Computing optical properties and photo-emission spectra : a new starting point

Igor Reshetnyak

Thèse présentée pour obtenir le grade de

DOCTEUR DE L'ÉCOLE POLYTECHNIQUE

par

IGOR RESHETNYAK

Calcul des proprietées optiques et des spectres de photo-émission : un nouveau point de départ.

Dr.	Lucia REINING	Directrice de thèse
Prof.	Silvana BOTTI	Rapporteur
Prof.	Feliciano GIUSTINO	Rapporteur
Prof.	Silke BIERMANN	Examinateur
Dr.	Jean-Pascal RUEFF	Examinateur
Prof.	Carsten ULLRICH	Examinateur

Abstract

When a material is irradiated by particles or light, it responds with the excitation of electrons and nuclei. Because of the Coulomb interaction, this gives rise to interesting many-body effects, that cannot be explained in a single-particle picture. In this thesis we are interested in their contribution to the electronic spectra. In particular we will be looking into excitonic effects. These phenomena are due to excitations that can be described as electron-hole pairs that interact.

The Bethe-Salpeter Equation (BSE) for the two-particle Green's function, in an approximation based on the GW approximation to the self-energy, is a well established approach for accounting for excitonic effects in theoretical spectroscopy. However, in its current formulation it is computationally heavy, as its starting point requires the knowledge of the interacting single particle Green's function. Moreover, the existing implementations give access to only the diagonal parts of the microscopic screening function $\varepsilon^{-1}(\mathbf{q},\omega)_{\mathbf{G},\mathbf{G}'}$ and Dynamic Structure Factor $S(\mathbf{q},\omega)_{\mathbf{G},\mathbf{G}'}$, both of which, in their full form, are dense matrices in reciprocal lattice vectors $\mathbf{G}$ and $\mathbf{G}'$. In inhomogeneous systems these off-diagonal elements can be important and, thus, it is highly desirable to be able to describe them.

In this work, on the one hand, we try to make the Bethe-Salpeter Equation approach more efficient. To this end we study the possibility of deriving alternative equations for the two-particle Green's function and modifying the standard Bethe-Salpeter Equation. In particular, we use the fact that the shifts of spectral weight induced by the GW correction to the single-particle energies and by the electron-hole interaction cancel at least partially. The idea is to incorporate these cancelation effects, and moreover to use insight from Time-dependent Density Functional Theory, to render our calculations lighter. Furthermore, based on detailed analysis and comparison of different approaches to theoretical spectroscopy we discuss the importance of various ingredients contained in them.

On the other hand we extend the Bethe-Salpeter Equation to the off-diagonal elements of the microscopic screening function and Dynamic Structure Factor. This allows us, first of all, to reproduce available Coherent Inelastic X-ray Scattering results and make theoretical prediction for new ones. Second, this gives us the possibility to calculate the

induced charge distributions due to excitons when the material is subject to an external perturbation. And, third, we demonstrate the existence of exciton satellites, alongside the plasmon ones, in photo-emission spectra of wide gap insulators.

Contents

1 Introduction **9**

1.1 General framework 9

1.1.1 Description of Matter 10

1.1.2 Description of Crystals 11

1.1.3 Dielectric permittivity 12

1.2 Spectroscopy 15

1.2.1 Photoemission and inverse photoemission spectroscopy 15

1.2.2 Loss spectroscopy and absorption 17

1.2.3 The excitation zoo 20

1.3 Ab-initio methods 21

1.3.1 Density Functional Theory 21

1.3.2 Density Functional Theory in practice 23

1.3.3 TDDFT 23

1.3.4 Many Body Perturbation Theory. Green's Functions 24

1.3.5 Many Body Perturbation Theory. Hedin's equations 29

1.3.6 Single-electron limit 31

2 The Bethe Salpeter Equation **33**

2.1 The Standard BSE in condensed matter ab initio calculation 33

2.1.1 Derivation of BSE 34

2.1.2 Application to Silicon 36

2.1.3 Discussion 41

2.2 New BSE. First attempts 43

2.2.1 Derivation 43

2.2.2 Analysis 45

2.3 New BSE. Iterating the two particle correlation function L 46

2.3.1 Derivation 46

2.3.2 Analysis 51

2.4 Self-consistent system of equations . . . 52
2.4.1 Derivation . . . 52
2.4.2 Approximation 1 . . . 55
2.4.3 Approximation 2 . . . 55
2.4.4 Approximation 3 . . . 56
2.5 Approximating the L^{-1} . . . 57
2.5.1 Derivation . . . 57
2.6 Conclusions . . . 60

3 Comparing and combining TDLDA and BSE 63
3.1 The Standard TDLDA . . . 64
3.1.1 TDLDA, an approximation to TDDFT . . . 64
3.1.2 TDLDA and BSE equations . . . 65
3.1.3 The two-particle hamiltonian . . . 66
3.1.4 General statements about the effective Hamiltonian . . . 68
3.1.5 Discussion . . . 69
3.2 Comparing TDLDA and BSE . . . 70
3.2.1 Preliminary analysis . . . 71
3.2.2 Eigenvectors . . . 72
3.2.3 Joint Density of States . . . 74
3.2.4 Interference . . . 78
3.2.5 Coupling . . . 79
3.2.6 Importance of the electron-hole exchange . . . 80
3.3 Ideas for new methods . . . 82
3.3.1 Changing the working space . . . 82
3.3.2 Perturbative coupling . . . 83
3.3.3 Replacing parts of GW-BSE by TDLDA . . . 84
3.4 Discussions and conclusions . . . 84

4 Dynamic Structure Factor 87
4.1 Definition . . . 87
4.1.1 Relation to the susceptibility . . . 88
4.1.2 Relation to Inelastic X-ray Scattering . . . 89
4.2 Diagonal and off-diagonal elements . . . 90
4.2.1 Discussion . . . 90
4.2.2 Numerical results. Silicon . . . 91
4.2.3 Numerical results. Lithium Fluoride . . . 92
4.3 Induced charges . . . 93

4.4 Discussion 99

5 Excitonic satellites **103**
5.1 Ab-initio description of satellites 103
5.2 Lithium Fluoride: a prototypical material 105
5.2.1 Ground state properties 105
5.2.2 Self-energy correction 107
5.3 Lithium Fluoride: spectral function 109
5.3.1 Screening model 109
5.3.2 Screening function from the BSE 110
5.4 Lithium Fluoride: satellites 112
5.5 Discussion 120

6 Conclusions **125**

Appendices **127**

A **129**
A.1 Convergence: Silicon ground state 129
A.2 Convergence: BSE on Silicon 129

B **137**
B.1 Extending EXC with coupling to MPI 137

C **143**
C.1 Green's function definitions 143
C.2 SC equations in frequency space 144
C.3 BSE using L^{-1} in frequency space 145

D **149**
D.1 LiF. Convergence 149

Chapter 1

Introduction

The main topic of this thesis is spectroscopy, the response of matter to an external perturbation.

In this chapter we introduce the basic concepts and approaches used to describe the interaction of matter with probe particles. We also give a brief overview of experimental techniques used to this end, and of the effects that one can observe.

1.1 General framework

The interaction of particles and matter can be described using different approaches at various levels of approximation. One of the simplest and oldest examples of such a description of the light-matter interaction is given in classical optics, where materials are characterized by a single quantity: the refraction index n, defined by $n = \frac{c}{v}$, where c is the speed of light in vacuum and v is the phase velocity of light in the medium. Taking into account the fact that materials absorb light leads to the introduction of a second quantity, the absorption coefficient α and the Beer-Lambert law $\Phi^t = \Phi^i e^{-\alpha z}$ which relates the transmitted light flux Φ^t and the incident one Φ^i, with the distance travelled by light equal to z [1].

The discovery of Maxwell's theory of electro-magnetism [2] gave rise to a fundamental description of interacting fields, charges and currents:

$$\nabla \cdot \mathbf{D} = n_{\mathrm{f}} \tag{1.1}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{1.2}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{1.3}$$

$$\nabla \times \mathbf{H} = \mathbf{J}_{\mathrm{f}} + \frac{\partial \mathbf{D}}{\partial t} \tag{1.4}$$

Matter is now characterized by the permittivity ε and the permeability μ that relate the electric $\mathbf{E}$ and magnetic $\mathbf{B}$ fields to the displacement field $\mathbf{D}$ and the magnetizing field $\mathbf{H}$

$$\mathbf{D} = \varepsilon \mathbf{E} \tag{1.5}$$

$$\mathbf{H} = \frac{1}{\mu}\mathbf{B} \tag{1.6}$$

which all enter the Maxwell's equations, together with the free charge density n_{f} and the free current density $\mathbf{J}_{\mathrm{f}}$.

In this work we will be mostly interested in processes involving the response of matter to external electric fields and therefore in the permittivity ε, while the permeability will be assumed equal to μ_0.

1.1.1 Description of Matter

Maxwell's equations give a classical description of the electro-magnetic fields, which is enough for most applications we are interested in. On the contrary, a classical description of matter is in most cases not sufficient. Therefore, we will proceed with a quantum-mechanical description of it. For this we use the non-relativistic hamiltonian, describing a material composed of electrons and ions, that interact via the Coulomb potential:

$$\hat{\mathbb{H}}_{total} = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_e} + \sum_{I=1}^{N_I} \frac{\mathbf{P}_I^2}{2M_I} + \frac{1}{2}\sum_{I\neq J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} + \frac{1}{2}\sum_{i\neq j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i,I} \frac{eZ_I}{|\mathbf{r}_i - \mathbf{R}_I|}.$$

Here $\mathbf{p}_i$ and $\mathbf{P}_J$ are the i-th electron and J-th nucleus momenta, $\mathbf{r}_i$ and $\mathbf{R}_J$ - their positions in space, m_e and M_J their masses, Z_I the charge of the I-th nucleus. In the following we will set the electron charge (e), the electron mass (m_e), and the Plank's constant ($\hbar$) to 1.

We will be interested in materials far from phase transitions, at low temperature and in the linear response regime. Therefore we expect the Born-Oppenheimer approximation

[3] to be valid, where we can separate the motion of ions from the motion of the electrons. Moreover, we restrict ourselves to the study of electrons, supposing that the ions are de-facto "frozen". In second quantization the hamiltonian for just the electrons reads:

$$\hat{\mathbb{H}}_{electron} = \int d\mathbf{x}_1 \Psi^\dagger(\mathbf{x}_1) h_0(\mathbf{x}_1) \Psi(\mathbf{x}_1) + \frac{1}{2} \int d\mathbf{x}_1 d\mathbf{x}_2 \Psi^\dagger(\mathbf{x}_1) \Psi^\dagger(\mathbf{x}_2) v_c(\mathbf{x}_1, \mathbf{x}_2) \Psi(\mathbf{x}_2) \Psi(\mathbf{x}_1). \tag{1.7}$$

Here h_0 is the one-electron hamiltonian, that includes the kinetic energy, the potential due to nuclei and/or other external potentials, the function $v_c(\mathbf{x}_1, \mathbf{x}_2)$ is the the electron-electron Coulomb interaction, and $\Psi^\dagger(\mathbf{x}_i)(\Psi(\mathbf{x}_i))$ are the electron field operators that create(annihilate) electrons at position $\mathbf{x}_i$. The arguments $\mathbf{x}_i$, are a short hand notation for r_i, s_i i.e. a space coordinate and a spin.

1.1.2 Description of Crystals

In the previous sub-section we have presented the general quantum-mechanical description of electronic systems. Here we will give a brief introduction to a specific case, when the system is crystalline. It is then defined by its unit cell, which is replicated through space using the corresponding translation symmetry group. In three dimensions, this symmetry group is defined by three primitive vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$. Taking these vectors with integer coefficients gives us the vectors $\mathbf{R}_i$ that are elements of the translation symmetry group.

One can define reciprocal space vectors via the relation $\mathbf{b}_i \cdot \mathbf{a}_j = 2\pi\delta_{i,j}$. Solving it gives

$$\begin{aligned}
\mathbf{b_1} &= 2\pi \frac{\mathbf{a_2} \times \mathbf{a_3}}{\mathbf{a_1} \cdot (\mathbf{a_2} \times \mathbf{a_3})} \\
\mathbf{b_2} &= 2\pi \frac{\mathbf{a_3} \times \mathbf{a_1}}{\mathbf{a_2} \cdot (\mathbf{a_3} \times \mathbf{a_1})} \\
\mathbf{b_3} &= 2\pi \frac{\mathbf{a_1} \times \mathbf{a_2}}{\mathbf{a_3} \cdot (\mathbf{a_1} \times \mathbf{a_2})}.
\end{aligned}$$

Similarly using linear combinations of $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ with integer coefficients one obtains a linear space. Its vectors are commonly denoted by $\mathbf{G}_i$ and called reciprocal lattice vectors. Due to the fact that the coefficients are integer their number is countable.

This definition of the reciprocal (or inverse) lattice allows us to perform decompositions of functions into plane-waves. This new basis is compatible with the translation symmetry: a plane-wave $f(\mathbf{r}) = e^{i\mathbf{v}\cdot\mathbf{r}}$ is periodic, $f(\mathbf{r}) = f(\mathbf{r} + \mathbf{R})$ if and only if $\mathbf{v} = \mathbf{G}_i$, since by definition $\mathbf{G}_i\mathbf{R} = 2\pi$. Therefore any periodic function $F(\mathbf{r})$ can be expanded as

$F(\mathbf{r}) = \sum_i e^{i\mathbf{G}_i\mathbf{r}} F_i$.

In many cases it is useful to restrict the number of summands in the decomposition to those belonging to a certain number of closed-shells, that is to reciprocal lattice vectors defined by $\mathbf{G}_i^2 \leq \mathbf{G}_{max}^2$. This is an example of the usage of a "cutoff" $\mathbf{G}_{max}$, and it is an important parameter in the calculations presented later.

1.1.3 Dielectric permittivity

In the present work we are interested in the *ab-initio* description of the particle-matter interaction. In practice this means that we want to be able to calculate, for example, the macroscopic permittivity ε, which enters the constructive relation Eqn. 1.5. This quantity is a tensor that relates two vectors, the displacement field and the total electric field.

$$\mathbf{D}(\mathbf{q},\omega) = \begin{pmatrix} D^L(\mathbf{q},\omega) \\ D^T(\mathbf{q},\omega) \end{pmatrix} = \begin{pmatrix} \varepsilon_M^{LL}(\mathbf{q},\omega) & \varepsilon_M^{LT}(\mathbf{q},\omega) \\ \varepsilon_M^{TL}(\mathbf{q},\omega) & \varepsilon_M^{T}(\mathbf{q},\omega) \end{pmatrix} \begin{pmatrix} E^L(\mathbf{q},\omega) \\ E^T(\mathbf{q},\omega) \end{pmatrix}. \tag{1.8}$$

Here the longitudinal and transverse components of a vector $\mathbf{F}(\mathbf{q})$ are defined by:

$$\begin{aligned} \mathbf{F}(\mathbf{q}) &= \mathbf{F}^T(\mathbf{q}) + \mathbf{F}^L(\mathbf{q}) \\ \mathbf{q}\cdot\mathbf{F}^T(\mathbf{q}) &= 0 \\ \mathbf{q}\times\mathbf{F}^L(\mathbf{q}) &= \mathbf{0}. \end{aligned}$$

The subscript "M" underlines the fact that these quantities are macroscopic, or, in other words, averaged over unit cells. These relations have been written down for the Fourier transforms of the vectors.

In general Eqn. 1.8 means that a longitudinal field can induce a transverse one, and vice-versa (for details see [4]). However, in highly symmetric cases, or when $\mathbf{q} \to 0$ the problem decouples and moreover, it is possible to still limit oneself to the case of longitudinal fields even to describe absorption of light. We will not display the superscript "L" in the following.

Up till now, we have been describing the macroscopic quantities. However, in practice, it is the microscopic ones that are the most straight-forward to calculate.

Let us explain what we mean by this and make the connection between the macroscopic and the microscopic description. Consider a general microscopic potential $v(\mathbf{r},\omega)$.

It can be expanded:

$$v(\mathbf{r},\omega) = \sum_{\mathbf{q}} e^{i\mathbf{qr}} \sum_{\mathbf{G}} v(\mathbf{q}+\mathbf{G},\omega) e^{i\mathbf{Gr}} \equiv \sum_{\mathbf{q}} e^{i\mathbf{qr}} v(\mathbf{q},\mathbf{r},\omega). \tag{1.9}$$

The macroscopic potential is the average over the volume of a unit cell:

$$v_M(\mathbf{q},\omega) = \frac{1}{V_{unitcell}} \int d\mathbf{r}\, v(\mathbf{q},\mathbf{r},\omega). \tag{1.10}$$

Substituting $v(\mathbf{q},\mathbf{r},\omega)$ we obtain:

$$v_M(\mathbf{q},\omega) = v(\mathbf{q}+\mathbf{0},\omega). \tag{1.11}$$

This means that the macroscopic averaged potential is given by the $\mathbf{G}=0$ component of the microscopic one. Let us now calculate the relation between the macroscopic perturbing potential v_M^{ext} and the total macroscopic v_M^{tot} that contains the response of the system.

We start from a perturbed microscopic hamiltonian:

$$\hat{\mathbb{H}} = \hat{\mathbb{H}}_{electron} + \hat{\mathbb{H}}_{int}(t),$$

where $\hat{\mathbb{H}}_{int}(t)$ is a time-dependent perturbation.

In linear response the first order variation of the electron density δn, which is the induced charge, reads:

$$\delta n(\mathbf{r},t) = \int dt' d\mathbf{r}' \chi(\mathbf{r},\mathbf{r}',t-t') v_{ext}(\mathbf{r}',t'). \tag{1.12}$$

The linear coefficient χ is the density-density response function, or the susceptibility. Note that in linear response and for a static $\hat{\mathbb{H}}_{electron}$, the response function depends only on the time difference $t-t'$, and not on two separate times.

As a consequence of the induced charge, the total classical potential felt by the electrons or test charges becomes $v_{tot} = v_{ext} + v_{ind}$, where the last term is the induced potential $v_{ind} = v_c \delta n$. This means that the external potential applied to the system is actually screened. To describe this, one introduces the microscopic screening function ε^{-1}:

$$v_{tot}(\mathbf{r},t) = \int dt' \int d\mathbf{r}' \varepsilon^{-1}(\mathbf{r},\mathbf{r}',t-t') v_{ext}(\mathbf{r}',t'). \tag{1.13}$$

Putting the previous relations together we get:

$$\varepsilon^{-1}(\mathbf{r},\mathbf{r}',t-t') = \delta(\mathbf{r}-\mathbf{r}')\delta(t-t') + \int d\mathbf{r}'' v_c(\mathbf{r}-\mathbf{r}'')\chi(\mathbf{r}',\mathbf{r}'',t-t'). \tag{1.14}$$

We are mostly interested in crystals. For functions of two space-time variables, the translation symmetry implies: $A(\mathbf{r},t;\mathbf{r}',t') = A(\mathbf{r}+\mathbf{R},t;\mathbf{r}'+\mathbf{R},t')$, with $\mathbf{R}$ a vector of the translation symmetry group $\mathbf{R} = c_1\mathbf{a_1} + c_2\mathbf{a_2} + c_3\mathbf{a_3}$. In reciprocal space this symmetry results in the fact that the Fourier transform $\hat{A}$ is not a function of $\mathbf{k}_1, \mathbf{k}_2$, but of a single $\mathbf{q}$, belonging to the Irreducible Brillion Zone (IBZ) and two vectors of the inverse lattice $\mathbf{G}_1, \mathbf{G}_2$. Therefore it can be written as a matrix $\hat{A}(\mathbf{q})_{\mathbf{G}_1,\mathbf{G}_2}$. This can be shown, taking as an example $\varepsilon(\mathbf{k},\mathbf{k}')$. We have:

$$\varepsilon(\mathbf{k},\mathbf{k}') = \frac{1}{V_{total}}\int d\mathbf{r}d\mathbf{r}'\varepsilon(\mathbf{r},\mathbf{r}')e^{i\mathbf{k}\mathbf{r}}e^{-i\mathbf{k}'\mathbf{r}'} = \frac{1}{V_{total}}\int d\mathbf{r}d\mathbf{r}'\varepsilon(\mathbf{r}+\mathbf{R},\mathbf{r}'+\mathbf{R})e^{i\mathbf{k}(\mathbf{r}+\mathbf{R})}e^{-i\mathbf{k}'(\mathbf{r}'+\mathbf{R})}$$

Using the relation $\varepsilon(\mathbf{r},\mathbf{r}') = \varepsilon(\mathbf{r}+\mathbf{R},\mathbf{r}'+\mathbf{R})$ we get:

$$\int d\mathbf{r}d\mathbf{r}'\varepsilon(\mathbf{r},\mathbf{r}')e^{i\mathbf{k}\mathbf{r}}e^{-i\mathbf{k}'\mathbf{r}'} = \int d\mathbf{r}d\mathbf{r}'\varepsilon(\mathbf{r},\mathbf{r}')e^{i\mathbf{k}(\mathbf{r}+\mathbf{R})}e^{-i\mathbf{k}'(\mathbf{r}'+\mathbf{R})}$$

Therefore $\exp(i\mathbf{k}\mathbf{R})\exp(-i\mathbf{k}'\mathbf{R}) = 1$ and either $\mathbf{k}-\mathbf{k}' = \mathbf{G}_0$ or $\varepsilon(\mathbf{k},\mathbf{k}') = 0$. This is equivalent to saying that our function $\varepsilon(\mathbf{k},\mathbf{k}')$ depends on $\mathbf{q},\mathbf{G},\mathbf{G}'$, where $\mathbf{q}$ is in the first Brillouin zone.

In a similar manner the homogeneity of time results in $\hat{A}$ depending not on two, but one frequency ω. [1]

We can thus rewrite Eqn. 1.14:

$$\varepsilon^{-1}_{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega) = \delta_{\mathbf{G},\mathbf{G}'} + v^c_{\mathbf{G}}(\mathbf{q})\chi_{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega) \tag{1.15}$$

and also Eqn. 1.13:

$$v_{tot}(\mathbf{q}+\mathbf{G},\omega) = \sum_{\mathbf{G}'}\varepsilon^{-1}_{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega)v_{ext}(\mathbf{q}+\mathbf{G}',\omega)$$

Inserting this into Eqn. 1.11, and using the fact that a macroscopic external potential has only the $\mathbf{G} = 0$ component we obtain:

1. From here on, we will not always put a "hat" on Fourier transforms of functions, as it can be deduced from the parameters on which the function depends.

$$v_{tot,M}(\mathbf{q},\omega) = \varepsilon^{-1}_{\mathbf{G}=0,\mathbf{G}'=0}(\mathbf{q},\omega) v_{ext,M}(\mathbf{q},\omega) \tag{1.16}$$

The macroscopic dielectric function is defined by the relation between the macroscopic external and total macroscopic potential

$$v_{tot,M} = \varepsilon_M^{-1} v_{ext,M}. \tag{1.17}$$

Therefore, we obtain:

$$\frac{1}{\varepsilon_M(\mathbf{q})} = \varepsilon(\mathbf{q})^{-1}_{\mathbf{G}=0,\mathbf{G}'=0}. \tag{1.18}$$

The first works in this formalism were by Noziere and Pines [5] and Ehrenreich and Cohen [6], who modeled the dielectric constant in the Random Phase Approximation (RPA). An accurate description of the relation between the microscopic and macroscopic susceptibilities was given by [7] and [8]. Note that the macroscopic dielectric function $\varepsilon_M(\mathbf{q}) \neq \varepsilon(\mathbf{q})_{\mathbf{G}'=0,\mathbf{G}=0}$ when the matrix ε is not diagonal. Off-diagonal elements will exist in materials that are not homogeneous ($\varepsilon(\mathbf{r},\mathbf{r}') \neq \varepsilon(\mathbf{r}-\mathbf{r}')$). Their effect is known as the Crystal Local Field Effects, often simply called Local Field Effects(LFE).

1.2 Spectroscopy

In this section we will give a brief introduction to the experimental techniques used to study matter and its interaction with particles and light. Furthermore we will discuss elementary excitations that play a role in their interpretation.

1.2.1 Photoemission and inverse photoemission spectroscopy

Photoemission spectroscopy (PES) and the related inverse photoemission spectroscopy (IPES) are experimental techniques used to study electronic structure of matter. They exist in a variety of different setups.

The basic idea of Photoemission spectroscopy is to illuminate matter with photons, so as to extract electrons from it. This process is governed by the Einstein law [9]: $E_k = \omega - E_B$, where E_k is the energy of the out-going electron, ω is the energy of the incoming photon, and E_B is the binding energy of the electron in the material. This relation expresses energy conservation. By measuring the energy of the out-going electrons, at fixed incident photon frequency, one can obtain information about the binding energy

of electrons in the material. In the independent electron picture, one would expect, for each momentum **k**, a sharp peak corresponding to the band energies of the electron in the solid. An example of such an experimental spectra for Lithium Fluoride is shown in Fig. 1.1 , taken from [10]. We see that the peaks, corresponding to the emission of different electrons are not very sharp, but broadened. This is not surprising for the dispersing valance band, but not obvious for the undispersing core levels. Moreover, we see the appearance of other structures in the spectra. These effects are due to the fact that the simple single-electron picture is not completely valid. The main peaks correspond to electrons, that are "dressed" by the interactions with other excitations. They are similar to single particle peaks and are called quasiparticles. A discussion of the various other structures present in the spectra will be performed in chapter 5.

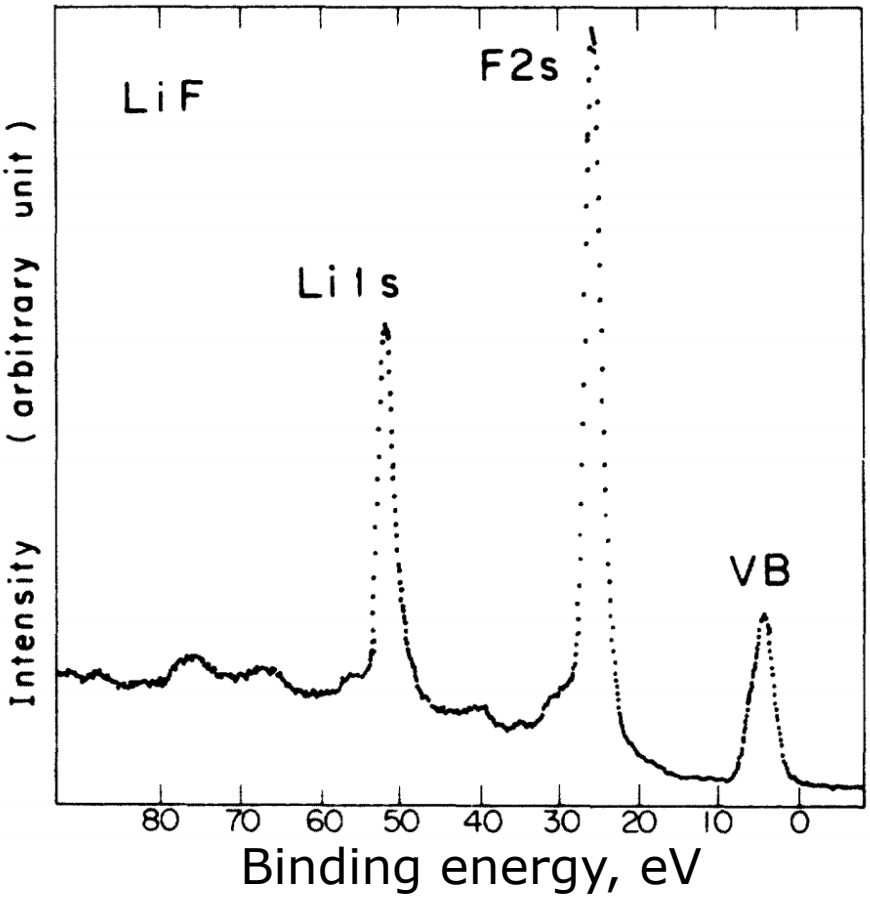


Figure 1.1 – Angle integrated photo-emission spectra of Lithium Fluoride, reprinted from [10].

The inverse photoemission is the reverse process, in which one bombards matter with electrons, which couple to unoccupied electronic states and then decay to low-lying unoccupied states. Some of these decay processes will cause photo-emission and one can

then measure its spectra, thus obtaining information on unoccupied states.

In a more elaborate setup one can measure not only the energy, but also the angle at which the electron is emitted. These types of experiments are called Angular Resolved Photoemission Spectroscopy (ARPES) [11]. A typical setup of such an experiment is shown in Fig. 1.2.

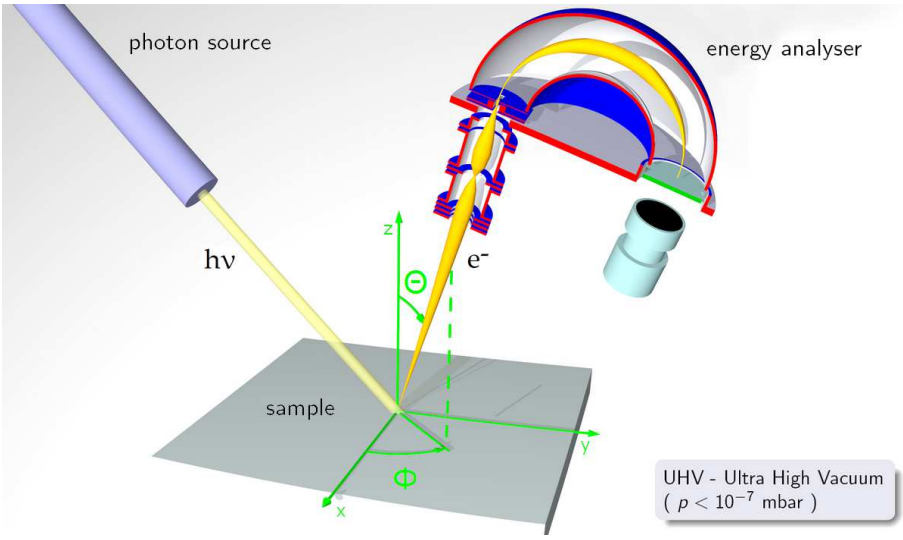


Figure 1.2 – Setup of an ARPES experiment, from [12].

Using momentum conservation, one can get information not only on the binding energy, but also the momenta $\mathbf{k}$ of the electrons in the material. This concerns only the momenta parallel to the surface: the momenta perpendicular to it is not conserved, due to the breaking of translational symmetry by the existence of the surface.

An even more elaborate technique that is used to study material dynamics is time resolved photoemission spectroscopy or even time and angular resolved photoemission spectroscopy [13, 14, 15]. In this case two pulses are used: one to excite the material, and the second one, at a later time, to extract electrons. This technique allows one to measure the excited states of matter and their decay towards the ground state as a function of time [16].

1.2.2 Loss spectroscopy and absorption

In the previous sub-section we have considered methods that involve either the addition or the removal of charged particles from the system. An alternative situation is when the excitation of the system is charge neutral. Among these methods one can separate two groups: when the probe is charged, for example transmission electron microscopy (TEM) [17] and when the probe is neutral, for example inelastic X-ray scattering (IXS) [18]. In

the first group of methods we find the electron energy loss spectroscopy (EELS). It can be performed in an electron microscope where matter is bombarded by electrons and angle and energy of the electrons after their interaction with the sample is recorded. This can be combined with transmission electron microscopy to allow one to obtain results with high spatial and spectral resolution. A possible realisation of such a setup is the energy-filtered transmission electron microscope (EFTEM) [19].

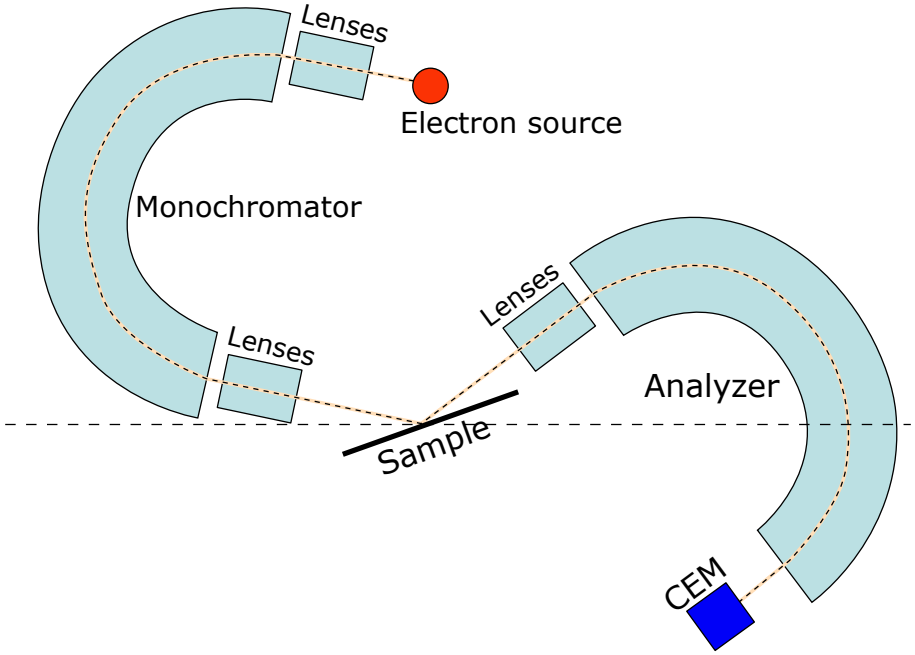


Figure 1.3 – Setup of an EELS experiment, from [12].

In the present work we are interested in EELS spectra of valence electrons. A schematic setup of such an experiment is shown in Fig. 1.3: an electron beam creates a field that interacts and excites matter, hence it loses energy. By measuring the energy difference we obtain information about the excitations of matter. One possible setup is the scanning transmission electron microscope [20] with an electron spectrometer.

Let us look at a simple theoretical description of such an experiment with electrons. Consider a fast moving charged particle, with velocity $\mathbf{v}$. The corresponding density is $n_f(\mathbf{r},t) = e\delta(\mathbf{r}-\mathbf{v}t)$. One can solve the macroscopic Maxwell equation $\nabla \cdot \mathbf{D} = n_f$. In reciprocal space the solution reads:

$$\mathbf{D}(\mathbf{q},\omega) = \frac{-i\mathbf{q}}{q^2}\frac{e}{(2\pi)^3}\delta(\omega - \mathbf{q}\mathbf{v}).$$

Work performed by the system, or alternatively the loss of energy per unit of time is given

by $\frac{dA}{dt} = \int dr \mathbf{j}(r,t) \cdot \mathbf{E}(r,t)$. Using the expression for the current $\mathbf{j} = -\mathbf{v}n_f = -e\mathbf{v}\delta(\mathbf{r} - \mathbf{v}t)$ and the fact that $E(\mathbf{q},\omega) = \varepsilon_M^{-1}(\mathbf{q},\omega)D(\mathbf{q},\omega)$ we get:

$$\begin{aligned}\frac{dA}{dt} &= \int dr \delta(\mathbf{r} - \mathbf{v}t) \int d\mathbf{q} d\omega \, e^{i(\mathbf{q}\cdot\mathbf{r} - \omega t)} \frac{i\mathbf{v} \cdot \varepsilon_M^{-1}(\mathbf{q},\omega)\mathbf{q}}{\mathbf{q}^2} \frac{e^2}{(2\pi)^3} \delta(\omega - \mathbf{q}\mathbf{v}) \\ &= \int d\mathbf{q} d\omega \, \frac{\omega}{q^2} \varepsilon_M^{-1}(\mathbf{q},\omega) \, \delta(\omega - \mathbf{q}\mathbf{v}) \frac{ie^2}{(2\pi)^3}.\end{aligned}$$

Using the relation that $\int_{-\infty}^{\infty} d\omega \, \omega \, \varepsilon^{-1}(\omega) = 2i \int_0^{\infty} d\omega \, \omega \, \Im\left[\varepsilon^{-1}(\omega)\right]$, obtained using the symmetry consistent with the Kramers-Kronig relations [21], we get

$$\frac{dA}{dt} = -\frac{e^2}{4\pi^3} \int d\mathbf{q} \frac{\mathbf{q}\mathbf{v}}{q^2} \Im\left[\varepsilon_M^{-1}(\mathbf{q}, \mathbf{q}\mathbf{v})\right].$$

This is the total energy loss. The energy loss probability P, that is defined by $\frac{dA}{dt} = \int_0^{\infty} d\omega \, \omega P(\omega)$, reads:

$$P(\omega) = -\frac{1}{4\pi^3} \int dq \frac{e^2}{q^2} \Im\left[\varepsilon_M^{-1}(\mathbf{q},\omega)\right] \delta(\omega - \mathbf{q}\mathbf{v}).$$

We see that the quantity that we thus measure is the imaginary part of the inverse dielectric function $\Im\left[\varepsilon_M^{-1}(\mathbf{q},\omega)\right]$ scaled by a factor $1/q^2$. The negative of $\Im\left[\varepsilon_M^{-1}(\mathbf{q},\omega)\right]$ is called the Loss Function. Note that we can also perturb the system at shorter wavelengths and hence measure $\Im\left[\varepsilon^{-1}_{\mathbf{G}=\mathbf{G}'}(\mathbf{q},\omega)\right]$.

In the second group of methods, we illuminate the sample with high energy photons and then measure the number of scattered photons per unit angle. In this case, the result is similar, but without the $1/q^2$ pre-factor [18]. Therefore these loss and scattering experiments can be complementary, in the sense that they are suitable for small, or large $\mathbf{q}$ respectively.

Finally, if the probe photons are prepared in a particular manner, that is in a coherent superposition of two plane waves:

$$\begin{aligned}A_0(\mathbf{r}) &= A_0 e^{i\mathbf{K}_0\mathbf{r}} \\ A_h(\mathbf{r}) &= A_h e^{i\mathbf{K}_h\mathbf{r}},\end{aligned}$$

with A being the electro-magnetic 4-potential that is related to the electro-magnetic fields $\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}$; $\mathbf{B} = \nabla \times \mathbf{A}$, and $\mathbf{K}_h = \mathbf{K}_0 + \mathbf{G}$, with $\mathbf{G}$ being a reciprocal lattice vector of the material we want to study, one can access also off-diagonal elements $\mathbf{G} \neq \mathbf{G}'$ of the microscopic $\varepsilon(\mathbf{q},\omega)_{\mathbf{G},\mathbf{G}'}$ matrix [22]. This kind of experiment is called Coherent Inelastic

X-ray Scattering (CIXS) Spectroscopy. It not has been used often up till now, but it is an important technique as there is no simple experimental alternative to access the full dielectric matrix. We will discuss this last type of experiments in detail in Chapter 4, where we will also present *ab-initio* results for them.

To complete this subsection on the spectroscopy of neutral excitations, it is important to link absorption experiments to the dielectric function. In principle, one would have to consider the transverse components, but in the approximation of infinite wavelength of light, which corresponds to $\mathbf{q} \to 0$ and if one can make a principal axis transformation of the dielectric tensor, an absorption spectrum is given by $\Im\left[\varepsilon_M(\mathbf{q} \to 0)\right]$, with ε_M defined by Eqn. 1.18. This limit is non-analytic and the direction $\mathbf{q} \to 0$ corresponds to the polarization of light. This is the quantity that is evaluated in most *ab-initio* calculations [23].

1.2.3 The excitation zoo

As mentioned previously the excitations of matter can be neutral or charged. In the case of a system of a finite number N of non-interacting particles in a volume V one can define single particle, two-particle, n-particle excitations. These correspond simply to states that are slater determinants with different occupations of the single-particle orbitals. As one adds interaction between particles the excitations start to couple, however in many cases one can still assign a certain character to every one of them. The simplest are the quasi-particle excitations, which are single-particle excitations "dressed" by the interaction. A good example of such a situation is the polaron [24], which is an electron excitation dressed by the interactions with phonons, the elementary lattice excitations. A more elaborate example of an excitation is an exciton [25, 26]. In its simplest case, known as Wannier exciton [27, 28], it is a bound electron-hole pair, that is similar to a hydrogen atom, with the proton being replaced by a positively charged hole (a positively charged excitation, corresponding to an electron missing in the N-particle system). Finally a macroscopic displacement of charges in materials gives rise to charge density excitations, known as plasmons.

The plasmons are collective excitations present in the electron gas. They appear as soon as one starts to consider the long-range Coulomb interaction. From a purely classical point of view they can be seen as collective oscillations that appear when we displace all the electrons by an infinitesimal distance with respect to nuclei. Alternatively they can be viewed as one of the wave types propagating in an electron-hole plasma. To see this one can consider the equation for wave propagation in a medium with a dielectric function

that is dominated by a characteristic excitation at frequency ω_p, the plasmon frequency: $\varepsilon(\omega) = \varepsilon_0 \left[1 - \frac{\omega_p^2}{\omega^2}\right]$. For simplicity, assume $\mu = \mu_0$. Taking the curl of the 3rd Maxwell equation Eqn. 1.3, and inserting the 4th Maxwell equation Eqn. 1.4 we get in the absence of free current:

$$\nabla \times (\nabla \times \mathbf{E}) = -\frac{\partial}{\partial t} \nabla \times \mathbf{B} = -\mu_0 \frac{\partial^2 \mathbf{D}}{\partial t^2}.$$

Using the first Maxwell equation Eqn. 1.1 we replace $\nabla \times (\nabla \times \mathbf{E})$ by $-\nabla^2 \mathbf{E}$. We then assume a wave-like excitation $\mathbf{E} = \mathbf{E}_0 e^{i(\omega t - \mathbf{kr})}$, use the relation of $\mathbf{D} = \varepsilon \mathbf{E}$ and go to reciprocal space. This gives us:

$$\left[\mathbf{k}^2 - \mu_0 \varepsilon_0 (\omega^2 - \omega_p^2)\right] \mathbf{E} = 0.$$

The non-trivial wave solution has a dispersion relation: $\omega^2 = k^2 c^2 + \omega_p^2$, where we used the fact that $c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$. This is the plasmon dispersion. As usual, when one goes to the quantum description these waves become quantized. Moreover one has to consider them together with other elementary excitations. This makes them no-longer sharp, but gives them a broadening. Finally, due to the interactions with the electron-hole continuum the spectra of the plasmon becomes asymetric. This effect is also known as Fano asymetry. A more in-depth discussion of excitons and plasmons in simple models can be found in [29, 30].

1.3 Ab-initio methods

The key quantity that emerges from the previous sections is the ω-dependent dielectric function. Here we will give an introduction to *ab-initio* methods used to calculate this quantity, and to reproduce and predict results of the various types of experiments, that have been presented in the second section.

1.3.1 Density Functional Theory

In principle, the Schrödinger equation based on the hamiltonian Eqn. 1.7 for electrons contains all the information we need, and if one were able to solve it - one could compute all the interesting physical quantities. However, in practice the solution of such a problem for a reasonable number of electrons is impossible, and even if it were possible one could not store the wave-function of $3N$ coordinates [31]. Therefore, different methods and approximations to them were developed to overcome this difficulty. The simplest are the

Hartree and Hartree-Fock approximations. They contain the classical electrostatic effects (Hartree) and the information that the electrons are fermions (Fock), but this is in general not sufficient. A break-through came from the observation that knowing the ground state electron density is, in principle, enough to describe the relevant physical observables. This is the basis of Density Function Theory. The practical realization relies on the fact that one can map the real interacting system into a fictitious system of non-interacting particles, such that the two systems have exactly the same density. This is the Kohn-Sham approach [32]. An in-depth description of this approach is given in [31].

Here, we will just give a brief overview of the foundations given by the Hohenberg-Kohn theorems [33]. The first theorem states that *the ground state density of a system of interacting particles in an external potential v_{ext} uniquely determines this potential (up to a constant) and hence the entire system*. From this, one derives that the energy of the system is also uniquely determined by its ground state density. This allows one to introduce the energy functional $E[n]$. Unfortunately the exact functional $E[n]$ is not known. The second theorem states, with the help of the variational formulation of the Schrödinger equation Eqn. 1.19 that this functional is minimised by the exact ground state density,

$$\Phi \text{ is the ground state} \equiv \Phi \text{minimizes} \langle \Phi | \hat{H} | \Phi \rangle - E \langle \Phi | \Phi \rangle \tag{1.19}$$

This finding was then followed by the seminal work of Kohn and Sham (KS), who have shown that one can find a system of non-interacting particles that will give the same density as the physical system [32]. This leads to the Kohn-Sham equations.

$$\begin{aligned} \left(-\nabla^2 + v_{eff}(\mathbf{r})\right) \phi_i(\mathbf{r}) &= \varepsilon_i \phi_i(\mathbf{r}) \\ v_{eff}(\mathbf{r}) &= v_{ext}(\mathbf{r}) + v_h(\mathbf{r}) + v_{xc}(\mathbf{r}), \end{aligned} \tag{1.20}$$

where $v_h(\mathbf{r})$ is the Hartree potential, and $v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})}$ is the exchange-correlation potential, that is also a functional of n, which means that it depends on the density all points in space. The exchange-correlation energy E_{xc} is defined as the difference: $E_{xc} = E - T_S - E_{ext} - E_H$, where T_S is the non-interacting kinetic energy:

$$T_s[n] = \frac{1}{2} \sum_{i=1}^{N} \int d\mathbf{r}\, \phi_i^*(\mathbf{r}) \nabla^2 \phi_i(\mathbf{r}),$$

$E_{ext} = \int d\mathbf{r} n(\mathbf{r}) v_{ext}(\mathbf{r})$ and E_H is the Hartree energy. Since $E[n]$ is not known, E_{xc} and v_{xc} are not known either, but useful approximations exist.

