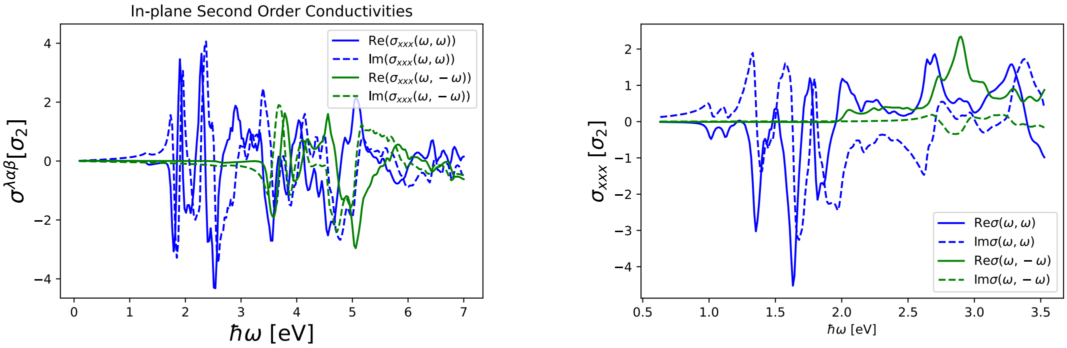
(d) BSE second order conductivity with 40×40 **k**-points. $\hbar\eta = 0.05$eV

Figure E.4: P_4 (monolayer black phosphorous) TB-model using atomic orbitals. Its point group is P2/m, yielding non-zero first order components $T_{xx}, T_{xz}, T_{yy}, T_{zx}, T_{zz}$ and second order components $T_{xxx}, T_{xxz}, T_{xyy}, T_{xzx}, T_{xzz}, T_{yxy}, T_{yyx}, T_{yyz}, T_{yzy}, T_{zxx}, T_{zxz}, T_{zyy}, T_{zzx}, T_{zzz}$. $\Delta = 1.1$ eV, 42×60 **k**-points, 3 valence and 3 conduction bands. Γ-centered grid.



(a) MoS_2 band structure

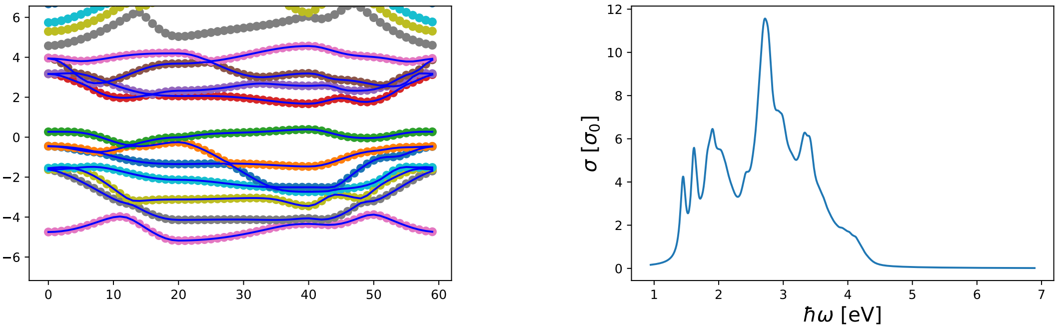
(b) BSE first order conductivity using two valence and conduction bands



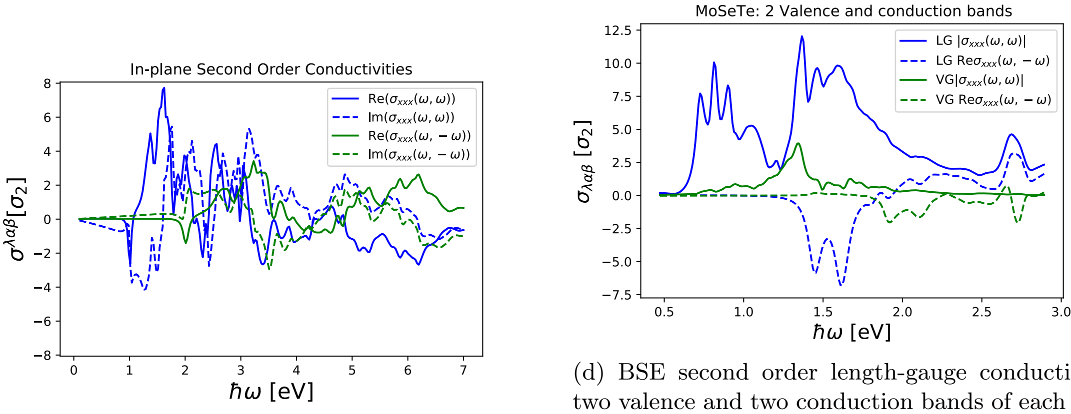
(c) .