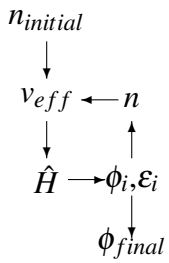


Figure 1.4 – The Kohn-Sham scheme in practice

1.3.2 Density Functional Theory in practice

In the simplest Local-Density Approximation (LDA)[32] one supposes that v_{xc} is a local function of the local density, $v_{xc}([n],\mathbf{r}) \to v_{xc}^{LDA}(n(\mathbf{r}),\mathbf{r})$.

This function is taken from the homogeneous electron gas. The final computational method that uses this approach in practice is depicted in Fig. 1.4. Starting from an initial approximation to the density $n_{initial}$ we compute the effective external potential for the non-interacting system $v_{eff} = v_{ext} + v_H + v_{xc}$ and solve the corresponding hamiltonian problem. The obtained wave-functions ϕ_i are then used to construct a new guess for the density $n(\mathbf{r}) = \sum_{occupied} |\phi_i(\mathbf{r})|^2$ and the scheme is repeated, until it converges to a given precision.

It is worth noting that the wave functions ϕ_i and energies ε_i obtained in this scheme are not the real wave functions or energies of the system. However, in some cases they are not such a bad guess [34] and we will use them often in this work as a first approximation. For a more in-depth discussion of Density Functional Theory, we refer the reader to [35]. For a discussion of a wide class of more advanced functionals one can look into [36].

A further simplification comes from the fact that one doesn't require the knowledge of the energies and wavefunctions of the electrons that are deep in the core of the atom, as they are strongly bound and do not have any effect in the range of energies we will be interested in. Therefore, to avoid their calculation, we use pseudopotentials, that mimic the behavior of the atom core, screened by the strongly bound electrons. A discussion of the accuracy of this approximation in the context of spectroscopy is found in [37, 38].

1.3.3 TDDFT

Density Functional Theory can be extended to the case of time-dependent external potentials. An equivalent of the Hohenberg-Kohn theorem in this case is the Runge-Gross theorem [39]. Together with the Kohn-Sham approach it leads to a time-dependent ef-

fective single-particle Schrödinger equations in strict analogy to Eqn. 1.20. The response of the system to a weak potential v_{ext} can be derived using time dependent perturbation theory. In the linear regime $n(\mathbf{r},t) = \int d\mathbf{r}'dt'\chi(\mathbf{r},\mathbf{r}',t-t')v_{ext}(\mathbf{r}',t')$ (Eqn. 1.12). Hence, schematically,

$$\frac{\delta n}{\delta V_{ext}} = \frac{\delta n}{\delta v_{eff}}\frac{\delta v_{eff}}{\delta V_{ext}}, \tag{1.21}$$

from which we obtain:

$$\chi(\mathbf{r}_1t_1,\mathbf{r}_2t_2) = \chi_{KS}(\mathbf{r_1}t_1,\mathbf{r}_2t_2) + \chi_{KS}(\mathbf{r_1}t_1,\mathbf{r}_2't_2')\left(\frac{1}{|\mathbf{r}_2'-\mathbf{r}_1'|} + f_{xc}(\mathbf{r}_2't_2',\mathbf{r}_1't_1')\right)\chi(\mathbf{r}_1't_1',\mathbf{r}_2t_2), \tag{1.22}$$

where, repeated indices are integrated over and $\chi_{KS} = \frac{\delta n}{\delta v_{eff}}$, is the independent particle (Kohn-Sham) susceptibility and $f_{xc} = \frac{\delta v_{xc}}{\delta n}$ is the variation of the exchange-correlation potential v_{xc}. We will come back to the formal derivation of this equation in chapter 3. An extensive review of TDDFT can be found in [40]. It is worth noting that the first calculations in this formalism were performed even before a rigorous theory was developed [41].

Though being formally exact, Time Dependent Density Functional Theory is known to have problems describing some types of excitations [42, 43]. One usually, except for some model cases [44], makes approximations to the f_{xc} kernel. The simplest approximation is the Adiabatic Local Density Approximation (ALDA)

$$f_{xc}^{ALDA}(\mathbf{r},t,\mathbf{r}',t') = \delta(\mathbf{r}-\mathbf{r}')\delta(t-t')\frac{dv_{xc}^{LDA}(n(\mathbf{r}),\mathbf{r})}{dn(\mathbf{r})}$$

that doesn't have the correct long-range behavior and misses the bound excitons. More advanced kernels have been suggested, for example in [45, 46]. The possibility of combining some elements of Many Body Perturbation Theory and TDDFT has also been suggested [47, 48, 49, 50, 51].

1.3.4 Many Body Perturbation Theory. Green's Functions

An alternative approach to the density-functional based methods described in the previous sub-section is Many-Body Perturbation Theory. The key objects in these approaches are the Green's functions. The N-particle Green's function (N being the total number of electrons in the system) is a Green's function in the most straightforward mathematical sense, namely $G_N(z) = (z-H)^{-1}$ with z a complex frequency, of the full hamiltonian Eqn. 1.7. This quantity has the same dimension as the full hamiltonian and is thus, in

practice, not very useful. However, here we are only interested in the information contained in the one- and the two-particle Green's functions, which are defined as follows:

$$G(1;2) = (-i)\left\langle \mathbb{T}\Psi(1)\Psi^{\dagger}(2)\right\rangle \tag{1.23}$$

$$\begin{aligned} G_2(1,2;3,4) &= (-i)^2\left\langle \mathbb{T}\Psi(1)\Psi(2)\Psi^{\dagger}(4)\Psi^{\dagger}(3)\right\rangle \\ &\equiv G(2,4)G(1,3) - L(1,2;3,4). \end{aligned} \tag{1.24}$$

Here, the numbers 1,2, etc stand for the space, spin and time coordinates i.e. $1 = (\mathbf{x}_1, t_1) = (\mathbf{r}_1, s_1, t_1)$. The symbol $\mathbb{T}$ stands for time ordering, and Ψ is the electron field operator in the Heisenberg picture. At zero temperature and fixed particle number, the averaging is performed over the ground state $|N\rangle$. This is the case considered in the following.

The single particle Green's function is the probability amplitude of an additional particle going from space, spin and time point 2 to 1(Fig. 1.5(a)). The diagonal part of the single particle Green's function gives us the density:

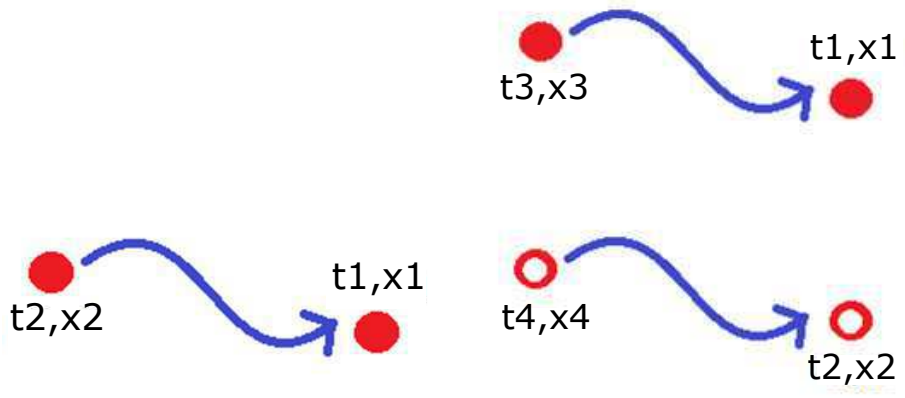
$$n(\mathbf{x}) = \left\langle \Psi^{\dagger}(\mathbf{x},t)\Psi(\mathbf{x},t)\right\rangle = -iG(\mathbf{x},\mathbf{x},t,t^{+}) \tag{1.25}$$

Similarly the two-particle Green's function describes a process involving two particles going from points 3,4 to points 1,2 (Fig. 1.5(b)). This is also the definition of L, in which we have subtracted from G_2 the trivial case where the movement of the two particles is independent. This is a key object, as its diagonal gives the density-density correlation function, which in its turn, through the Fluctuation Dissipation Theorem [52, 53], is proportional to the susceptibility:

$$\chi(1,2) = -iL(1,2^{+};1^{+},2). \tag{1.26}$$

The physical meaning of these Green's functions can be better understood, if we rewrite them in the Lehmann representation [54]. Let us do this for the single particle G. First we write out its definition Eqn. 1.23 using Heaviside step functions ($\Theta(t) = 1$ if $t > 0$ and $\Theta(t) = 0$ if $t < 0$):

$$\begin{aligned} iG(\mathbf{x}_1,t_1;\mathbf{x}_2,t_2) = \Theta(t_1-t_2)\left\langle \Phi^N\right| \Psi(\mathbf{x}_1,t_1)\Psi^{\dagger}(\mathbf{x}_2,t_2)\left|\Phi^N\right\rangle - \\ - \Theta(t_2-t_1)\left\langle \Phi^N\right| \Psi^{\dagger}(\mathbf{x}_2,t_2)\Psi(\mathbf{x}_1,t_1)\left|\Phi^N\right\rangle \end{aligned}$$



(a) Single-particle G (b) Two-particle G

Figure 1.5 – Schematic representation of Green's functions

By inserting a complete set of many body-states $\sum_k \left|\Phi_k^{N+1}\right\rangle \left\langle\Phi_k^{N+1}\right|$ and $\sum_k \left|\Phi_k^{N-1}\right\rangle \left\langle\Phi_k^{N-1}\right|$, between the field operators, we obtain:

$$\begin{aligned} iG(\mathbf{x}_1,t_1;\mathbf{x}_2,t_2) = \theta(t_1-t_2)\sum_k exp\left(i(E_0^N - E_k^{N+1})(t_1-t_2)\right) g_k(\mathbf{x}_1)g_k^*(\mathbf{x}_2) - \\ - \theta(t_2-t_1)\sum_k exp\left(i(E_0^N - E_k^{N-1})(t_2-t_1)\right) f_k(\mathbf{x}_1)f_k^*(\mathbf{x}_2), \end{aligned}$$

where E_0^N s the ground state total energy, and E_k are the energies of excited many-body states. Here f_k and g_k are the Lehmann amplitudes:

$$\begin{aligned} f_k &= \left\langle\Phi_k^{N-1}\right|\Psi(x)\left|\Phi^N\right\rangle \\ g_k &= \left\langle\Phi^N\right|\Psi(x)\left|\Phi_k^{N+1}\right\rangle, \end{aligned}$$

where $\Psi(x)$ are the field operators in the Schrödinger picture.

In frequency space the exponentials $e^{i\Delta E\Delta t}$ become fractions $\frac{1}{z-\Delta E}$. This means that the single particle Green's function will has poles at the energies corresponding to electron addition and removal energies. To make this even more explicit one can look at the spectral function $A(\omega)$, defined as:

$$A(\mathbf{x}_1,\mathbf{x}_2,\omega) = \lim_{\eta\to 0^+}\frac{i}{2\pi}\left[G(\mathbf{x}_1,\mathbf{x}_2,\omega+i\eta)-G(\mathbf{x}_1,\mathbf{x}_2,\omega-i\eta)\right] = \\ = \sum_k f_k(\mathbf{x}_1)f_k^*(\mathbf{x}_2)\delta(\omega-(E_0^N-E_k^{N-1}))+ \\ + \sum_k g_k(\mathbf{x}_1)g_k^*(\mathbf{x}_2)\delta(\omega+(E_0^N-E_k^{N+1})). \tag{1.27}$$

We see that the spectral function has peaks at electron addition and removal energies, weighted by the Lehmann amplitudes. The spectral function can be directly related to experiments. In particular, to first approximation, the photo-current $I(\mathbf{k},\omega)$, for an outgoing electron with momentum $\mathbf{k}$, is directly related to the diagonal element of the spectral function $A_{\mathbf{k},\mathbf{k}}(\omega)$.

It is important to point out that for an interacting system the state $\Psi^\dagger\left|\Phi^N\right\rangle$ where an electron is added to the N-particle system is in general not an eigenstate of the N+1 particle system. The same is true for the electron removal. The existence of this difference will result, among other things, in the appearance of satellites in spectra. These satellites will be discussed in more details in chapter 5.

Let us now derive an equation relating these Green's functions, the first equation of the Bogolyubov hierarchy [55] for this problem. Returning to our hamiltonian (Eqn. 1.7) and using that in the Heisenberg picture $\partial_t A(1) = [A(1),H(1)]$, we can write down the equations of motion (EOM) for Ψ and $\Psi^\dagger$,

$$i\partial_t\Psi(1) = [\Psi(1),H(1)] = h_0(1)\Psi(1) + \int dx'\Psi^\dagger(1')v(1,1')\Psi(1')\Psi(1)$$
$$i\partial_t\Psi^\dagger(1) = \left[\Psi^\dagger(1),H(1)\right] = \Psi^\dagger(1)h_0(1) + \int dx'\Psi^\dagger(1)\Psi^\dagger(1')v(1,1')\Psi(1')$$

,

where $v(1,1') = v_c(\mathbf{r}_1,\mathbf{r}_1')\delta(t_1'-t_1)$.

With the time ordering in the definition of $G(1,2)$ (Eqn. 1.23) given by the Heaviside theta-functions we find the EOM for the single-particle Green's function:

$$
\begin{aligned}
i\partial_t G(1,2) = \delta(1-2) + h_0(1)G(1,2) + \\
+ i\Theta(t_1-t_2)\left\langle \int dx_1' \Psi^\dagger(1')\Psi(1')v(1,1')\Psi(1)\Psi^\dagger(2) - \right. \\
\left. - i\,\Theta(t_2-t_1)\int dx_1' \Psi^\dagger(1')\Psi(1')v(1,1')\Psi^\dagger(2)\Psi(1)\right\rangle
\end{aligned}
$$

The part that is integrated can be identified with a two particle Green's function. However care must be taken, because G_2 not only depends on the 4 spatial coordinates, but also on the 4 times, whereas the part that is integrated here depends only on two times, as the Coulomb interaction is instantaneous. The correct time-ordered expression will be $i\int dx_1' G_2(1,1';2,1'')v(1,1')$; here $1 := (x_1,t)$, $1' := (x_1',t+\varepsilon)$, $1'' := (x_1',t+2\varepsilon)$.

One can then rewrite our equation as follows, using the non-interacting Green's function G_0, that satisfies $i\partial_t G_0(1,2) = \delta(1-2) + h_0 G_0(1,2)$:

$$
G_0^{-1}(1,1')G(1',2) + iv(1,1')G_2(1,1',2,1'^+) = \delta(1-2)
$$

Here, primed quantities are integrated over. Using the relation between G_2 and L (Eqn. 1.24) one obtains:

$$
G(1,2) = G_0(1,2) - iG_0(1,1')v(1',2')G(2',2'^+)G(1',2) + iG_0(1,1')v(1',2')L(1',2';2,2'^+) \tag{1.28}
$$

With $-iG(1,1^+) = n(1)$ (Eqn. 1.25) we see that the second term in the r.h.s. contains the Hartree potential.

The equation Eqn. 1.28 shows that the propagation of a particle in the system is modified by the classical electrostatic Hartree potential, and by the correlation of the particle with other particles in the system, contained in L. Finally one can write Eqn. 1.28 in the form of a Dyson equation:

$$
G(1,2) = G_0(1,2) + G_0(1,1')v_H(1')G(1',2) + G_0(1,1')\Sigma_{XC}(1',2')G(2',2), \tag{1.29}
$$

where we have defined the self-energy Σ_{XC}, as

$$
\Sigma_{XC}(1,2) = v(1,1')L(1,1';2',1'^+)G^{-1}(2',2). \tag{1.30}
$$

1.3.5 Many Body Perturbation Theory. Hedin's equations

The equation for the Green's function Eqn. 1.28 cannot be solved, as a priori L is unknown. However, according to Eqn. 1.26, its diagonal, using the Fluctuation Dissipation Theorem [53], can be written as

$$-iL(1,2^+;1^+,2) = \chi(1,2) = \frac{\delta n(1)}{\delta V_{ext}(2)} = -i\frac{\delta G(1,1^+)}{\delta U_{ext}(2,2^+)}, \tag{1.31}$$

where we introduced a general non-local potential U_{ext}, with $U_{ext}(2,2^+) = V_{ext}(2)$. This relation can be generalized to the full $L(1,2,3,4)$. It can be written using functional integrals [56], by adding a non-local external potential $\Delta H = \int dx_1 dx_2 U_{ext}(1,2)\Psi^\dagger(1)\Psi(2)$ into the electron hamiltonian Eqn. 1.7, which leads to

$$L(1,4;2,3) = \frac{\delta G(1,2)}{\delta U_{ext}(3,4)}. \tag{1.32}$$

Using these relations and the equations of motion for the Green's function, one can derive an integral equation for L analogous to Eqn. 1.22. This equation is called the Bethe-Salpeter equation. We will perform this derivation in the second chapter. Here, we just note its schematic form:

$$L = GG + GG\frac{\delta\Sigma}{\delta G}L. \tag{1.33}$$

From this one can derive a set of equations known as the Hedin's equation [57]. First we define a vertex function $\Gamma = LG^{-1}G^{-1}$. Eqn. 1.33 then leads to Eqn. 1.38. We, furthermore, introduce the polarizability P (Eqn. 1.37), the screened Coulomb interaction W (Eqn. 1.35), and, finally the self-energy Σ (Eqn. 1.36). Comparing these definitions with Eqn. 1.28 we obtain the Dyson equation for the Green's function G (Eqn. 1.34), the same as Eqn. 1.29. The full set of equations reads:

$$G(1,2) = G_0(1,2) + G_0(1,1')\Sigma(1',2')G(2',2) \tag{1.34}$$

$$W(1,2) = v(1,2) + v(1,1')P(1',2')W(2',2) \tag{1.35}$$

$$\Sigma(1,2) = i\Gamma(2';1,1')G(1',2)W(2',2) \tag{1.36}$$

$$P(1,2) = -i\Gamma(1;2',2'')G(2,2')G(2'',2) \tag{1.37}$$

$$\Gamma(1;2,3) = \delta(1,2)\delta(1,3) + \quad \Gamma(1;2',3')G(2'',2')G(3',3'')\frac{\delta\Sigma(2,3)}{\delta G(2'',3'')}. \tag{1.38}$$

Note that P yields the dielectric function, $\varepsilon(1,2) = \delta(1-2) - v_c(1,1')P(1',2)$ and

$W(1,2) = \varepsilon^{-1}(1,1')v_c(1',2)$.

In the simplest approximation, when the vertex function is given by the product of two delta functions (i.e. in Eqn. 1.38 the second term is neglected), this gives the widely used *GW* approximation [58] for the self-energy $\Sigma(1,2) = iG(1,2)W(1,2)$. In this approximation Eqn. 1.37 becomes $P(1,2) = -iG(1,2)G(2,1)$, which is the Random Phase Approximation (RPA). The GW approximation is the simplest non-trivial approximation used to describe the photo-emission processes and for the calculation of *ab-initio* band-structures [59, 60, 61]. To understand this, using the solution of Eqn. 1.34 and Eqn. 1.27, one can rewrite the spectral function $A(\omega)$ in terms of the self-energy:

$$A(\mathbf{k},\omega) = \frac{1}{\pi}\frac{\Im\Sigma(\mathbf{k},\omega)}{\left[\omega - \varepsilon^0_{\mathbf{k}} - \Re\Sigma(\mathbf{k},\omega)\right]^2 + \left[\Im\Sigma(\mathbf{k},\omega)\right]^2}. \tag{1.39}$$

With the GW approximation $\Sigma(1,2) = iG(1,2)W(1,2)$ this gives an approximation to the photo-emission spectra.

At this point we can give a more precise definition of quasi-particles, in particular we see that for $\omega - \varepsilon^0_{\mathbf{k}} - \Re\Sigma(\mathbf{k},\omega) = 0$ the spectral function will contain peaks. These solutions ω_k will give us the quasi-particle energies. The other option for peaks in $A(\mathbf{k},\omega)$, is when there is a maximum in $\Im\Sigma(\mathbf{k},\omega)$. These can be related in the *GW* approximation to the peaks the loss function $\Im\varepsilon^{-1}(\omega)$, which is as shown in subsection 1.2.2, related to the Electron Energy Loss Spectroscopy measurements. These peaks are called satellites and physically correspond to the situation, where the hole created in the photoemission process excites the system, and hence the outgoing electron has less energy.

The properties of Hedin's equation and the GW approximation have been analyzed for real materials, or using simple models. Among the simplest are the one-point model [62, 63] for the Hedin's equations, and the two-electrons on a sphere model, for the GW approximation [64]. The first realistic GW calculations were carried out for the simple semi-conductors silicon and diamond. They solved the so-called Kohn-Sham "Gap problem" [65, 59]. Even today, most of the *GW* calculations concentrate only on the quasi-particle part of the spectrum, calculating the band structure [66]. Satellites are much less studied; they are one of the topics of this thesis.

An in-depth overview and derivation of the Many-Body Perturbation theory approach can be found in [67]. An approach combining some elements of the Green's function approach with Density Functional Theory, or rather the broader class of mean field theories, is the Dynamical Mean Field Theory [68]. The limitations of the *GW* approach and some possible extensions have been discussed for example in [69, 70].

1.3.6 Single-electron limit

It is always useful to have a simple toy-model, for which we know the exact results. To this end, let us consider the limit of one electron i.e. we take the ground state to be $|N=1\rangle$ and write down some of the quantities defined in the previous sections of this chapter. Spin is omitted here.

To do this we will first of all introduce a set of notations. The vacuum, and the single-electron excitations in the interacting and non-interacting case are the same, as well as the ground state.

	H	H_0
Vacuum	$\|N=0\rangle, 0$	$\|N=0\rangle, 0$
Ground	$\|N=1\rangle, \varepsilon_1$	$\|N=1\rangle, \varepsilon_1$
1-particle	$\|N=1,k,\nu\rangle, \varepsilon_{1,k,\nu}$	$\|N=1,k,\nu\rangle, \varepsilon_{1,k,\nu}$
2-particle	$\|N=2,k,\nu\rangle, \varepsilon_{2,k,\nu}$	$\|N=2,k,\nu\rangle_0, \varepsilon_{1,k,\nu}+\varepsilon_1$

If we define the amplitudes:

$$\begin{aligned}
\psi^e_{k,\nu}(\mathbf{r}) &= \langle N=1|\,\Psi(\mathbf{r})\,|N=2;\nu,k\rangle \\
\psi^{*h}(\mathbf{r}) &= \langle N=1|\,\Psi^\dagger(\mathbf{r})\,|N=0\rangle \\
\phi_{k,\nu}(\mathbf{r}) &= \langle N=0|\,\Psi(\mathbf{r})\,|N=1;\nu,k\rangle \\
\phi^e_{k,\nu}(\mathbf{r}) &= \langle N=1|\,\Psi(\mathbf{r})\,|N=2;\nu,k\rangle_0 = \phi_{k,\nu}(\mathbf{r}) \\
\phi^{*h}(\mathbf{r}) &= \langle N=1|\,\Psi^\dagger(\mathbf{r})\,|N=0\rangle
\end{aligned}$$

corresponding to transitions between various, we obtain:

$$\begin{aligned}
G(\mathbf{r}_1,\mathbf{r}_2,\omega) &= \left[\sum_{\nu,k}\frac{\psi^e_{k,\nu}(\mathbf{r}_1)\psi^{e*}_{k,\nu}(\mathbf{r}_2)}{\omega-(\varepsilon_{2,k,\nu}-\varepsilon_1)+i\delta}+\frac{\psi^{h*}(\mathbf{r}_2)\psi^h(\mathbf{r}_1)}{\omega-\varepsilon_1-i\delta}\right] \\
G_0(\mathbf{r}_1,\mathbf{r}_2,\omega) &= \left[\sum_{\nu,k}\frac{\phi^e_{k,\nu}(\mathbf{r}_1)\phi^{e*}_{k,\nu}(\mathbf{r}_2)}{\omega-\varepsilon_{1,k,\nu}+i\delta}+\frac{\phi^{h*}(\mathbf{r}_2)\phi^h(\mathbf{r}_1)}{\omega-\varepsilon_1-i\delta}\right]
\end{aligned}$$

Taking into account the fact that $\psi^h(r)=\phi^h(r)$ we obtain that the hole-part is the same for both Green's functions, but there is a difference in the electron addition part. The densities in both cases are:

$$n(r_1) = -iG(1,1^+) = -(-i)^2\Theta(1^+-1)\left\langle\Psi^\dagger(1)\Psi(1)\right\rangle = \psi^{h*}(r_1)\psi^h(r_1) = \phi^{h*}(r_1)\phi^h(r_1)$$

We can also define the electron and hole energies:

$$\begin{aligned} \varepsilon^e_{\nu,k} &= \varepsilon_{2,k,\nu} - \varepsilon_1 \\ \varepsilon^h &= \varepsilon_1 . \end{aligned} \tag{1.40}$$

These are the energies that one would measure in direct or inverse photoemission.

Let us now also compute the susceptibility χ. To do this we express the diagonal of the two-particle Green's function:

$$\begin{aligned} G(1,2;1^+,2^+) = (-i)^2 \left\langle \mathbb{T}\Psi(1)\Psi^\dagger(1)\Psi(2)\Psi^\dagger(2) \right\rangle = \\ - \Theta(t_1 - t_2) \sum_{k,\nu} \langle N=1 | \Psi(1)\Psi^\dagger(1) | k,\nu \rangle \langle k,\nu | \Psi(2)\Psi^\dagger(2) | N=1 \rangle - \\ - \Theta(t_2 - t_1) \sum_{k,\nu} \langle N=1 | \Psi(2)\Psi^\dagger(2) | k,\nu \rangle \langle k,\nu | \Psi(1)\Psi^\dagger(1) | N=1 \rangle . \end{aligned}$$

We then go to the contracted L, which gives us what we want via the relation $\chi(1,2) = \frac{\delta n(1)}{\delta v_{ext}(2)} = -iL(1,2,1^+,2^+)$, derived from the Fluctuation Dissipation theorem.

$$\chi(\mathbf{r}_1,\mathbf{r}_2,\omega) = \left[\sum_{k,\nu}{}' \frac{\phi^*(\mathbf{r}_1)\phi_{k,\nu}(\mathbf{r}_1)\phi(\mathbf{r}_2)\phi^*_{k,\nu}(\mathbf{r}_2)}{\omega + (\varepsilon_1 - \varepsilon_{k,\nu}) + i\delta} - \sum_{k,\nu}{}' \frac{\phi^*(\mathbf{r}_2)\phi_{k,\nu}(\mathbf{r}_2)\phi(\mathbf{r}_1)\phi^*_{k,\nu}(\mathbf{r}_1)}{\omega - (\varepsilon_1 - \varepsilon_{k,\nu}) - i\delta} \right] \tag{1.41}$$

Note that the signs of the small infinitesimal parts $i\delta$ indicate the time-ordered form of χ. The prime at the sum expresses the fact that the ground state is not summed over. It is worth noting the fact that this is the same result as the one that we would have obtained by taking $\chi = G_0 G_0$. This might sound trivial, since there cannot be any contribution of the Coulomb interaction in the excitation of a single electron. However, we will discuss the implications of this in the next chapter.

Chapter 2

The Bethe Salpeter Equation

In the Bethe-Salpeter Equation Eqn. 1.33 the one-particle Green's function and self-energy appear. In order to obtain results one must therefore have a good approximation for these two quantities. One can obtain them, for example, by using the GW approximation.

The Bethe-Salpeter Equation (BSE) based on the GW approximation to the self-energy is a well established approach for accounting for excitonic effects in optical properties and photo-absorption spectra (see, for example [71, 23]). However, in its current formulation it is both computationally heavy and displays cancelation effects not accounted for analytically.
In this chapter we introduce the Bethe-Salpeter Equation (BSE), discuss the sources of cancelations [72] based on simple models and the possibility of putting them forward explicitly. Furthermore, we suggest alternative formulations and sets of approximations to the Bethe-Salpeter Equation. Finally, we discuss the possibility of using these new approaches for real systems.

2.1 The Standard BSE in condensed matter ab initio calculation

In the present section we derive an equation for L, the two-particle correlation function, known as the Bethe-Salpeter equation. Together with the equations of motion for the single-particle Greens' function they form a closed set of equation. We then briefly discuss the application of this method to a prototypical semi-conductor, Silicon, and the problems that one encounters when using the standard Bethe-Salpeter equation.

2.1.1 Derivation of BSE

The Bethe-Salpeter Equation was first introduced in the context of high energy physics [73, 74]. It was then extended to condensed matter physics [75, 76]. Let us sketch the derivation. To do this, we consider a general hamiltonian for electrons, as introduced in Eqn. 1.7.

$$\hat{H} = \int d\mathbf{x}_1 \Psi^\dagger(\mathbf{x}_1) h_0(\mathbf{x}_1) \Psi(\mathbf{x}_1) + \frac{1}{2} \int d\mathbf{x}_1 d\mathbf{x}_2 \Psi^\dagger(\mathbf{x}_1) \Psi^\dagger(\mathbf{x}_2) v_c(\mathbf{x}_1, \mathbf{x}_2) \Psi(\mathbf{x}_2) \Psi(\mathbf{x}_1) \tag{2.1}$$

One can then consider a non-local perturbation $\Delta H = \int dx_1 dx_2 U_{ext}(1,2) \Psi^\dagger(1) \Psi(2)$, and calculate in first order perturbation theory in ΔH the variation of G, $\frac{\delta G}{\delta U_{ext}}$. One obtains a term with four field operators that equals L. We start our considerations from this relation:

$$L(1,4;2,3) = \frac{\delta G(1,2)}{\delta U_{ext}(3,4)}.$$

Taking the variational derivative of $G^{-1}G = 1$ one obtains:

$$\frac{\delta G(1,2)}{\delta U_{ext}(3,4)} = -G(1,1') \frac{\delta G^{-1}(1',2')}{\delta U_{ext}(3,4)} G(2',2). \tag{2.2}$$

As before, and further in this section the primed quantities are integrated over.

Substituting the variational derivative of G from this equation into the previous one and then replacing G^{-1} from the Dyson equation $G^{-1} = G_0^{-1} - U_{ext} - \Sigma$, one obtains:

$$L(1,4;2,3) = G(1,3)G(4,2) + G(1,1') \frac{\delta \Sigma(1',2')}{\delta U_{ext}(3,4)} G(2',2).$$

Finally, using the chain rule and defining $L_0(1,2;3,4) = G(1,4)G(2,3)$:

$$L(1,4;2,3) = L_0(1,4;2,3) + L_0(1,2';2,1') \frac{\delta \Sigma(1',2')}{\delta G(1'',2'')} L(1'',4;2'',3). \tag{2.3}$$

After some relabeling of variables this equation becomes:

$$L(1,2;3,4) = L_0(1,2;3,4) + L_0(1,2';3,1') \frac{\delta \Sigma(1',2')}{\delta G(1'',2'')} L(1'',2;2'',4).$$

If we define

$$\Xi(1',2'',2',1'') = \frac{\delta \Sigma(1',2')}{\delta G(1'',2'')},$$

we obtain:

$$L(1,2;3,4) = L_0(1,2;3,4) + L_0(1,2';3,1')\Xi(1',2'',2',1'')L(1'',2;2'',4). \tag{2.4}$$

Taking the self-energy Σ to be the sum of a Hartree term v_H and of the GW correction, assuming W to be independent of G and static, we get the standard approximation to the Bethe-Salpeter equation,

$$\begin{aligned} L(x_1,x_3,x_2,x_3|\omega) = L_0(x_1,x_3,x_2,x_3|\omega) + L_0(x_1,x_3',x_2,x_3'|\omega)v(x_3',x_5')L(x_5',x_3,x_5',x_3|\omega) - \\ - L_0(x_1,x_3',x_2,x_5'|\omega)W(x_3',x_5')L(x_5',x_3,x_3',x_3|\omega). \end{aligned} \tag{2.5}$$

Here, we have also reduced our discussion to the electron-hole part of L, which means that we set $t_3 = t_1 + \eta$ and $t_4 = t_2 + \eta$. This corresponds to the time-ordering present in the density-density correlation function.

When the W term is neglected, one obtains the Random Phase Approximation, introduced in the first chapter. Diagrammatically the Bethe-Salpeter equation can be visualized as depicted in Fig. 2.1.

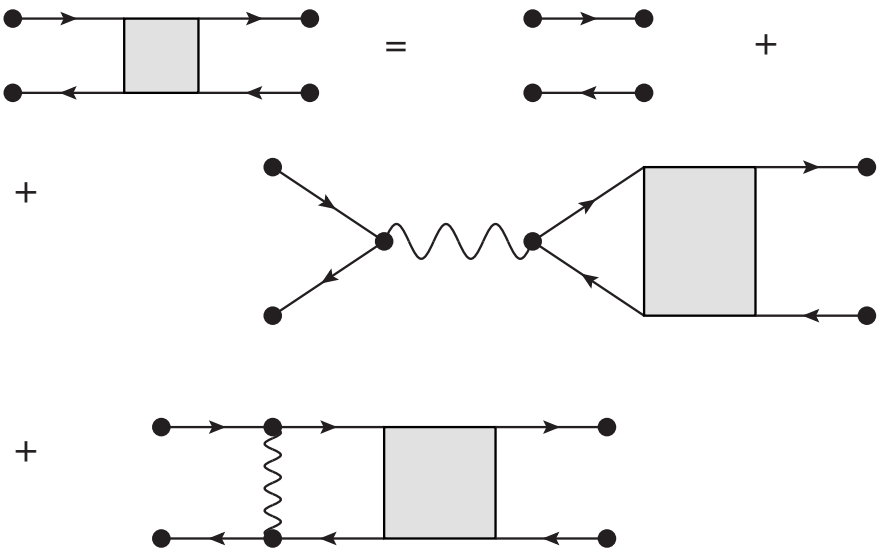


Figure 2.1 – The diagrammatic representation of the standard Bethe-Salpeter equation

Here the straight lines represent single-particle Green's functions, the horizontal wiggly line is the Coulomb interaction v_c, the vertical wiggly line is the screened coulomb interaction W and the square is L.

This equation has been widely applied to calculate properties of condensed matter systems. It is useful to rewrite the static Bethe-Salpeter equation in a transition basis.

Matrix elements in the basis of orbitals $\psi_n^*(\mathbf{x})$ gives:

$$ {}^4\chi_{n_1,n_2,n_3,n_4} = \psi_{n_1}(\mathbf{x}_1')\psi_{n_2}^*(\mathbf{x}_2')L(\mathbf{x}_1',\mathbf{x}_4';\mathbf{x}_2',\mathbf{x}_3'|\omega)\psi_{n_3}^*(\mathbf{x}_3')\psi_{n_4}(\mathbf{x}_4'). \tag{2.6} $$

With this Eqn. 2.1 becomes:

$$ {}^4\chi_{n_1,n_2,n_3,n_4} = {}^4\chi^0_{n_1,n_2,n_3,n_4} + {}^4\chi^0_{n_1,n_2,n_3',n_4'}\mathbb{K}_{n_3',n_4',n_1',n_2'}{}^4\chi_{n_1',n_2',n_3,n_4}, \tag{2.7} $$

where

$$ \begin{aligned} \mathbb{K}_{n_3',n_4',n_1',n_2'} = \psi_{n_3'}(\mathbf{x}_3')\psi_{n_4'}^*(\mathbf{x}_4')\left[v(\mathbf{x}_1',\mathbf{x}_3')\delta(\mathbf{x}_1'-\mathbf{x}_2')\delta(\mathbf{x}_3'-\mathbf{x}_4') + \right. \\ \left. + W(\mathbf{x}_1',\mathbf{x}_3')\delta(\mathbf{x}_1'-\mathbf{x}_4')\delta(\mathbf{x}_3'-\mathbf{x}_2')\right]\psi_{n_1'}^*(\mathbf{x}_1')\psi_{n_2'}(\mathbf{x}_2'). \end{aligned} \tag{2.8} $$

Note that the two summands that enter the Kernel have two delta functions, but with different arguments. The difference explains why there is a horizontal and a vertical line in Fig. 2.1. Without the vertical line (W), one could close the end-points of the diagrams and have a closed equation for χ. Because of the presence of W, this not possible.

The Bethe-Salpeter equation in its matrix form is the basis for most ab-initio calculations of optical properties. In semiconductors and insulators one can show that only couples (n_1,n_2) and (n_3,n_4) where one state is occupied, and the other unoccupied contribute to ${}^4\chi$ and ${}^4\chi_0$. These couples are also called transitions. Dependent on whether the first element of the couple is occupied or unoccupied, they correspond to transitions of positive and negative frequency. Because of the matrix inversions, all these couples mix. This is called a coupling of transitions. Often one neglects the mixing of transitions of positive and negative frequencies: this is called the Tamm-Dancoff Approximation. If one further restricts oneself to transitions of positive frequency only this is called the resonant-only approximation.

2.1.2 Application to Silicon

In the present subsection we apply the Bethe-Salpeter Equation using state of the art codes to a prototypical semi-conductor: Silicon (Fig. 2.1.2). A much more detailed analysis of various results for this material is left to the next chapter. Here we just use this example to point out some basic properties. The two codes we shall be using to calculate the spectra are DP and EXC [77]. The ground state calculation was performed in Abinit [78]. For computational parameters we refer to Appendix A.1.

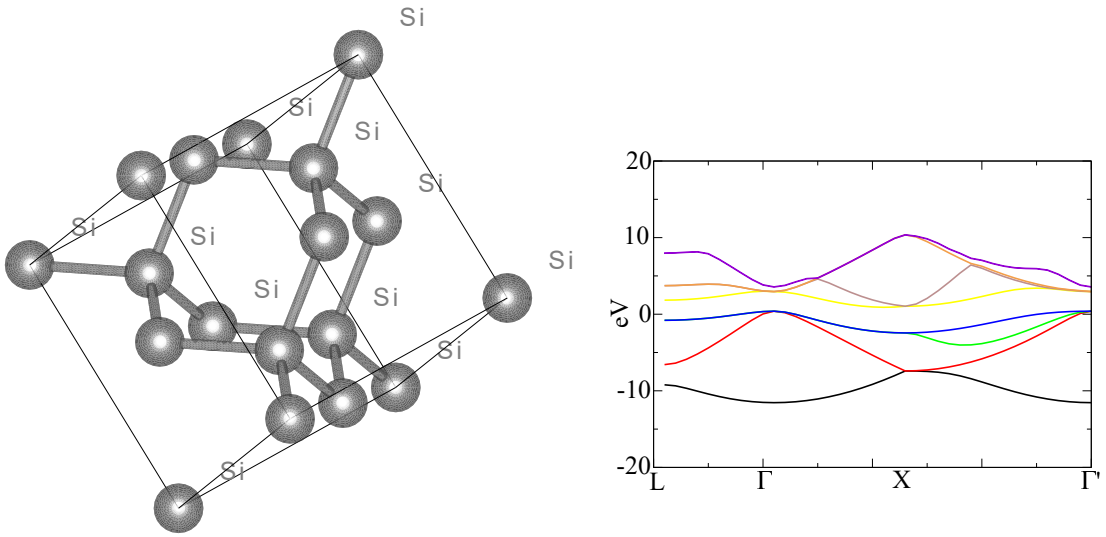


Figure 2.2 – Crystalline structure of bulk Silicon (left); Band structure of Silicon(right)

Ground state

The properties of the ground state were calculated within the Local Density Approximation to Density Functional Theory. We thus obtained the independent particle Kohn-Sham Green's Function, using the wave functions and energies, of the Kohn-Sham system. However, the Bethe-Salpeter Equation, contains the interacting Green's function. To obtain a good approximation to it, we calculated the self-energy corrections to the quasiparticle energies within the $\Sigma = iGW$ approximation. This was done, using the Abinit code. The results obtained were in good agreement with previous results from literature [79]. As some of the calculations in the following require a large number of bands, the effect of the self-energy was replaced by a simple shift of the valence with respect to the conduction bands. This is also known as the scissor operator approximation to the self-energy. In the following we are interested in the loss function. For Silicon, it is dominated by one strong plasmon excitation a small $\mathbf{q}$, that decays into the electron-hole continuum at larger $\mathbf{q}$, as observed in Figures 2.3 to 2.7, where we show the calculated loss function $-\Im\left[\varepsilon_M^{-1}(\mathbf{q},\omega)\right]$ for different values of $\mathbf{q}$.

Fig. 2.3 shows the loss function calculated in the RPA using GW energies. At $\mathbf{q} = 0$ the plasmon is close to 18 eV and it moves to higher energies with increasing $\mathbf{q}$, consistently with experimental observations. The figure also shows the importance of including the lowest valence band. Note that this is different from calculations of the $\Im[\varepsilon_M]$ where much less conduction and valence bands are required.

Finally, we perform a Bethe-Salpeter calculation, based on the GW approximation to

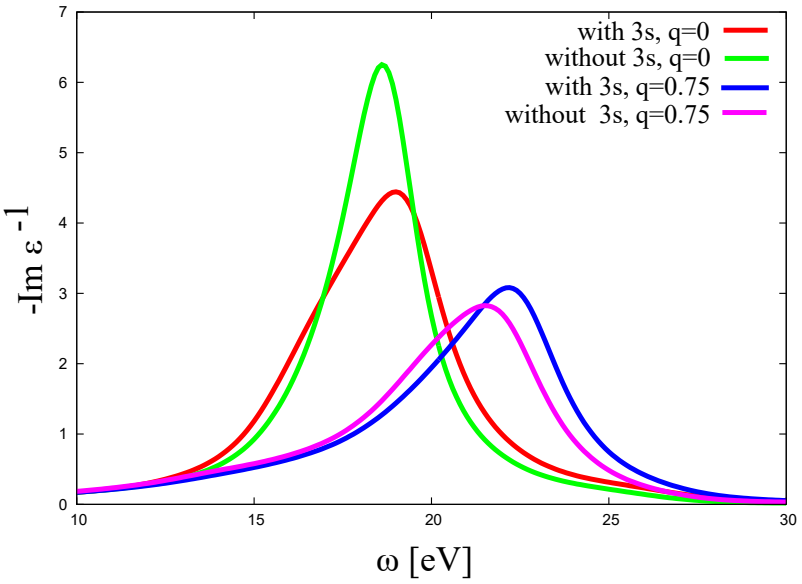


Figure 2.3 – Bulk Silicon: loss function within the RPA using GW energies with and without the 3s band for $q = 0$ and $q = 0.75$

the self-energy. An extensive convergence study is given in details in Appendix A.2.The results for $q = 0$ are shown in Fig. 2.4. They agree with [80]. In particular the BSE shifts spectral weight to lower energies compared to the RPA.

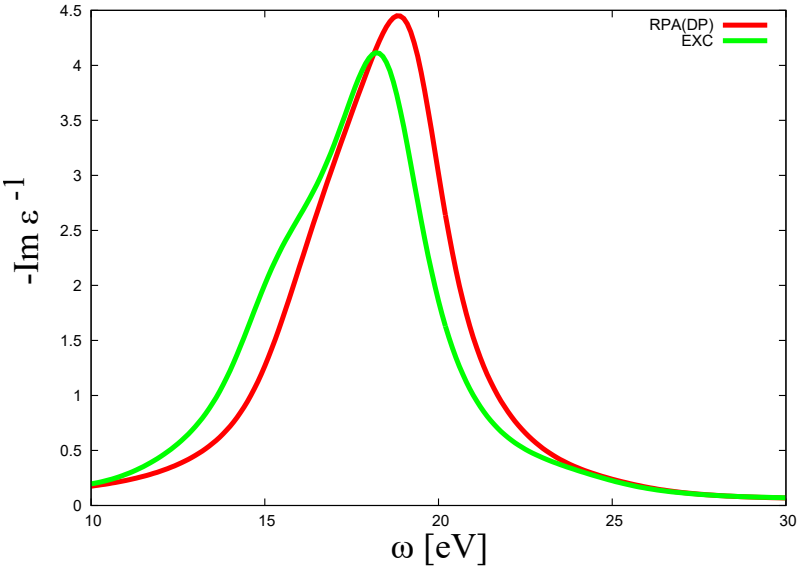
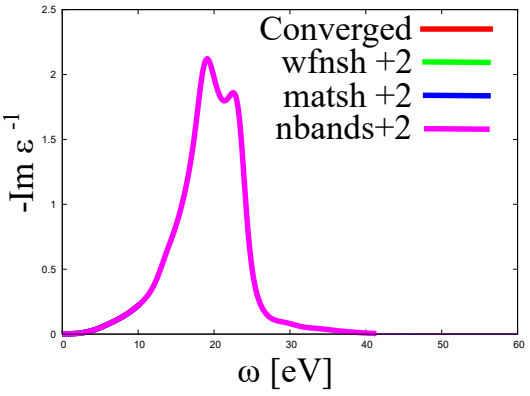


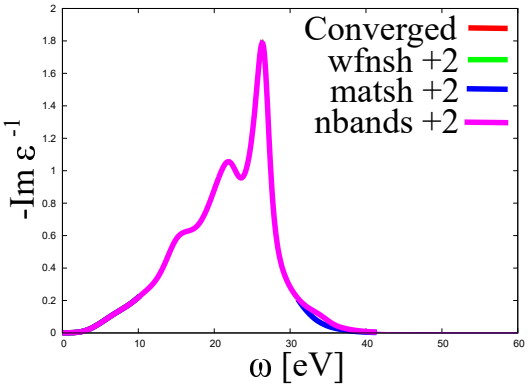
Figure 2.4 – Loss function of Silicon, comparing RPA to the BSE(EXC) at $q = 0$ (30 bands)

We, furthermore perform the BSE calculation for $q \neq 0$, which has become possible due to recent developments of the EXC code [81]. The results of this calculation are

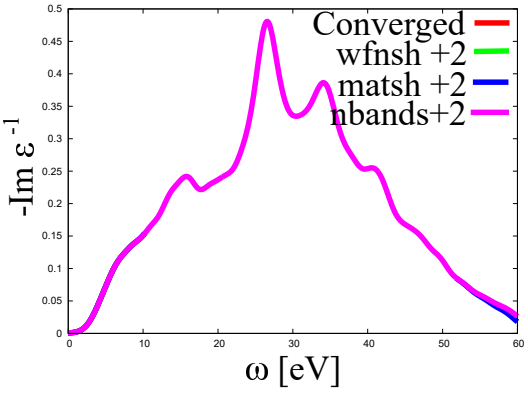
shown in Fig. 2.5(a), Fig. 2.5(b) and Fig. 2.5(c). We have also plotted the spectra obtained by increasing the main convergence parameters. The fact that there is little difference means that all these spectra are converged.



(a) EELS $q = 0.5$



(b) EELS $q = 0.75$



(c) EELS $q = 1.25$

Figure 2.5 – Loss Function in BSE(EXC) in the resonant-only approximation for different q in the [1,1,1] direction. Apart from the final spectra, we have also plotted the spectra, obtained by increasing the number of plane waves used to describe the wavefunction(wfnsh), the dielectric matrix (matsh) and the number of bands(nband)

However, these results have been obtained by neglecting the coupling between the positive and negative transition frequencies. For realistic comparison with experiments, one has to perform the full Bethe-Salpeter calculation including this coupling. This is discussed in the next chapter. To this end the current implementation was parallelized. For details see Appendix B.1.

One would have to take take into account the fact that, even within the simple $\Sigma = iGW$ approximation the quasi-particles acquire a finite life-time. This life-time (LT) effect has been extensively discussed in [82] in the framework of TDDFT. Here, for Silicon, we are

interested in some questions of principle rather than pushing the results close to experiment. Therefore we do not add lifetimes, although this can be done in the BSE framework [83].

Fig. 2.6 shows results obtained in EXC with and without coupling for q=0.75 in the [1,1,1] direction (this corresponds to a transferred momenta of 80 a.u.) differ considerably. This shows a failure of the resonant-only approximation to the Bethe-Salpeter Equation for the calculation of the loss function.

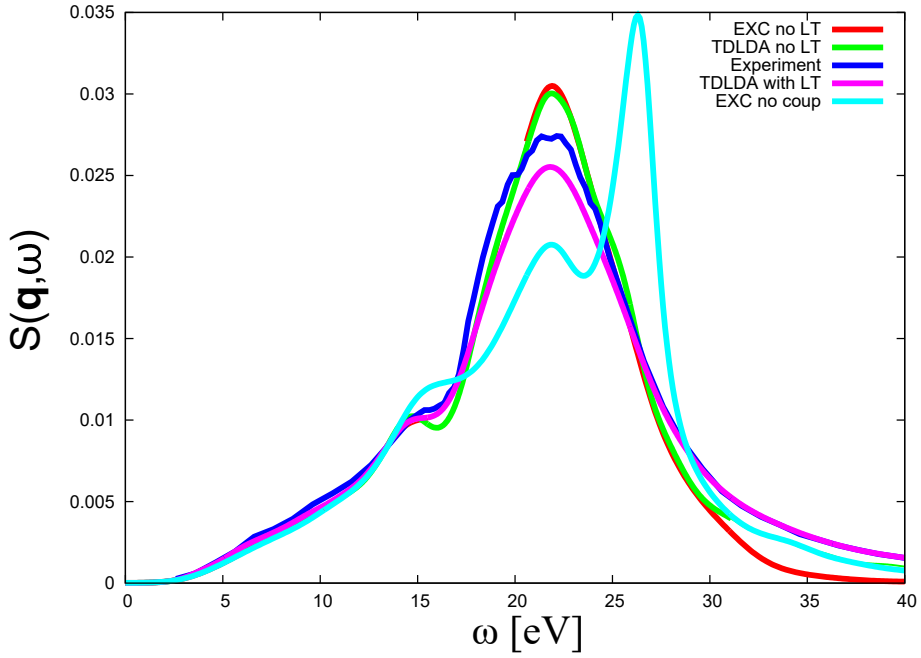


Figure 2.6 – Dynamic structure factor of Si at q=0.75 in the [1,1,1] direction using different methods: BSE(EXC) with and without coupling and without life-time effect,TDLDA with and without life-time effect, experiment from [82]

Indeed, the figure shows that the BSE results including the coupling are much closer to the experiment. The figure also contains TDLDA results.

Surprisingly, the agreement between TDLDA and BSE with coupling is very good, apart from the higher energy region, which is due to the limited number of occupied bands taken into account. The closeness of BSE and TDLDA is even better seen in Fig. 2.7, where we compare two spectra computed within the TDLDA and BSE formalisms, but with the same set of convergence parameters. We will further discuss the implication of this in the next chapter. When lifetime effects are included in the TDLDA, the agreement with experiments is very good. This strongly suggests they should be also included in the BSE in the future. It is worth noting, that inserting an imaginary part computed within the

GW approximation to the self energy into the denominator G, to account for the life-time effects, is more natural, than using a modified f_{xc}^{LT} in TDLDA, as in [82].

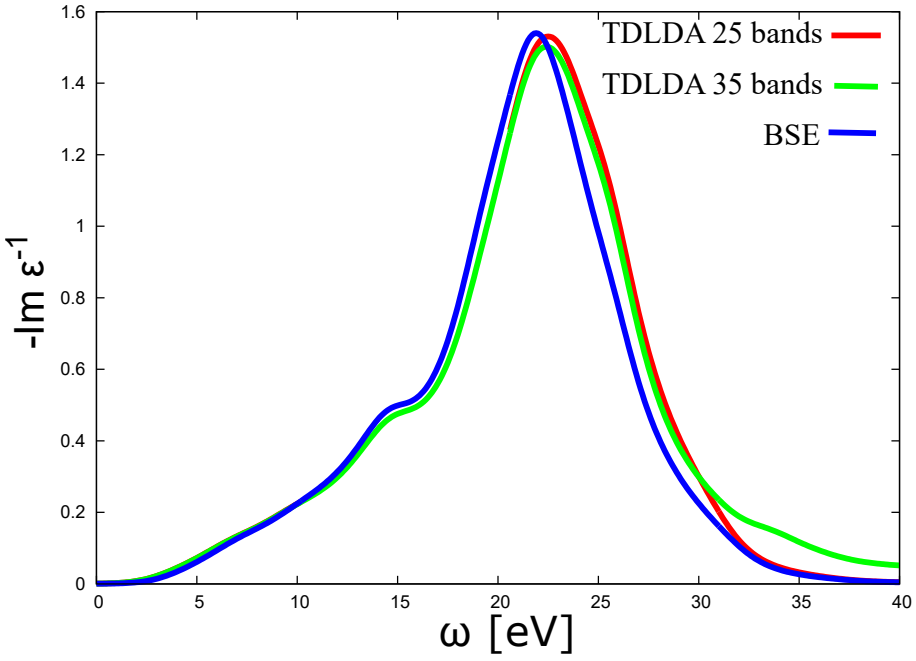


Figure 2.7 – BSE(EXC) and TDLDA Loss functions of bulk Si for q=0.75 in [1,1,1] direction.

2.1.3 Discussion

From the above results we see that in the case of loss spectra of silicon two very different approaches, TDLDA and BSE, give very similar results. This is not always true, but since TDLDA is computationally much simpler it is important to analyze this situation, This shall be done in Chapter 3.

For now, let us look into how the Bethe-Salpeter approach works. This can be done by visualizing the first order diagrams for χ, that are all representatives of much larger classes of diagrams. The diagrams are shown in Fig. 2.8.

The first step we do is a DFT (LDA) calculation. It can be thought of as an elaborate Hartree calculation, corresponding to the lower right diagrams. The next step we do is a GW calculation. It gives us roughly a screened Hartree-Fock correction to the wave-function and energies. This is already a much more demanding calculation. It corresponds to the lower left diagrams. Finally we take into account the excitonic effects via the BS equation which accounts for the two upper diagrams. This again is a rather cumbersome calculation. Moreover, at each of these steps we make different approximations, and no

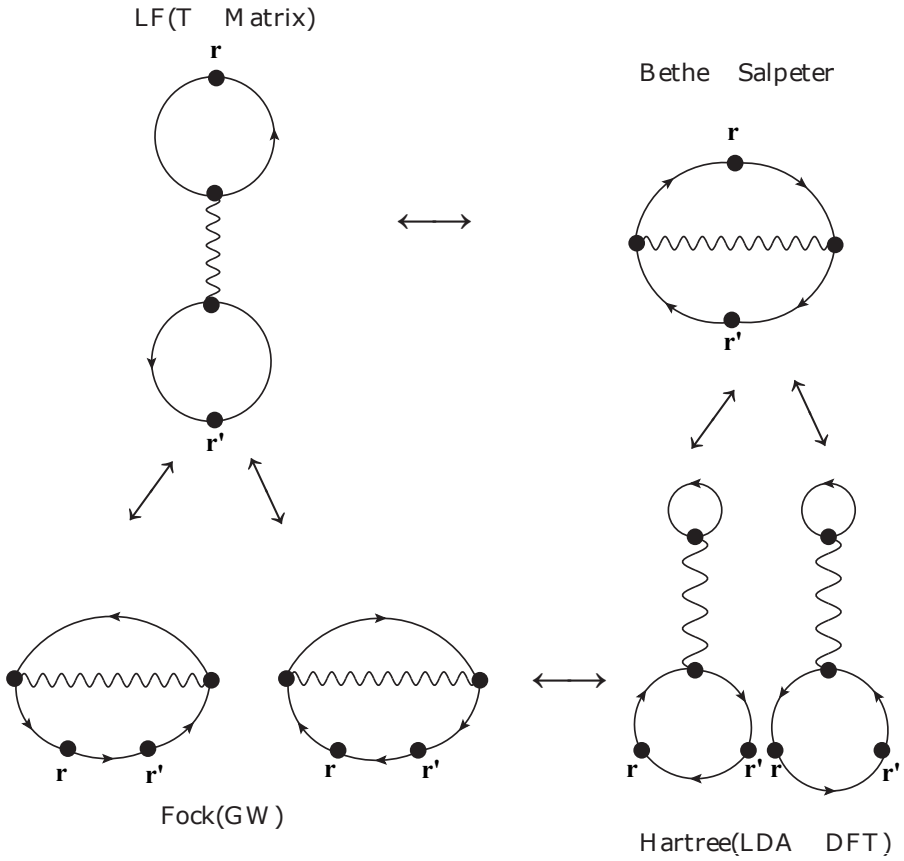


Figure 2.8 – 1st order diagrams for χ. All the wiggly lines are bare Coulomb interactions.

one guaranties us, that they will all be consistent.

One other drawback of such a complicated approach can be easily seen in the one-electron limit presented in the first chapter. Let us assume that we want to describe the excitation of this system (Fig. 2.9) in the Bethe-Salpeter framework.

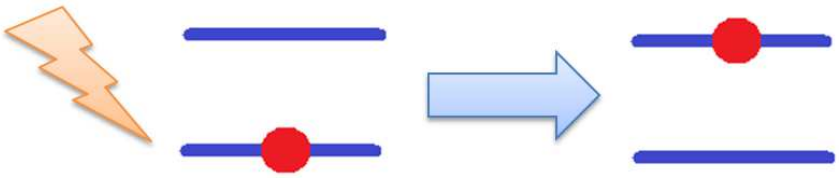


Figure 2.9 – Simple excitation of a one-electron system

To this end we will first calculate the single particle Green's function. Its poles describe the addition and removal processes (Fig. 2.10). It is important to note that in the case of the electron addition process, electron repulsion and screening occurs, on the one

hand, the transition frequency will differ from that, of a non-interacting system. On the other hand, we know that in the exact result Eqn. 1.41 the susceptibility of the single electron system doesn't depend on the electron-electron interaction. This means that the Bethe-Salpeter Kernel cancel the effect of the interaction in the electron addition part. In other words, when applying the Bethe-Salpeter Equation formalism to the single electron system we perform two complicated calculations, that in the end cancel each other. In the following sections we will derive alternative equations for the connected part of the two-particles Green's function, which could account for this cancelation, an thus avoid useless calculations.

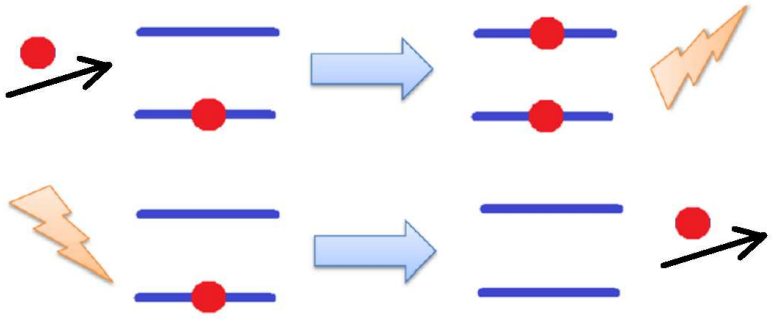


Figure 2.10 – Processes described by the single particle Green's function

2.2 New BSE. First attempts

Our goal in the present section is to try and find new alternative equations for the two-particle correlation function, L. As a first option we consider ways of writing an equation similar to 2.4, but with a Kernel that is easier to approximate.

2.2.1 Derivation

Let us now consider different schemes of modifying the standard derivation of the Bethe-Salpeter equation, which was described in the first section of this chapter:

— The first one is not to pass through the inversion of G (Eqn. 2.2)

— The second lies in the subdivision of the full Kernel of the Dyson equation $\Sigma_U = U_{ext} + \Sigma$ into two parts, namely U_{ext} and the rest, solving the Dyson equation for the first part with solution G_U, and, finally, taking $G^{-1} = G_U^{-1} - \Sigma$

— The third and last one relies on the subdivision $\Sigma_1 = U_{ext} + v_H$ and $\Sigma_2 = \Sigma_{XC}$, to which we apply the same two step procedure as in the second case.

Case 1

Taking the derivative of the Dyson equation $G = G_0 + G_0 U_{ext} G + G_0 v_H G + i G_0 \Sigma_{XC} G$, with respect to U_{ext}, and then setting U_{ext} to zero we get:

$$L(1,4;2,3) = \frac{\delta G(1,2)}{\delta U_{ext}(3,4)} = G_0(1,3)G(4,2) + G_0(1,1')\Xi(1',4',2',3')L(3',4,4',3)G(2',2) + \\ + G_0(1,1')\Sigma(1',2')L(2',4,2,3) \tag{2.9}$$

This is the first alternative formulation of the Bethe-Salpeter Equation.

Case 2

Writing down the two coupled Dyson equations:

$$G_U = G_0 + G_0 U_{ext} G_U \tag{2.10}$$
$$G = G_U + G_U \Sigma G, \tag{2.11}$$

from Eqn. 2.10 one obtains that:

$$\frac{\delta G_U^{-1}(1,2)}{\delta U_{ext}(3,4)} = -\delta(1-3)\delta(2-4)$$
$$\frac{\delta G_U(1,2)}{\delta U_{ext}(3,4)} = G_U(1,3)G_U(4,2).$$

Using this and taking the derivative of G from Eqn. 2.11, and finally setting $U_{ext} = 0$, we obtain:

$$L(1,4;2,3) = \frac{\delta G(1,2)}{\delta U_{ext}(3,4)} = G_0(1,3)G_0(4,2) + G_0(1,3)G_0(4,1')\Sigma(1',2')G(2',2) + \\ + G_0(1,1')\Xi(1',2'',2',1'')L(1'',4,2'',3)G(2',2) + G_0(1,1')\Sigma(1',2')L(2',4,2,3) \tag{2.12}$$

Case 3

Similarly, as in the previous case we obtain that

$$\frac{\delta G_H^{-1}(1,2)}{\delta U_{ext}(3,4)} = -\delta(1-3)\delta(2-4) + iv(1,2')L(2',4,2',3)\delta(1-2)$$
$$\frac{\delta G_H(1,2)}{\delta U_{ext}(3,4)} = G_H(1,3)G_H(4,2) - iG_H(1,1')G_H(1',2)v(1',2')L(2',4,2',3)$$

Using these relations we obtain, for L:

$$L = \frac{\delta G}{\delta U_{ext}} = G_H G_H - iG_H G_H vL + (G_H G_H - iG_H G_H vL)\Sigma_{XC} G + \\ + G_H \Xi_{XC} LG + G_H \Sigma_{XC} L \tag{2.13}$$

One again we have set $U_{ext} = 0$ in the last equation. Here we have omitted the arguments in order to display the formula in a compact way.

2.2.2 Analysis

In the first chapter, when looking at the single-electron limit of the electron-hole part of L, we noted that the interacting χ was equal to the one obtained by simply convoluting two non-interacting Green's function. In practice this means that in the BSE equation Eqn. 2.4 the effect of going beyond G_0G_0 in L_0 i.e. $G_0 \to G$ is fully canceled by the second term, containing the BSE Kernel. Let us look at whether this cancelation is explicit in our new equations for the two-particle correlation function.

To this end, let us analyze the structure of equations Eqn. 2.3, Eqn. 2.9, Eqn. 2.12, Eqn. 2.13, or more precisely the structure of poles on the left hand side (exact L) and right hand side (equation fo L) for the case of one electron. To make this structure apparent, we set $t_3 = t_4 + \eta$ and $t_2 = t_1 + \eta$. These equations will then depend on two times, or equivalently one frequency. On the left we have poles for $\omega = \varepsilon_1 - \varepsilon_{k,v}$, the differences between the ground state and the other 1-particle excited states. The right hand side should therefore also have only these. However L in Eqn. 2.3 has poles of the type $\varepsilon^e_{v,k} - \varepsilon^h$, that contain energies of two interacting electrons(Eqn. 1.40), which should be canceled by the more complicated term containing Ξ. A similar situation happens for Eqn. 2.9 and Eqn. 2.13, where we have to cancel the poles of convolutions of G and G_0 or G_H. Only in the case Eqn. 2.12 do we have an expression G_0G_0, which in the frequency domain has the same non-interacting poles, but still the interacting G appears in the equation.

Therefore, none of these equations are simpler and only one of them seems to make

the 0-order term explicit, but without eliminating the need to calculate the interacting G. We, therefore, proceed to more elaborate possibilities of rewriting the Bethe-Salpeter Equation.

2.3 New BSE. Iterating the two particle correlation function L

Instead of trying to find a simple expression involving Ξ, let us consider an alternative formulation of the Bethe-Salpeter equation, as a functional differential equation.

2.3.1 Derivation

To do this we will use the definition $L = \frac{\delta G}{\delta U_{ext}}$. Before doing this we will partially solve the Dyson equation Eqn. 1.29 using the same splitting of the self-energy, as in the previous section. As there are no simple ways for solving equations with functional derivatives, we will then iterate the obtained equations.

Case 2

Taking the derivative of G, defined by Eqn. 2.11, as we did in Eqn. 2.12, without setting $U_{ext} = 0$, we obtain the following expression:

$$\begin{aligned} L(1,4;2,3) = \frac{\delta G(1,2)}{\delta U_{ext}(3,4)} &= G_U(1,3)G_U(4,2) - \\ &- G_U(1,3)G_U(4,1')iv(1',2')G(2',2'^+)G(1',2) - \\ &- G_U(1,1')iv(1',2')L(2',4;2'^+,3)G(1',2) - \\ &- G_U(1,1')iv(1',2')G(2',2'^+)L(1',4;2,3) + \\ &+ G_U(1,3)G_U(4,1')iv(1',2')L(1',2';2,2') + \\ &+ G_U(1,1')iv(1',2')\frac{\delta L(1',2';2,2')}{\delta U_{ext}(3,4)}. \end{aligned} \tag{2.14}$$

As already stated, there is, unfortunately no simple expression for the functional derivative contained in this equation and we cannot solve the functional differential equation. Therefore we iterate the equation, to obtain approximations.

In Eqn. 2.14 we can set $3 = 4^+$. Setting $\frac{\delta L}{\delta U_{ext}}$ and L to zero in the r.h.s. we obtain the

first approximation for L:

$$L^{(0)}(1,3;2,3^+) = G_U(1,3^+)G_U(3,2) - G_U(1,3^+)G_U(3,1')iv(1',2')G(2',2')G(1',2).$$

The derivative yields

$$\begin{aligned}\frac{\delta L^{(0)}}{\delta U_{ext}} &= G_U(1,4^+)G_U(4,3^+)G_U(3,2) + G_U(1,3^+)G_U(3,4^+)G_U(4,2) - \\ &- G_U(1,4^+)G_U(4,3^+)G_U(3,1')iv(1',2')G(2',2')G(1',2) - \\ &- G_U(1,3^+)G_U(3,4^+)G_U(4,1')iv(1',2')G(2',2')G(1',2) + \\ &+ G_U(1,3^+)G_U(3,1')v(1',2')\chi^{(0)}(2',4)G(1',2) - \\ &- G_U(1,3^+)G_U(3,1')iv(1',2')G(2',2')L^{(0)}(1',4,2,4^+).\end{aligned}$$

Substituting these into Eqn. 2.14 we obtain L up to second order. However this does not give any simple expressions and the expansions we obtain are not in full orders of v.

The other option is to take $G_U G_U$ as the 0-order approximation:

$$\begin{aligned}L^{(0)}(1,3;2,3^+) &= G_U(1,3^+)G_U(3,2) \\ \frac{\delta L^{(0)}}{\delta U_{ext}} &= G_U(1,4^+)G_U(4,3^+)G_U(3,2) + G_U(1,3^+)G_U(3,4^+)G_U(4,2).\end{aligned}$$

In this case the first order is quite simple:

$$\begin{aligned}L^{(1)}(1,3;2,3^+) &= L^{(0)}(1,3;2,3^+) - G_U(1,3^+)G_U(3,1')iv(1',2')G(2',2'^+)G(1',2) - \\ &- G_U(1,1')iv(1',2')G_U(2',3^+)G_U(3,2'^+)G(1',2) - \\ &- G_U(1,1')iv(1',2')G(2',2'^+)G_U(1',3^+)G_U(3,2) + \\ &+ G_U(1,3^+)G_U(3,1')iv(1',2')G_U(1',2'^+)G_U(2',2) + \\ &+ G_U(1,1')iv(1',2')\left[G_U(1',3^+)G_U(3,2'^+)G_U(2',2) + G_U(1',2'^+)G_U(2',3^+)G_U(3,2)\right].\end{aligned}$$

If we now send $U \to 0$, take the first order in v in all the functions and look at the susceptibility we get:

$$
\begin{aligned}
L^{(1)}(1,3;1^+,3^+) = L^{(0)}(1,3;1^+,3^+) - G_0(1,3^+)G_0(3,1')iv(1',2')G_0(2',2'^+)G_0(1',1^+) - \\
- G_0(1,1')iv(1',2')G_0(2',3^+)G_0(3,2'^+)G_0(1',1^+) - \\
- G_0(1,1')iv(1',2')G_0(2',2'^+)G_0(1',3^+)G_0(3,1^+) + \\
+ G_0(1,3^+)G_0(3,1')iv(1',2')G_0(1',2'^+)G_0(2',1^+) + \\
+ G_0(1,1')iv(1',2')\left[G_0(1',3^+)G_0(3,2'^+)G_0(2',1^+) + G_0(1',2'^+)G_0(2',3^+)G_0(3,1^+)\right].
\end{aligned}
$$

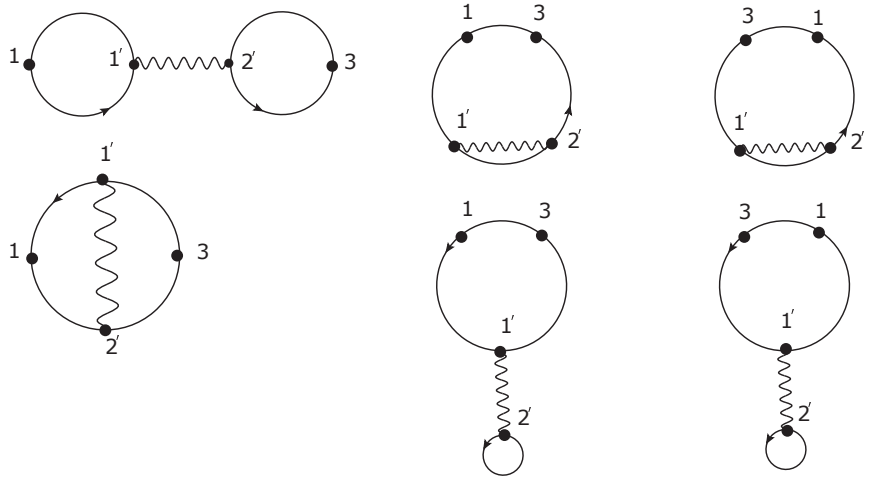


Figure 2.11 – 1st order diagrams in the 2nd case. The wiggly lines are bare Coulomb interactions.

The graphical representation of these terms can be seen in Fig. 2.11. The terms can be written down explicitly.Since all the Green's functions are non-interacting. Taking into account the ordering of times ($t_{1'} + \varepsilon = t_{2'}$) and going to the single-electron limit we obtain that the contribution of these terms is zero. This is not surprising since we know that the final result should be χ_0, but is a good sign with respect to the new formulations suggested in the previous section, where this cancelation was not explicit.

Let us now go into a deeper analysis of which terms cancel which. The first thing to note is that χ is symmetric with respect to the change $1 \leftrightarrow 3$. This results in a symmetry on the diagrammatic level i.e. for each diagram we can find its dual under this transformation. On Fig. 2.11 the diagrams in the leftmost column are self-dual, and the ones in the remaining two are dual to each other. It is interesting to note that these diagrams might come from different terms e.g. the last two terms in the first line come from $G_U G_U vL$ and $G_U v\frac{\delta L}{\delta U_{ext}}$ respectivly. This therefore can give us insight on what things can be found in the derivative term without computing it explicitly. Moreover all the diagrams can be

split into two groups (self-dual terms will be in both), with cancelations inside of each one of the groups, but not between groups. This points out that it might be interesting to have a symmetric Bethe-Salpeter equation, where the two points will be equivalent. It is important to note, however, that this might cause some other exact properties to break down [84]. In particular the approximation may not fulfill conservation laws.

One more thing worth noting is that the terms in the first column are terms that come from the electron-hole interaction and the other two can be thought of as simply the renormalization of the fermionic propagator, where the upper terms correspond to an exchange term $\Sigma_{G_U v} = iG_U v$ and the lower ones to a Hartree-type correction $\Sigma_{H_0} = vn_U$.

If one now evaluates the second order in v one gets new terms contributing to L. Most of them can be thought of as a simple renormalizations of the 1st order terms. These come from the already mentioned corrections to the fermion propagator, but also for the first time the interaction starts to be corrected, and one finds an effective:

$$W_1(1,2) = v(1,2) + v(1,1')G_U(1',2')G_U(2',1')v(2',2)$$

. This introduces screening.

Case 3

Let us write down a new equation for L, obtained after solving the Dyson equation for $\Sigma_H = U_{ext} + v_H$ and then using this solution, as a starting point. This gives us the following formula for the derivative of the Hartree Green's function:

$$\frac{\delta G_H(1,2)}{\delta U_{ext}(3,4)} = G_H(1,3)G_H(4,2) - iG_H(1,1')G_H(1',2)v(1',2')L(2',4,2',3).$$

Using this and the Dyson equation for the interaction Green's function, we obtain:

$$\begin{aligned} L(1,3;2,3^+) &= G_H(1,3^+)G_H(3,2) + G_H(1,1')G_H(1',2)v(1',2')\chi(2',3) + \\ &+ \left[G_H(1,3^+)G_H(3,1') + G_H(1,3')G_H(3',1')v(3',4')\chi(4',3)\right] iv(1',2')L(1',2';2,2') + \\ &+ G_H(1,1')iv(1',2')\frac{\delta L(1',2';2,2')}{\delta U_{ext}(3)}. \end{aligned} \tag{2.15}$$

Here we have already contracted 3 and 4, as in the previous case. Note here that for the first time the interacting one-particle G doesn't appear. The 0-order approximation, starting from the Hartree Green's function G_H is rather simple:

$$
\begin{aligned}
L^{(0)} &= G_H(1,3^+)G_H(3,2) \\
\frac{\delta L^{(0)}}{\delta U_{ext}} &= G_H(1,4^+)G_H(4,3^+)G_H(3,2) + G_H(1,3^+)G_H(3,4^+)G_H(4,2) + \\
&\quad + G_H(1,1')G_H(1',3^+)v(1',2')\chi(2',4)G_H(3,2) + \\
&\quad + G_H(1,3^+)G_H(3,1')G_H(1',2)v(1',2')\chi(2',4).
\end{aligned}
$$

One can also rewrite everything in terms of a screened function,

$$
L_H^S(1,3,2,3^+) = G_H(1,3^+)G_H(3,2) + G_H(1,1')G_H(1',2)v(1',2')\chi(2',3), \tag{2.16}
$$

for which Eqn. 2.15 becomes:

$$
\begin{aligned}
L(1,3,2,3^+) &= L_H^S(1,3,2,3^+) + L_H^S(1,3,1',3^+)iv(1',2')L(1',2';2,2'^+) + \\
&\quad + G_H(1,1')iv(1',2')\frac{\delta L(1',2';2,2'^+)}{\delta U_{ext}(3)}.
\end{aligned} \tag{2.17}
$$

Note that for $2 = 1^+$ the self-consistent solution of Eqn. 2.16 alone corresponds to a time-dependent Hartree approximation.

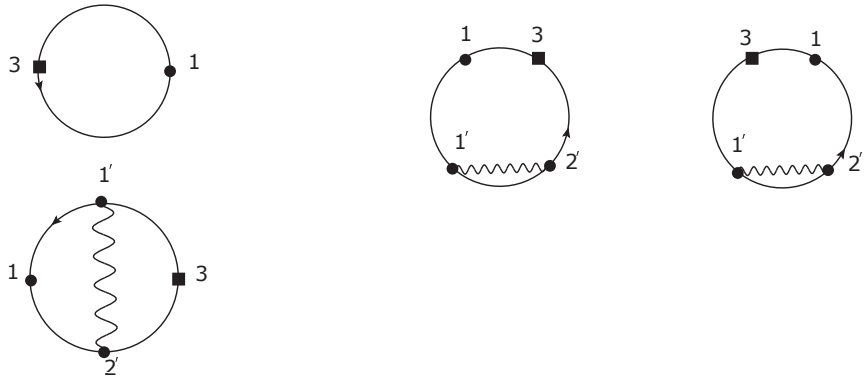


Figure 2.12 – 1st iteration diagrams in the 3rd case, for Eqn. 2.17 with $2 = 1^+$

The graphical representation of the first iteration can be seen in Fig. 2.12. Here we have introduced a notation corresponding to Eqn. 2.16, which in the first iteration can be represented as Fig. 2.13. In both these figures the lines are no longer the bare propagators, but the Hartree ones.

If we look at the first order we see that the diagrams no-longer present are exactly the two rightmost diagrams in the last row of Fig. 2.11, which as mentioned earlier correspond to Σ_{H_0}. The advantage of this equation is the fact that it does not have any explicit

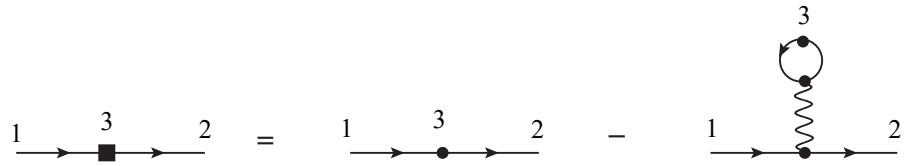


Figure 2.13 – $L_H^S(1,3,2,3^+)$ to 1st order. The square corresponds to the contacted coordinate 3.

dependence on G. This is very important, since we know that G is itself a complicated object. Moreover, as seen in the previous section, effects of going from G_0 to G are at least partially canceled in the electron-hole Bethe-Salpeter Equation. This is physically motivated, as in an absorption process there shouldn't be any electron addition or removal. It is worth noting, that as soon as one tries to go beyond the self-energy containing the Hartree term and terms with convolutions of Hartree Green's functions, the full G reapers in the equation for the two-particle Green's function. Therefore the new set of equations Eqn. 2.17 and Eqn. 2.16 derived here can be considered as the best one can get without having to go through the step of calculating G.

As in the previous case we can find the second iteration and the second order L. It will have less diagrams due to the fact that it now contains G_H instead of G_0.

An alternative approximation to iterating our equations Eqn. 2.16 and Eqn. 2.17 is to abandon self-consistency between the two of them, for example approximating χ in the first one of them, and then obtain the corresponding L from the second one.

2.3.2 Analysis

Our final goal is to find ε_M or alternatively χ. To do this we are brought to the fact that we have to make some approximations to Eqn. 2.16 and Eqn. 2.17, as we cannot solve them exactly. For instance, we can look at the 0-order approximation and use $\chi = \chi_H$ in the first equation. This will give us χ_0^H and χ_{RPA}^H, which can be used as starting points for further approximations. As a test material we can look at the absorption spectra of Silicon. To understand how good or bad are our results are, we compare them to the standard LDA and GW from Abinit.

To find the exact G_H, one needs the exact (interacting) density, that enters the corresponding self-energy. We can obtain a good approximation to it due to the fact that DFT [32, 33] gives us in principle the exact density. In Fig. 2.14 we see the resulting absorption spectra obtained using the DP code [77] within the Random Phase Approximation. The spectra are plotted for a self-consistent LDA-DFT calculation, a self-consistent Hartree calculation ($v_{xc} = 0$) and non-self-consistent Hartree with an LDA-DFT self-consistent

density. The calculations are performed with a coarse k-point grid and should be used for the comparison of methods done here. As it is known from previous calculations [23] the LDA-RPA spectrum is at too low energies. The Hartree results are even worse. In particular, Hartree with a LDA-DFT density puts spectral weight at very low energy.

To understand the origin of this, Fig. 2.15 shows the band structures. In the Hartree approximation the indirect gap negative open and therefore there is nothing surprising in the incorrectness of the spectra. Note that the low energy peak should in principle also be present in the spectra of the non-self-consistent Hartree calculation, but it is absent due to the fact that the sampling is not dense enough, to capture the excitations around the very small region around the Γ point.

To open the gap, we must take into account some self-energy corrections. They do not have to be as large as the *GW* ones, as we know that there is cancelation, however being in a metallic state is clearly a problem.

2.4 Self-consistent system of equations

In the previous section we have derived alternative equations for the two particle correlation function, L. The most promising of them are the couple Eqn. 2.17 and Eqn. 2.16. Here we rewrite them in a different manner, so as to put forward some quantities we already know. Furthermore, we then make approximations to these equations and analyse them.

2.4.1 Derivation

We start from equations Eqn. 2.17 and Eqn. 2.16. We note that Eqn. 2.16 can be rewritten using $\varepsilon^{-1} = 1 + v\chi$:

$$L_H^S(1,3,2,3^+) = G_H(1,1')\varepsilon^{-1}(1',3)G_H(1',2).$$

If we multiply this by $v(3,3')$ and integrate we get W, the screened potential instead of ε^{-1}. We note that the same kind of integration occurs in Eqn. 2.17. Therefore defining a new quantity $M(1,3,2) = -iv(3,3')L(1,3';2,3')$, which in the limit $1 \to 2$ becomes $v\chi$ we get the following set of equations:

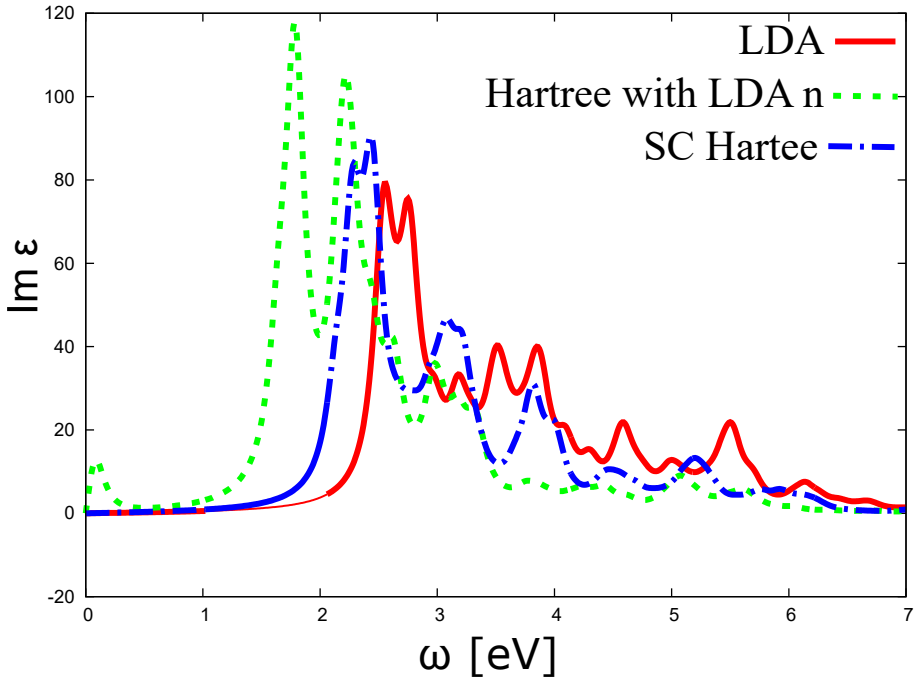


Figure 2.14 – Bulk Silicon Im ε (q=0) in the RPA approximation, including the local fields effect, obtained using the DP code, for LDA, self-consistent Hartree and Hartree with LDA density

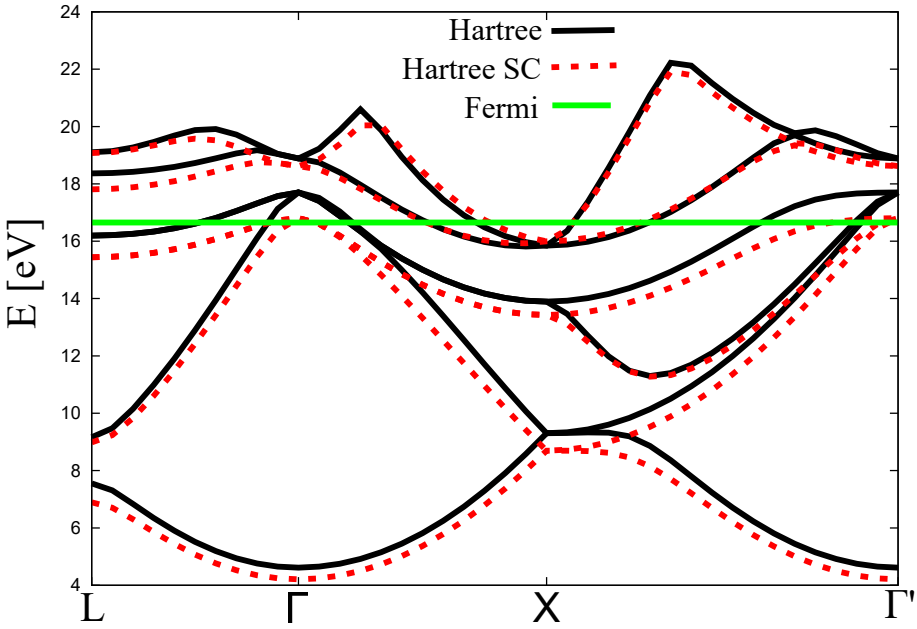


Figure 2.15 – *Ab-initio* bulk Silicon band structure within the Hartree and Hartree self-consistent approximations. The Fermi levels differ by a couple of hundreds of meV.

$$
\begin{aligned}
W(1,2) &= v(1,2) + v(1,1')M(1',2,1'^{+}) \\
M_H^S(1,3,2) &= -iG_H(1,1')W(1',3)G_H(1',2) \\
M(1,3,2) &= M_H^S(1,3,2) - M_H^S(1,3,1')M(1',1',2) - G_H(1,1')\frac{\delta M(1',1',2)}{\delta \tilde{U}(3)}.
\end{aligned}
$$

Here the functional derivative is taken with respect to $\tilde{U}$ in which we absorbed $v(3,3')$, defining $\tilde{U}$ using ($-iv\frac{\delta M}{\delta U_{ext}} = \frac{\delta M}{\delta \tilde{U}}$), as v is independent of U.

We note that this would not have been possible in the standard Bethe-Salpeter equation Eqn. 2.3, as there the structure of indices in the term containing the Kernel is different. This set of equations can be itself rewritten if we note that we can send $3 \to 1$ in the last one of them and get a closed equation for this new $S^W(1,2) = M(1,1,2)$. A different contraction defines $M(1,2) = M(1,2,1^+) = v(1,1')\chi(1',2)$. Finally, we get the following set of equations:

$$W(1,2) = v(1,2) + M(1,1')v(1',2) \tag{2.18-a}$$

$$S_0^W(1,2) = -iG_H(1,1')W(1',1)G_H(1',2) \tag{2.18-b}$$

$$S^W(1,2) = S_0^W(1,2) - S_0^W(1,1')S^W(1',2) - G_H(1,1')\frac{\delta S^W(1',2)}{\delta \tilde{U}(1)} \tag{2.18-c}$$

$$
\begin{aligned}
M(1,2) = &-iG_H(1,1')W(1',2)G_H(1',1^+) + iG_H(1,1')W(1',2)G_H(1',2')S^W(2',1^+) + \\
&- G_H(1,1')\frac{\delta S^W(1',1^+)}{\delta \tilde{U}(2)}.
\end{aligned} \tag{2.18-d}
$$

These equations have the advantage of containing only 2-point quantities. Moreover, the interacting Green's function, G, doesn't appear. The quantities that enter these equations can by interpreted. The first equation (Eqn. 2.18-a), is just a rewriting of the definition of the screened Coulomb potential, with M, thus being similar to a susceptibility. The Eqn. 2.18-c looks similar to a functional differential equation for G, S_W that enters it is similar to a self-energy, but integrated with G. Finally, S_0^W in Eqn. 2.18-b corresponds to an integration of a non-self-consistent GW self-energy with a G_H Green's function.

We can now make some approximations on the derivatives, not forgetting relations of the type $\frac{\delta G}{\delta U_{ext}} = -G\frac{\delta G^{-1}}{\delta U_{ext}}G$.

2.4.2 Approximation 1

To start with, we can neglect all terms except the first one in Eqn. 2.18-c and Eqn. 2.18-d. This will give us:

$$\begin{aligned} W(1,2) &= v(1,2) + M(1,1')v(1',2) \\ S_0^W(1,2) &= -iG_H(1,1')W(1',1)G_H(1',2) \\ S^W(1,2) &= S_0^W(1,2) \\ M(1,2) &= -iG_H(1,1')W(1',2)G_H(1',1^+). \end{aligned}$$

The second and the third equation are no longer needed. Substituting the first into the last we obtain:

$$M(1,2) = (-i)G_H(1,1')v(1',2)G_H(1',1^+) + (-i)G_H(1,1')M(1',2')v(2',2)G_H(1',1^+),$$

which is the equation Eqn. 2.16, multiplied by $-iv$ and contracted, of which we already know the solution: the RPA limit, calculated with Hartree Green's functions. This corresponds to the results depicted in Fig. 2.14. Let us now make a more elaborate approximation.