(d)

Figure E.5: MoS_2 SCDM-**k**. No spin-orbit coupling included. M-centered grid.



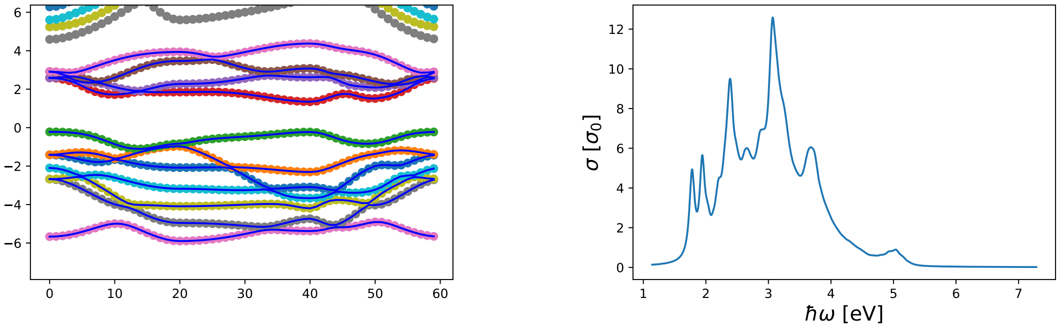
(a) MoSeTe bandstructure (b) BSE first order conductivity



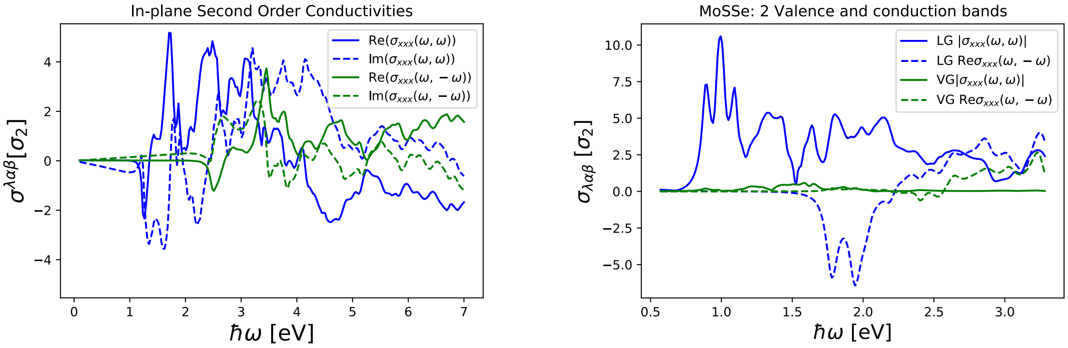
(c) IPA second order length-gauge conductivity

(d) BSE second order length-gauge conductivity. two valence and two conduction bands of each spin used.

Figure E.6: MoSeTe TB-model using atomic orbitals. Its point-group is P3m1, giving first order elements $T_{xx} = T_{yy}, T_{zz}$ and second order elements $T_{xxx} = -T_{yyx} = -T_{yxy} = -T_{xyy}$ and $T_{xxz} = T_{yyz}$, $T_{xzx} = T_{yzy}$, $T_{zxx} = T_{zyy}$. Tensor elements containing z-indecies are assumed zero as . M-centered grid.



(a) MoSSe bandstructure (b) BSE first order conductivity



(c) IPA second order length-gauge conductivity. No SOC.

(d) BSE second order length-gauge conductivity. Two valence and conduction bands of each spin.

Figure E.7: MoSSe TB-model using atomic orbitals. M-centered grid.