2.4.3 Approximation 2

If we neglect only the derivative terms in Eqn. 2.18-c and Eqn. 2.18-d we can make a closed approximation for our set of equations. This will correspond exactly to neglecting the derivative in Eqn. 2.17. The set of equations we obtain is:

$$W(1,2) = v(1,2) + M(1,1')v(1',2) \tag{2.19-a}$$

$$S_0^W(1,2) = -iG_H(1,1')W(1',1)G_H(1',2) \tag{2.19-b}$$

$$S^W(1,2) = S_0^W(1,2) - S_0^W(1,1')S^W(1',2) \tag{2.19-c}$$

$$M(1,2) = -iG_H(1,1')W(1',2)G_H(1',1^+) + iG_H(1,1')W(1',2)G_H(1',2')S^W(2',1^+) \tag{2.19-d}$$

Substituting Eqn. 2.19-b into Eqn. 2.19-c we get a closed equation for S^W that we can solve:

$$S^W(1,2) = \left[\delta + S_0^W\right]^{-1}(1,1')S_0^W(1',2).$$

This, in its turn, can be plugged into the last equation (Eqn. 2.19-d):

$$M(1,2) = -iG_H(1,1')W(1',2)G_H(1',2')\left[\delta + S_0^W\right]^{-1}(2',1^+)$$

From this form we clearly see that our approximation is equivalent to the usage of a G_HW self-energy insertion and can be thought of as $M(1,2) = -iG_H(1,1')W(1',2)G(1',2)$, with $G(1,2) = G_H(1,2) + iG(1,2')G_H(2',1')W(1',1)G_H(1',2)$. Note that this self-energy would be constructed using a Hartree Green's function. Moreover the approximation we discussed in the previous subsection is simply the first step, where $G = G_H$.

To make a link with earlier attempts to rewrite the BSE, note that if we make a further approximation to replace $W \to v$ in Eqn. 2.19-b and Eqn. 2.19-d and use $G = G_H + G_HvL$ we would get a $L = L_H$ approximation, where $L_H = G_HG$, such an expression has been found in [85]. This is seen easier if one takes Eqn. 2.17 and plugs in $L_H^S = G_HG_H$. Then the first two terms will give exactly $L = G_HG$. However, for our purpose it seems more promising to maintain the equations symmetric, as pointed out previously.

2.4.4 Approximation 3

In the third approximation, we replace the last term in last equation Eqn. 2.18-d by

$$iG_H(1,2')S^W(2',1')W(1',2)G_H(1',1^+),$$

which comes out if we insert the zeroth order approximation into the functional derivative. In an analogous manner, the derivative term in Eqn. 2.18-c is replaced. We thus obtain a closed set of equations giving us all the diagrams of $\chi = GG$, with G being the solution of a Dyson equation with $\Sigma = \Sigma_{Hartee} + \Sigma_{G_HW}$ up to second order:

$$\begin{aligned}
W(1,2) &= v(1,2) + v(1,1')M(1',2) \\
S_0^W(1,2) &= -iG_H(1,1')W(1',1)G_H(1',2) \\
S^W(1,2) &= S_0^W(1,2) - S_0^W(1,1')S^W(1',2) + iG_H(1,1')S^W(1',2')W(2',1^+)G_H(2',2) \\
M(1,2) &= -iG_H(1,1')W(1',2)G_H(1',1^+) + iG_H(1,1')W(1',2)G_H(1',2')S^W(2',1^+) + \\
&\quad + iG_H(1,1')S^W(1',2')W(2',2)G_H(2',1^+).
\end{aligned} \tag{2.20}$$

The 1-electron limit of these equations gives all the 1st order diagrams except the one

in the lower left of Fig. 2.8.

Moreover, this approximation contains more than G_0G_0 diagrams with the simple G_HW self-energy insertion. For in second iteration it will give for instance the diagram Fig. 2.16(a). However, it does not contain all the diagrams. For instance, Fig. 2.16(b) in missing.

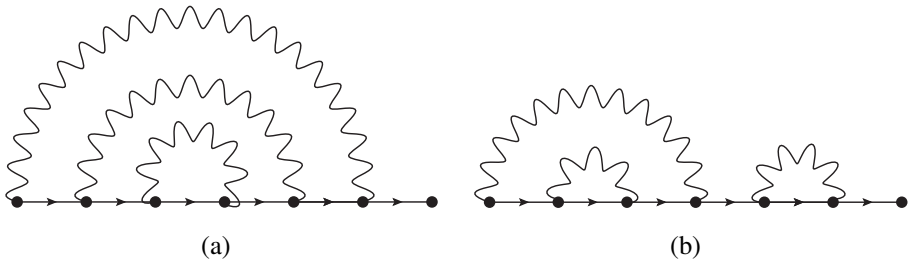


Figure 2.16 – Third order diagrams for S^W

This approximation is still insufficient, as we want to have at least all the diagrams of order 1 in v, to obtain exact cancelations in the single electron limit. The only hope for this approximation is that the contribution of the diagram in the lower left of Fig. 2.8 will not be very important. For application purposes, these equations can be rewritten in frequency space (See appendix: C.2).

2.5 Approximating the L^{-1}

In the previous sections we have tried to approximate the functional derivative using various iteration schemes or approximations for the derivative. In this section we will approximate it using an expression similar to Eqn. 2.2, but for L.

2.5.1 Derivation

Let us take again Eqn. 2.4.1

$$G_H(1,1')\varepsilon^{-1}(1',3)G_H(1',2) = L_H^S(1,3;2,3^+),$$

as well as the Eqn. 2.17:

$$L(1,3,2,3^+) = L_H^S(1,3,2,3^+) + L_H^S(1,3,1',3^+)iv(1',2')L(1',2';2,2'^+) + \\ + G_H(1,1')iv(1',2')\frac{\delta L(1',2';2,2'^+)}{\delta U_{ext}(3)}.$$

If one, now, views $L(1,3;2,3^+)$ as a function of 1,2, with $3,3^+$ as parameters, then one can apply a trick equivalent to Eqn. 2.2, where the inversion is defined by:

$$L(1,3,1',3^+)L^{-1}(1',3,2,3^+) = \delta(1-2) \text{ for all } 3,3^+.$$

This choice of variables is motivated by the fact that they are not contracted, and gives:

$$\frac{\delta L(1,3;2,3^+)}{\delta U_{ext}(4^+,4)} = -L(1,3;1',3^+)\frac{\delta L^{-1}(1',3;2',3^+)}{\delta U_{ext}(4^+,4)}L(2',3;2,3^+).$$

This leads us to a BSE-like equation, with a Kernel, that contains terms of type L^{-1}, that have to be approximated.

$$L(1,3,2,3^+) = L_H^S(1,3,2,3^+) + L_H^S(1,3,1',3^+)iv(1',2')L(1',2';2,2'^+) + \\ - G_H(1,1')iv(1',2')L(1',2';4',2'^+)\frac{\delta L^{-1}(4',2';5',2'^+)}{\delta U(3)}L(5',2';2,2'^+). \tag{2.21}$$

This is a new Bethe-Salpeter Equation, where only Hartree Green's functions appear. It is a self-consistent, non-linear expression, since L and χ appear in the Kernel. One has to find approximations for this Kernel. Similar to the Dyson equation for the single particle Green's function, we start by approximating L^{-1} in the functional derivative.

As a first approximation one can take $L^{-1} = {L_H^S}^{-1}$. This last quantity can be calculated:

$${L_H^S}^{-1}(1,3;2,3^+) = G_H^{-1}(1,1')\frac{1}{\varepsilon^{-1}(1',3)}G_H^{-1}(1',2).$$

Its derivative is simple if we make the approximation $\frac{\delta\varepsilon^{-1}}{\delta U} = 0$ which is similar to the approximation $\frac{\delta W}{\delta U} = 0$ used in the standard Bethe-Salpeter Equation. Then,

$$\frac{\delta {L_H^S}^{-1}(1,4;2,4^+)}{\delta U(3^+,3)} = -\frac{\varepsilon^{-1}(1,3)}{\varepsilon^{-1}(1,4)} G_H^{-1}(1,2) - G_H^{-1}(1,2)\frac{\varepsilon^{-1}(2,3)}{\varepsilon^{-1}(2,4)}.$$

The approximate BSE reads then:

$$\begin{aligned} L(1,3;2,3^+) &= L_H^S(1,3;2,3^+) + L_H^S(1,3;1',3^+) iv(1',2') L(1',2';2,2'^+) + \\ &+ G_H(1,1') iv(1',2') L(1',2';3'^+,2'^+) \frac{\varepsilon^{-1}(3',3)}{\varepsilon^{-1}(3',2')} G_H^{-1}(3',4') L(4',2';2,2'^+) \\ &+ G_H(1,1') iv(1',2') L(1',2';3'^+,2'^+) G_H^{-1}(3',4') \frac{\varepsilon^{-1}(4',3)}{\varepsilon^{-1}(4',2')} L(4',2';2,2'^+) \end{aligned} \tag{2.22}$$

If one assumes $\frac{\varepsilon^{-1}(1,2)}{\varepsilon^{-1}(1,3)} \approx 1$, which is valid when ε is slow varying, it becomes even simpler, since:

$$\frac{\delta {L_H^S}^{-1}(1,4;2,4^+)}{\delta U(3^+,3)} = -2G_H^{-1}(1,2),$$

and we get

$$\begin{aligned} L(1,3,2,3^+) &= L_H^S(1,3,2,3^+) + L_H^S(1,3,1',3^+) iv(1',2') L(1',2';2,2'^+) + \\ &+ 2G_H(1,1') iv(1',2') L(1',2';3',2'^+) G_H^{-1}(3',4') L(4',2';2,2'^+) \end{aligned} \tag{2.23}$$

If one linearizes in the non-linear term setting $L \approx L_H^S$ in a symmetric way, we get two terms instead of one:

$$\begin{aligned} L(1,3,2,3^+) &= L_H^S(1,3,2,3^+) + L_H^S(1,3,1',3^+) iv(1',2') L(1',2';2,2'^+) + \\ &+ G_H(1,1') iv(1',2') G_H(1',3') \varepsilon^{-1}(3',2') L(3',2';2,2'^+) + \\ &+ G_H(1,1') iv(1',2') L(1',2';3',2'^+) \varepsilon^{-1}(3',2') G_H(3',2). \end{aligned} \tag{2.24}$$

Both of these equations 2.23 and 2.24 can now be analysed to see what kind of results we can get. Their frequency structure differs from that of a standard Dyson equation, and from that of a standard Bethe-Salpeter Equation. For details see Appendix C.3. The simplest forms are also derived in that section, and shows that they miss one of the 1-

st order diagrams in v, just as Eqn. 2.20. However, this is not an intrinsic property of Eqn. 2.23. All the first order terms are present in it.

2.6 Conclusions

The aim of this chapter was to reformulate the current approach to the calculation of the correlation function L and propose an alternative to the Bethe-Salpeter Equation based on the GW approximation to the self energy. Particular importance was attached to the possibility of writing down equations that do not contain the fully interacting G, as it is both difficult to calculate and contains exchange-correlation effects that are canceled by the Kernel Ξ.

To this end we combined in different ways two equations that contain L:
— The self-energy $\Sigma_{XC} = iv(1,2')L(1,2',3',2'^{+})G^{-1}(3',2)$,
— The response function $\chi(1,2) = -iL(1,2;1^{+},2^{+})$.

The key result that was found here are the equations Eqn. 2.16 and Eqn. 2.17 and from this, the new Dyson Eqn. 2.21 for the correlation function L. These equations are exact and contain not the fully interacting G, but the Hatree-Green's function G_H, which can, in principle, be obtained easily using the electron density n, obtained from DFT.

My analysis of these equations shows that they still contain a self-energy-like correction (the term $L_H^S(1,3,1',3^{+})iv(1',2')L(1',2';2,2'^{+})$ in the left hand side of Eqn. 2.17). However, its explicit form is given; no approximation is needed to evaluate it (contrary to the self-energy in general), besides an approximation to ε^{-1}. Often it is easier to approximate ε^{-1} than to find an approximation to G (the best example is the one-electron limit), which suggests that this equation is a better starting point for approximations than the standard Bethe-Salpeter Equation. The screened Coulomb interaction naturally appears in the equations, and they can be rewritten so as to have only two-point quantities in them (Eqn. 2.18-a-Eqn. 2.18-d). This is physical and makes a link to Hedin's equations.

A drawback of the formulation might be that Hartree and exchange contributions are not treated on the same footing (which would be the case if we were to use the Hartree-Fock Green's function G_{HF} instead of G_H as starting quantity) which bears the risk for example of self-interaction errors in practical approximations. We will have to be careful with this. For the same reason it still difficult to have approximations with the correct one

electron limit. On the other hand, going beyond the Hartree Green's functions makes the fully interacting Green's function reappear: we have found no way, for example, to work with the Hartree-Fock Green's functions G_{HF} alone.

Having written down different exact equations, we have made approximations to them, in particular to the functional derivative term, that enters Eqn. 2.17, or alternatively Eqn. 2.18-c and Eqn. 2.18-d. The simple approximations were covered in subsections 2.4.2-2.4.4. A more elaborate reformulation came from the rewriting Eqn. 2.17 as a Dyson equation, and approximating the functional derivative of L^{-1}, instead of L. This gave rise to Eqn. 2.23 and Eqn. 2.24. Even with the drastic approximations we used, such as $\frac{\delta\varepsilon}{\delta U} = 0$, they are non-linear and contain non-trivial physics.

A drawback of this non-linearity is that there might be multiple solutions, and this has to be investigated. We know this is a general problem in MBPT [86], and the problem is avoided if the equations are solved by iteration. Furthermore other approximations to the functional derivatives of L and L^{-1} have to be investigated.

To resume, we have seen that writing alternatives to the Bethe-Salpeter Equation for the correlation function L is difficult. Nevertheless my work indicates promising directions with first suggestions for approximations. In particular Eqn. 2.21 demonstrates that it is possible to formulate a Dyson equation for L where the fully interacting single-particle Green's function G doesn't appear.

Chapter 3

Comparing and combining TDLDA and BSE

An approach alternative to the Bethe-Salpeter Equation based on the GW approximation to the self-energy is Time-Dependent Density Functional Theory (TDDFT), based on an extension of the Kohn-Sham DFT. The simplest approximation, which is widely used is the Adiabatic Local Density Approximation (TDLDA). However it is known to have problems, for example with describing bound excitons.
Both the Bethe-Salpeter Equation and Time-dependent Density Functional Theory are in principle exact approaches to describe electron excitation spectra. Moreover, the currently used approximations to them can be formulated in a very similar manner: as an eigenvalue problem. Nevertheless this does not guaranty that the corresponding electron-hole hamiltonians will be the same and have similar eigenfunctions, even when resulting spectra are the same.
In this chapter we introduce the Adiabatic Local Density Approximation to Time-dependent Density Functional Theory. We, then, rewrite it as an eigenvalue problem, in close analogy to the Bethe-Salpeter Equation. We compare the results obtained using these two methods for different physical quantities and experiments. In the cases where both methods give similar spectra in agreement with experiment, we analyze and compare different ingredients that contribute to give the final spectra. Based on this we discuss to which extent different features of spectra can be attributed to transitions in the band structure and how one should proceed in their analysis. Finally we suggest new methods based on this analysis.

3.1 The Standard TDLDA

In this section we look into and discuss one of the state of the art approaches to the calculation of spectra: the Time-Dependent Density Function Theory in the Adiabatic Local Density Approximation. Furthermore we show that this approach and the Bethe-Salpeter equation, discussed in the previous chapter can formulated in a very similar manner.

3.1.1 TDLDA, an approximation to TDDFT

In linear response, the equation for the susceptibility χ in TDDFT, as stated in the first chapter (Eqn. 1.22), reads:

$$\chi(1,2) = \chi_{KS}(1,2) + \chi_{KS}(1,1')\left(v(1',2') + f_{xc}(1',2')\right)\chi(2',2). \tag{3.1}$$

Here, χ_{KS} is the Kohn-Sham independent particle response function. As in the case of Density Functional Theory DFT, which is exact, but one doesn't know the exchange-correlation potential v_{xc}, this equation is also exact, but one doesn't know f_{xc}, the exchange-correlation Kernel. An extensive overview of different kernels is available, for example in [87, 23].

Let us now derive Eqn. 3.1. For this we will follow the lines of [88]. The retarded linear response function is given by:

$$\chi(1,2) = \left.\frac{\delta n(1)}{\delta V_{ext}(2)}\right|_{V_{ext}=0} \tag{3.2}$$

Here V_{ext} is an additional external potential applied to the system.

In the same manner the linear response of a fictitious Kohn-Sham system gives rise to:

$$\chi_{KS}(1,2) = \left.\frac{\delta n(1)}{\delta V_{eff}(2)}\right|_{V_{eff}=V_{KS}} \tag{3.3}$$

Here V_{eff}, as in DFT is equal to $V_{ext} + v_{ext} + v_H + v_{xc}$, with the terms being the time-dependent additional external, the original external, the Hartree and the exchange-correlation potentials respectively.

We now use the chain rule. Schematically:

$$\frac{\delta n}{\delta V_{ext}} = \frac{\delta n}{\delta V_{eff}}\frac{\delta V_{eff}}{\delta V_{ext}} = \chi_{KS}\frac{\delta V_{eff}}{\delta V_{ext}} = \chi_{KS}\left(1 + v\chi + \frac{\delta v_{xc}}{\delta V_{ext}}\right) \tag{3.4}$$

Using the chain rule one more time, we obtain the required Eqn. 3.1, with

$$f_{xc}(1,2) = \frac{\delta v_{xc}([n],1)}{\delta n(2)}$$

In the Adiabatic Local Density Approximation we suppose that the exchange-correlation potential is equal to the one of the Local Density Approximation to Density Functional Theory, taken at the instantaneous density:

$$v_{xc}([n],r,t) = v_{xc}^{LDA}(n(r,t),r,t)$$

Then f_{xc} is local and static $f_{xc}(1,2) = \frac{dv_{xc}^{LDA}}{dn}\Big|_{n=n(r,t)} \delta(\mathbf{r}-\mathbf{r}')\delta(t-t')$.

Let us note that the derivation in Eqn. 3.4 is closely analogous to the derivation of the Bethe-Salpeter Equation (from Eqn. 2.2 to Eqn. 2.4). Schematically:

$$L = \frac{\delta G}{\delta U} = \frac{\delta G}{\delta U_{tot}}\frac{\delta U_{tot}}{\delta U} = GG\frac{\delta\left[U+V_H+\Sigma\right]}{\delta U} = GG\left[1+vL+\frac{\delta\Sigma}{\delta G}\frac{\delta G}{\delta U}\right].$$

3.1.2 TDLDA and BSE equations

In the previous sub-section we have derived the linear response TDDFT equation. It can be rewritten in frequency space as follows:

$$\chi(\mathbf{x}_1,\mathbf{x}_2|\omega) = \chi_{KS}(\mathbf{x}_1,\mathbf{x}_2|\omega) + \chi_{KS}(\mathbf{x}_1,\mathbf{x}_1'|\omega)\mathbb{K}(\mathbf{x}_1',\mathbf{x}_2'|\omega)\chi(\mathbf{x}_2',\mathbf{x}_2|\omega), \tag{3.5}$$

with the kernel $\mathbb{K}(\mathbf{x}_1',\mathbf{x}_2'|\omega) = v(\mathbf{x}_1',\mathbf{x}_2') + f_{xc}(\mathbf{x}_1',\mathbf{x}_2'|\omega)$. In TDLDA, the kernel is static. Then, the equation looks similar to the static Bethe-Salpeter Equation Eqn. 2.5, apart from the fact that it is not a 4-point, but a 2-point equation.

To make this even more precise, let us consider the Random Phase Approximation (RPA), where the BSE L_0 is G_0G_0, and there is no W term. In this case the equation Eqn. 2.5 can be further simplified, by contracting the variable $x_1 = x_2$ and thus we will directly obtain an equation for the susceptibility χ. This equation is the same as the RPA approximation to the TDDFT, if one uses Kohn-Sham Green's functions for G_0.

Let us now introduce a fictitious L^{TDLDA}, that satisfies

$$L^{TDLDA}_{n_1,n_2,n_3,n_4} = L^{KS}_{n_1,n_2,n_3,n_4} + L^{KS}_{n_1,n_2,n_3',n_4'}\mathbb{K}^{TDLDA}_{n_3',n_4',n_1',n_2'}L^{TDLDA}_{n_1',n_2',n_3,n_4}, \tag{3.6}$$

with $L^{KS} = G_{KS}G_{KS}$ and the Kernel:

$$\mathbb{K}^{TDLDA}_{n3',n4',n1',n2'} = \psi_{n3'}(\mathbf{x}'_3)\psi^*_{n4'}(\mathbf{x}'_4)\left[v(\mathbf{x}'_1,\mathbf{x}'_3) + \right. \\ \left. + f_{xc}(\mathbf{x}'_1,\mathbf{x}'_3)\right]\delta(\mathbf{x}'_1-\mathbf{x}'_2)\delta(\mathbf{x}'_3-\mathbf{x}'_4)\psi^*_{n1'}(\mathbf{x}'_1)\psi_{n2'}(\mathbf{x}'_2), \tag{3.7}$$

in close analogy with Eqn. 2.7. This definition guarantees that

$$\chi(\mathbf{x}_1,\mathbf{x}_2|\omega) = \sum_{n1,n2,n3,n4} \psi^*_{n1}(\mathbf{x}_1)\psi_{n2}(\mathbf{x}_1)L^{TDLDA}_{n1,n2,n3,n4}(\omega)\psi_{n3}(\mathbf{x}_2)\psi^*_{n4}(\mathbf{x}_2)$$

satisfies the TDLDA Eqn. 3.5. We thus, see that both the TDLDA and the BSE can be written in the same form of a 4-point Dyson equation.

In some cases, as we will see further in this chapter, the two approaches to spectroscopy, TDLDA and BSE give very similar results. However, the L^{TDLDA} is a fictitious object, unlike the L, that enters the Bethe-Salpeter equation. Only its diagonal χ has a physical meaning. It is, thus, interesting to analyse how different these two object are.

Note that the exact TDDFT can also be written in the form Eqn. 3.6, even though $\mathbb{K}$ depends on the frequency ω in that case. Still, even in this case there is no reason for the L^{TDDFT} to be equal to the exact L, as the variational principle requires only its diagonal χ to be, in principle, exact.

3.1.3 The two-particle hamiltonian

In the previous sub-section we have shown that the Time-Dependent Local Density Approximation, just as the Bethe-Salpeter Equation can be written in the form:

$${}^4\chi = {}^4\chi^0 + {}^4\chi^0\mathbb{K}^4\chi. \tag{3.8}$$

For the two cases we will be interested in, we have, schematically:

— TDLDA: The 4-point susceptibility, ${}^4\chi$ is L^{TDLDA}, ${}^4\chi^0 = G_{KS}G_{KS}$, with G_{KS}, the Kohn-Sham Green's functions in the LDA and the Kernel $\mathbb{K} = v + f_{xc}$

— BSE: The 4-point susceptibility, ${}^4\chi$ is L, ${}^4\chi^0 = GG$, with G being the interacting Green's function which is used in practice in a quasi-particle approximation, and the most widely used approximation to the Kernel $\mathbb{K} = v - W$

Its formal solution of Eqn. 3.8, in both of these cases is given by

$$ {}^4\chi = \left[1 - {}^4\chi^0 \mathbb{K}\right]^{-1} {}^4\chi^0. $$

With the above approximations, this inversion can be written in terms of an effective hamiltonian defined as:

$$ \mathbb{H}^{2p}_{n_1,n_2,n_3,n_4} = (\varepsilon_{n_1} - \varepsilon_{n_2})\,\delta_{n_1,n_3}\delta_{n_2,n_4} + (f_{n_1} - f_{n_2})\,\mathbb{K}_{n_1,n_2,n_3,n_4}, \tag{3.9} $$

where f_n are the occupation numbers. Here we have already used the fact that [1]

$$ {}^4\chi^0_{n_1,n_2,n_3,n_4} = \frac{(f_{n_1} - f_{n_2})\,\delta_{n_1,n_3}\delta_{n_2,n_4}}{\varepsilon_{n_1} - \varepsilon_{n_2} - \omega} \tag{3.10} $$

Using this definition we can write

$$ {}^4\chi_{n_1,n_2,n_3,n_4} = \left[\mathbb{H}^{2p} - \mathbb{I}\omega\right]^{-1}_{n_1,n_2,n_3,n_4} (f_{n_1} - f_{n_2}) \tag{3.11} $$

Now if we solve the eigenvalue problem for this effective hamiltonian, the inversion gives us the following formula:

$$ \left[\mathbb{H}^{2p} - \mathbb{I}\omega\right]^{-1}_{n_1,n_2,n_3,n_4} = \sum_{\lambda,\lambda'} \frac{V^{n_1,n_2}_{\lambda} S^{-1}_{\lambda,\lambda'} V^{*n_3,n_4}_{\lambda'}}{E_\lambda - \omega} \tag{3.12} $$

Here we have introduced the eigenvectors and eigenvalues of $\mathbb{H}^{2p}$ via $\mathbb{H}^{2p} V_\lambda = \varepsilon_\lambda V_\lambda$, and the overlap matrix S, between them.

It is worth noting that writing the TDLDA as an eigenvalue problem is also known as the Cassida formalism. It was extensively discussed in [89, 90, 91].

In the Bethe-Salpeter case, the effective hamiltonian has a clear meaning. Let's look into it. Suppose that we have just two levels $E_c = E_g$ and $E_v = 0$, their wave functions being real $\phi_c(\mathbf{r})$ and $\phi_v(\mathbf{r})$ respectively. The hamiltonian can then we written down explicitly:

$$ H_{2p} = \begin{pmatrix} E_g + V - W_1 & V - W_2 \\ -V + W_2 & -E_g - V + W_1 \end{pmatrix} $$

1. For simplicity, here and in the following we do not write down explicitly the small imaginary part in the denominator, that determines whether the Green's functions are retarded, time-ordered or advanced

Here we have introduced $V = 2\phi_v(\mathbf{r}_1)\phi_c(\mathbf{r}_1)v(\mathbf{r}_1,\mathbf{r}_2)\phi_v(\mathbf{r}_2)\phi_c(\mathbf{r}_2)$, $W_1 = W(\mathbf{r}_1,\mathbf{r}_2)\phi_v^2(\mathbf{r}_1)\phi_c^2(\mathbf{r}_2)$ and $W_2 = \phi_v(\mathbf{r}_1)\phi_c(\mathbf{r}_1)W(\mathbf{r}_1,2)\phi_v(\mathbf{r}_2)\phi_c(\mathbf{r}_2)$. First of all we see that if the off-diagonal elements are neglected - then the inclusion of W will shift the spectra to lower energies with respect to the RPA result. This is something we have already seen in the spectra. It corresponds to the fact that W is the screened electron-hole attraction.

3.1.4 General statements about the effective Hamiltonian

Because of the difference of occupation numbers $f_{n_1} - f_{n_2}$ contained in Eqn. 3.10, in semiconductors and insulators at zero temperature only pairs of wave functions containing one occupied (v) and one unoccupied state (c) contribute to the final result. This justifies the fact that we call them transitions. For $(n_1,n_2) = (c,v)$ the transition is called resonant, and (v,c) is called antiresonant.

The effective two-particle hamiltonian defined in the previous section can be written in this resonant-antiresonant space as

$$\mathbb{H}^{2p} = \begin{pmatrix} H_{res} & H_c \\ -H_c^{\dagger} & H_{antires} \end{pmatrix} \tag{3.13}$$

It is non-hermitian so, its eigenvectors need not be orthogonal, which explains the appearance of the overlap matrix $S_{\lambda,\lambda'} = \sum_{n1,n2} V_{\lambda}^{*n1,n2} V_{\lambda'}^{n1,n2}$.

Let us now define $\tilde{H}$ such that $\mathbb{H}^{2p} = \tilde{H}B$, where B is a matrix

$$\begin{pmatrix} 1_N & 0_N \\ 0_N & -1_N \end{pmatrix}. \tag{3.14}$$

Here X_N are diagonal matrices of dimension N with X on the diagonal. This new hamiltonian $\tilde{H}$ is hermitian. The original one on the other hand obeys a skew-symmetry relation $\mathbb{H}^{2p\dagger}B = B\mathbb{H}^{2p}$ as pointed out in [92] and [93].

The eigenvalue problem for H^{2p} is stated as follows:

$$\mathbb{H}^{2p}V_i = \varepsilon_i V_i$$

It can be rewritten as follows:

$$\tilde{H}BV_i = \varepsilon_i V_i$$

Or if one takes the hermitian conjugate

$$V_j^\dagger B\tilde{H} = \varepsilon_j^* V_j^\dagger$$

Multiplying the first equation by $V_j^\dagger B$ and the second one by BV_i and then taking the difference one gets:

$$V_j^\dagger BV_i \left(\varepsilon_i - \varepsilon_j^*\right) = 0$$

If we exclude the pathological $V_i^\dagger BV_i = 0$ for $i = j$ we get $\varepsilon_i = \varepsilon_i^* \Rightarrow \varepsilon \in R$. Whereas for $i \neq j$ if there are no degeneracies in energy we have $V_j^\dagger BV_i = 0$. This proves the following statement:

Statement Let V_i, ε_i be the set of eigenpairs of $\mathbb{H}^{2p}$. If the problem is non-pathological ($V_i^\dagger BV_i \neq 0$) and non-degenerate ($\varepsilon_i \neq \varepsilon_j$ for $\forall\, i \neq j$) the eigenvalues are real and $V_j^\dagger BV_i = 0$ for $\forall\, i \neq j$

Under the same conditions we can rewrite things differently:

$$U^\dagger BU = N$$

Here U is the matrix of eigenvectors, and N is the Norm Matrix: $N_{ij} = V_i^\dagger BV_i \delta_{ij}$ from which we get $N^{-1}U^\dagger BU = 1$. Defining Λ to be a diagonal matrix, with ε_i on it diagonal, we get:

Statement Under the same conditions $\mathbb{H}^{2p} = U\Lambda N^{-1}U^\dagger B$

We see that, even though the effective two-particle hamiltonian is non-hermitian, we can still define its eigenvectors, that shall be orthogonal, but with a modified scalar product $V_j^\dagger BV_i$. This allows us to treat them as normal eigenvectors for many purposes, as wee shall do in the following sections. Moreover, one can define a spectral decomposition for it, using a diagonal matrix of its eigenvalues, matrices of eigenvectors and two auxiliary matrices. This allows one in principle to formulate alternatives to Eqn. 3.12. Nevertheless, in the following we will stay with the formula Eqn. 3.12, which is less elegant, but suitable for our purposes.

3.1.5 Discussion

One of the main goals of theoretical spectroscopy is not just to reproduce experimental results, but to gain additional information from them. Of particular interest is the assignment of peaks in spectra. To this end, one would like to decompose the peaks into contributions from different transitions. Their mixing is given by the coefficients of the eigenvectors of the hamiltonian Eqn. 3.9 in transition space, as seen from Eqn. 3.11 and

Eqn. 3.12, which leads to:

$$\chi(\mathbf{x}_1,\mathbf{x}_2|\omega) \sim \sum_{\lambda,\lambda'} \frac{\left(\sum_{n_1,n_2} \tilde{\rho}_{n_1,n_2}(\mathbf{x}_1)\cdot V_{\lambda}^{n1,n2}\right)^* S_{\lambda,\lambda'} \sum_{n_3,n_4}\left(V_{\lambda'}^{n_3,n_4}\cdot \tilde{\rho}_{n3,n4}(\mathbf{x}_2)\right)}{\omega-\varepsilon_{\lambda}}. \quad (3.15)$$

Here $\tilde{\rho}_{n_1,n_2}(\mathbf{x}_1) = \phi_{n_1}(\mathbf{x}_1)\phi^*_{n_2}(\mathbf{x}_1)$, V^{λ} and ε^{λ} are the eigenvectors and eigenvalues of the Hamiltonian Eqn. 3.9.

Momentum resolved spectra are given by the Fourier transform

$$\chi(\mathbf{q},\omega) \sim \int\int d\mathbf{x}_1 d\mathbf{x}_2 \chi(\mathbf{x}_1,\mathbf{x}_2|\omega) e^{-i\mathbf{q}(\mathbf{x}_1-\mathbf{x}_2)}. \quad (3.16)$$

Due to the momentum conservation in a crystal only pairs $(v,\mathbf{k}_1;c,\mathbf{k}_2)$ with $\mathbf{k}_2 = \mathbf{k}_1+\mathbf{q}$ contribute.

The Bethe-Salpeter equation in transition space provides a clean framework for the spectral analysis: the coefficients V_{λ} express the mixing of formally independent transitions. On the other hand TDLDA is a much cheaper calculation and therefore one would like to be able to extract some useful information out of it. However, as discussed in the previous subsections, the L^{TDLDA} is not a physical quantity and the same would be true even for the exact L^{TDDFT}. Therefore the question that arises is whether an analysis based on eigenvectors and eigenvalues of the TDLDA effective two-particle hamiltonian is meaningful.

3.2 Comparing TDLDA and BSE

In the previous section we have shown that TDLDA can be written as an eigenvalue problem. The question that arises is to which extent the eigenvectors and eigenvalues of the corresponding problem are similar to the ones of the standard Bethe-Salpeter Equation, based on the GW approximation to the self-energy. We will partially answer it in this section and, furthermore, discuss other ingredients that enter the final formula for the dielectric permittivity, that is closely related to Photo-emission and Electron Energy Loss spectra. It is important to point out, that some partial comparisons have already been performed for small systems [94, 95, 96].

3.2.1 Preliminary analysis

Before proceeding further we explore whether the eigenvectors in the calculation with and without coupling H_c, between the resonant and antiresonant sectors of the hamiltonian Eqn. 3.13 are similar. The easiest way to see this is to compute scalar products of eigenvectors for positive eigenvalues, ordered by eigenvalue magnitude:

$$O_{\lambda,\lambda'} = \left\langle v_{\lambda}^{(1)} \middle| \middle| v_{\lambda'}^{(2)} \right\rangle = \sum_{n1,n2} V_{\lambda}^{(1),n1,n2*} V_{\lambda'}^{(2),n1,n2} \tag{3.17}$$

The upper indices correspond to two different hamiltonians, whose eigenvectors we are comparing. Unless specified otherwise, in the following, we will be working with loss spectra of Silicon. For q=0 the data is summarized in Table 3.1. The "coup" and "ncoup" labels correspond to the BSE calculations with and without coupling respectively; the "tdlda" and "rtdlda" labels correspond to the full TDLDA, and the one without coupling in the BSE form. In the first row we have written down the minimum overlap (Eqn. 3.17) between two eigenvectors for different couples of hamiltonians constructed using 32k points in the Brillouin zone, the second row shows the average overlaps for the same system, the third row - minimum overlaps for systems with 4k points, and the forth the average overlaps for these last systems.

	coup vs tdlda	coup vs ncoup	ncoup vs tdlda	ncoup vs rtdlda
minimum overlap 32k	0.012	0.025	0.009	0.004
average overlap 32k	0.465	0.95	0.486	0.457
minimum overlap 4k	0.076	0.326	0.012	0.038
average overlap 4k	0.605	0.964	0.622	0.59

Table 3.1 – Overlaps of eigenvectors for q=0. For the definitions see text.

One notes that the overlap between the eigenvectors with and without coupling is large. Even the one between the BSE without coupling (ncoup) and the full TDLDA (tdlda) is larger than that between the full BSE (coup) and full TDLDA. This is surprising, as the full BSE and TDLDA give similar plasmon positions and shapes, whereas the BSE without coupling is quite different from them. The corresponding spectra are plotted in Fig. 3.1.

The fact that the overlap between the calculations with and without coupling is very big can be used to improve the diagonalization procedure, through passing to a basis where H_{nocoup} is diagonal. This will be further discussed later section. The overlaps for $q \neq 0$ are smaller, but still large enough to say that the eigenvectors are quite similar.

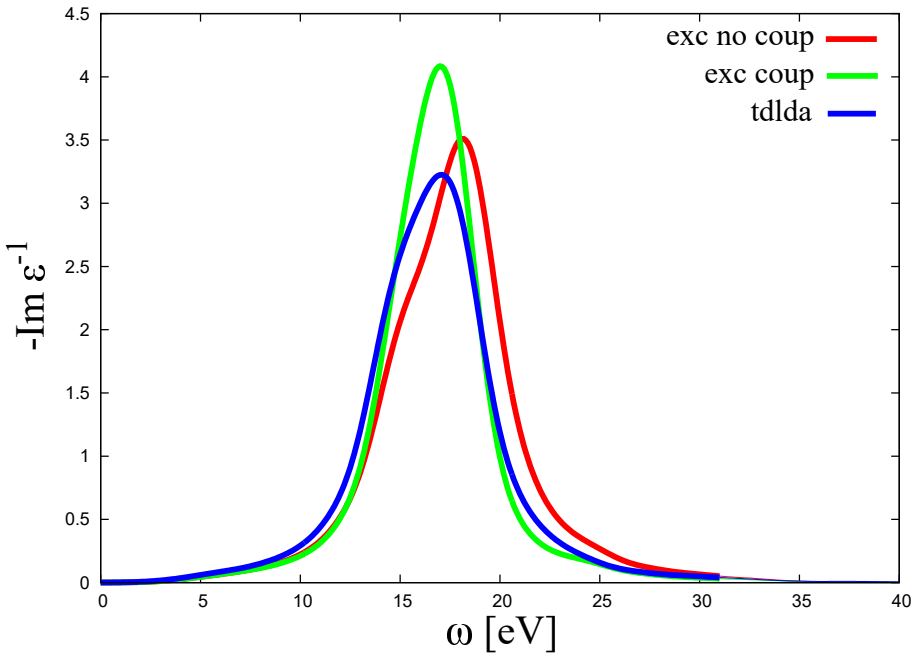


Figure 3.1 – Bulk Silicon loss spectra at q=0 for a 32k-points IBZ sampling. Comparison of different methods: BSE(EXC) with and without coupling, TDLDA

3.2.2 Eigenvectors

In the previous sub-section we have performed a preliminary analysis of the differences between the eigenvectors, corresponding to different approximations. Let us look deeper into it. To this end we first of all look at a given eigenvector of the BSE hamiltonian V_λ. The absolute values of its projections on the transition space vectors is given in Fig. 3.2. These projections can be viewed as overlaps (Eqn. 3.17) between the eigenvectors of the BSE and those of non-interacting BSE hamiltonian ($\mathbb{K} = 0$). We see that their distribution is rather sharp, which allows one to identify specific eigenvectors of the non-interacting hamiltonian with positive eigenvalues, with transitions from valence to conduction band, that are responsible for a given transition in the final spectrum.

In the same manner we can look at the overlaps between the eigenvectors of effective hamiltonians corresponding to different approaches, keeping an eigenvector of one hamiltonian fixed, and varying eigenvectors of the other one. We have done this to compare eigenvectors of the BSE hamiltonian with and without coupling. The corresponding figure is Fig. 3.2.2. One sees that the distribution of overlaps remains sharp-peaked, around one specific eigenvector.

To see that this is not something related to specific eigenvectors we chose to calculate the logarithms of absolute values of all the elements of the matrix $O_{\lambda,\lambda'}$ (Eqn. 3.17), where we have chosen the two sets of eigenvectors to be those of the Bethe-Selpeter effective

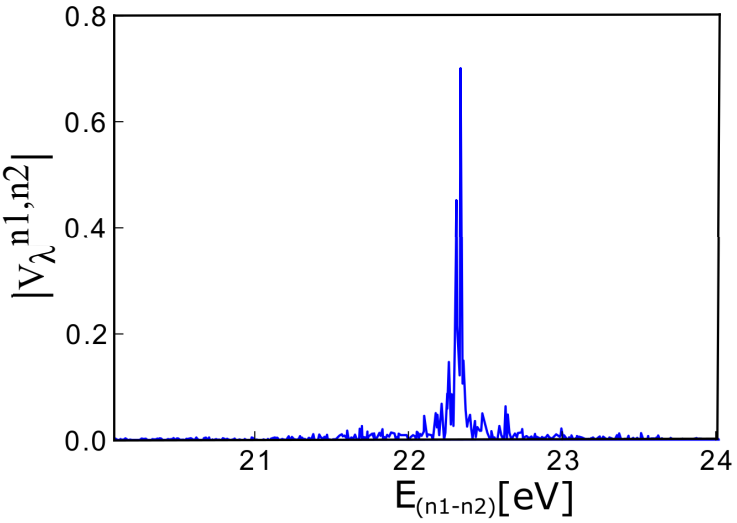


Figure 3.2 – The absolute value of coefficients $V_\lambda^{n1,n2}$ in terms of the LDA energy $E_{(n1-n2)} = \varepsilon_{n_2} - \varepsilon_{n_1}$, for a transition given by its energy $\varepsilon_\lambda = 22.2$ eV

hamiltonian without coupling respectively, and those of the full TDLDA. The result is shown in Fig. 3.4. We see that the overlap drops drastically as soon as we move away from the diagonal $O_{\lambda=\lambda'}$. From this we understand that the matrix that transforms the eigenvectors of one effective hamiltonian to the eigenvectors of the other one is also close to being diagonal. We will use this in the construction of new calculation methods in this chapter.

A good way of estimating how many eigenvectors of one system we need to construct the eigenvectors of the other is to calculate sums of squares of absolute values of the elements of the matrix of overlaps:

$$S_{\lambda,\lambda'} = \sum_{\lambda'' \in [\lambda'-dn, \lambda'+dn]} \left| O_{\lambda'',\lambda} \right|^2, \tag{3.18}$$

or alternatively, if we define the range using energies:

$$S_{\lambda,\lambda'} = \sum_{\lambda'' : |\varepsilon_{\lambda''} - \varepsilon_\lambda| < de} \left| O_{\lambda'',\lambda} \right|^2. \tag{3.19}$$

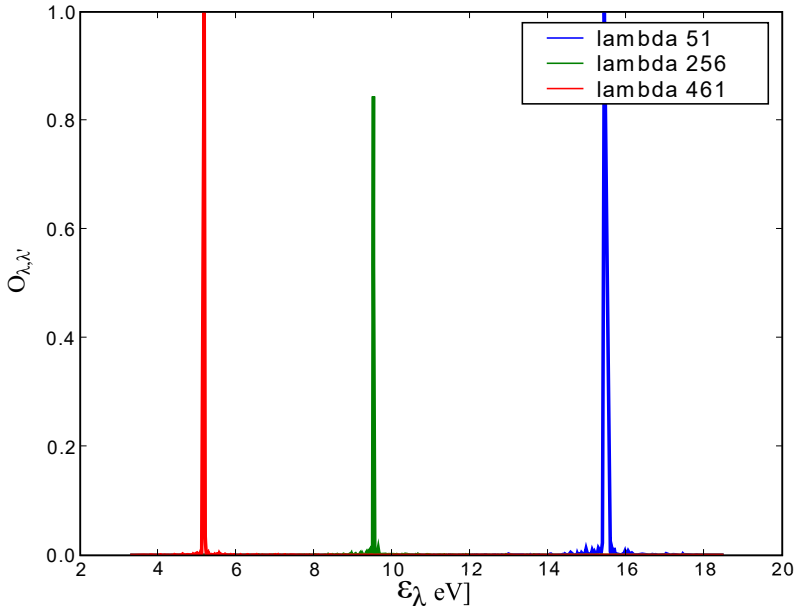


Figure 3.3 – Absolute values of overlaps for selected eigenvectors of the effective hamiltonian of the Bethe-Salpeter Equation with coupling, with those of the effective hamiltonian of the Bethe-Salpeter Equation without coupling. Calculation performed for Bulk Silicon 32k-points IBZ sampling, with 8 bands.

The results of such a calculation are seen in Fig. 3.5. We see that taking an energy range of $de = 2eV$, or $dn = 20$ gives almost 1, for both sums Eqn. 3.18 and Eqn. 3.19.

Finally, we can indirectly compare our results with those available in literature. Computing the absolute values of the coefficients of the eigenvectors in transition space Fig. 3.6(a) Fig. 3.6(b), we note that in the TDLDA case they are spread over a larger range of elementary transitions than in the case of BSE, in agreement with results in [97]

3.2.3 Joint Density of States

Coupled to the analysis of eigenvectors is the analysis of the Joint Density of States (JDOS). For the non-interacting case it results from simple energy differences $jdos(\omega) = \sum_{v,c,k} \delta(\varepsilon_{c,k} - \varepsilon_{v,k} - \omega)$, whereas after the addition of electron-hole interaction it becomes the spectral function of the effective excitonic hamiltonian neglecting matrix elements, $\sum_\lambda \delta(\varepsilon_\lambda - \omega)$, since if one sets the numerator of Eqn. 3.15 to one and takes its imaginary part the resulting expression is

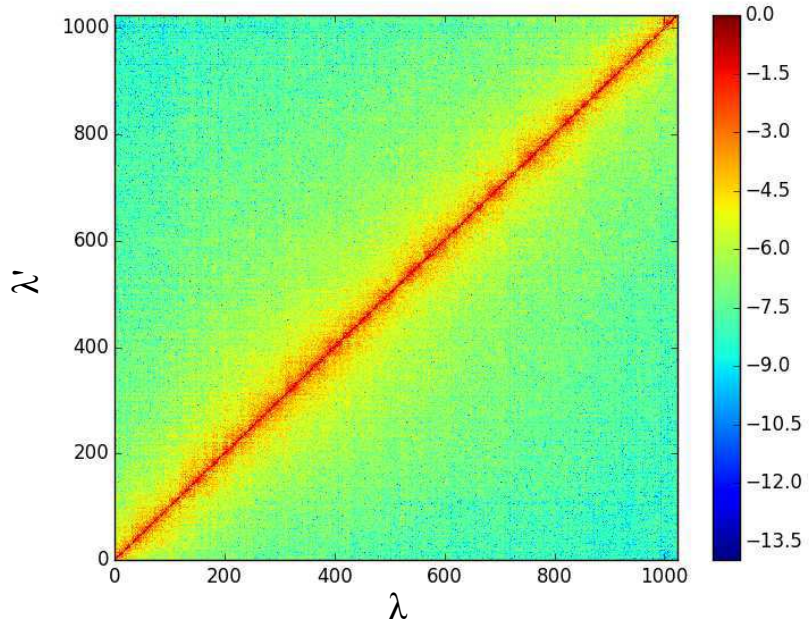


Figure 3.4 – Logarithm of the absolute value of overlaps between eigenvectors of effective hamiltonians for the BSE without coupling and TDLDA. Calculation performed for Bulk Silicon 32k-points IBZ sampling, with 12 bands.

$$Im\sum_{\lambda}\frac{1}{\omega-\varepsilon^{\lambda}}\sim\sum_{\lambda}\delta\left(\omega-\varepsilon^{\lambda}\right). \tag{3.20}$$

The result is shown in Fig. 3.7. Here we see that the JDOS of the effective hamiltonians of the BSE with and without coupling coincide, and are both shifted with respect to the TDLDA one.

This is expected because the effective hamiltonians in the case of Silicon are dominated by their diagonals. From a perturbative point of view, this means that the eigenvalues and eigenvectors will be close to those of a diagonal-only hamiltonian. Furthermore, the corrections to the energy are of second order, and not of first order, as the corrections to the eigenvectors. Looking at the diagonal of Eqn. 3.9, we see that an important ingredient are the energy differences $\varepsilon_{n_1}-\varepsilon_{n_2}$. For the TDLDA, they will be simply equal to the differences of Kohn-Sham energies, whereas in the case of BSE they will correspond to GW corrected energy differences. This explains the rigid shift.

It is worth noting, that this doesn't change when we go to cruder Brillouin zone sampling, as seen in Fig. 3.8(a). We see that the Joint Density of States for the effective

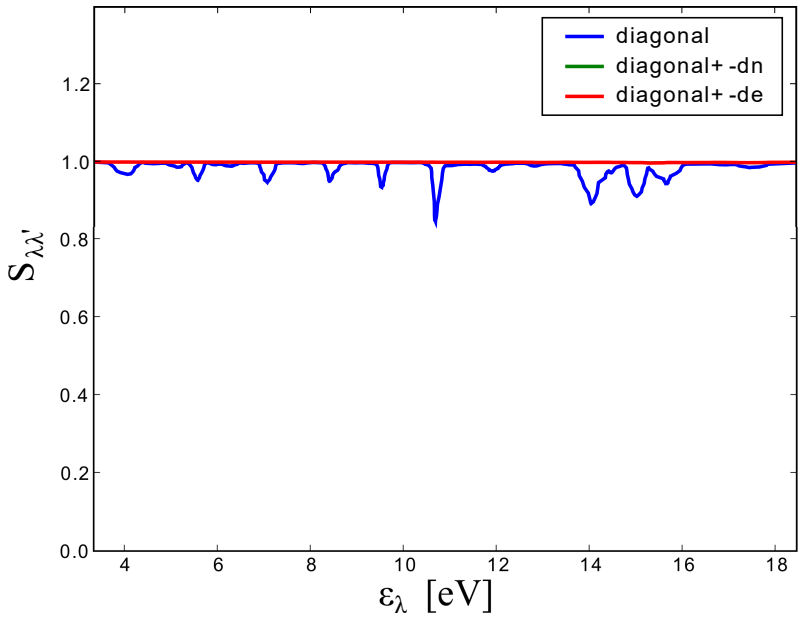
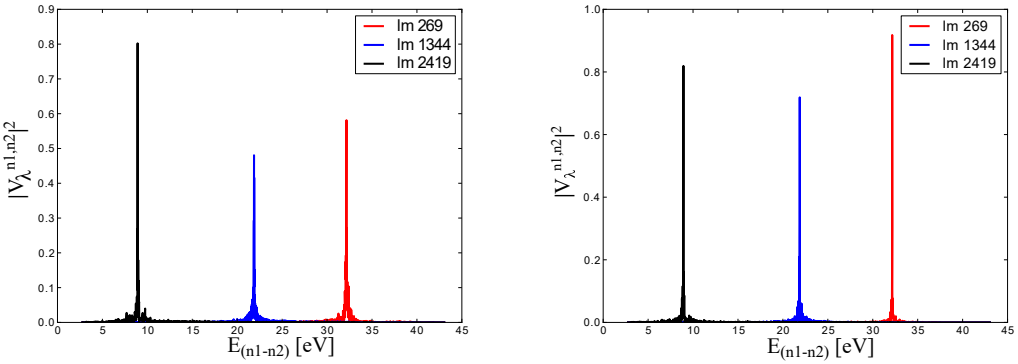


Figure 3.5 – Square overlap between eigenvectors of effective hamiltonians for the BSE and TDLDA (blue). And the sum of squares of overlaps over several eigenvectors around the diagonal (green and red). Calculation performed for Bulk Silicon 32k-points IBZ sampling, with 8 bands.



(a) TDLDA with coupling

(b) BSE without coupling

Figure 3.6 – Bulk Silicon. The square absolute values of the effective hamiltonian eigenvectors' coefficients in the transition basis

hamilitonians with and without coupling coincide perfectly. Furthermore (Fig. 3.8(b)), if we compare the TDLDA joint-density of states, it coincides with the LDA one, and the Bethe-Salpeter one, coincides with the LDA, shifted using a scissor correction to single-particle energies. We thus confirm, what was stated previously, that the main contribution

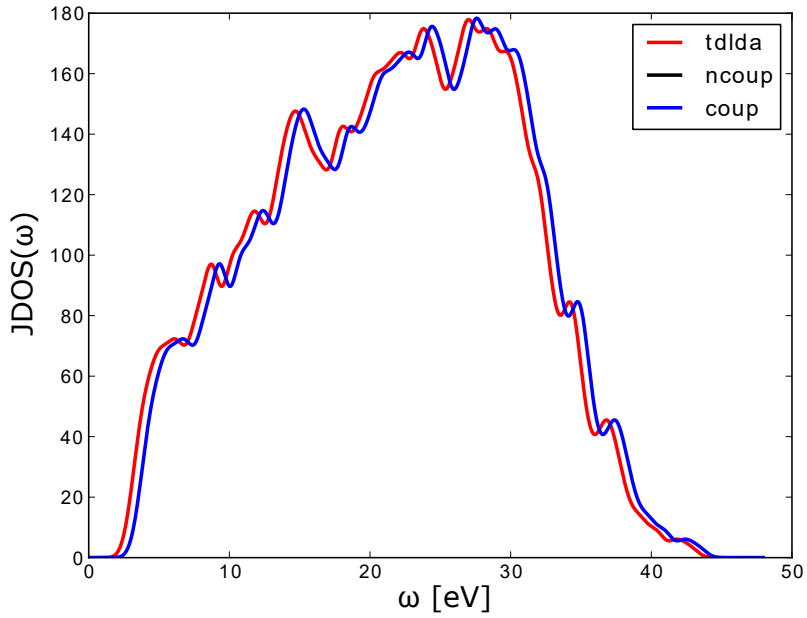
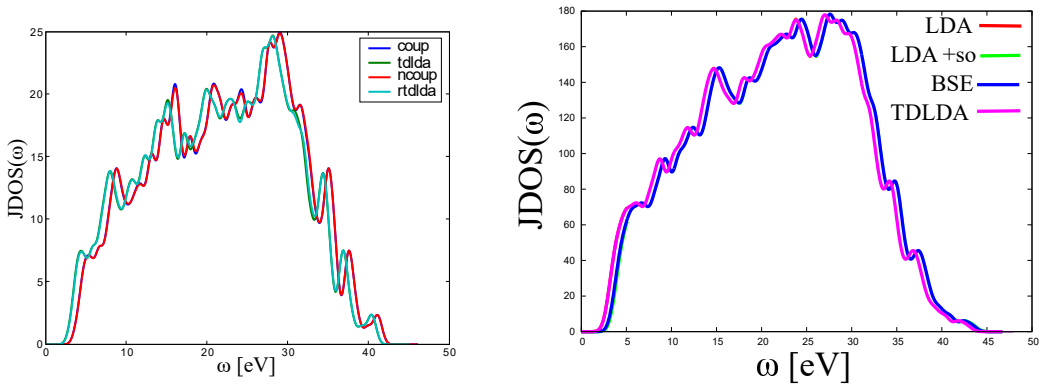


Figure 3.7 – Bulk Silicon. Joint density of states for different approaches. BSE with coupling(coup) and without coupling(ncoup), TDLDA

to the transition energies comes from the diagonal, that is given by the effective hamiltonian with the Kernel $\mathbb{K}$ set to 0.



(a) 4k in IBZ

(b) 32k in IBZ, 25 bands

Figure 3.8 – Bulk Silicon. The joint density of states corresponding to the effective hamiltonians for different approaches: BSE with coupling(coup), without coupling(ncoup), TDLDA (tdlda), resonant TDLDA (rtdlda).

3.2.4 Interference

In the previous sub-sections we have looked into the eigenvectors and eigenvalues that are obtained from the effective two-particle hamiltonian. We have noticed that the eigenvectors are similar, and that the Bethe-Salpeter JDOS is shifted with respect to the Time-Dependent Density Functional Theory JDOS. This last remark, will be especially visible at low energies, for example, for optical spectra. This is one of the reasons the optical spectra a different as illustrated in Fig. 3.9.

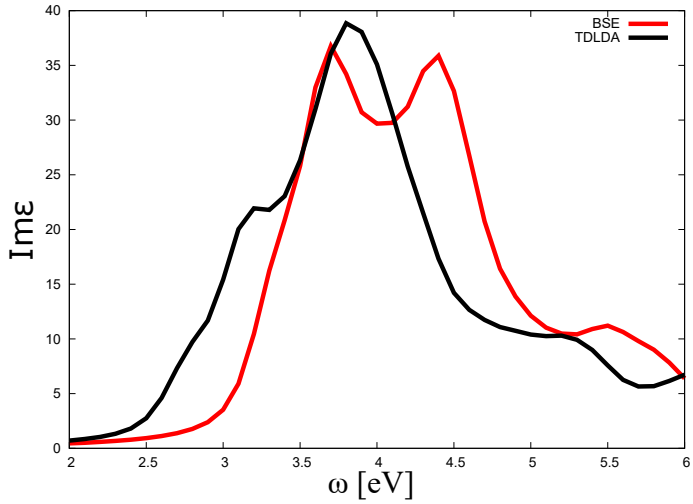


Figure 3.9 – Optical Spectra of Bulk Silicon obtained using different theoretical approaches: BSE and TDLDA.

Surprisingly, even though the two Joint Densities of States are shifted, the Loss spectra obtained obtained using TDLDA and BSE (Fig. 3.10) are very close. In Fig. 3.10 we have also plotted the experimental and theoretical results from [82]. The difference between the two TDLDA calculations being due to different k-point sampling. For consistency, we chose to use the same k-point sampling between the two new BSE and TDLDA calculations.

The fact that the two spectra are so close means that the second ingredient that one gets from the BSE/TDLDA effective hamiltonian, the eigenvectors, compensate for the shift of JDOS. To understand how this might happen, one should not forget that the eigenvectors do not enter our expressions directly, but through a scalar product with $\tilde{\rho}$. This means that there might be interference effects that enter. This is actually what happens. One can plot partial sums (Eqn. 3.21) in terms of $Tmax$, the transition up to which we perform the

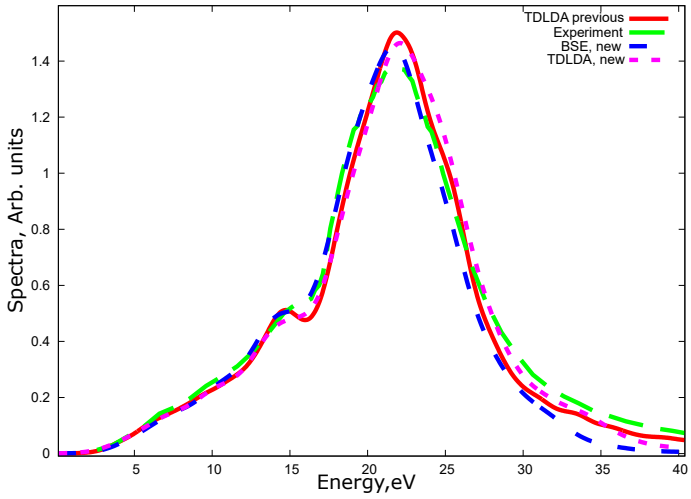


Figure 3.10 – Loss Function of Bulk Silicon at q = 0.8 a.u. in the [1,1,1] direction, using different theoretical approaches (BSE, TDLDA), compared to TDLDA and experimental results from [82]

.

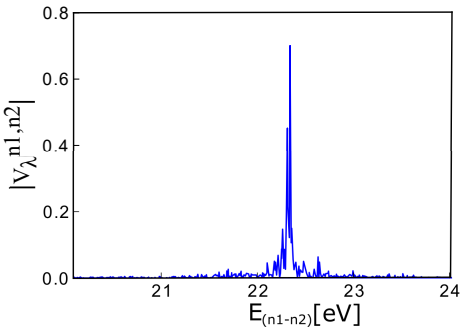
summation, as it was suggested in [98]:

$$R(Tmax) = \frac{\left|\sum_{t=(n1,n2)}^{Tmax} \rho_t V_t^{\lambda}\right|^2}{\left|\sum_{t=(n1,n2)} \rho_t V_t^{\lambda}\right|^2}. \tag{3.21}$$

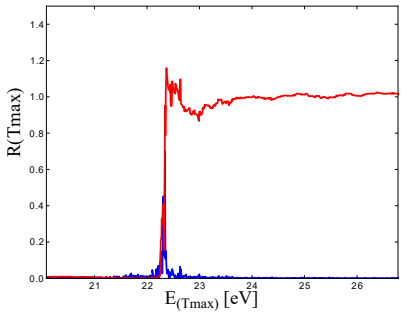
We show the result as a function of the $E(Tmax)$ in Fig. 3.11(b). Comparison to the absolute values of the coefficients $V_{n1,n2}^{\lambda}$ of the eigenvectors that are shown in Fig. 3.11(a), reveals that the contribution does not come exclusively from the peak, but also from the tails (Fig. 3.11(b)). This amounts to an extension of the energy range required for converge of the scalar product $\sum_{n_1,n_2} \tilde{\rho}_{n_1,n_2}(\mathbf{x}_1) \cdot V_{\lambda}^{n1,n2}$ that enters Eqn. 3.15. One sees that the number of transitions over which one has to sum is larger, than what one could have thought from analyzing Fig. 3.6.

3.2.5 Coupling

Another important element contained in Eqn. 3.15 is the overlap matrix S. As shown in Fig. 3.12 one sees that the effect of coupling is important for $q \neq 0$. This is in agreement with what was pointed out in [80], but for $q = 0$. Moreover, the overlap matrix has a significant contribution. Finally, the results suggest that this it is actually this inclusion of the coupling term that accounts for the cancelation of the shift induced by the single-particle energy corrections (scissor operator or GW correction).



(a) Coefficients of an eigenvector of the effective hamiltonian in transition space, for a transition given by its energy $\varepsilon_\lambda = 22.2$ eV.



(b) Coefficients of an eigenvector of the effective hamiltonian in transition space and partial sums of these coefficients Eqn. 3.21

Figure 3.11 – Bulk Silicon. The contribution of different transitions to the spectra and the impact of interference.

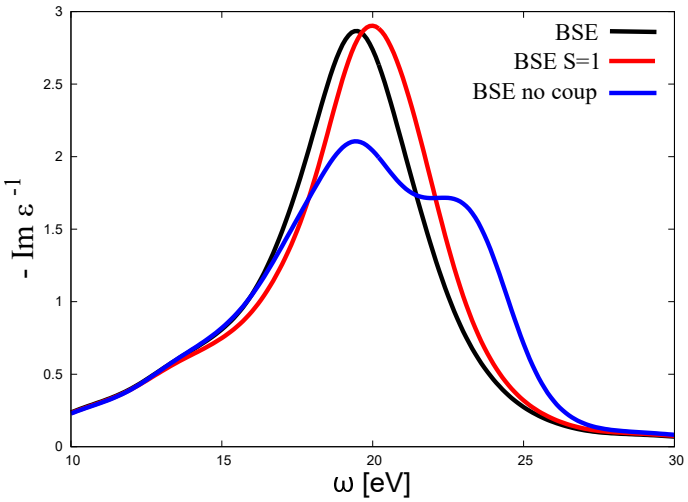
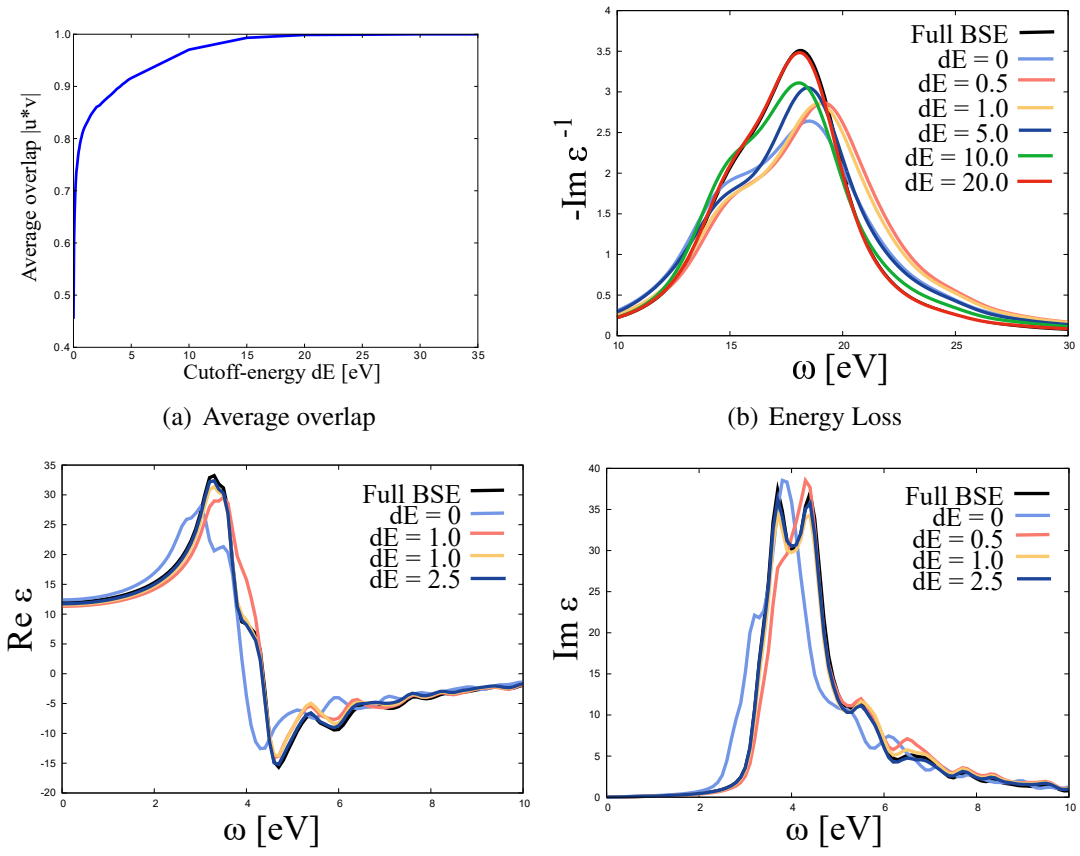


Figure 3.12 – Bulk Silicon loss function at q=0.75 in the (1,1,1) direction: importance of the coupling. (blue) - no coupling included; (red) - coupling included in the effective hamiltonian diagonalisation, the eigenvector non-orthogonality not taken into account; (black) - coupling included everywhere

3.2.6 Importance of the electron-hole exchange

In this last sub-section we want to analyze the importance of the v_c term in the two particle effective hamiltonian Eqn. 3.9. To this end we look at the average overlap between the eigenvectors of the effective hamiltonians. It turns out that this is a good estimate of the proximity of the spectra, if we are replacing elements of one hamiltonian by the other.

To understand this more in details, we can compare Fig. 3.13(a) and Fig. 3.13(b). We see that when the average overlap becomes close to 1, the two energy loss spectra coincide almost perfectly.



(a) Average overlap

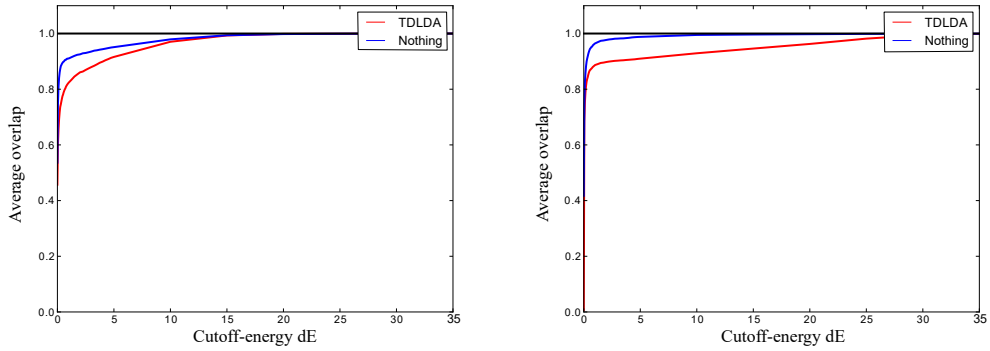
(b) Energy Loss

(c) Re ε

(d) Im ε

Figure 3.13 – Bulk Silicon: the average overlap between the eigenvectors of the effective hamiltonian of the Bethe-Salpeter Equation without coupling and the eigenvectors of the effective hamiltonian for the Bethe-Salpeter Equation without coupling, where parts of the hamiltonian were replaced their TDLDA counterparts in terms of the amount of replacements(upper left); the evolution of the loss function (upper right), Re ε (lower left) and Im ε (lower right) as we increase the number of elements of the hamiltonian that are replaced.

Let us now use this to analyze the importance of v_c. To this end we start replacing parts of the effective hamiltonian for the Bethe-Salpeter Equation, with elements of the TDLDA effective hamiltonian. What we see is that if both effective hamiltonians contain the v_c term (Fig. 3.14(a)), the average overlap converges to 1, rather rapidly. However, if we remove the v_c terms from both hamiltonians, the convergence becomes much worse (Fig. 3.14(b)).



(a) With the coulomb term in $\mathbb{K}$ (b) Without the coulomb term in $\mathbb{K}$

Figure 3.14 – Bulk Silicon: the average overlap between the eigenvectors of the effective hamiltonian of the Bethe-Salpeter Equation and the eigenvectors of the effective hamiltonian for the Bethe-Salpeter Equation, where parts of the hamiltonian were replaced their TDLDA counterparts or set to 0 (Nothing).

This also shows that the f_{xc} and W terms, that enter the respective effective hamiltonians for TDLDA and BSE, function differently, and for the results to be close one absolutely needs the v_c term. In other words, this gives a hint why TDLDA works much better for loss than for absorption spectra.

3.3 Ideas for new methods

In the previous sections we have analyzed different ingredients that enter the reformulation of TDLDA and BSE as eigenvalue problems, and the properties of the effective hamiltonian in general. Here we suggest new approaches based on this.

3.3.1 Changing the working space

In [92] and [93] the authors have modified the scalar product $V_i^\dagger V_j \Rightarrow V_i^\dagger H_{ncoup} V_j$ to make the eigenvalue problem $\mathbb{H}^{2p} V_\lambda = \varepsilon_\lambda V_\lambda$ hermitian. Alternatively a perturbation procedure has been used in [80]. Here we suggest a reformulation of the latter approach, which can either be used to find approximate solutions or exact iterative, but with a better convergence rate, in some cases. It is based on the fact that in many cases $\|H_{res}\|, \|H_{antires}\| \gg \|H_c\|$. We can first solve the problem for the simplified $H_{reduced} = \begin{pmatrix} H_{res} & 0 \\ 0 & H_{antires} \end{pmatrix}$. This can be separated into two problems for the resonant and anti-resonant part. Denoting V_r and V_a their respective eigenvector matricies we can organise

them into $\begin{pmatrix} V_r & 0 \\ 0 & V_a \end{pmatrix}$. Using this last matrix to make a transformation of basis we will get

$$H_{full} = \begin{pmatrix} \Lambda_r & 0 \\ 0 & \Lambda_a \end{pmatrix} + \begin{pmatrix} 0 & V_r^\dagger H_c V_a \\ -V_a^\dagger H_c^\dagger V_r & 0 \end{pmatrix}. \tag{3.22}$$

Here Λ_r and Λ_a are diagonal matrices composed of eigenvalues of the H_{res} and $H_{antires}$ respectively. One can then analyse $V_r^\dagger H_c V_a$ and find the largest elements. In this new basis the idea of [92] and [93] becomes also numerically simpler as now the modification $V_i^\dagger V_j \Rightarrow V_i^\dagger diag\Lambda_r\Lambda_a V_j$, where $diag\Lambda_r\Lambda_a$ is a diagonal matrix with the eigenvalues of the resonant and antiresonant sectors on its diagonal. This means that one needs much less computational effort to calculate the matrix products. Moreover one could think of using the matrix B (Eqn. 3.14) as the modified scalar product to write down a new Haydock algorithm in this case, so as to get rid of negative norms that might cause instability of the algorithm. The way the problem has been rewritten also offers a possibility of reducing the memory space usage. (We now have only $N \times N$ dense matrices and not $2N \times 2N$ where N is the number of transitions).

3.3.2 Perturbative coupling

Let us suppose that we have a system in the form described in the previous sub-section: a diagonal part and an off-diagonal anti-hermitian coupling part. Let us now suppose that we want to do a perturbation theory on it. It turns out to be rather complicated as the levels with the increase of number of k-points become quasi-degenerate. This means that standard non-degenerate perturbation theory will start failing. Still we know that the overlap between the full eigenvectors and eigenvectors in the case we neglect coupling is large. Moreover if we start summing over an energy range it becomes essentially 1. Therefore we can apply Quasi-Degenerate perturbation theory trying to get rid of the contribution from coupling [99]. Let us note that in this formalism the matrices for the resonant and anti-resonant part remain hermitian. This is due to the fact that the non-hermitian H_c elements always come in pairs and their respective sign change will cancel. We can even extend this to the case when we have not made the basis change to the basis where H_r is diagonal as still the actual contributions will contain couples of H_c or simply H_r or H_a that are all in fact hermitian. In the end we need just to solve the eigenvalue problem for two Hermitian matrices and then perform a rotation backwards to the original basis via a rotation matrix e^S. The matrix S in the exponent being computed perturbationaly from H_c, H_r and H_a. The advantage of this stems from the fact that we use the symmetry underlying our problem reducing in the end the complexity from a $2N \times 2N$

problem to two $N \times N$ problems plus some matrix multiplications. In the case of $q = 0$ this will reduce even further to just one $N \times N$ matrix digitalization.

3.3.3 Replacing parts of GW-BSE by TDLDA

Finally, here we present a method that comes from the analysis performed in the previous section. Taking into account the fact that for Silicon the EELS spectra obtained using TDLDA and GW+BSE are very similar, and the eigenvectors seem to have quite a large overlap one can devise a method that combines the two, taking advantage of the strong sides of each: TDLDA is generally faster for solids, and in GW+BSE excitonic effects are accounted for. In fact one could split the problem of solving the 4-point equation into two. As a first step solving the equation with a local Kernel $\mathbb{K}_{local}$ that couples transitions that have similar initial (and final states). This part could be carried-out within the TDLDA framework. Then, to account for the coupling of transitions which are more distant, one would solve a second equation, this time, within BSE:

$$
\begin{aligned}
{}^4\chi_{local} &= {}^4\chi_0 + {}^4\chi_0 \mathbb{K}_{local} {}^4\chi_{local} \\
{}^4\chi &= {}^4\chi_{local} + {}^4\chi_{local} \mathbb{K}_{full-local} {}^4\chi
\end{aligned}
$$

These latter effects should be important in the case when the transition energies in different points in k-space are close of even quasi-degenerate, therefore one could also think of a method separating things in energy or frequency space as it is done Fig. 3.3.3. There we use the BSE+GW hamiltonian in transition space whenever the separation in energy is smaller than dE and the TDLDA one in the other case.

The variations in the spectra, when we replace parts of the TDLDA kernel by a BSE one are shown in Fig. 3.13

It is worth noting that this is not the only possibility for creating methods combining advantages of Many Body Perturbation Theory and Time Dependent Density Functional based methods. For example [51] use the Bethe-Salpeter equation to develop new kernels, whereas [100] suggest using TDDFT as an ingredient for simple models.

3.4 Discussions and conclusions

In the present chapter we have discussed the importance of different ingredients that contribute to *ab-initio* spectra. By using the same framework, the effective two particle hamiltonian, for the approximations to the Time-Dependent Density Functional Theory

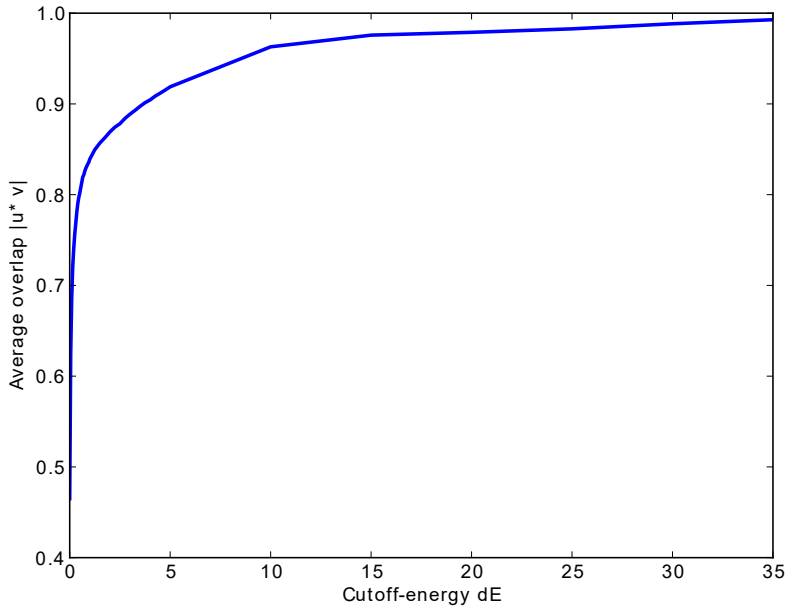


Figure 3.15 – Bulk Silicon: the average overlap between the eigenvectors of the effective hamiltonian of the Bethe-Salpeter Equation and the eigenvectors of the effective hamiltonian for the Bethe-Salpeter Equation, where parts of the hamiltonian were replaced their TDLDA counterparts as a function of the amount of replacements. 32k-points in IBZ, 25 bands, with coupling.

and the Bethe-Salpter equation we were able to see that the assignment of spectral structures to particular transitions is far from obvious. The interference between different transitions means, that even if two eigenvectors of the corresponding two-particle hamiltonians seem close, as their overlaps $\left\langle \Phi^{(1)}_{\lambda} \middle| \middle| \Phi^{(2)}_{\lambda'} \right\rangle$ are almost a delta function $\delta(\lambda - \lambda')$, this doesn't mean that the transitions they describe are the same. Moreover, when one expands the eigenvectors of one approximation in terms of the other, there can be a non-negligible contribution from the resonant part to the anti-resonant part and vice-versa. This means that even if the TDLDA and BSE spectra look the same, the analysis that comes out of them can be different.

In the second part of this chapter we have looked into possibilities of devising new methods, reducing the complexity of the Bethe-Salpeter equation. First, by taking care of the coupling between the resonant and anti-resonant of the effective two-particle hamiltonian separately. Second, by replacing parts of the Bethe-Salpeter equation by their TDLDA counter-parts. This became possible due to the fact that taking into account the

contribution from the bare Coulomb term already gives a good approximation to the final spectra in particular in Silicon. More work has to be performed to test these new approximations and calculation schemes.

It is worth noting, that to the best of our knowledge, this is the first analysis of the impact of the coupling on Electron Energy Loss Spectra in the BSE, for $q \neq 0$ [2] .

2. During the writing of the present work, independent result related to this topic appeared [101]

Chapter 4

Dynamical Structure Factor

The Dynamical Structure Factor (DSF), $S(\mathbf{q},\omega)_{\mathbf{G},\mathbf{G}'}$, and the related dielectric susceptibility $\chi(\mathbf{q},\omega)_{\mathbf{G},\mathbf{G}'}$ are crucial quantities for describing Inelastic X-ray Scattering (IXS) and Electron Energy Loss (EELS) experiments. They contain information about interband and collective electron excitations in materials. The dielectric function ε^{-1}, that can be obtained from the susceptibility χ is, moreover, a key ingredient for the interpretation of other spectroscopies such as Photo-emission (PES) and Inverse Photo-emission (IPES). In this chapter, on the one hand, we introduce and calculate the full Dynamical Structure Factor $S(\mathbf{q},\omega)_{G,G'}$ for transferred momenta $\mathbf{q}$ across the entire Brillouin zone and reciprocal lattice vectors $\mathbf{G},\mathbf{G}'$ including off-diagonal ($\mathbf{G}\neq\mathbf{G}'$). This quantity is then compared to existing experimental results obtained by Coherent X-ray Scattering (CIXS). On the other hand we calculate the full dielectric susceptibility $\chi(\mathbf{q},\omega)_{\mathbf{G},\mathbf{G}'}$ and from this the induced charge $n(\mathbf{r})$ due to different external potentials.

4.1 Definition

The diagonal elements of the Dynamical Structure Factor, $S(\mathbf{q},\omega)_{\mathbf{G}=\mathbf{G}'}$, are often measured by IXS and those of $\varepsilon^{-1}(\mathbf{q},\omega)_{\mathbf{G}=\mathbf{G}'}$ by EELS, as pointed out in subsection 1.2.2.

Back in 1982 it was shown that elements of $\varepsilon^{-1}(\mathbf{q},\omega)_{\mathbf{G},\mathbf{G}'}$ were measurable by means of inelastic scattering of X-ray photons from two different modes of the wave field [22]. Following this an experimental procedure was devised [102] and applied [103, 104]. In all of these, the key quantity of interest is the Dynamical Structure Factor. Let us derive its relation to the Coherent Inelastic X-ray Scattering experiments and the susceptibility or, alternatively the inverse dielectric function.

4.1.1 Relation to the susceptibility

In the previous chapter we have shown that the definition of the many-body dielectric susceptibility Eqn. 1.31 can be rewritten using the eigenvectors and eigenvalue of the effective two-particle hamiltonian Eqn. 3.15, in the static BSE approximation. The exact χ be written in a similar form, in particular its Fourier transform at zero temperature reads:

$$\chi(q,\omega)_{G,G'} = \sum_{\lambda} \tilde{\rho}\left(q,G'\right)^{*}_{\lambda} \tilde{\rho}\left(q,G\right)_{\lambda} \left(\frac{1}{\omega - (\varepsilon_{\lambda} - \varepsilon_{N}) + i\eta} - \frac{1}{\omega + (\varepsilon_{\lambda} - \varepsilon_{N}) + i\eta} \right), \tag{4.1}$$

where ε_{λ} is the energy of the excited state N-body state $|N_{\lambda}\rangle$, ε_{N} is the energy of the ground state $|N\rangle$, and $\tilde{\rho}\left(q,G\right)_{\lambda}$ is the Fourier transform of the electron-hole amplitude $\langle N_{\lambda}|\Psi^{\dagger}(\mathbf{r})\Psi(\mathbf{r})|N\rangle$.

The Dynamical Structure Factor, defined by its relation to the double differential cross-section of Electron Energy Loss [105], can be written as:

$$S(q,\omega)_{G,G'} = \sum_{\lambda} \tilde{\rho}\left(q,G'\right)^{*}_{\lambda} \tilde{\rho}\left(q,G\right)_{\lambda} \delta((\varepsilon_{\lambda} - \varepsilon_{N}) - \omega). \tag{4.2}$$

Using the relation:

$$\frac{1}{x+i\eta} = v.p.\frac{1}{x} - i\pi\delta(x), \tag{4.3}$$

where v.p. stands for the principal value, and the fact that $\Im(ab) = \Im(a)\Re(b) + \Im(b)\Re(a)$ and $\Re(ab) = \Re(a)\Re(b) - \Im(b)\Im(a)$ and finally that exchanging G with G' in $\tilde{\rho}\left(q,G'\right)^{*}_{\lambda} \tilde{\rho}\left(q,G\right)_{\lambda}$ is equivalent to complex conjugation, one obtains up to constant factors:

$$\Re(S(q,\omega)_{G,G'}) = \frac{\Im(\chi(q,\omega)_{G,G'}) + \Im(\chi(q,\omega)_{G',G})}{2\pi} \tag{4.4}$$

$$\Im(S(q,\omega)_{G,G'}) = \frac{\Re(\chi(q,\omega)_{G',G}) - \Re(\chi(q,\omega)_{G,G'})}{2\pi}. \tag{4.5}$$

We thus have related the Dynamical Structure factor to quantities we can calculate using *ab-initio* methods. Let us now relate it to experimental results.

4.1.2 Relation to Inelastic X-ray Scattering

As stated in the Introduction the Coherent Inelastic X-ray Scattering experiments involve the the superposition of two plane waves:

$$A_0(\mathbf{r}) = A_0 e^{i\mathbf{K}_0\mathbf{r}}$$
$$A_h(\mathbf{r}) = A_h e^{i\mathbf{K}_h\mathbf{r}}.$$

These waves, both with energy ω, are then scattered into a wave of frequency ω' and wave vector $\mathbf{K}_2$. The material, also changes its state from $|N\rangle \rightarrow |N_\lambda\rangle$.

The Inelastic X-ray scattering comes from the non-linear part of the interaction between light and matter, also known as the Thomson term:

$$\hat{\mathbb{H}}_{int} = \frac{e^2}{2mc^2}\sum_j A^2(\mathbf{r}_j). \tag{4.6}$$

Here $A(\mathbf{r}_j)$ is the vector potential operator, at the position of the electron j.

The fermi golden rule states that the transition probability is proportional to the square of the matrix element of this interaction, multiplied by the delta function describing energy conservation. It can be shown that the double differential cross-section is also proportional to this quantity:

$$\frac{d^2\sigma}{dEd\Omega} \sim \sum_F \left|\langle F|\,\hat{\mathbb{H}}_{int}\,|I\rangle\right|^2 \delta(E_F - E_I)$$

Here, F is the final state of the full system (material in state $|N_\lambda\rangle$, and photon with frequency ω'), with energy $E_F = \varepsilon_\lambda + \omega'$ and I is the initial state of the system (material in state $|N\rangle$ and photon with frequency ω), with energy $E_I = \varepsilon_N + \omega$.

Following the lines of [106] we write out the the vector potential that enters Eqn. 4.6:

$$A(\mathbf{r}) = \frac{1}{\sqrt{\omega}}\left[\hat{e}_0(\hat{a}_0^\dagger e^{i\mathbf{K}_0\mathbf{r}} + \hat{a}_0 e^{-i\mathbf{K}_0\mathbf{r}}) + \hat{e}_h(\hat{a}_h^\dagger e^{i\mathbf{K}_h\mathbf{r}} + \hat{a}_h e^{-i\mathbf{K}_h\mathbf{r}})\right] + \frac{1}{\sqrt{\omega'}}\left[\hat{e}_2(\hat{a}_2^\dagger e^{i\mathbf{K}_2\mathbf{r}} + \hat{a}_2 e^{-i\mathbf{K}_2\mathbf{r}})\right].$$

Here $\hat{e}_i$ is the polarization of the wave i and $\hat{a}_k, \hat{a}_k^\dagger$ are photon annihilation and creation operators. Furthermore, using the expressions $|I\rangle = |N\rangle|\Phi\rangle$, where $|\Phi\rangle$ is the initial state of the electro-magnetic field, and $|F\rangle = |N_\lambda\rangle|\Phi'\rangle$, where $|\Phi'\rangle$ is the final state of the

electro-magnetic field, we obtain that the double differential cross section is:

$$\begin{aligned}\frac{d^2\sigma}{dEd\Omega} \sim \big[&(\hat{e}_0\dot{\hat{e}}_2)^2 A_0^2 S(\mathbf{q}_0,\mathbf{q}_0,E)+\\ &+(\hat{e}_h\dot{\hat{e}}_2)^2 A_h^2 S(\mathbf{q}_h,\mathbf{q}_h,E)+\\ &+2(\hat{e}_h\dot{\hat{e}}_2)(\hat{e}_0\dot{\hat{e}}_2)cos(\Delta\phi)\Im\left(S(\mathbf{q}_0,\mathbf{q}_h,E)\right)\big]\end{aligned} \tag{4.7}$$

Here, $\mathbf{q}_h = \mathbf{K}_h - \mathbf{K}_2$, $\mathbf{q}_0 = \mathbf{K}_0 - \mathbf{K}_2 = \mathbf{q} + \mathbf{G}'$, $\Delta\phi$ is the phase shift between the two superimposed waves. The Dynamical Structure Factor S, as it is written in Eqn. 4.7 is related to the one introduced in the previous section as follows: $S(\mathbf{q}_0,\mathbf{q}_h,E) = S(\mathbf{q},\omega = E)_{\mathbf{G},\mathbf{G}'}$, where $\mathbf{q}_h = \mathbf{q} + \mathbf{G}'$ and $\mathbf{q}_0 = \mathbf{q} + \mathbf{G}$.

The relation between experiments and the elements of the Dynamical Structure Factor is, hence, more complicated than in the case of IXS and requires some post-treatment of experimental data. Nevertheless, we have a clear relation of CIXS to the susceptibility χ, which allows one to calculate ε^{-1}, the microscopic screening. The relation Eqn. 4.7 derived here is complicated, this explains the fact the very few experimental results are available in this case, and their error-bars are large. Let us now proceed to the next section where we describe the actual theoretical results obtained for this quantity.

4.2 Diagonal and off-diagonal elements

In the present section we will discuss why one must consider not only the diagonal, but also the off-diagonal elements of the Dynamical Structure Factor and present numerical results obtained for them.

4.2.1 Discussion

Most experimental techniques in condensed matter physics give information about the response of the system in frequency space, and not in time space, which is related to the measurement of energy. In the seminal work [107], the authors suggest to transform data from Inelastic X-ray Scattering (IXS) experiments to obtain the response of the system in real time. They, furthermore, apply it to image the induced charges in water, using the relation $n_{ind} \sim \chi v_{ext}$. However, their approach is limited by the fact that the IXS gives access to the diagonal part ($\mathbf{G} = \mathbf{G}'$) of $\chi(\mathbf{q},\omega)_{\mathbf{G},\mathbf{G}'}$, or alternatively of $S(\mathbf{q},\omega)_{\mathbf{G},\mathbf{G}'}$. Due to this, the response obtained in real space is a function of the distance $r - r'$, and thus the images of the induced charge, can only be seen as averages. Since then this study has

been extended to more complicated cases [108, 109].