F Improving on the Slow-Rapid Approximation

In the tight-binding scheme the wave function is on the form in (2.7) where χ_i are now the MLWF's which we denote by $\omega_i(\mathbf{r}-\mathbf{R})$. Inserting the tight-binding wave function in the formula for the screened Coulomb interaction yields the general form

$$\begin{aligned}\left\langle vc\mathbf{k}\right| W_C \left|v'c'\mathbf{k}'\right\rangle =&\frac{1}{N^2}\sum_{ijkl} c_i^{c*}c_j^{c'}c_k^{v}c_l^{v'*}\sum_{\mathbf{R}\mathbf{R}'\mathbf{R}''\mathbf{R}'''}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\mathbf{R}''-\mathbf{R})}\mathrm{e}^{\mathrm{i}\mathbf{k}'\cdot(\mathbf{R}'-\mathbf{R}''')}\\ &\times\int\int\omega_i(\mathbf{r}_1-\mathbf{R})\omega_j(\mathbf{r}_1-\mathbf{R}')W(\mathbf{r}_1-\mathbf{r}_2)\omega_k(\mathbf{r}_2-\mathbf{R}'')\omega_l(\mathbf{r}_2-\mathbf{R}''')\mathrm{d}^3\mathbf{r}_1\mathrm{d}^3\mathbf{r}_2.\end{aligned} \tag{F.1}$$

An approximation is needed to reduce the computational burden of this expression. Using the fact that the MLWF are constructed to be as tightly localised around their centers, the function $\omega_i(\mathbf{r}-\mathbf{R})\omega_j(\mathbf{r}-\mathbf{R}')$ should be small for $\mathbf{R}\neq\mathbf{R}'$ compared to the case $\mathbf{R}=\mathbf{R}'$. Taking all terms with $\mathbf{R}\neq\mathbf{R}'$ and $\mathbf{R}''\neq\mathbf{R}'''$ as zero, while also changing coordinates $\mathbf{r}_2\to\mathbf{r}_2+\mathbf{R}'', \mathbf{r}_1\to\mathbf{r}_1+\mathbf{R}''$ yields

$$\begin{aligned}\left\langle vc\mathbf{k}\right| W_C \left|v'c'\mathbf{k}'\right\rangle =&\frac{1}{N^2}\sum_{ijkl} c_i^{c*}c_j^{c'}c_k^{v}c_l^{v'*}\sum_{\mathbf{R}\mathbf{R}''}\mathrm{e}^{\mathrm{i}(\mathbf{k}-\mathbf{k}')\cdot(\mathbf{R}-\mathbf{R}'')}\\ &\times\int\int\omega_{ij}(\mathbf{r}_1-(\mathbf{R}-\mathbf{R}''))W(\mathbf{r}_1-\mathbf{r}_2)\omega_{kl}(\mathbf{r}_2)\mathrm{d}^3\mathbf{r}_1\mathrm{d}^3\mathbf{r}_2\end{aligned} \tag{F.2}$$

It can be seen the $\mathbf{R},\mathbf{R}''$-sum contains N identical terms, cancelling a factor $\frac{1}{N}$ outside the sum. Using $\mathrm{d}^2\mathbf{k}=\frac{(2\pi)^2}{A}=\frac{4\pi^2}{A_{\mathrm{UC}}N}$ writing the double integral as $W_{ijkl,\mathbf{R}}$ yields

$$\left\langle vc\mathbf{k}\right| W_C \left|v'c'\mathbf{k}'\right\rangle = \frac{A_{\mathrm{UC}}}{4\pi^2}\mathrm{d}^2\mathbf{k}\sum_{ijkl} c_i^{c*}c_j^{c'}c_k^{v}c_l^{v'*}\sum_{\mathbf{R}}\mathrm{e}^{\mathrm{i}(\mathbf{k}-\mathbf{k}')\cdot\mathbf{R}}W_{ijkl,\mathbf{R}}. \tag{F.3}$$

In the case of the exchange interaction the phases instead cancel as a result of the condition that $\mathbf{R}=\mathbf{R}''$ and $\mathbf{R}'=\mathbf{R}'''$, yielding

$$\left\langle vc\mathbf{k}\right| V_X \left|v'c'\mathbf{k}'\right\rangle = \frac{A_{\mathrm{UC}}}{4\pi^2}\mathrm{d}^2\mathbf{k}\sum_{ijkl} c_i^{c*}c_j^{c'}c_k^{v}c_l^{v'*}\sum_{\mathbf{R}}V_{ikjl,\mathbf{R}}. \tag{F.4}$$