In the present work we will study the full $S(q,\omega)_{G,G'}$ for transferred momenta $\mathbf{q}$ across the entire Brillouin zone and reciprocal lattice vectors $\mathbf{G},\mathbf{G}'$ including off-diagonal elements ($\mathbf{G} \neq \mathbf{G}'$). That can therefore be viewed as a matrix,

$$S(q,\omega)_{G,G'} = \begin{bmatrix} S(q,\omega)_{G_0,G_0} & S(q,\omega)_{G_0,G_1} & \cdots & S(q,\omega)_{G_0,G_N} \\ S(q,\omega)_{G_1,G_0} & S(q,\omega)_{G_1,G_1} & \cdots & S(q,\omega)_{G_1,G_N} \\ \vdots & \ddots & \ddots & \vdots \\ S(q,\omega)_{G_N,G_0} & S(q,\omega)_{G_N,G_1} & \cdots & S(q,\omega)_{G_N,G_N} \end{bmatrix}. \tag{4.8}$$

This full description gives access to the full $\chi(\mathbf{r},\mathbf{r}',\omega)$. When one applies the formalism of [107], one can calculate the induced charges on individual atoms, not averaged over different unit cells.

In this thesis it became possible to calculate the the off-diagonal elements of $\chi(\mathbf{q},\omega)_{\mathbf{G},\mathbf{G}'}$, including the electron-hole interaction, by extending the formula Eqn. 3.15, which had already been implemented in the EXC code. To be more precise, until recently, it was implemented in the form Eqn. 4.1, for $\mathbf{G} = \mathbf{G}'$. Moreover, we used the parallelization discussed in Appendix B.1 to make the calculation faster.

In the following all the coordinates of the vectors are given in terms of the reciprocal lattice translation vectors of an fcc crystal, which are given by:

$$\begin{aligned} \mathbf{b_1} &= \frac{2\pi}{a}(0,1/2,1/2) \\ \mathbf{b_2} &= \frac{2\pi}{a}(1/2,0,1/2) \\ \mathbf{b_3} &= \frac{2\pi}{a}(0,1/2,1/2), \end{aligned}$$

where a is the lattice constant. For example $\mathbf{q} = (1,0,0)$ corresponds to $\frac{2\pi}{a}(0,1/2,1/2)$.

4.2.2 Numerical results. Silicon

The first questions that we want to answer is: how do diagonal and off-diagonal elements (in $\mathbf{G},\mathbf{G}'$) of $S(\mathbf{q},\omega)_{\mathbf{G},\mathbf{G}'}$ compare to each-other in size, convergence and what is the additional information that can be learned from them? Previously, the results for the diagonal part for Silicon, have been largely discussed for TDLDA in [82] and very recently, but not in detail, for BSE [101]. The off-diagonal elements have been calculated in the Random Phase [103] and TDLDA approximations [110].

In Fig. 4.1 we show both a diagonal and an off-diagonal element for Silicon in RPA and TDLDA. One sees that diagonal and off-diagonal elements are very different: it is worth noting that the off-diagonal elements do not have to be positive, as they are not a sum of squares of absolute values of $\tilde{\rho}\,(q,G)^*_\lambda$. This means that they can be sensitive to different symmetries of the underlying excitations and distinguish between the excitation and its back-folded version. This was already noted for Si in the RPA [103].

Furthermore, one sees that even for rather large $G-G'$ the off-diagonal element remains non-negligible with respect to the diagonal one. One also sees that there is a correction that is beyond the simple Random Phase Approximation, with spectral weight shifted to lower energies.

As in the case of diagonal elements, also the off-diagonal elements calculated in the BSE are similar to the TDLDA results. Moreover, they are in good agreement with experiments for Si. This is shown in Fig. 4.2. Both TDLDA and BSE are within the error bar, therefore for this simple material one can use both of these methods, or our combined methods, as suggested in the previous chapter.

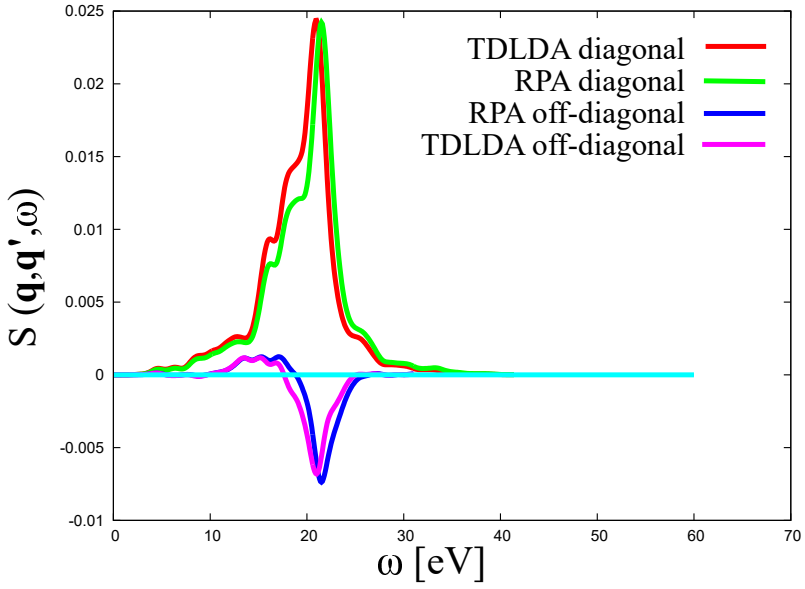
Concerning the convergence, for Si we found that there is no difference between the diagonal and the off-diagonal elements.

4.2.3 Numerical results. Lithium Fluoride

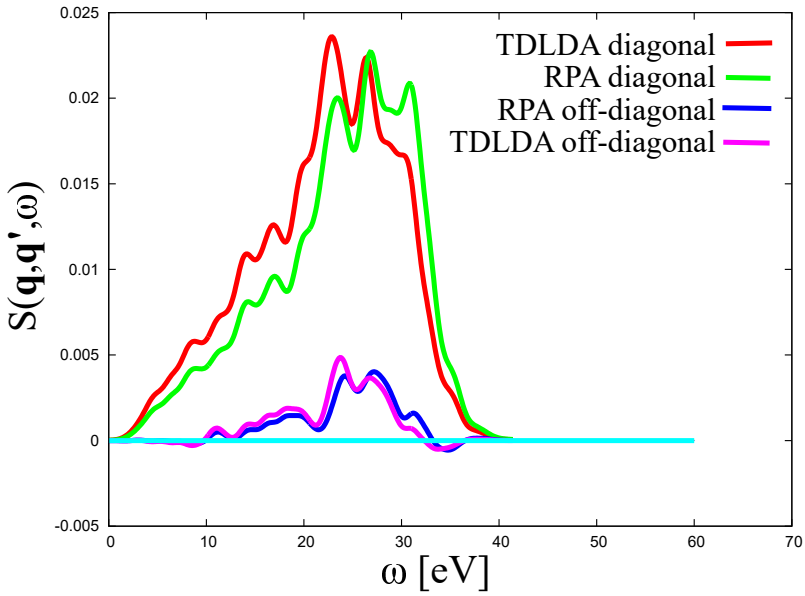
Let us now consider a more interesting material, Lithium Fluoride, which is known to have a bound exciton within the photoemission gap that is clearly visible in absorption or loss experiments [109, 111]. On the one hand we see in Fig. 4.3(a) that the TDLDA fails to produce a strong exciton. On the other hand, we see that BSE does give rise to such an exciton in the spectra. It has already been pointed out, that the BSE allows one to reproduce correctly the experimental exciton position and dispersion [81]. Interestingly, our work shows that this exciton is also well visible in the off-diagonal element in Fig. 4.3(b).

The amplitude $\Im\varepsilon$ at the excitonic peak can even be larger in some of the off-diagonal elements, than in the diagonal ones Fig. 4.4(a). This is, however, not always the case as seen in Fig. 4.4(b).

Finally, we can point out that there is not significant difference in convergence between the diagonal and off-diagonal elements for Lithium Fluoride, which means that this study can be performed for the diagonal part. More details concerning convergence for LiF can be found in the next chapter and in the Appendix D.1.



(a) RPA and TDLDA q=-0.1(5,6,4), q'=0.1(5,4,6)



(b) RPA and TDLDA q=-0.5(-1,3,1), q'=0.5(3,-1,1)

Figure 4.1 – Bulk Silicon: comparison of *ab-initio* diagonal and off-diagonal structure factors, $S(\mathbf{q},\mathbf{q}',\omega)$ obtained using TDLDA and RPA

4.3 Induced charges

We can now use the results obtained to access new interesting physics. Our first target is the induced charge due to some external potential V_{ext} (Eqn. 1.12), that can be written as follows:

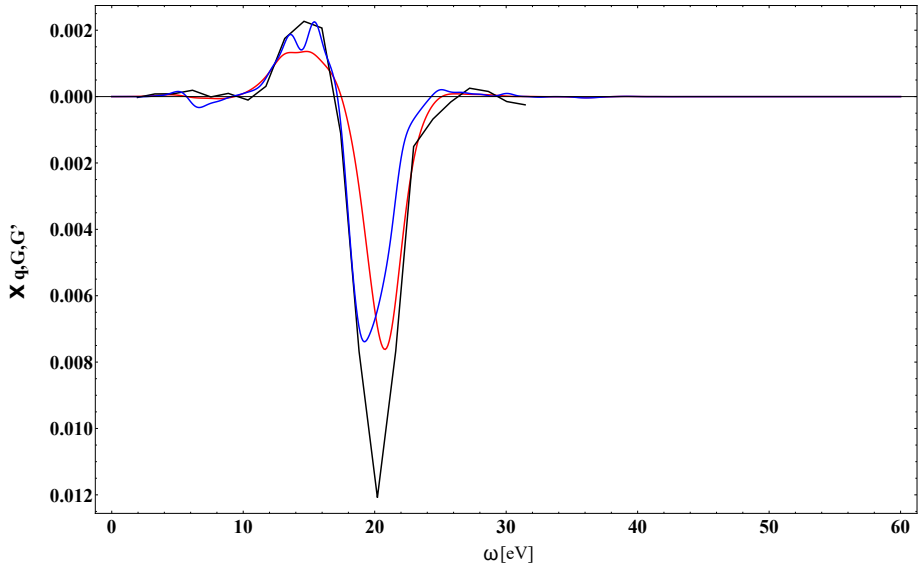


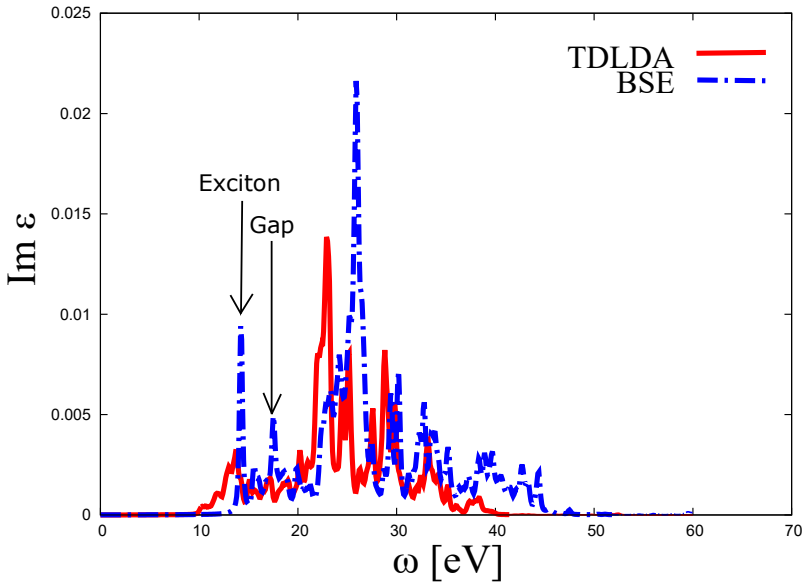
Figure 4.2 – Imaginary part of the off-diagonal element of the Bulk Silicon susceptibility obtained using the BSE (blue) and TDLDA (red) compared to the experiment (black) from [104]. Here q=0.5(1,1,1), q'=-0.5(1,1,1)

$$\delta n(\mathbf{r},t) = \frac{1}{N_q}\sum_{\mathbf{q}}\sum_{\mathbf{G},\mathbf{G}'}\int d\omega \chi(\mathbf{q},\omega)_{\mathbf{G},\mathbf{G}'} V_{ext}(\mathbf{q}+\mathbf{G}',\omega) e^{i(\mathbf{q}+\mathbf{G})r} e^{-i\omega t}, \tag{4.9}$$

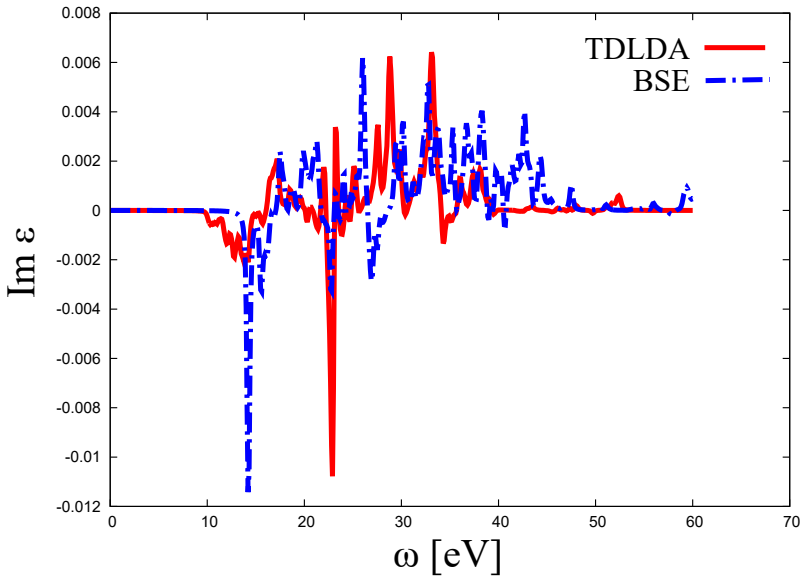
using the Fourier transforms of the susceptibility χ and the external potential V_{ext}.

Let us first forget about the frequency integral, setting $V_{ext}(\mathbf{q}+\mathbf{G}',\omega) = V_{ext}(\mathbf{q}+\mathbf{G}')\delta(\omega-\omega_0)$ and setting $t=0$. Then one will just have three sums: over G,G' and q. If one sends a plane wave, $V_{ext}(\mathbf{q}+\mathbf{G}') = \delta(\mathbf{q}-\mathbf{q}_0)\delta_{\mathbf{G}',\mathbf{G}_0}$ - then two of the three are lifted and one is left with a sum over G. One can then look at the convergence of this sum with respect to the number of G-vectors. This is not just a technical point: since the experiments are difficult it is important to make predictions about how many, and which, elements should be measured in order to obtain desired information. Of particular interest is the case when the frequency of the perturbation is close to the exciton dispersion relation $\omega_0 = \varepsilon_q$.

In the following, to visualize the charges, we make a cut through the real space distribution of charges. In both cases it is a [0,0,1] cut, with the (0,0) being located on a Silicon atom in the case of Bulk Silicon, and on a Lithium atom in the case of Lithium Fluoride. We choose $\mathbf{q}_0 = (0,-0.25,-0.25)$ in inverse lattice units. This is one of the directions in which the exciton in LiF is well visible in absorption. In the plots we have visualized two unit cells in each direction.



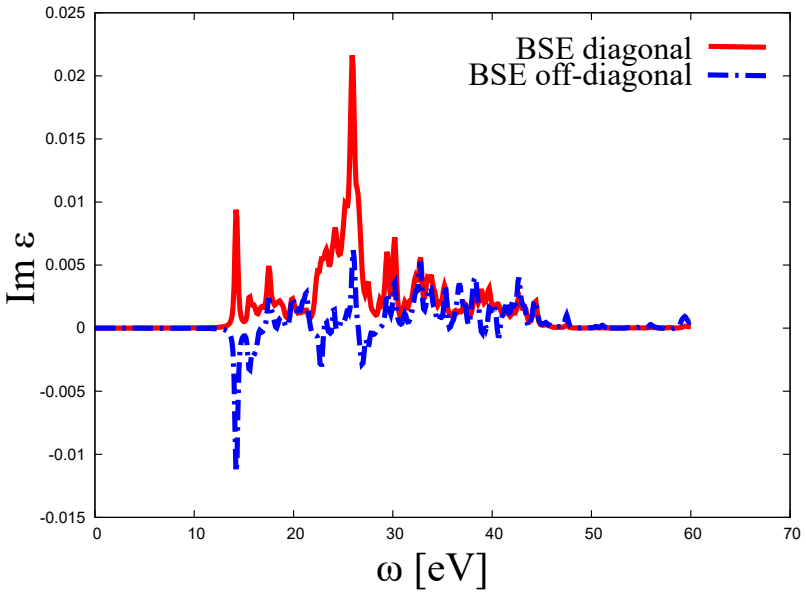
(a) Diagonal element



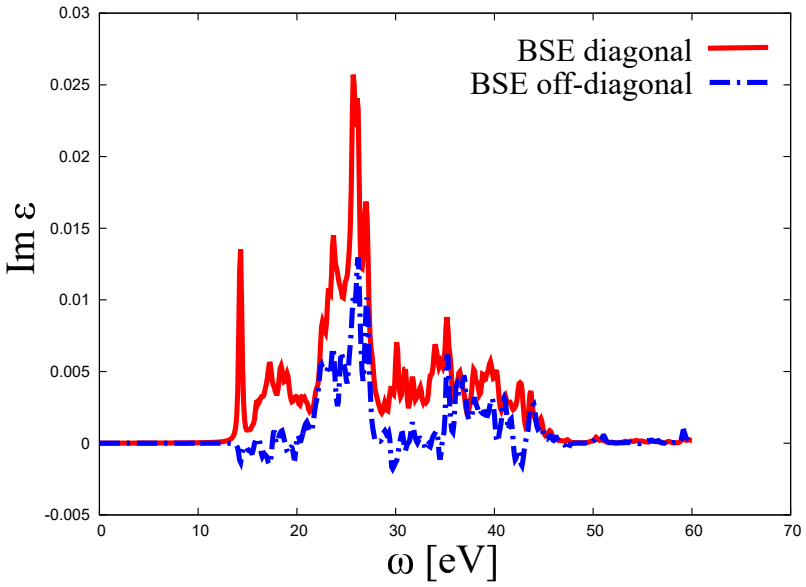
(b) Off-diagonal element

Figure 4.3 – Lithium Fluoride: comparison of diagonal and off-diagonal elements of $\Im\varepsilon$ for q = (0,-0.25,-0.25), G=0, G'=(1,1,1), obtained using the BSE and TDLDA

First we look at the results of the BSE calculation. Fig. 4.5(a) shows the induced charge obtained from Eqn. 4.9 for $V_{ext}(\mathbf{q}+\mathbf{G}') = \delta(\mathbf{q}-\mathbf{q}_0)\delta_{\mathbf{G},\mathbf{G}_0}$, for $\mathbf{q}_0 = (0,-0.25,-0.25)$ and $\mathbf{G}_0 = 0$, and only the $\mathbf{G} = 0$ element is taken into account in the sum over G. The result shows only the structure of the applied field, and not that of the material. However as soon as one starts to sum over some G vectors, thus including the off-diagonal elements, one



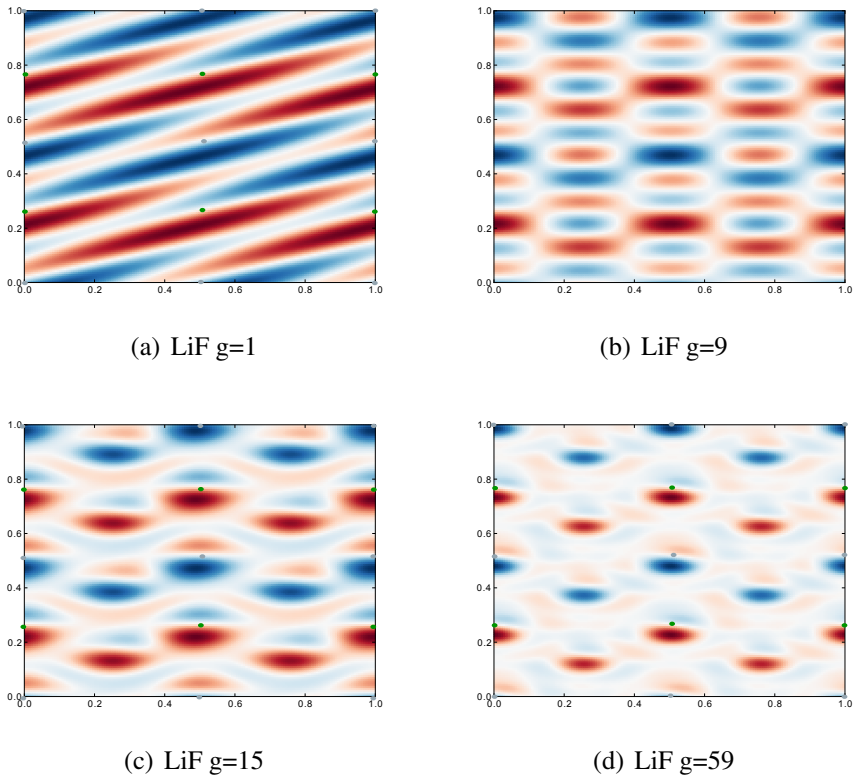
(a) q = (0,-0.25,-0.25) G=0 G'=(1,1,1)



(b) q = (0,0.25,-0.25) G=0 G'=(1,1,1)

Figure 4.4 – Lithium Fluoride: diagonal and off-diagonal elements of $\Im\varepsilon$ calculated in the BSE

gets structures, that reproduce the lattice periodicity (Fig. 4.5(b)). The fully converged results are obtained in the last figure Fig. 4.5(d). One starts getting most of the features as soon as one goes beyond the diagonal approximation. It is important to note that some of the contributions will be negligible. For example, the contribution of the exciton in Fig. 4.4(b). This allows one to restrict the sum even further. As pointed out above, this is



(a) LiF g=1 (b) LiF g=9

(c) LiF g=15 (d) LiF g=59

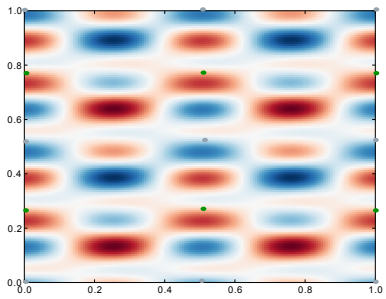
Figure 4.5 – Lithium Fluoride: number g of G vectors required to converge the induced charge density. V_{ext} is a plane wave with $\mathbf{q}_0 = (0, -0.25, -0.25)$ and $\mathbf{G}_0 = 0$. The number of G vectors is chosen such that the shell is closed. The response function used is the BSE one. The grey points are Lithium atoms, the green ones are Fluorine

important as, in experiments it is difficult to measure all the matrix elements.

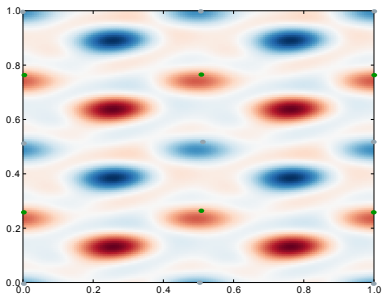
Next we study the case of a delta-like periodic perturbation $V_{ext}(r) = e^{i\mathbf{q}_0\mathbf{r}} \sum_{\mathbf{G}'} e^{i\mathbf{G}'\mathbf{r}}$, with the sum over G and G' up to a certain shell. The results are shown in Fig. 4.6. Here the induced density becomes localized even more. The convergence is similar to the previous case and for 3 shells, corresponding to 15 G vectors, most of the features of the converged induced density are reproduced.

It is also interesting to study a localized excitation. To this end one has to make a sum over the q vectors. It is worth noting that just summing over q vectors already localizes the density (Fig. 4.7).

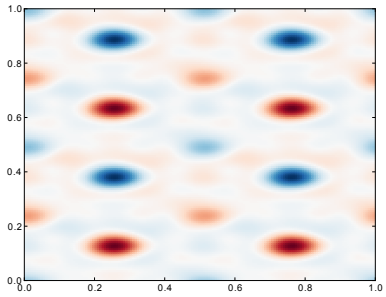
In fact even taking two elements in q already reproduces some localization, even if one takes just one G vector in each case. On the other hand some features of the spectra get smeared-out (Fig. 4.8). This smearing out is due to the fact that the exciton disperses, so for fixed ω_0 and varying $\mathbf{q}$ one mixes more or less localized excitations.



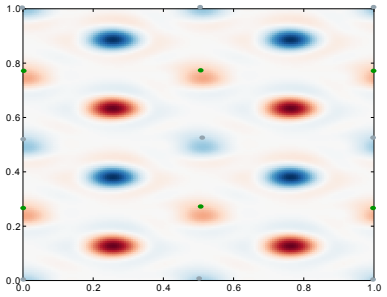
(a) LiF g=9 g'=9



(b) LiF g=15 g'=15



(c) LiF g=51 g'=51



(d) LiF g=65 g'=65

Figure 4.6 – Lithium Fluoride: number g of G vectors required to converge the induced charge density. $V_{ext}(r) = e^{i\mathbf{q}_0\mathbf{r}} \sum_{\mathbf{G}'} e^{i\mathbf{G}'\mathbf{r}}$ is a superposition of g' plane wave. The response function used is the BSE one.

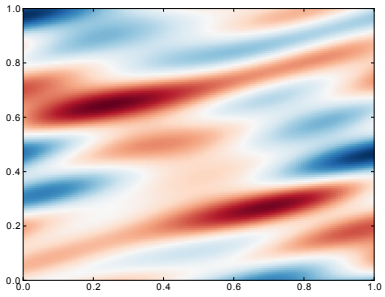
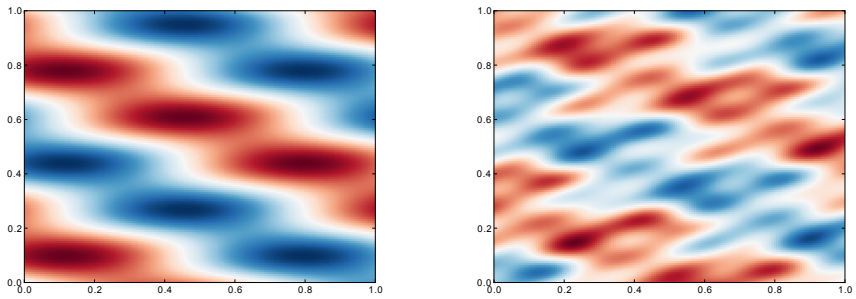


Figure 4.7 – Lithium Fluoride: the induced density, when $V_{ext}(r) = \sum_q e^{i\mathbf{qr}}$ is a superposition of 6 plane waves with different q-vectors, taken for 1 g and 1 g'. The response function used is the BSE one.

Finally one can compare the results of the Bethe-Salpeter treatment of the response with the TDLDA and RPA ones (Fig. 4.9). Only this comparison can tell us to which extent the localization is due to the external potential alone, or due to the electron-hole



(a) LiF g=1 q=1 (b) LiF g=15 q=2

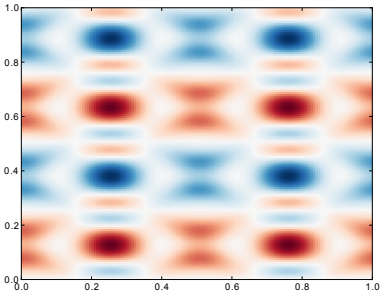
Figure 4.8 – Lithium Fluoride: study of the number of **G** vectors required to converge the induced charge density. V_{ext} is a superposition of q **q**-vectors. The response function used is the BSE one.

interaction. We see that the results obtained using a BSE susceptibility close to the exciton frequency (Fig. 4.9(e)) are very different from the TDLDA and RPA ones (Fig. 4.9(a), Fig. 4.9(c)), they present sharp features. However, if one goes closer to the plasmon frequency, the BSE and RPA results come closer (Fig. 4.9(b),Fig. 4.9(f)). One of the reasons they are different from the TDLDA ones is the difference in the plasmon peak position, clearly visible in Fig. 4.3(a). It is also worth noting that the differences between different susceptibilities become smaller as soon as we apply a more localized external potential, which corresponds to an integral over dispersing features (Fig. 4.10). In other words, in order to detect the excitonic contribution experimentally one should use a plane wave in space, rather than a localized potential.

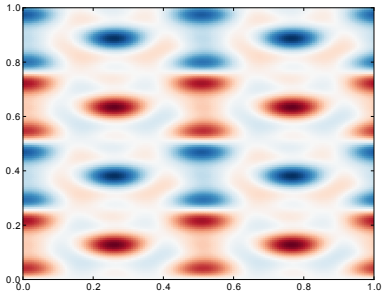
4.4 Discussion

In the present chapter we have calculated the susceptibility $\chi(\mathbf{q},\omega)_{\mathbf{G},\mathbf{G}'}$ for transferred momenta **q** across the entire Brillouin zone and reciprocal lattice vectors $\mathbf{G},\mathbf{G}'$ including off-diagonal elements ($\mathbf{G} \neq \mathbf{G}'$) for two prototypical materials, Silicon and Lithium Fluoride. This was possible thanks to the parallelization of the EXC code (See B.1 for details), and the extension of the formula for χ, in particular the treatment of the **G** dependence of $\tilde{\rho}(\mathbf{q},\mathbf{G})_{f,i} = \langle f | exp(-i(\mathbf{q}+\mathbf{G})r) | i \rangle$.

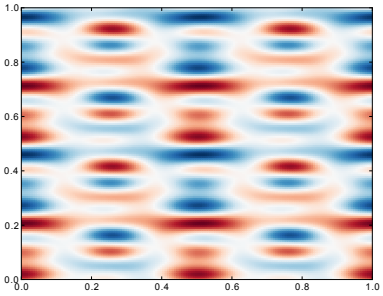
Off diagonal elements of $S(\mathbf{q},\omega)_{\mathbf{G},\mathbf{G}'}$ look qualitatively differently from the diagonal ones, and can even become negative. However, the electron-hole interaction acts in a similar manner and magnitude, both for the diagonal and off-diagonal elements, shifting



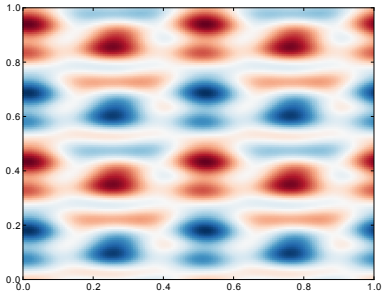
(a) RPA susceptibility, $\omega = 14eV$



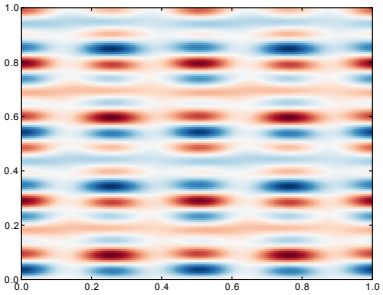
(b) RPA susceptibility, $\omega = 25.6eV$



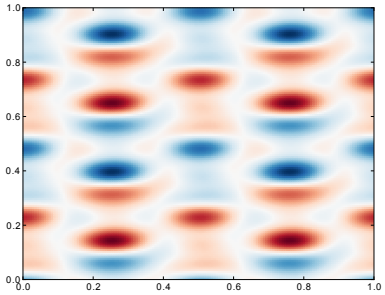
(c) TDLDA susceptibility, $\omega = 14eV$



(d) TDLDA susceptibility, $\omega = 25.6eV$



(e) BSE susceptibility, $\omega = 14eV$



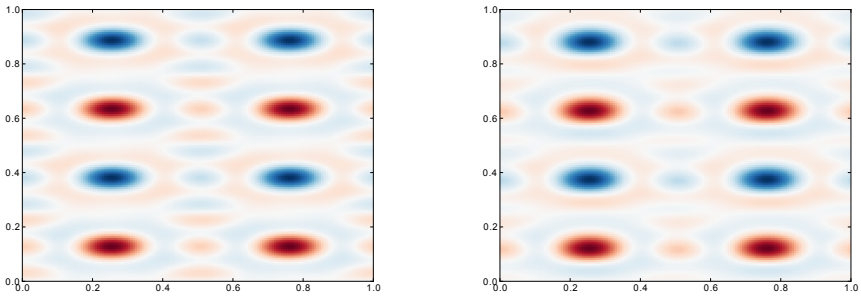
(f) BSE susceptibility, $\omega = 25.6eV$

Figure 4.9 – Lithium Fluoride: comparison of induced charge profiles for different methods and excitation frequencies ω. V_{ext} is a plane wave. $\mathbf{q}_0 = (0, -0.25, -0.25)$ and $\mathbf{G}_0 = 0$, the sum is over 27 $\mathbf{G}$ vectors.

the spectral weight to lower energies. The BSE and TDLDA seem to perform similarly for diagonal and off-diagonal elements.

The calculated $\chi(\mathbf{q}, \omega)_{\mathbf{G},\mathbf{G}'}$ was used to produce images of the induced charge. These can be used to model the real space and time evolution of different excitations in materials.

In the following chapter we will use these results to model the dielectric microscopic function, and obtain Photo-Emission Spectra.



(a) RPA susceptibility, $\omega = 14eV$ (b) BSE susceptibility, $\omega = 14eV$

Figure 4.10 – Lithium Fluoride: comparison of induced charge profiles for different methods and excitation frequencies ω. $V_{ext}(r) = e^{i\mathbf{q}_0\mathbf{r}} \sum_{\mathbf{G}'} e^{i\mathbf{G}'\mathbf{r}}$ is a superposition of 27 plane waves. $\mathbf{q}_0 = (0, -0.25, -0.25)$, the sum is over 27 $\mathbf{G}$ vectors.

Chapter 5

Excitonic satellites

Despite the progress of *ab initio* techniques, the correct description of satellite peaks in photoemission spectra still remains a challenge. Recently progress was made concerning plasmon satellites, where the hole left by the outgoing electron excites plasmons in the material. Our goal is to extend these methods to the description of satellite peaks arising from the interaction of quasiparticles and excitons. To this end, the dynamically screened Coulomb interaction W is calculated from the Bethe-Salpeter Equation (BSE), which gives a correct description of excitons. In contrast, this is an ingredient that is missing in the GW approximation that is based on the Random Phase approximation (RPA) for W.
This became possible, as we are now able to calculate not only the diagonal, but also the off-diagonal elements of the susceptibility $\chi_{\mathbf{G},\mathbf{G}'}$ which allows one to obtain the whole dielectric function $\varepsilon_{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega)$, which is a matrix in the reciprocal space, for all the $\mathbf{q}$ of the Brillouin zone.
This ε is then used to construct W which is an ingredient for the cumulant approach to the photo-emission spectra. Only including the excitonic effects can the experimental spectra of Lithium Flouride be explained, whereas the Random Phase Approximation is not sufficient to describe all the observed features. In this way, we demonstrate the existence of exciton satellites, alongside the plasmon ones, in wide gap insulators.

5.1 Ab-initio description of satellites

As seen in the introduction a typical photoemission spectra Fig. 1.1 will have quasiparticle peaks related to the solutions of the equation: $\omega - \varepsilon^0_{\mathbf{k}} - \Re\Sigma(\mathbf{k},\omega) = 0$, which enters the denominator of Eqn. 1.39. However, other structures are also clearly present, related to other types of excitations. To understand them, let us think in terms of an

independent quasi-particle picture, with different types of particles: electrons, excitons, plasmons. Each of these particles has its own excitation probability p_i and energy E_i.
Let us imagine a simple photo-emission experiment. If we illuminate our material with light of frequency ω superior to the binding energy ω_e of a given electron, we can extract it. Now, if the frequency of light is superior to the $\omega_e + \omega_b^i$, where ω_b^i is the excitation energy of some bosonic excitation i (plasmon, exciton,...) we can not only extract the electron, but also excite one boson. If the incident light frequency is superior to $\omega_e + 2\omega_b^i$, we can excite two bosons, etc.
Now assuming that the excitation probabilities are independent and that the probability of a photon with frequency ω of giving away energy ε is given by a Lorenzian, we obtain a phenomenological formula for the photo-electric current:

$$\begin{aligned} I(\omega) \sim & \frac{1}{(\omega-\omega_e)^2+\Gamma^2} + \sum_i \frac{a_i}{(\omega-\omega_e-\omega_b^i)^2+\Gamma^2} + \\ & + \frac{1}{2!}\sum_i \frac{a_i a_j}{(\omega-\omega_e-\omega_b^i-\omega_b^j)^2+\Gamma^2} + \\ & + \frac{1}{3!}\sum_i \frac{a_i a_j a_k}{(\omega-\omega_e-\omega_b^i-\omega_b^j-\omega_b^k)^2+\Gamma^2} \ldots \end{aligned} \tag{5.1}$$

Here the a_m are the probabilities of exciting a given boson m, the pre-factors $1/n!$ come from the fact that we are working with bosons, and Γ is some broadening.

Such an expression is the solution of an electron-boson coupling hamiltonian [112], and it is used in the cumulant approach to the calculation of photoemission spectra [113, 114, 115]. It has been already used to describe different features of photoemission spectra [116, 117, 118, 119, 120]. Here we summarize a schematic derivation, in the *ab-initio* context, following [114, 121]:

We start from the Dyson equation (Eqn. 1.28), rewritten, as a functional differential equation using an auxiliary external field U:

$$\begin{aligned} G(1,2) &= G_0(1,2) + G_0(1,1')u_{cl}G(1',2) + iG_0(1,1')v_c(1',2')\frac{\delta G(1',2)}{\delta U(2')} = \\ &= G_0(1,2) + G_0(1,1')u_{cl}G(1',2) + iG_0(1,1')W(1',2')\frac{\delta G(1',2)}{\delta u_{cl}(2')}, \end{aligned}$$

where we have introduced the total classical potential $u_{cl} = v_H + U$ and the screened Coulomb interaction $W = \varepsilon^{-1}v_c = \frac{\delta u_{cl}}{\delta U}v_c$. Supposing that $\frac{\delta \varepsilon}{\delta U} = 0$, and more importantly, that G and G_H are diagonal in the same basis, which is called the decoupling approxima-

tions, we obtain the following equation for every element $G_{i,i}$:

$$G_{i,i}(t_1,t_2) = G^0_{i,i}(t_1,t_2) + G^0_{i,i}(t_1,t_3)u_{cl}(t_3)G^0_{i,i}(t_3,t_2) + iG^0_{i,i}(t_1,t_3)W_{i,i,i,i}(t_3,t_4)\frac{\delta G_{i,i}(t_3,t_2)}{\delta u_{cl}(t_4)},$$

where $W_{i,i,i,i} = \int d\mathbf{r}d\mathbf{r}'|\phi_i(\mathbf{r})|^2|\phi_i(\mathbf{r}')|^2 W(\mathbf{r},t_1,\mathbf{r}',t_2)$, and $u_{cl}(t) = \int d\mathbf{r}u(\mathbf{r},t)|\phi_i(\mathbf{r})|^2$ (Repeated times are integrated over). The solution of this equation is:

$$G_{i,i}(t_1,t_2) = G^H_{i,i}(t_1,t_2)exp\left[-i\int_{t_1}^{t_2}dt'\int_{t'}^{t_2}dt''W_{i,i,i,i}(t',t'')\right]. \tag{5.2}$$

In this same decoupling approximation the G_0W_0 self-energy reads:

$$\Sigma_{i,i}(t_1,t_2) = -\Theta(t_2-t_i)e^{i\varepsilon_i(t_2-t_1)}W_{i,i,i,i}(t_1,t_2).$$

It is thus proportional to W, and we can modify the expression Eqn. 5.2:

$$G_{i,i}(t_1,t_2) = G^H_{i,i}(t_1,t_2)exp\left[i\int_{t_1}^{t_2}dt'\int_{t'}^{t_2}dt''\Sigma_{i,i}(t',t'')e^{i\varepsilon_i(t'-t'')}\right]. \tag{5.3}$$

Writing the self-energy Σ as sum of lorentzians, we obtain Eqn. 5.1. As can be seen from Eqn. 5.3, only GW ingredients are needed to obtain the cumulant spectrum.

It should be noted that the derivation does not suggest an RPA approximation for W, but rather the use of the full, measurable W. This speaks in favor of a BSE calculation for W and implies that excitonic effects could also be seen in satellites.

It is worth noting that similar approaches based on the DMFT have also been developed for the description of satellites [122].

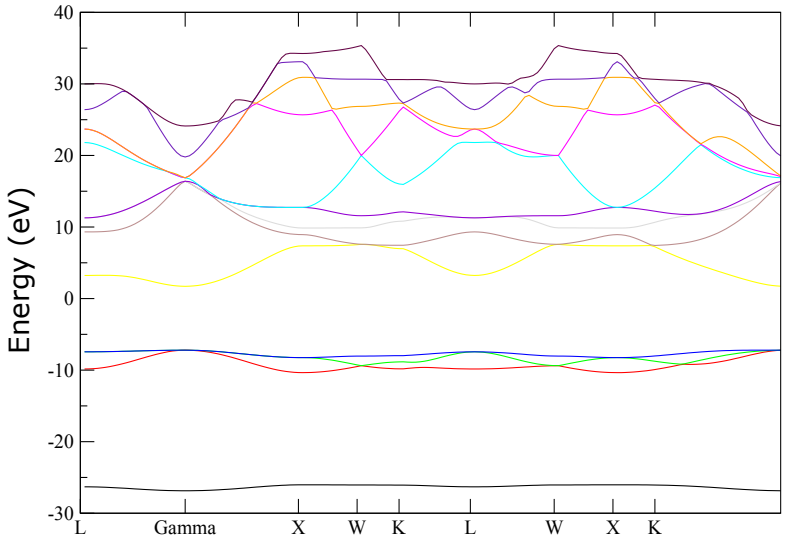
In the following we will use this cumulant expansion approach to model the photoemission spectra of Lithium Fluoride.

5.2 Lithium Fluoride: a prototypical material

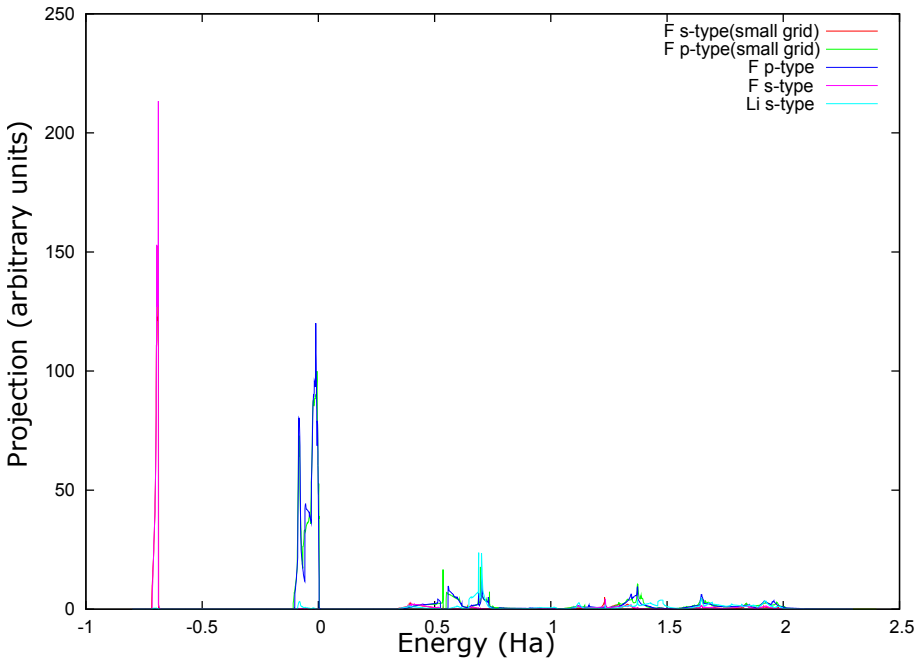
In the present section we will give a brief overview of known theoretical results for the prototypical large gap semi-conductor Lithium Fluoride.

5.2.1 Ground state properties

In the present sub-section we give details on the ground state properties of Lithium Fluoride.



(a) Ab-initio band structure in the LDA approximation



(b) Projected Density of States in the LDA approximation

Figure 5.1 – Ground state properties of Lithium Fluoride

The band structure of LiF is shown in Fig. 5.1(a). We have one lower valence band, which is essentially s-like (F 2s), and three upper valence bands which are p-like (F 2p) (Fig. 5.1(b)). We also note that the projected density of states doesn't vary much between a 4x4x4 and a 8x8x8 k-point grid. Therefore it is safe to say that the 4x4x4 grid is sufficient for integration over different q and k point in the first Brillouin zone. The primitive cell (shown in Fig. 5.2) used in the following is FCC with a lattice constant of 4.0173 angstroms. We use a 40 Ha cutoff for the plane-wave basis of the DFT calculation. The pseudo-potentials used for Li and F are standard Troullier-Martins Norm-conserving [123]. For details on convergence refer to appendix D.1.

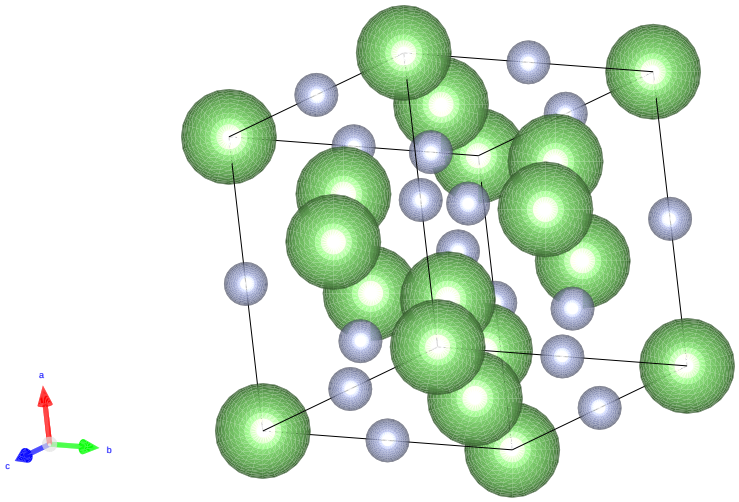


Figure 5.2 – Crystalline structure of Lithium Fluoride. Big green balls are lithium, and small grey ones - fluorine.

5.2.2 Self-energy correction

In the previous sub-section we have shortly discussed the ground state band structure obtained from Density Functional Theory. The direct photoemission gap obtained from this Local Density calculation for LiF is 8.9 eV, which is an underestimation for the experimental photoemission gap of 14.2 eV [124]. To correct it we proceed as in [125] and calculate a self-energy correction to the quasi-particle energies in the GW approximation [59, 60], taking W in the Random Phase Approximation:

$$\varepsilon_n^{QP} = \varepsilon_n^{KS} + Z_n \left\langle \Phi^{KS} \middle| \Sigma(\varepsilon_n^{KS}) - v_{xc} \middle| \Phi^{KS} \right\rangle$$
$$Z^{-1} = \left(1 - \left\langle \Phi^{KS} \middle| \left. \frac{\partial \Sigma(w)}{\partial \omega} \right|_{\omega=\varepsilon_n^{KS}} \middle| \Phi^{KS} \right\rangle \right)$$

Here Σ is the *GW* self-energy, ε_n^{QP} and ε_n^{KS} are the quasi-particle and Kohn-Sham energies respectively and Φ^{KS} is the Kohn-Sham wave function.

	Bandgap	Bandwidth	F(2p) F(2s) at Γ
LDA [126]	8.7	3.1	19.0
GW [126]	14.4	3.6	21.7
LDA [127]	8.3	3.1	21.2
GW [127]	14.3	3.5	24.8
LDA (this work)	8.9	3.1	19.4
All-electron LDA [128]	8.7	3.15	19.6
$G_{sc}W$ (this work)	14.1	3.4	21.5
$G_{sc}W_{sc}$ (this work)	15.8	3.5	22.5
so (this work)	14.15	3.5	22.5
experiment	14.2 [129]	3.5 [130]	24.9 [10]

Table 5.1 – Results for energy parameters from literature and this work

We perform this calculation using the ABINIT code [78]. The cut-off used for the screening calculation is of 30 Ha for the wave-function basis and of 12 Ha for the susceptibility matrix. The calculation is performed using the contour deformation technique, 60 frequencies on the real and 8 on the imaginary axis. The maximum real frequency used is 3 Ha which is much larger than the plasmon frequency of ≈1Ha. The cutoff used for the exchange part of the self-energy is of 44 Ha. The number of bands is 275. The calculation we have performed was self-consistent on the quasi-particle energies in *G* i.e. a GW_0 type one. The resulting self-energy correction to the gap is of 5.2 eV, which is in reasonable agreement with results from [126, 71, 127]. We have also performed a full *GW* calculation with quasi-particle energy self-consistency also in *W*. This gives an overestimation of the band-gap by 1.6 eV. On the other hand the *GW* calculation gives a F 2s F 2p splitting of 22.5 eV, which is closer to the experimental value of 24.9 eV [10] than the GW_0 which gives 21.7 eV as in [126]. This shows that it is not possible to have a unique recipe for *GW* calculation. One can imagine that in the valence and semi-core region the characteristic distances are different and therefore different contributions to the screening dominate. Since *W* is calculated in the RPA, errors cancel differently.

The self-energy corrections to the quasi-particle energies can be quite well mimicked by a scissor operator and a stretching of the valence and conduction bands $\varepsilon_n^{QP} = \varepsilon_n^{KS} \times (Stretching_n - 1) + Scissor_n$, n being the band-index. This is performed similarly to [81] and in the following we will use it whenever possible. The various LDA and GW results are regrouped in Table 5.1. It is important to note that our results are in agreement with the all-electron ones for LiF, therefore one can consider that the effect of the pseudo-potentials on the band structure is negligible. For older LDA and Hartree-Fock calculations and discussions one can look in [131, 132].

5.3 Lithium Fluoride: spectral function

In the previous section we have looked into ground state and single-particle excited states properties for Lithium Fluoride. However, our final goal is to calculate Photoemission Spectra of this material. A key quantity, that is required for this is the spectral function, related to the Green's function $A(\omega) = -\frac{1}{\pi}\Im[G(\omega)]$. In turn, the fully interacting Green's function, can be obtained once we know the self-energy Σ. Finally, in the *GW* approximation the self-energy requires the knowledge of the screening function, to which we dedicate the present section.

5.3.1 Screening model

There are different approaches to the calculation of the self-energy in the GW approximation such as the Plasmon-pole models [133, 60] or the Contour-deformation technique [125]. Moreover GW itself exists in different flavours depending on whether we do a self-consistent calculation on G, W or both.

First of all even though the Plasmon-pole model gives a reasonable approximation for the GW gap (14.0 eV) as compared to the one-shot GW using contour-deformation (13.4 eV) it is quite far away from the corresponding results for the spectral function. This is predictable because it only contains one plasmon, which is insufficient to model the screening to a degree of accuracy required for the spectral function.

The second point is that the result obtained in the G_0W_0 approximation is not satisfactory as neither the gap, nor the F 2s F 2p splitting is well reproduced. Therefore one has to go to he next level of approximation i.e. do a self-consistent calculation on both G and W or at least one of them. It turns out that a self-consistent calculation on G gives results that are reasonably close in shape to the fully quasi-particle self-consistent GW (Fig. 5.3),

even though the F 2s F 2p splitting is not well reproduced. Therefore if one is interested in the satellite structures and not the exact positions of the peaks on can limit oneself to this approximation. A discussion of the effect of self-consistency on the quasi-particle energies can be found in [134].

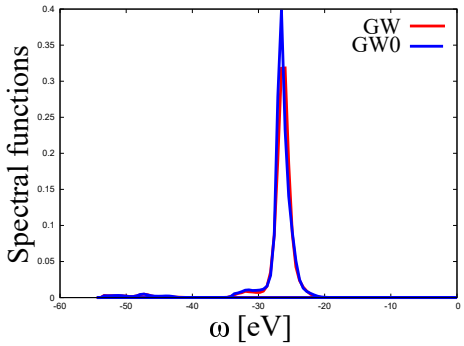
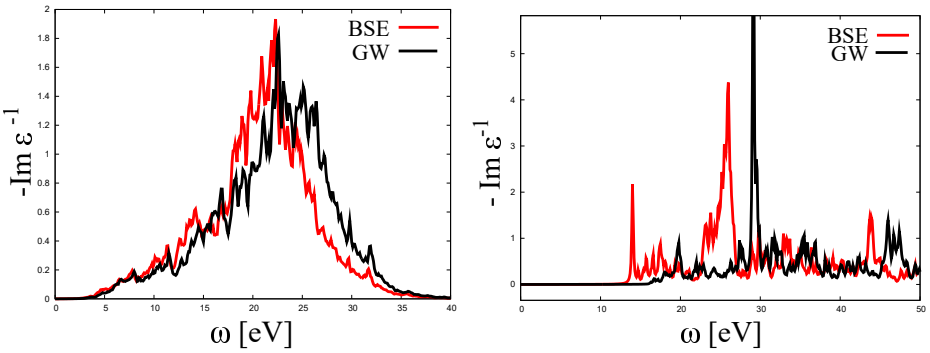


Figure 5.3 – Lithium Fluoride: the spectral functions of the lower valence band in different approximations to the self-energy. Self-consistency on G and W vs self-consistency on G only.

5.3.2 Screening function from the BSE

In the present section we will present the results obtained for LiF using the newly developed extension of the EXC code. First of all it is worth noting that in the case of LiF ε^{-1} and hence W have a non-trivial dependency on ω. This can readily be seen when we look at the EELS spectra of Si (Fig. 5.4(a)) and LiF (Fig. 5.4(b)). In Silicon we have a spectrum that is dominated by a single plasmon, in LiF we have different structures that will be identified with an exciton, a plasmon and inter-band transitions. This means that a simple plasmon-pole model approach is not be sufficient to describe the dielectric function of LiF. We will therefore use the contour-deformation technique, that allows one to capture more details of the spectra. Moreover, we can expect the results in satellite spectra are much richer and interesting.

In the case of Si, the spectral function in the contour-deformation method does not change much if we vary the number of real frequencies (See Fig. 5.5). This confirms the fact that if we have very broad structures in W, it is sufficient to use a coarse sampling or even a plasmon-pole model. In the case of LiF, we have both broad structures (plasmon) and sharp structures (exciton) and it turns out that the results for the Spectral function requires a large number of frequencies on the real axis (about 600, with respect to 30 for Silicon). Another important fact is that it requires a large number of empty states in the Green's function (about 200) to obtain good results, whereas to obtain a decent position of



(a) EELS spectra of Si q=(0.75,0.75,0.75) (b) EELS spectra of LiF q=(0.25,0.0,0.0)

Figure 5.4 – Loss Functions of bulk Silicon and Lithium Fluoride in the RPA and BSE approaches baes on the GW approximation to the self-energy

the exciton this is not needed [81]. All together, it should be observed that the calculations presented here are huge, with memory occupation in the Gigabyte range and hundreds of thousands of CPU hours to obtain a typical spectrum.

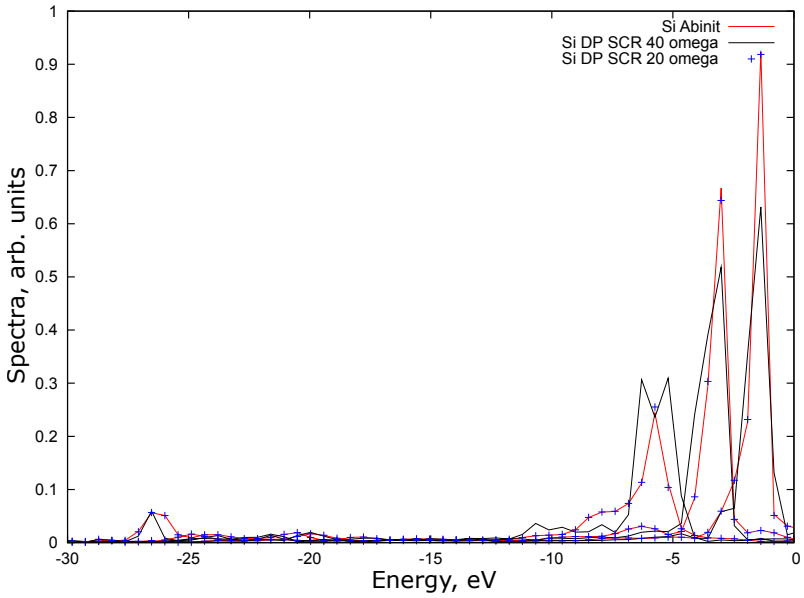
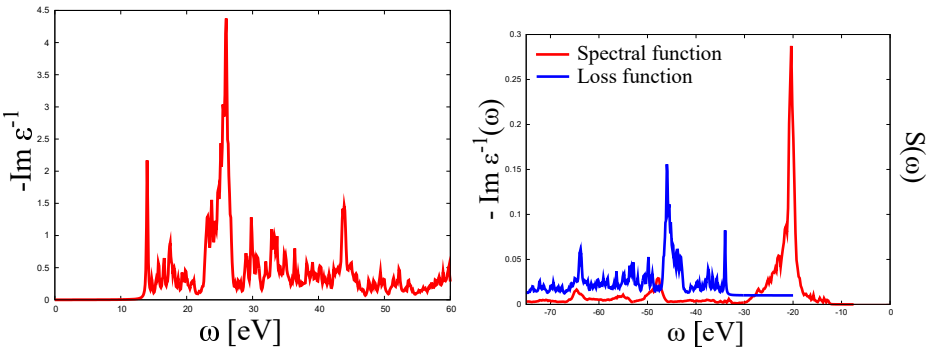


Figure 5.5 – Spectral function for bulk Silicon, computed using Abinit and DP codes, and form different numbers of frequencies in the contour deformation method.

Before we continue let us see whether the calculated spectra can be understood. For this we first calculate an EELS spectrum using the EXC code (Fig. 5.6(a)). Then we compare it with the *GW* spectral function (Fig. 5.6(b)). Here we have aligned the EELS spectrum to the quasi-particle position. For sharp excitations there should be satellites in the spectral function at every peak of EELS. We see that some of the structures in

the Spectral function correspond to the structures in the EELS. This can be seen for the plasmon-satellites that seem to be at the correct position. One however still needs to do a more elaborate cumulant expansion [113] to be sure that we are looking at plasmon satellites, and not a fake plasmarons [114], quasiparticle-like excitation introduced in the *GW* approximation.



(a) Loss function for q=0.25,0,0

(b) The loss function, shifted by the quasi-particle energy, and the spectral function

Figure 5.6 – Lithium Flouride: loss and spectral functions

What is already clear at this stage is that the exciton is not very visible. This is due to the fact that in the calculation of the spectral function we are summing over a whole set of q-points, for which the exciton is at different positions (Fig. 5.7). This is consistent with the results obtained previously in section 4.3 for the induced charge density.

5.4 Lithium Fluoride: satellites

In the previous section we have obtained the Lithium Fluoride self-energy, corresponding to the *GW* approximation. Now, in this section we will apply the state of the art cumulant approach [114], that we have extended to account for the excitonic effects to the calculation of the Photo-emission Spectra.

First of all we point out that the spectral function (Eqn. 1.39) calculated using the GW approximation for the self-energy can give results that are incorrect. For instance it can be seen if we look at the real and imaginary parts of the self energy, for the lower valence band of Lithium Fluoride (Fig. 5.8). The zero of $\omega - Re\Sigma - E_{hartree}$ give rise not only to the quasi-particle peak (around 20 eV), but also to a second crossing of the real part, which is not damped by the imaginary part. This yields an unphysical plasmaron at 50 eV. This is clearly visible in Fig. 5.9.

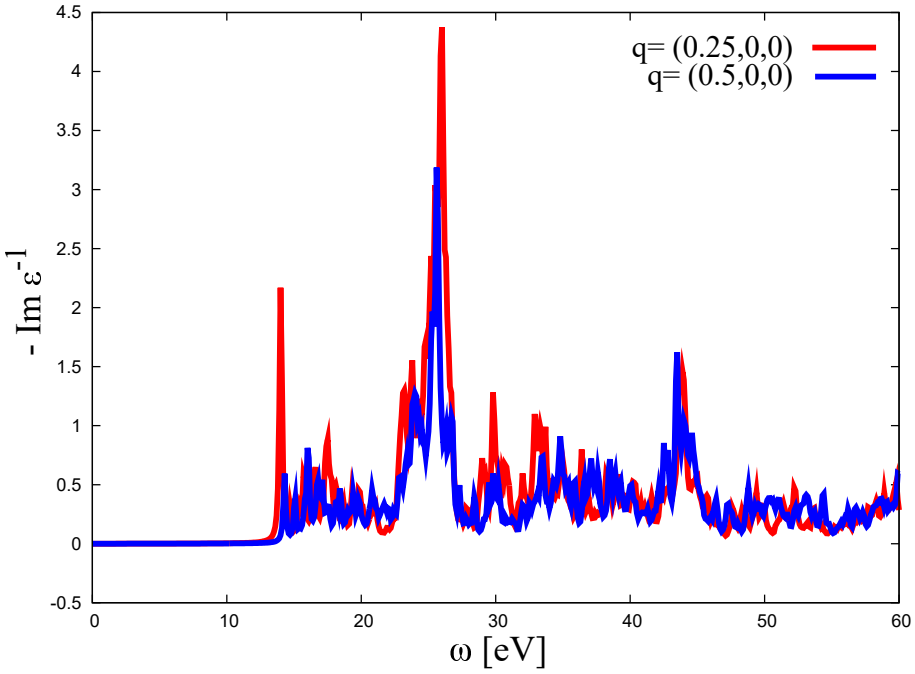


Figure 5.7 – Loss function of Lithium Fluoride for two different transferred momenta q.

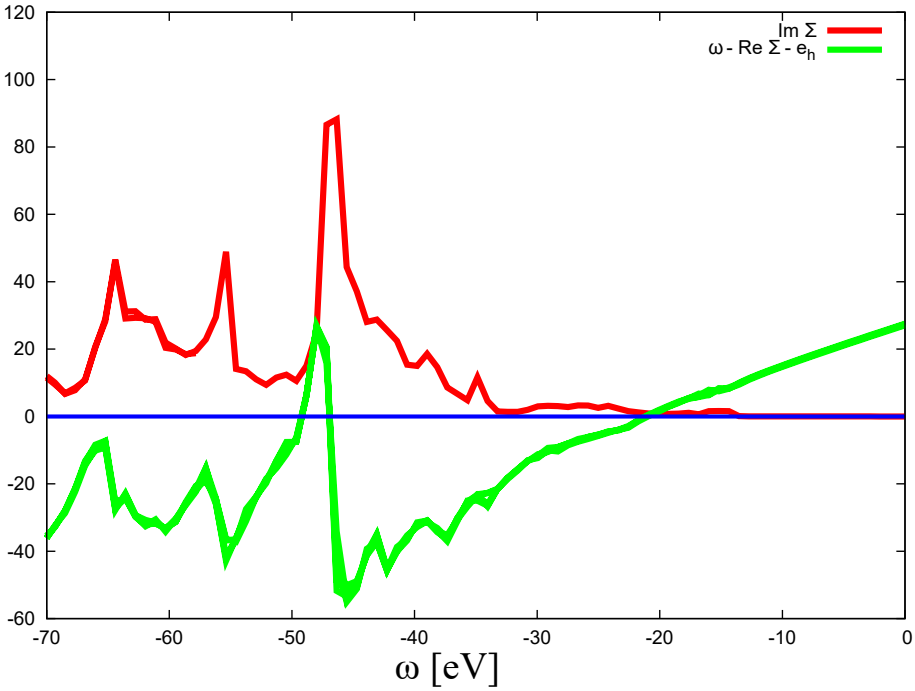


Figure 5.8 – Lithium Fluoride: the real and imaginary parts of the self-energy (Sigma), in combinations which can give rise to peaks in the spectral function.

This problem of *GW* can be cured using the cumulant expansion for the spectral function. The result of its application is compared to the GW one in Fig. 5.9. We see that

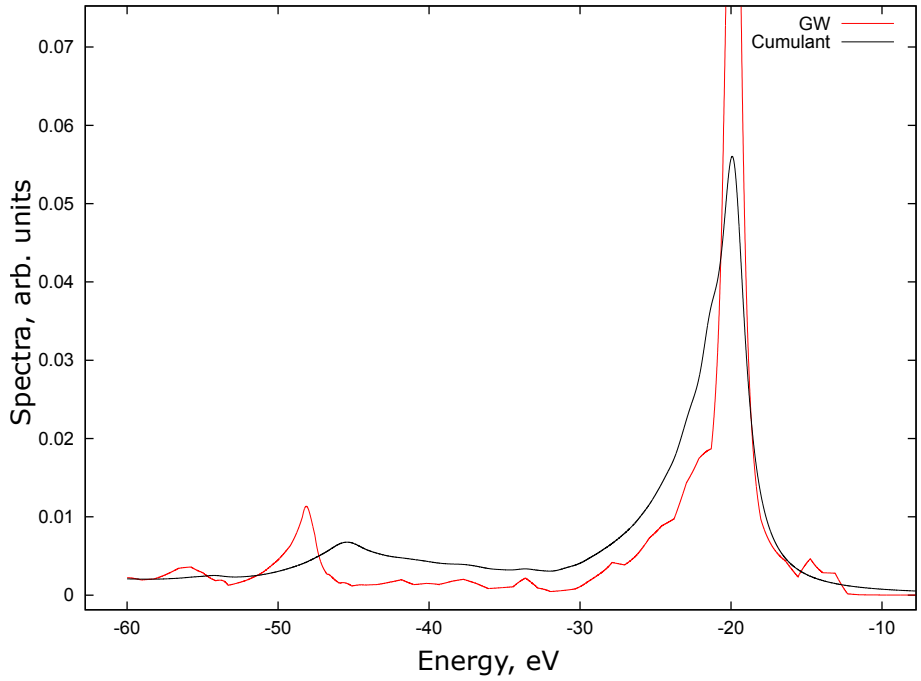
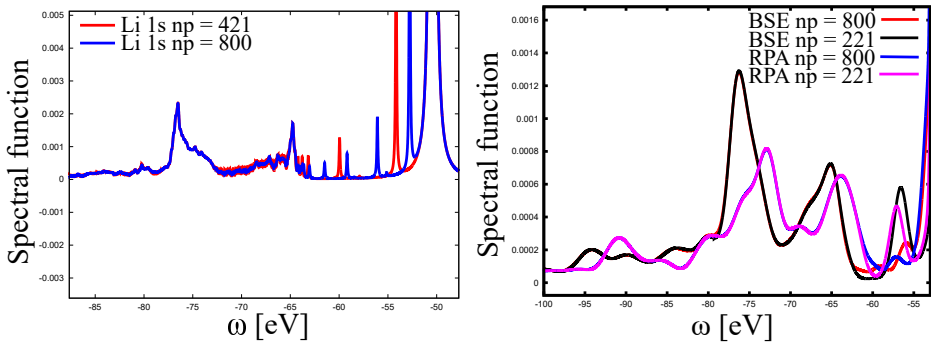


Figure 5.9 – Spectral functions of the lower valence band of Lithium Fluoride obtained with screenings generated using Abinit and EXC codes. In one case in the GW approximation, in the other using the cumulant expansion.

the non-physical plasmaron has disappeared, and has been replaced by a smooth plasmon satellite at the expected distance from the quasi-particle peak.

The key ingredient that is needed for the cumulant expansion of the spectral function is a well sampled screened coulomb interaction W, and the self-energy Σ. One element of this sampling is the number of real frequencies used to describe W (Fig. 5.6). The other one is the number of poles used in the cumulant expansion. This last quantity is particularly important for the deep bands. To illustrate this, we look at the spectral function of the Li1s band (Fig. 5.10). We see that, close to the quasi-particle peak we have some not well converged structures (Fig. 5.10(a)). However, by comparing the RPA and BSE spectral functions (Fig. 5.10(b)), we see that these structures are not due to some physical effect, but are a result of slow convergence of the spectra close to the quasi-particle peak. The remainder of the spectra doesn't change when we increase the number of poles from 221 to 800.

In the previous sections, when considering different approximations for the self-energy, we have noted that using a self-consistent G in the self-energy Σ was the most suited of them, as it reproduces well the gap, and the shape of quasi-particle peaks. Similarly to



(a) Cumulant approach with varying number of poles.

(b) Comparing RPA and BSE cases

Figure 5.10 – Spectral function of the Li1s band of Lithium Fluoride: comparison of the results in the cumulant approach based on the RPA and static BSE approximations to W, for different numbers of poles

the fact that the Green's function enters the Kernel of its Dyson equation, the susceptibility enters the Kernel (through the screening of the Coulomb interaction) of its Dyson equation, the BSE (Eqn. 2.5). We can, therefore, iterate also the screening calculation towards self-consistency. The results of the second iteration, where we reinsert the BSE screening into the BSE equation, are given in Fig. 5.11. The spectral function acquires an extra shift, of 0.25 eV. Such a shift has been seen in quartz in [135]. However, in the present work the full screening was used, whereas in [135]only a renormalization of the screening constant was applied. We further observe that a matrix dimension of 65 inverse lattice vectors is sufficient for convergence, as we see no difference when we increase the matrix dimension to 256.

To be able to compare the present calculations to experiments, one has to take into account the X-ray cross-sections which differ for s and p type electrons. The lower valence is essentially s-type, corresponding to the F 2s electrons, the upper valence is itself F 2p (Fig. 5.12(a)). For the experimental photon energy around 1500 eV the corresponding cross-sections relate to each other as 28:6 Fig. 5.12(b). This is seen in the experiment where the quasi-particle peak corresponding to the lower-valence(F 2s) band is more intensive than the upper valence one.

Furthermore, experimental results also include a background. Instead of subtracting it from the experiment, as it is performed in many cases, we add an *ab-initio* background to our results, using an approach suggested in [136]. In this approach the background $B(\omega)$ is obtained using a simple formula $B(\omega) = \int_{\omega}^{E_{fermi}} A(\omega)$. The background then assumes a step-like behavior, with most of the additional weight added at quasi-particle or satellite peak positions (Fig. 5.13).

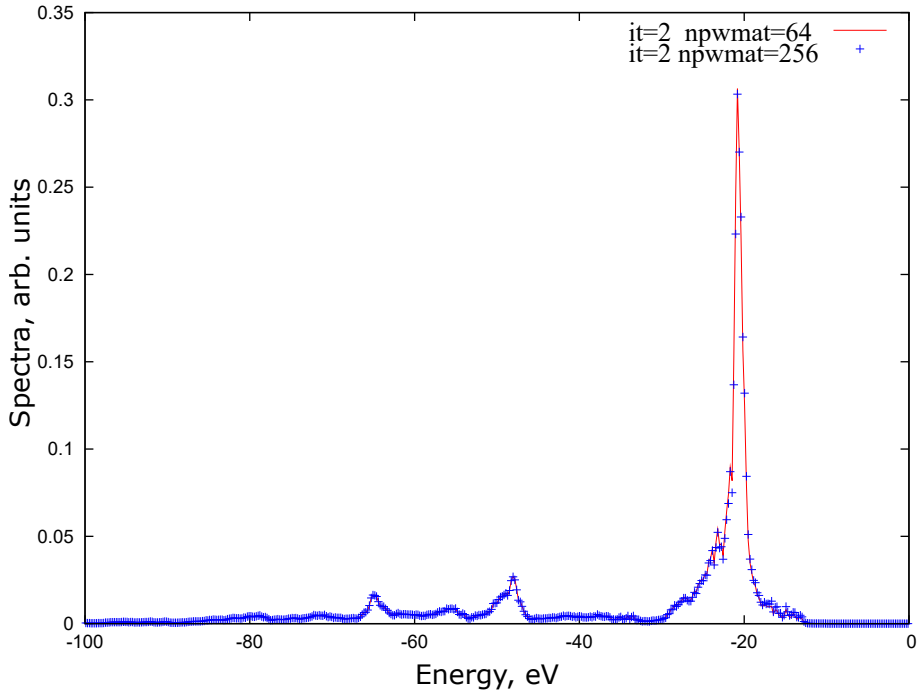
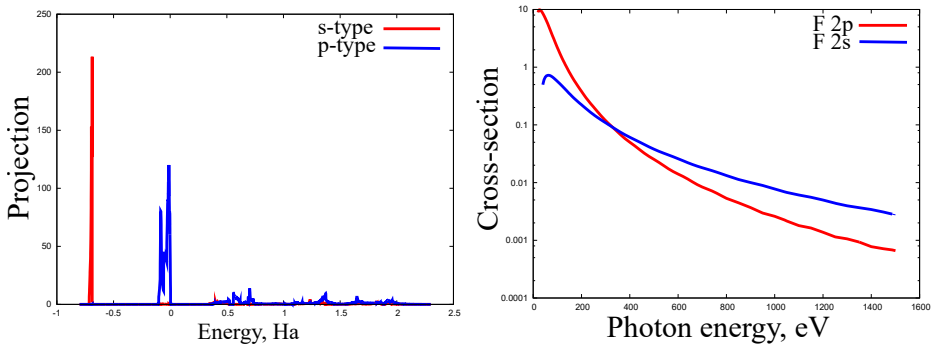


Figure 5.11 – Spectral function of Lithium Fluoride: towards self consistency in W in the BSE equation.



(a) Projected density of States (F atom) (b) Cross-section for F 2s and 2p elections

Figure 5.12 – Lithium Fluoride. The projected density of states and the photo-absorption cross sections for Lithium and Fluoride.

Having all this in mind we can proceed to the comparison of the full spectra and identification of specific features in it with various types of excitations without the background (Fig. 5.14) and with it (Fig. 5.15). The calculations performed are based on a self-consistent G, that was used for the starting point of the self-consistent BSE. The resulting W and self-consistent G were then used as ingredients for the cumulant expansion of the spectral function. We can see that there is a rather good general agreement with

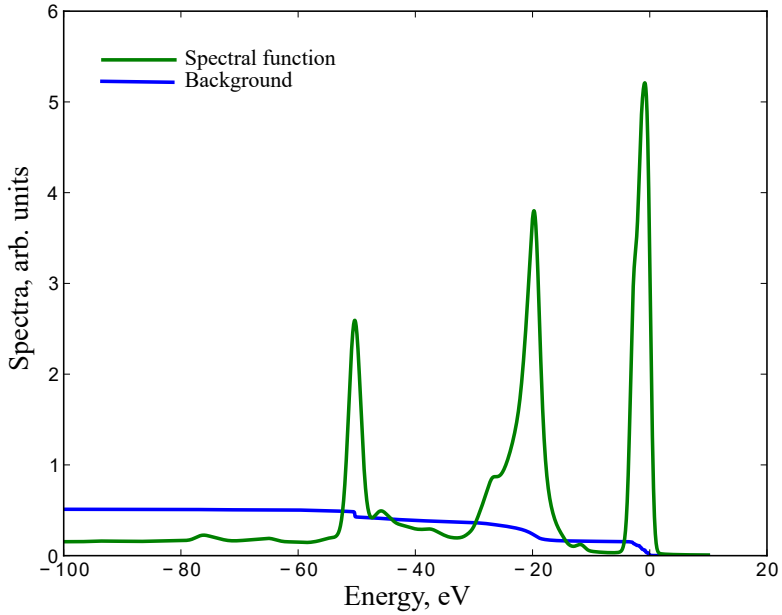


Figure 5.13 – Lithium Fluoride: theoretical prediction of the experimental background using the Shirley formula [136].

both of the experiments even though the results do not match completely. However, in view of the discrepancy between different experimental results, it would be difficult to discuss the remaining differences.

We can now proceed to the analysis. To do this we plot separately the contributions that come from different valence bands. This allows one to separate the structures. For the F2p band (Fig. 5.16), there is a first excitonic satellite around -13 eV. There is then a plasmon satellite that comes in at -27 eV and interband transitions from -35 eV to -45 eV that also come from the first peak. For the second valence band (Fig. 5.17) we have the same series. An exciton satellite which is around -30 eV, a plasmon (-45 eV) and inter-band satellites starting from -53 eV.

The nature of the inter-band transitions can be identified using the loss-function calculations. In Fig. 5.18 we have plotted the BSE and RPA spectra on top of each other. We see see that the spectral weight between 30 and 38 eV is quite similar and it is quite safe to expect that it comes from similar transitions. As in the RPA, the transitions are completely independent, we identify them to be F 2p $\rightarrow$ Li 2s and Li 2p by eliminating selected transitions in Fig. 5.18 for illustration.

The W we have used does not include the Li 1s band, that is located at -50 eV, as its contribution to the screening is not expected to be large. Nevertheless we still calculate the self-energy Σ for this band and then, applying the cumulant expansion, get a

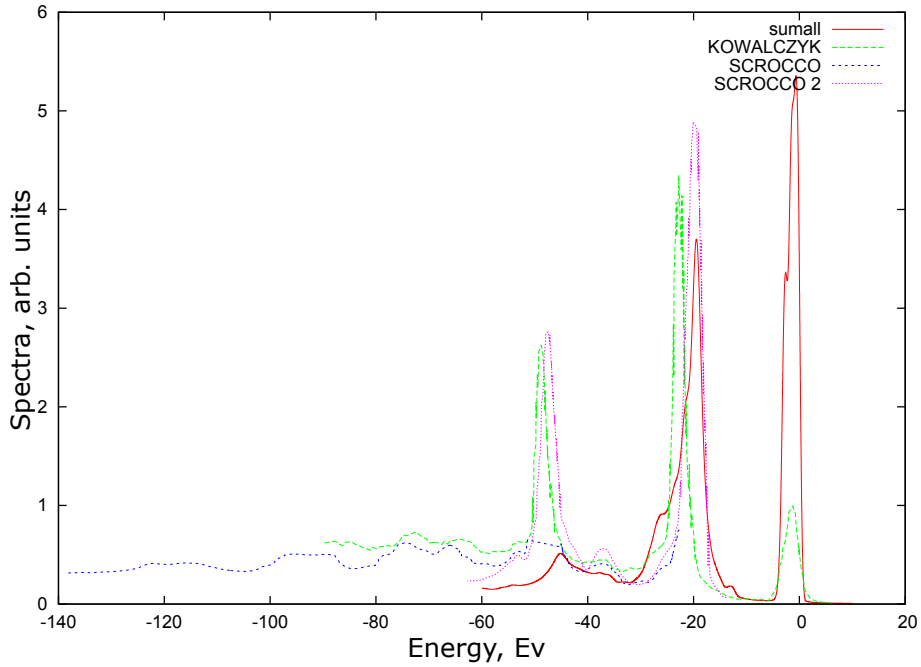


Figure 5.14 – Lithium Fluoride: full *ab-initio* spectral function (sumall) compared to experimental results from [137] and [138]

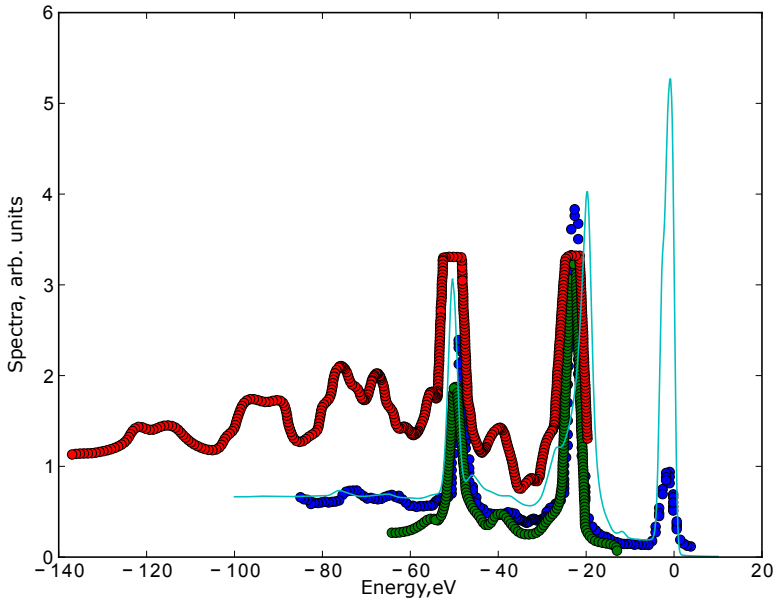


Figure 5.15 – Lithium Fluoride: comparison or the full *ab-initio* spectra with background (solid line) with experimental data (circle) from [137] and [138]

reasonable description of the satellites. We distinguish the main structures: the excitonic

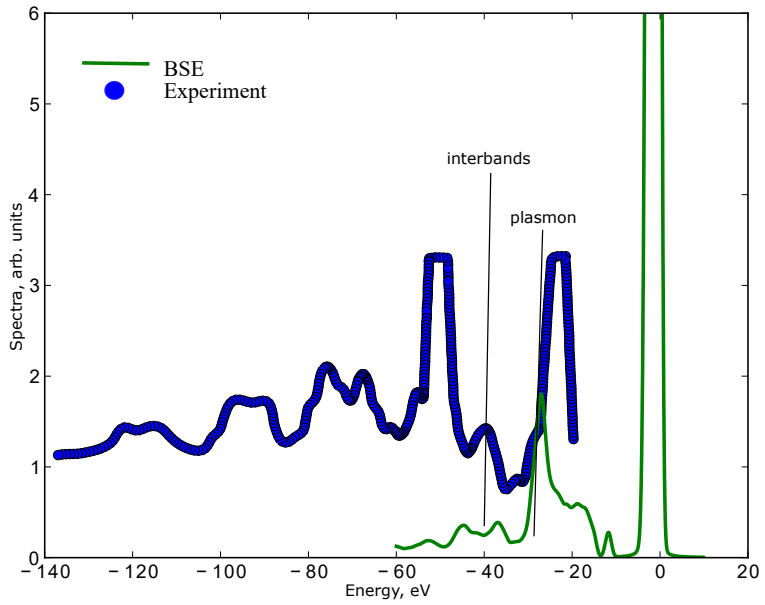


Figure 5.16 – Lithium Fluoride: assignment of peaks in the experimental spectra to corresponding satellites of the F2p quasi-particle peak

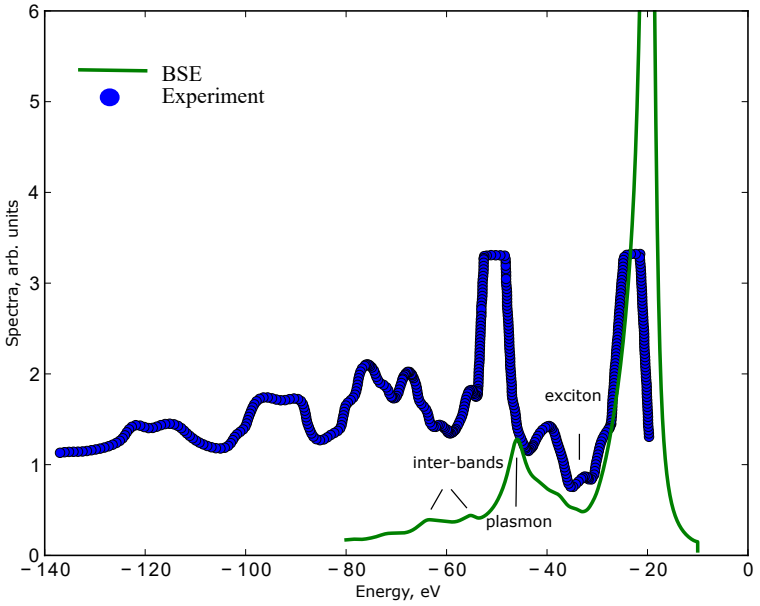


Figure 5.17 – Lithium Fluoride: assignment of peaks in the experimental spectra to corresponding satellites of the F2s quasi-particle peak

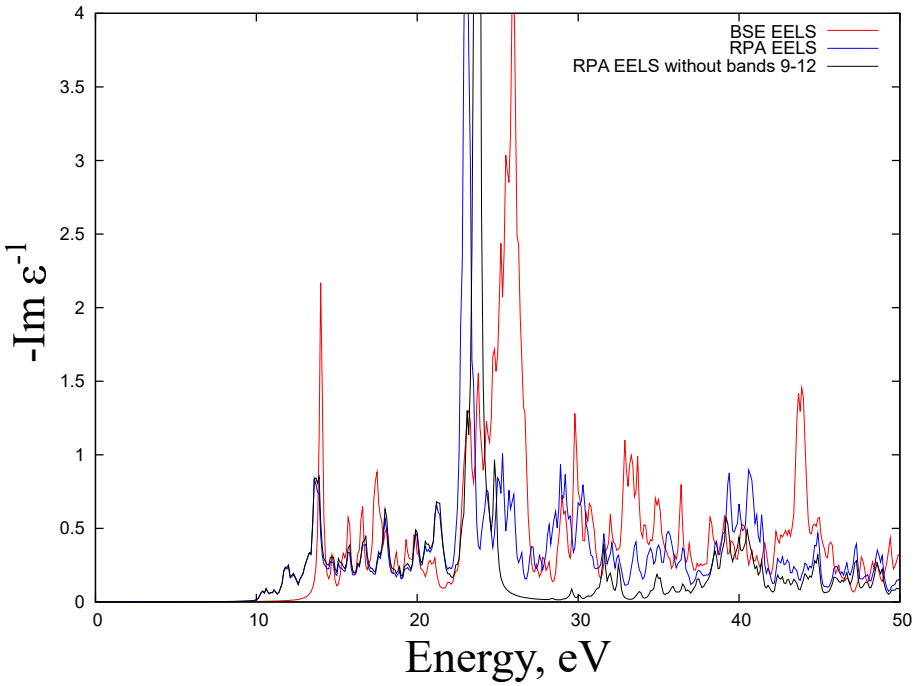


Figure 5.18 – Loss function of Lithium Fluoride: in the static BSE approximation, in the RPA approximation, in the RPA with excited states 9 to 12 removed.

satellite around -64 eV and a plasmon at -76 eV, a second excitonic satellite at -80 eV, and interbands starting from -85 eV (Fig. 5.19).

Finally, we compare the two spectra for the Li1s and F2p bands, as their corresponding quasi-particle peaks are sharp (Fig. 5.20). It allows us to give on the one hand an extra confirmation of the correct peak assignment and on the other hand it allows one to see to which extent the superposition of spectral features can contribute to the masking of some of the features. For instance, the second excitonic satellite that is visible in the Li1s spectra is almost completely masked by the broad plasmon satellite for the F2p band that is located at an energy that is close by. This assignment of peaks would not have been possible within the RPA approximation (Fig. 5.21).

5.5 Discussion

In this chapter we have shown that in order to describe the Photo-Emission Spectra of large gap semi-conductors it is important to go beyond the Random-Phase Approximation on the one hand, and apply the cumulant approach on the other. Using the screening $\varepsilon_{G,G'}(\mathbf{q},\omega)$ obtained from the Bethe-Salpeter equation, including its off-diagonal elements $G \neq G'$, we have assigned the various peaks in the experimental spectra of LiF

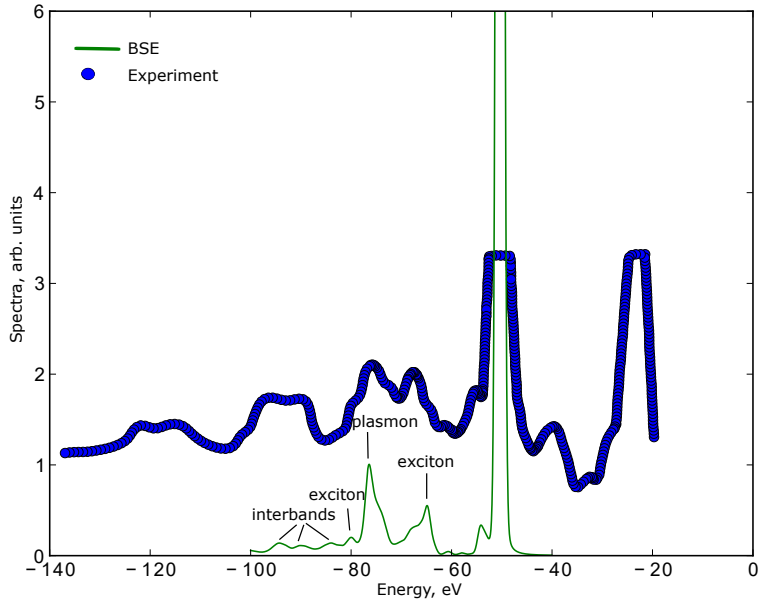


Figure 5.19 – Lithium Fluoride: assignment of peaks in the experimental spectra to corresponding satellites of the Li1s quasi-particle peak

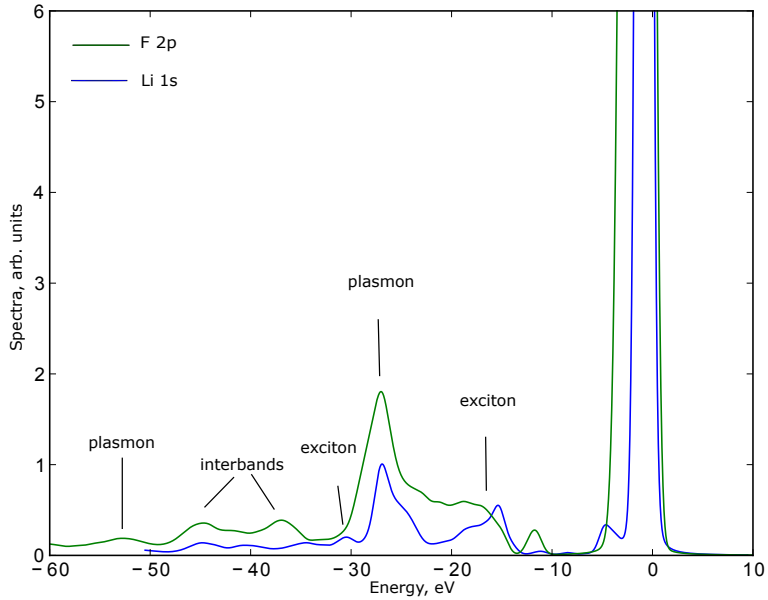


Figure 5.20 – Lithium Fluoride: comparison of the satellite structure of the Li1s and F2p bands to the spectral function

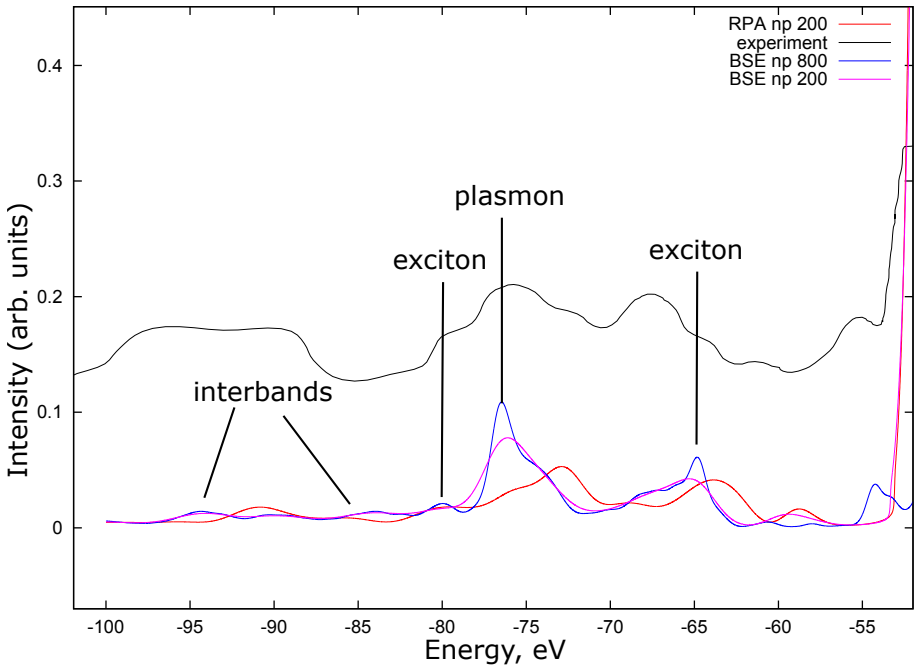


Figure 5.21 – Lithium Fluoride: comparison of the satellite structure of the Li1s obtained using the cumulant expansion with an RPA and BSE W.

to their theoretical counterparts. This is the first time that an *ab-itinio* proof of existence of spectral features arising from the coupling of excitons and quasiparticles was accomplished.