The integrals $W_{ijkl,\mathbf{R}}$ and $V_{ijkl,\mathbf{R}}$ must be evaluated numerically. The convolution part of the integral can be evaluated as

$$\int W(\mathbf{r}_1-\mathbf{r}_2)\omega_{kl}(\mathbf{r}_2)\mathrm{d}^3\mathbf{r}_2=\mathcal{F}^{-1}\left[W(\mathbf{k})\omega_{kl}(\mathbf{k})\right](\mathbf{r}_1) \tag{F.5}$$

Analytical forms of both potentials in reciprocal space are known, meaning only the Fourier transform $\omega_{kl}(\mathbf{k})$ needs calculation.

F.1 Including Contributions From Neighbouring Cells

Previously, we approximated that $\mathbf{R} = \mathbf{R}'$ and $\mathbf{R}'' = \mathbf{R}'''$; however, it is possible to remove a factor $1/N$ without this approximation. Starting again from (F.1), we focus on

$$\begin{aligned}\frac{1}{N^2} \sum_{\mathbf{R}\mathbf{R}'\mathbf{R}''\mathbf{R}'''} & \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\mathbf{R}''-\mathbf{R})} \mathrm{e}^{\mathrm{i}\mathbf{k}'\cdot(\mathbf{R}'-\mathbf{R}''')} \\ & \times \int\int \omega_i(\mathbf{r}_1 - \mathbf{R})\omega_j(\mathbf{r}_1 - \mathbf{R}')W(\mathbf{r}_1 - \mathbf{r}_2)\omega_k(\mathbf{r}_2 - \mathbf{R}'')\omega_l(\mathbf{r}_2 - \mathbf{R}''')\mathrm{d}^3\mathbf{r}_1\mathrm{d}^3\mathbf{r}_2.\end{aligned} \tag{F.6}$$

With the substitutions $\mathbf{r}_1 - \mathbf{R} \to \mathbf{r}_1$ and then $\mathbf{r}_2 - \mathbf{R} \to \mathbf{r}_2$, this becomes

$$\begin{aligned}\frac{1}{N^2} \sum_{\mathbf{R}\mathbf{R}'\mathbf{R}''\mathbf{R}'''} & \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\mathbf{R}''-\mathbf{R})} \mathrm{e}^{\mathrm{i}\mathbf{k}'\cdot\left((\mathbf{R}'-\mathbf{R})-(\mathbf{R}'''-\mathbf{R})\right)} \\ \times \int\int & \omega_i(\mathbf{r}_1)\omega_j(\mathbf{r}_1 + \mathbf{R} - \mathbf{R}')W(\mathbf{r}_1 - \mathbf{r}_2)\omega_k(\mathbf{r}_2 + \mathbf{R} - \mathbf{R}'')\omega_l(\mathbf{r}_2 + \mathbf{R} - \mathbf{R}''')\mathrm{d}^3\mathbf{r}_1\mathrm{d}^3\mathbf{r}_2.\end{aligned} \tag{F.7}$$

We see that the terms in the sum actually only rely on the three lattice vectors $\mathbf{R}_1 = \mathbf{R}' - \mathbf{R}$, $\mathbf{R}_2 = \mathbf{R}'' - \mathbf{R}$ and $\mathbf{R}_3 = \mathbf{R}''' - \mathbf{R}$ instead of the previous four lattice vectors. Performing the entire sum will then yield N duplicate terms, and we can write

$$\begin{aligned}\frac{1}{N} \sum_{\mathbf{R}_1\mathbf{R}_2\mathbf{R}_3} & \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{R}_2} \mathrm{e}^{\mathrm{i}\mathbf{k}'\cdot(\mathbf{R}_1-\mathbf{R}_3)} \\ & \times \int\int \omega_i(\mathbf{r}_1)\omega_j(\mathbf{r}_1 - \mathbf{R}_1)W(\mathbf{r}_1 - \mathbf{r}_2)\omega_k(\mathbf{r}_2 - \mathbf{R}_2)\omega_l(\mathbf{r}_2 - \mathbf{R}_3)\mathrm{d}^3\mathbf{r}_1\mathrm{d}^3\mathbf{r}_2.\end{aligned} \tag{F.8}$$

From here, the last $1/N$ can be removed through the **k**-integration in the BSE.