The way the calculations were performed and the "added value" of this thesis is resumed in Fig. 5.22.

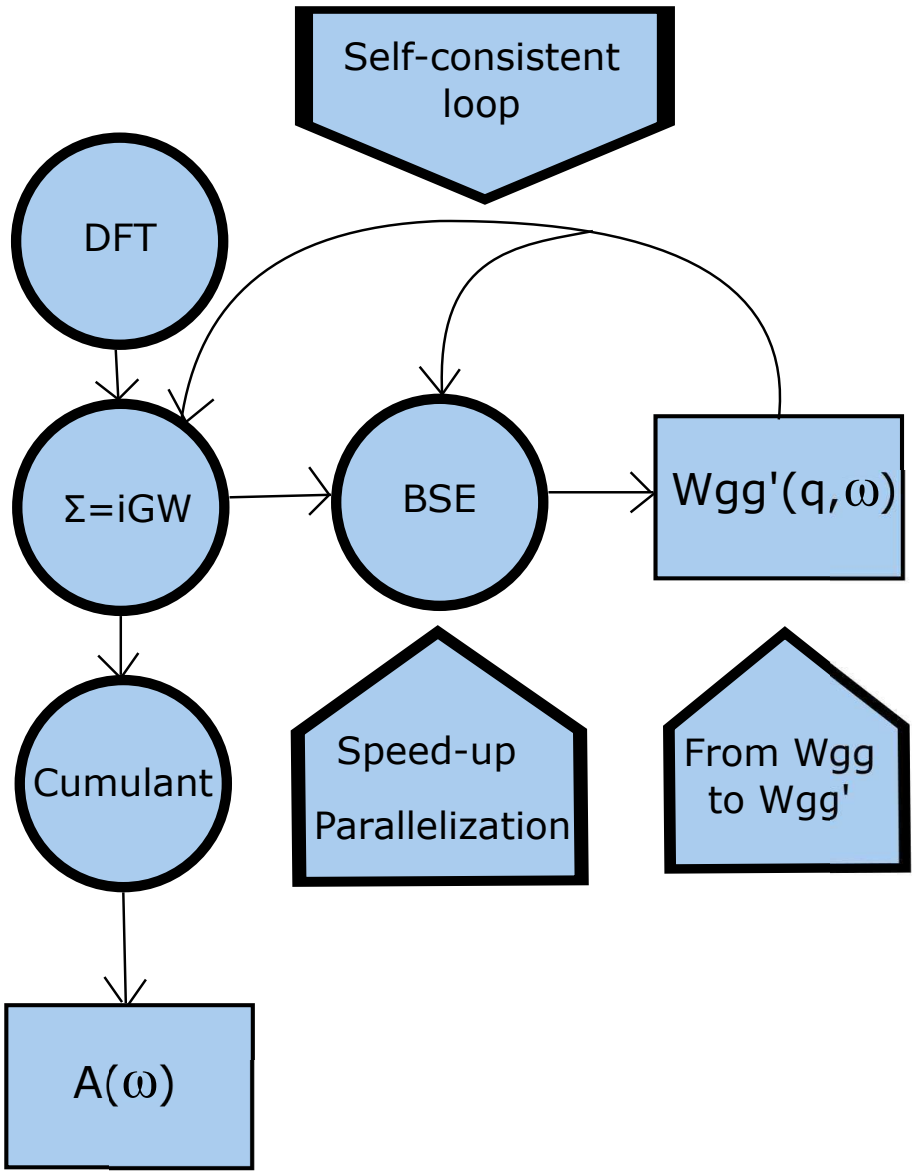


Figure 5.22 – Scheme of calculation of the spectral function. Circles - methods and approximations. Rectangles - quantities of interest. Rectangles with arrows - the contribution of this thesis.

Chapter 6

Conclusions

In the present work we have put forward several problems of existing approaches to *ab-initio* spectroscopy, discussed the possibilities of overcoming them and advanced in the solutions to some of them. The problems stated can be split into two groups: numerical and physical. The first one contains problems related to slow convergence of algorithms and their instability, the second one contains intrinsic problems of approaches, that is effects that cannot be correctly accounted for. In the present work we have, on the one hand extended the parallelization of the existing algorithms, and studied the possibilities of replacing the particularly time-consuming parts of them by simpler approximations. Moreover, we have extended the existing Bethe-Salpeter code to calculate not only the diagonal elements, but also the off-diagonal $\mathbf{G} \neq \mathbf{G}'$ of the susceptibility $\chi_{\mathbf{G},\mathbf{G}'}$, where $\mathbf{G}, \mathbf{G}'$ are reciprocal lattice vectors. On the other hand, we have studied the possibilities of modifying the state of the art approximations and approaches to the *ab-initio* calculation of spectra. More precisely we have looked into the various reformulations of the Bethe-Salpeter equation using different starting points and approximations for the functional derivative, that enters the exact functional differential equation for the two-particle Green's function. Moreover, we have also studied the possibility of combining the Time-dependent Density functional Theory based of the Adiabatic Local Density Approximation and the Bethe-Salpeter equation to achieve deeper insight into analysis of *ab-initio* spectra and possibly more efficient approximations to both of these approaches.

Using the advances described above we were able for the first time to calculate the off-diagonal elements of the Dynamic Structure Factor (DSF), the susceptibility and screened Coulomb interaction, which are all matrices in reciprocal space, for values of the momentum transfer $\mathbf{q}$ in the whole Brillouin zone. These measurable quantities were compared to experimental results, and used to reproduce and predict induced charges, photo-emission

and electron energy loss spectra. Excitonic effects could be detected in the off-diagonal elements and in the resulting localization of induced charges. Finally, in the present work we have demonstrated, to my knowledge, for the first time in the *ab-initio* framework the existence of excitonic satellites in wide gap semi-conductors such as Lithium Fluoride.

Appendices

Appendix A

A.1 Convergence: Silicon ground state

Before doing any TDDFT or BSE calculation one needs a starting point. This starting point is the Kohn-Sham band structure which we get from the Abinit code. Therefore we start from converging the LDA+GW result, so as to have an idea of the parameters we might use for the Bethe-Salpeter calculations later on and have input files for a convergence study in DP. The parameters considered are:

- ecutsigx(GW), the cutoff used for the calculation of the exchange part of the GW self energy, dE=0.01 reached for 7.0 Ha
- ecuteps(GW, Screening), the cutoff used for the screening, that enters the GW self-energy, dE=0.01 reached for 7.0 Ha
- nbands (LDA, GW, Screening), number of bands used in the LDA, GW and screening calculation, dE=0.01 reached for 192+4=196
- ecut(LDA, GW, Screening), cutoff in the number of plane waves used to describe wave-functions, dE=0.01 reached for 16 Ha
- ngkpt(LDA, GW, Screening), number of k-points used, dE=0.01 reached for a 8 8 8 grid.

The other relevant parameters are: nbands=11 (For the SC-LDA), toldfe = 1.0d-10, toldwf = 1.0d-16. The grid that is used is a Monkhorst-Pack [139] one which corresponds to 4 shifts.

All these parameters as will be seen in the following can be in fact lowered, as a small precision is required for a broadened EELS spectra.

A.2 Convergence: BSE on Silicon

For the EXC part a Python code, which I have written, was used to make convergence studies easier. The resulting parameters we get in the case $q = 0$ are summarized below.

The thing worth noting is that the main computational overhead comes from the diagonalization of the excitonic hamiltonian. For a fixed number of occupied bands, its size is proportional to the number of k-points times the number of unoccupied bands. Furthermore its size is doubled if one includes the coupling. First, we perform convergence studies neglecting setting the coupling to zero. As we see from Fig. A.1(d) and Fig. A.1(e) the number of k-points and the number of unoccupied bands have to be chosen quite large to converge completely for a non-broadened spectra. However if one applies broadening a 4 4 4 grid becomes already converged which allows one to increase the number of bands by a factor $864/256 \approx 3$

- ecuteps(Screening) reached for 6.0 Ha
- nbands (LDA,Screening) reached for 96+4=100
- ecut(LDA, Screening) reached for 16 Ha Fig. A.1(c)
- ngkpt(LDA, Screening) reached for 6 6 6 or 4 4 4 if a 1.1 eV broadening is applied Fig. A.1(d)
- wfnsh (DP) 14 Fig. A.1(a)
- matsh (DP) 6 Fig. A.1(b)
- nbands(DP) more than 20. Fig. A.1(e)
- broad=$d\omega$ (DP) 0.1

In the case $q \neq 0$ More exactly $1/2, 1/2, 1/2$ or 0.53 a.u. in the [1,1,1] direction. We find the following values are required for convergence:

- ecuteps(Screening) reached for 5.0 Ha Fig. A.2(e)
- nbands (LDA,Screening) reached for 64+4=68
- ecut(LDA, Screening) reached for 16 Ha Fig. A.2(d)
- ngkpt(LDA, Screening) reached for 4 4 4 in the screened case Fig. A.2(f)
- wfnsh (EXC) 11 Fig. A.2(a)
- matsh (EXC) 6 Fig. A.2(b)
- nbands(EXC) 20 Fig. A.2(c)
- broad=2 $d\omega$ (DP) 0.1

One sees that the values required for convergence on the Abinit part are decreased when we go away from $q = 0$. One can therefore safely use the screening and KSS files generated for $q = 0$ and perform only a convergence study with respect to EXC parameters. This can be done not for a full set of points, but for a 2 2 2 or 4 4 4 grid. Therefore we generate one test set of SCR and KSS files, and one set for a fully-converged calculation. (Table. A.1)

The resulting convergence studies are summarized in Fig. A.3,Fig. A.4, Fig. A.5. If one were interested in the BSE with coupling one sees that in fact the convergence study might not need to be performed as the converged parameters are roughly the same, see for

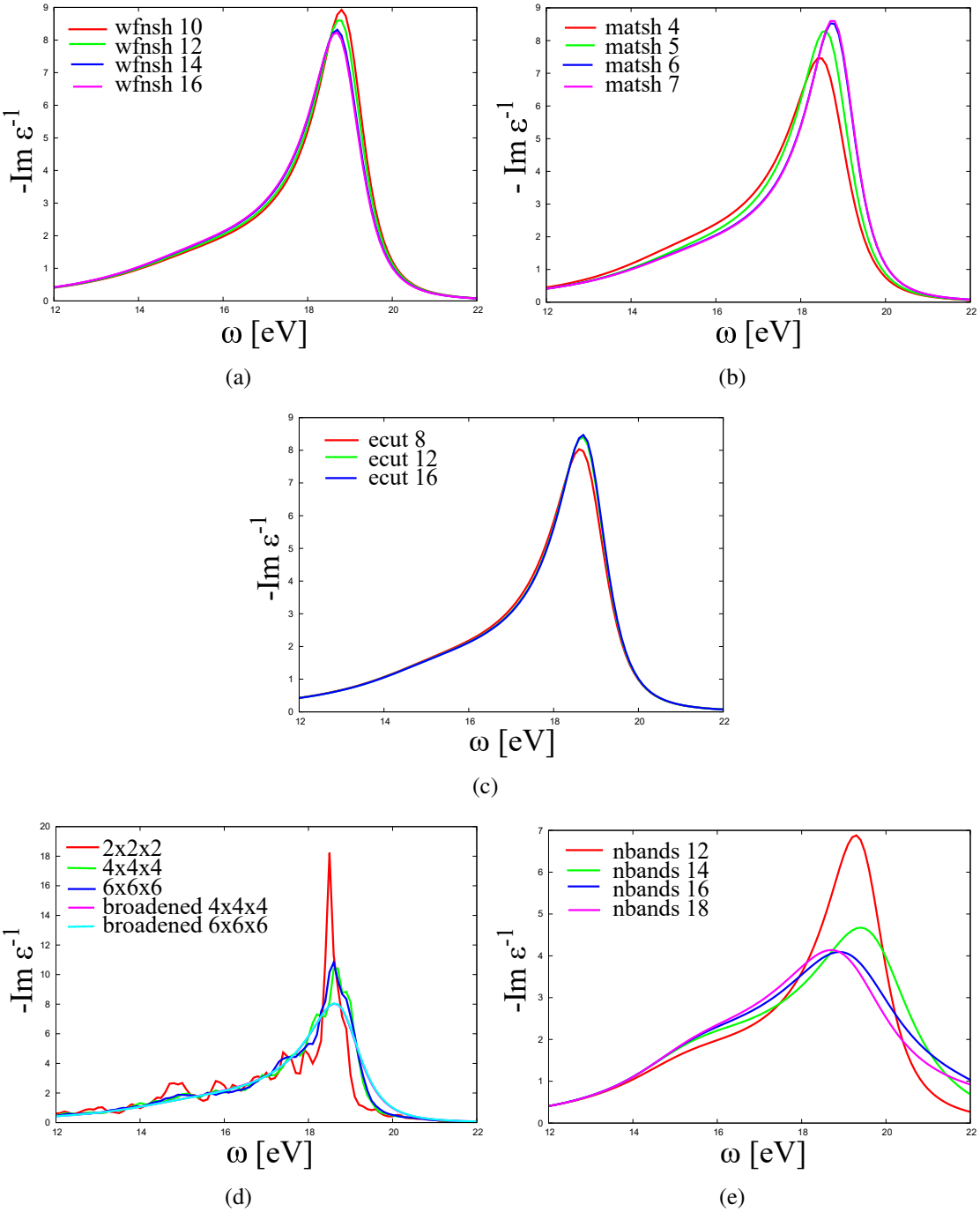


Figure A.1 – Loss spectrum of Silicon: BSE without coupling convergence study for $q = 0$

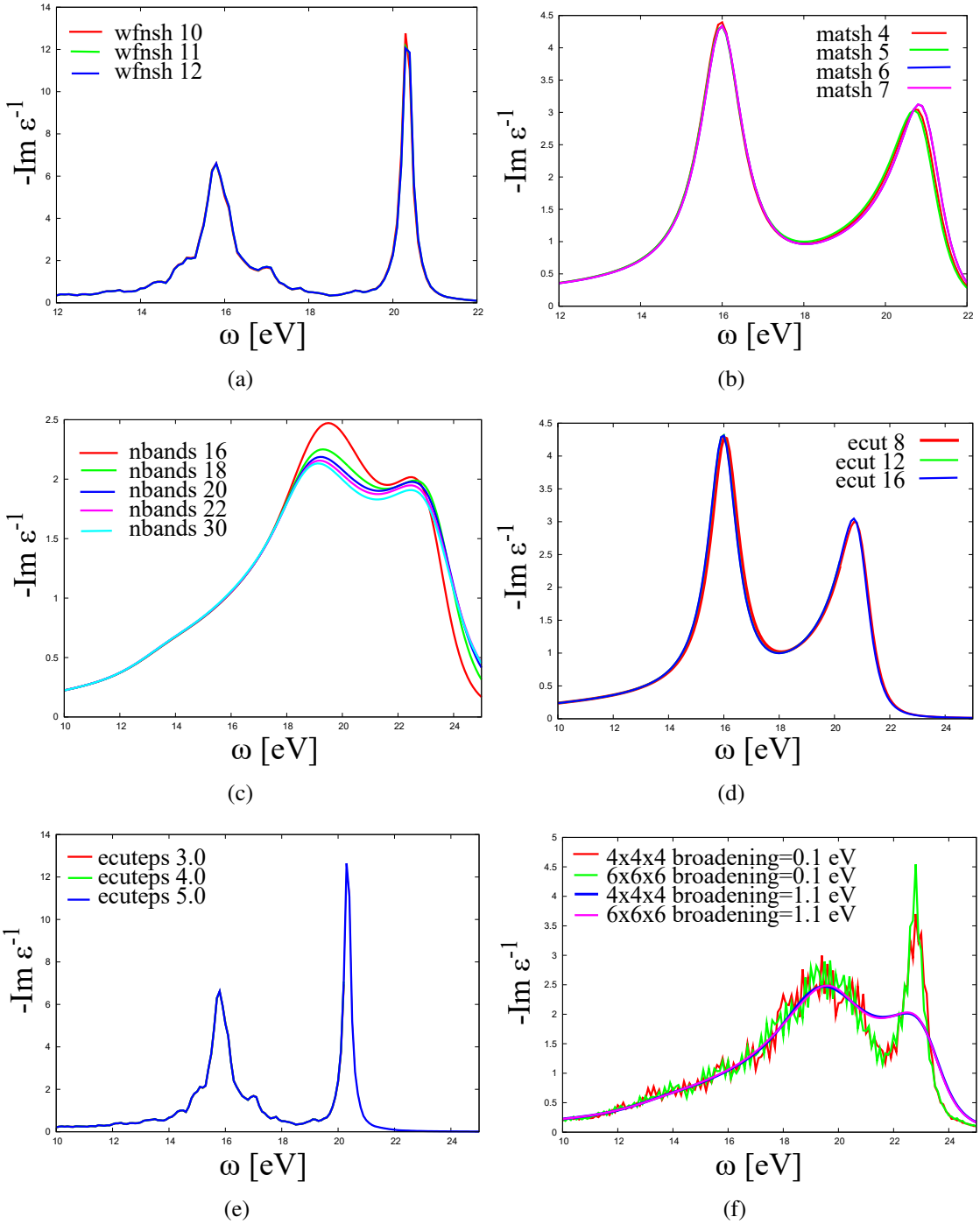


Figure A.2 – Loss spectrum of Silicon: BSE without coupling convergence study for $q = (1/2, 1/2, 1/2)$ or 0.53 a.u. in the [1,1,1] direction.

Parameter	Test run	Final run
ecut	12 (26 shells)	16 (39 shells)
grid	2 2 2	4 4 4
nbands	68	100
ecuteps	4.0 (9 shells)	5.0 (10 shells)

Table A.1 – Abinit parameters

example Fig. A.6 and Fig. A.3. One could even suppose that RPA gives already a rough estimate of the number of valance bands required for the convergence.

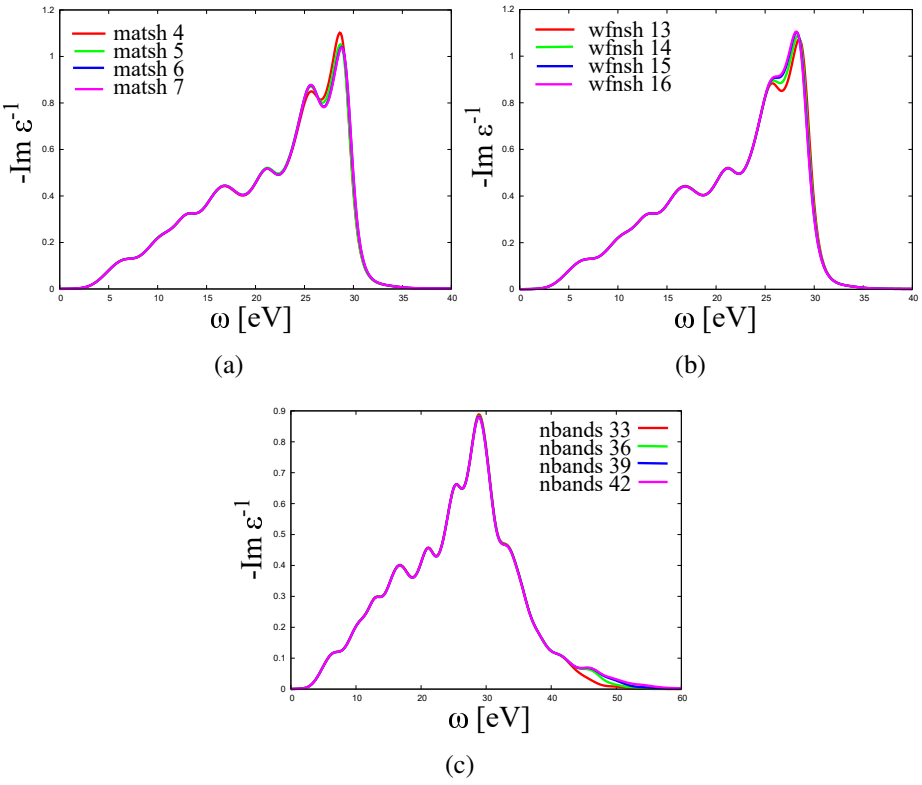


Figure A.3 – Loss spectrum of Silicon: BSE without coupling convergence study for $q = (1,1,1)$.

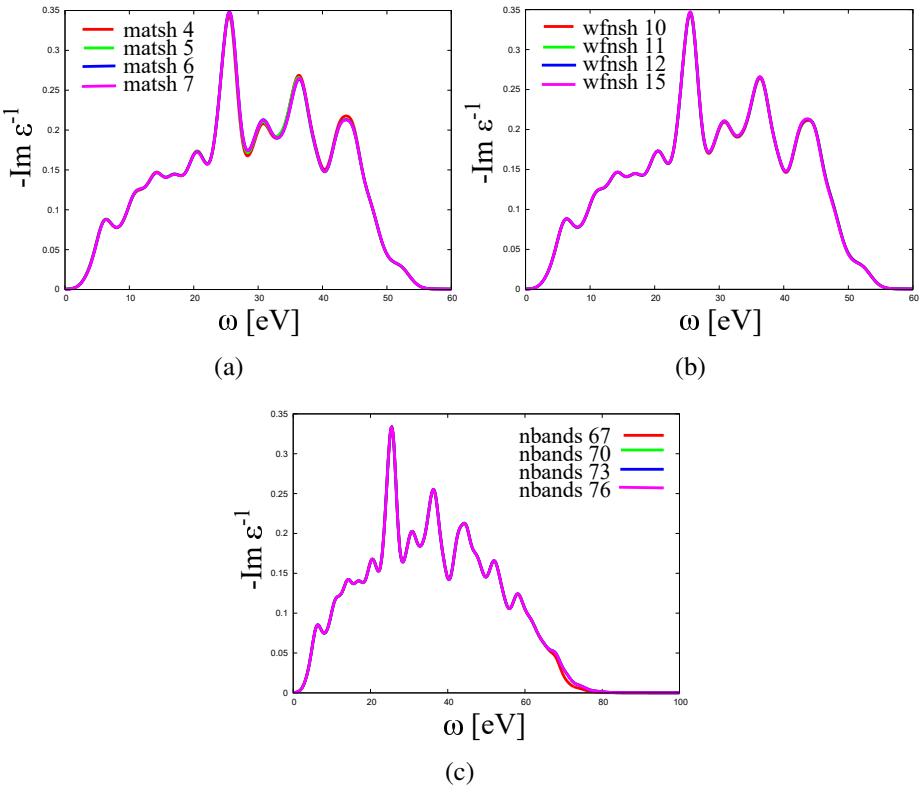


Figure A.4 – Loss spectrum of Silicon: BSE without coupling convergence study for $q = (1.5, 1.5, 1.5)$.

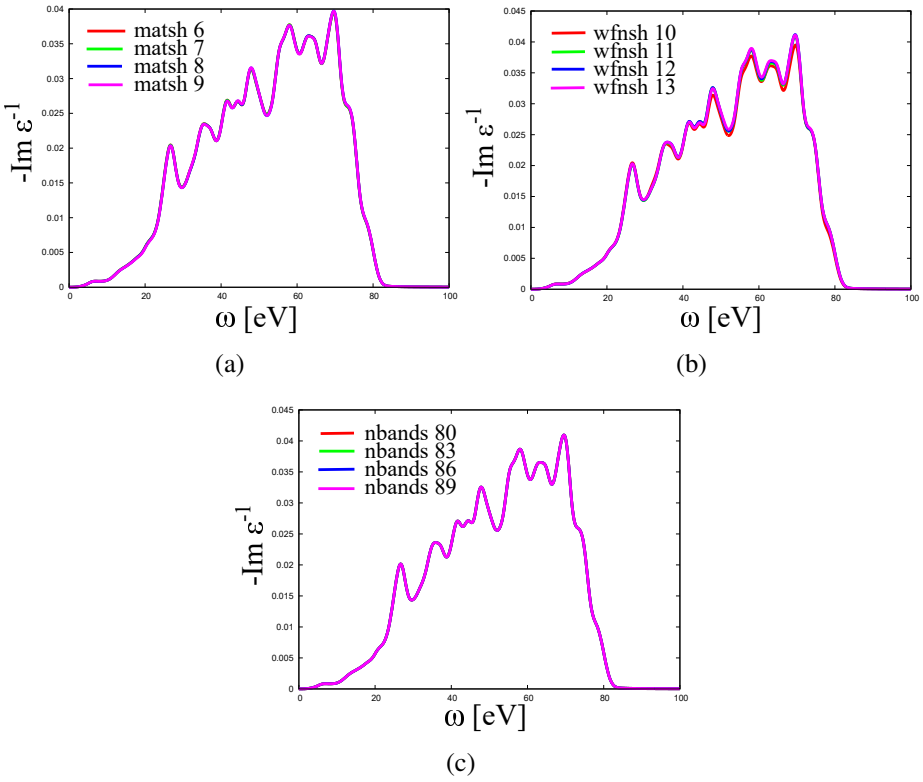


Figure A.5 – Loss spectrum of Silicon: BSE without coupling convergence study for $q = (2.5, 2.5, 2.5)$.

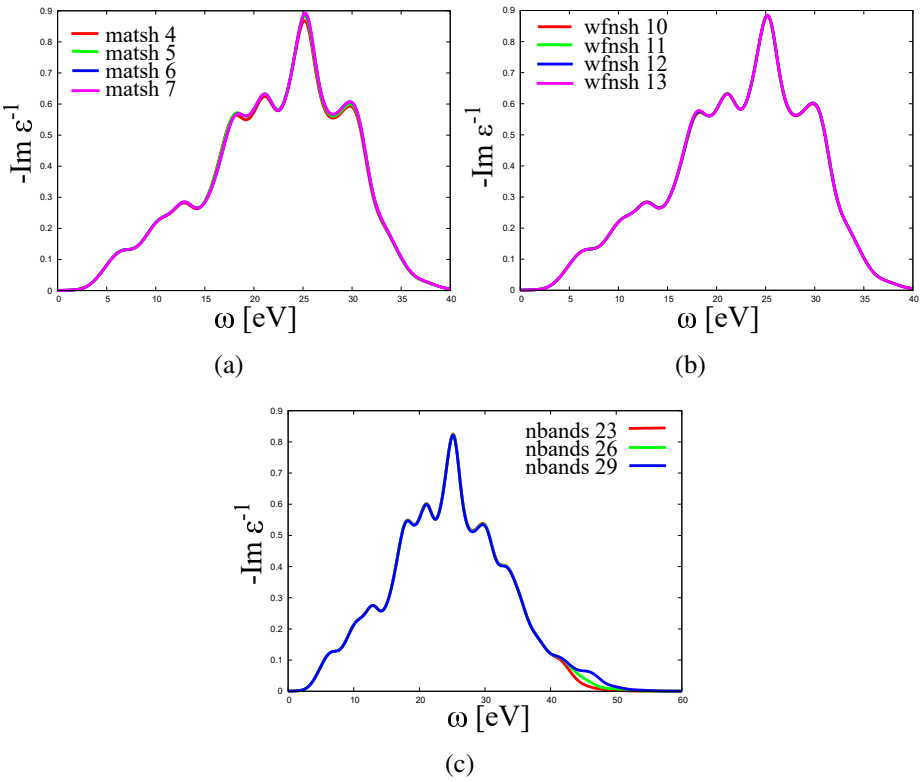


Figure A.6 – Loss spectrum of Silicon: BSE with coupling convergence study for $q = (1,1,1)$

Appendix B

B.1 Extending EXC with coupling to MPI

As shown on Fig. B.1 and Fig. B.2 on cannot neglect the effect of coupling when computing the loss function of Silicon. Its importance is increased as we move to higher values of transferred momenta **q**. Due to the large matrix sizes involved in the solution of the Bethe-Salpeter equation with coupling (the hamiltonian matrix is 2 times bigger and non-hermitian) one is brought to the fact that one needs to perform some of the steps in parallel. For instance two crucial steps: the solution of the eigenvalue problem for the hamiltonian and the inversion of the overlap matrix. The LAPACK sequential approach to these problems is simply too time-consuming.

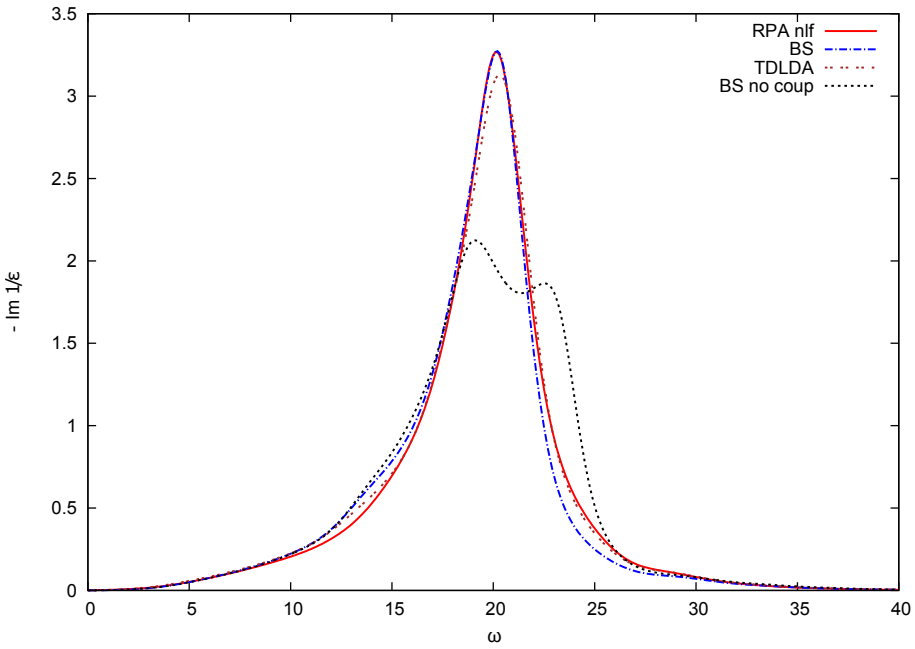


Figure B.1 – Coupling effect q= 1/2 [1,1,1]

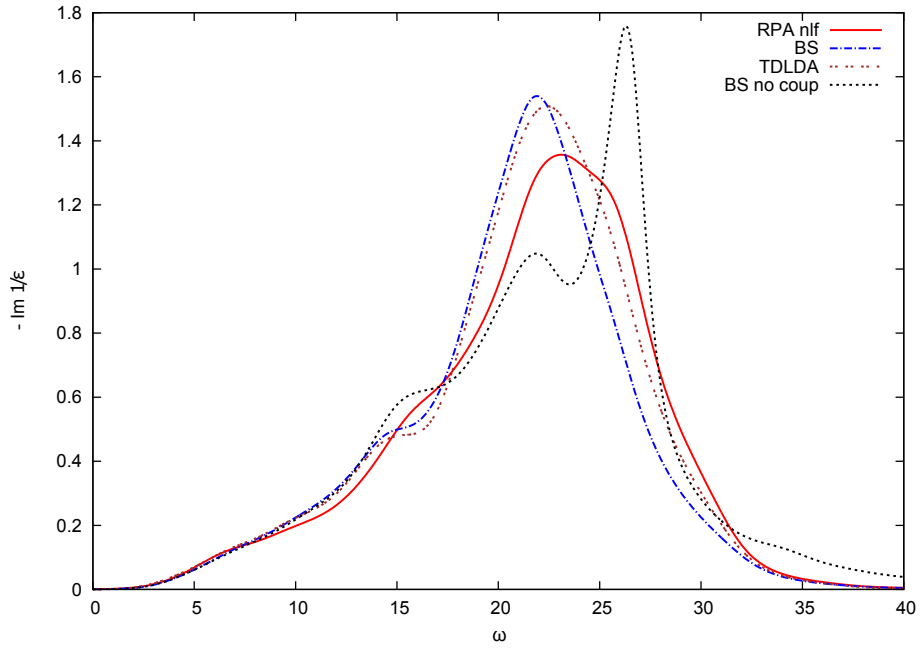


Figure B.2 – Coupling effect q=3/4 [1,1,1]

Let us first consider the inversion of the overlap matrix S. As pointed out in [80] often the off-diagonal coupling blocks are smaller than the resonant diagonal blocks. In other words the Hamiltonian is not that non-hermitian. Therefore the overlap matrix S is rather close to 1. In fact its diagonal is 1, and the norm of the remainder $||S-1||/||1|| = \sqrt{\sum_{i,j}|S-1|^2_{i,j}}/\sqrt{N}$ for example for Si is of the order of 3×10^{-3}. Therefore the corrections that come from the off-diagonal part S^{-1} are also going to be rather small. This can be clearly seen on Fig. B.3, where several direct and iterative methods are compared to the simple $S^{-1}=1$ and $S^{-1}=2-S$. Keeping this in mind one clearly sees that an iterative inversion method starting from an initial guess such as $S^{-1}=1$ is more favorable over direct solvers.

The method of choice for us here is the simple Conjugate-gradient solution of equations $SS_i^{-1}=e^i$, where $e^i=(\delta_{i,j})$ is the i-th basis vectors in the transition space. Setting the tolerance to a reasonable value 10^{-4} one obtains convergence quite rapidly (Fig. B.4). In fact the maximal number of steps in this case is smaller then 10, that goes well with the smallness of the deviation from 1.

The next point to cover is the solution of the eigenvalue and eigenvector problem for H. Here again we could think of starting the iterations from a point close to the exact result. For example one could solve the problem without coupling and use its eigenvectors and eigenvalues as a restart subspace. In general this poses a question that would be interesting to answer not only from the point of view of our current problem, but it would also be

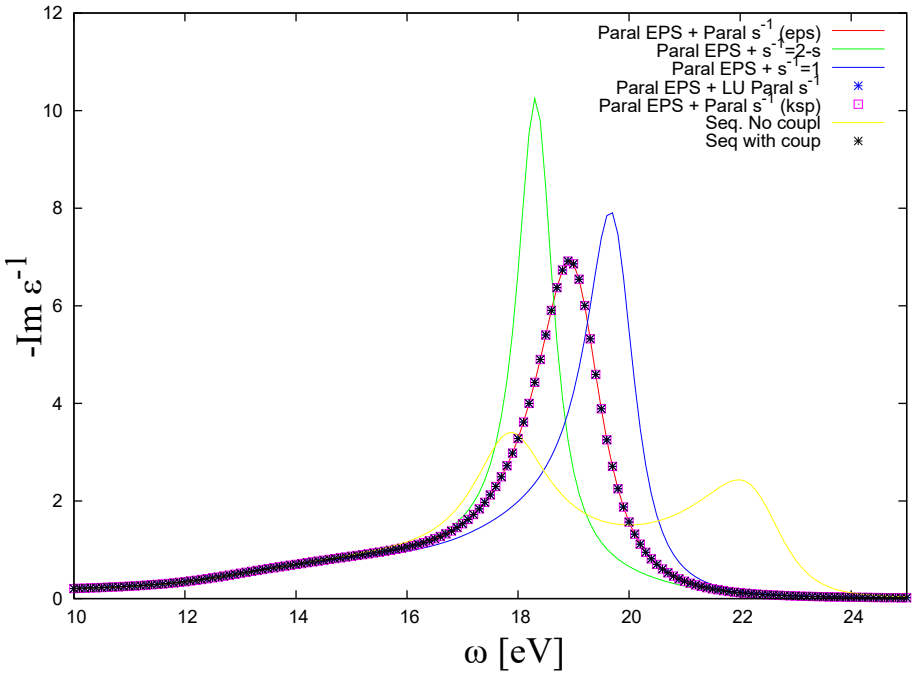


Figure B.3 – Comparison of results obtained using different approximations to the overlap matrix and verification that the results obtained after parallelization are the same as in the sequential case.

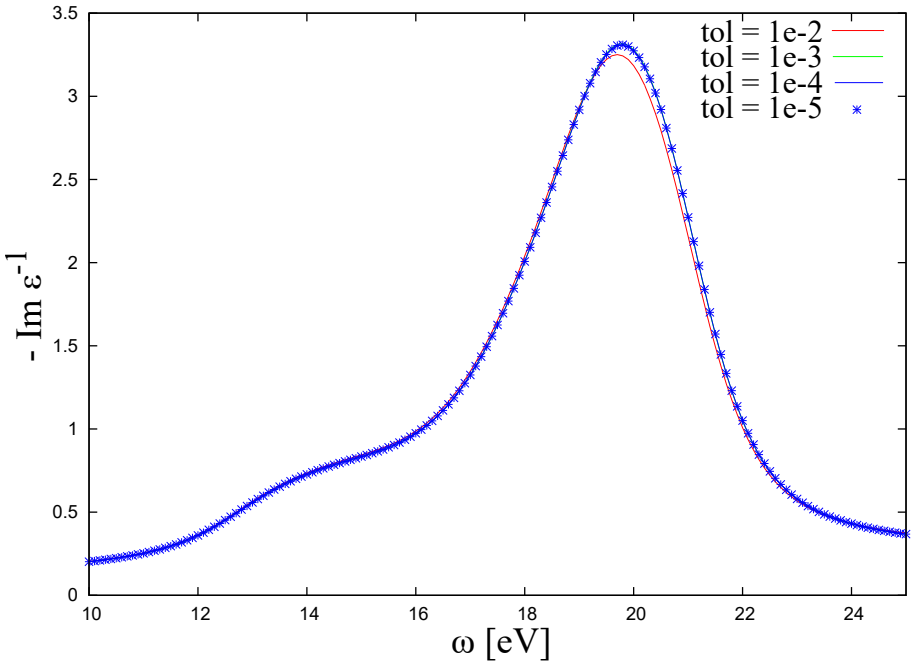


Figure B.4 – Convergence with respect to maximal error on elements of the inverse matrix.

interesting to compare the eigenvectors of the TDLDA with eigenvectors of the Bethe-Salpeter equation both of them with and without coupling. At first sight the eigenvectors

of TDLDA might not have any meaning as they appear only in the Cassida formulation, when one adds extra delta-functions to the Kernel (see discussion in chapter 3). However this has not been verified in more detail and this could give another low-cost starting point for the general eigenvector problem. For now we have implemented the Krylov-Schur approach from the PETSc + SLEPc libraries. This already gives us a parallel MPI implementation that we can use to get an increase in calculation speed. The comparison of parallel and sequential run-times is give in table B.1. For the sequential run the time of the eigenvalue problem solution and the inversion of the overlap matrix were not measured separately.

	seq	2proc	4proc	8proc	32proc
Create H	320	206	147	129	804
Read and Send H	-	20	20	20	20
Solve H	6100	4800	3715	2750	1740
Create S	-//-	360	200	115	35
Compute S^{-1}	-//-	1500	850	450	100
End	4	4	4	4	4

Table B.1 – Calculation speed for lomo = 1, nbands=25, 32 k-points, q=1/2 (1,1,1)

If now one doubles the size of the k-point grid the run-time should increase approximately 8-fold for both of the time-consuming calculations have a complexity of n^3. However if one were to increase the number of processors one would also get a linear decrease in the required time at least until saturation, when the communication costs start to dominate the computational ones. As our goal is to compute for a 4 4 4 MP grid, which contains 256 k-points, which is 8 times 32, it will require up to 60 hours on a 256 - core machine, or 15 on a 1024 core one. As one sees that the creation of hamiltonian times does not scale well one can pregenerate the hamiltonian using dp compiled for OpenMP, and then use them as input for MPI calculations.

	4 proc	32 proc	2*32 proc
Read and Send H	84	90	227
Solve H	28600	12300	11800
Create S	1400	250	210
Compute S^{-1}	5900	1100	900
Post-processing	20	20	20

Table B.2 – Calculation speed for lomo = 1, nbands=25, 64 k-points, q=1/2 (1,1,1)

The Krylov-Schurr can be viewed as an implicit-restart Arnoldi method [140]. Other

alternatives exist in principle to find the eigenvalues and eigenvectors of the hamiltonian in question. First of all there exists parallel implementations for dense matrices in ScaLAPACK, these in principle should be compared against. Then there are other variants of the iterative algorithms for the eigenproblem solution. They include different variants of generalized Lanczos, such as Arnoldi and bi-orthogonalization Lanczos (e.g. Iurii Timrov's thesis [141]), and block iteration methods, such as the Davidson-type methods. It can be shown that bi-orthogonalization Lanczos is actually equivalent to a simple Lanczos, but with an modified scalar product [142]. In our case it turns out to be rather simple... one easily sees that the eigenvectors considered are actually orthogonal with respect to a diagonal metric: $(1_n; -1_n)$. Unfortunately in the last case the convergence seems to be unstable and as for the other Arnoldi-type algorithms they have been shown to perform worse than the SLEPc implementation [140]. One other alternative modification of the scalar product is to follow [93] and then apply Lanczos, however this seems to have the same draw-backs as the previous diagonal metric modification, as one has a high chance of encountering convergence difficulties and actually one is not guaranteed to have a positive-definite metric here either.

In conclusion, our choice is to use the Krylov-Schur approach from the PETSc + SLEPc libraries to perform parallel diagonalization of the full Bethe-Salpeter hamiltonian. We construct the hamiltonian using the shared-memory OpenMP version the EXC code, and then solve it using our newly developed MPI-parallelization.

Appendix C

C.1 Green's function definitions

Let us derive a different equation for the connected part of the two-particles Green's function.

Single particle G

$$\begin{aligned} G(1,2) &= (-i)\langle N|\, T\Psi(x_1,t_1)\Psi^\dagger(x_2,t_2)\,|N\rangle = \\ &= (-i)\left\{\Theta(t_1-t_2)\langle N|\,\Psi(x_1,t_1)\Psi^\dagger(x_2,t_2)\,|N\rangle - \Theta(t_2-t_1)\langle N|\,\Psi^\dagger(x_2,t_2)\Psi(x_1,t_1)\,|N\rangle\right\} = \\ &= (-i)\Theta(t_1-t_2)G_e(1,2) - (-i)\Theta(t_2-t_1)G_h(1,2) \end{aligned}$$

Particle density n

$$n(1) = |N\rangle\,\Psi^\dagger(1)\Psi(1)\,\langle N| = -iG(1,1^+)$$

Two particle G

$$G_2(1,2;3,4) = (-i)^2\langle N|\, T\Psi(x_1,t_1)\Psi(x_2,t_2)\Psi^\dagger(x_4,t_4)\Psi^\dagger(x_3,t_3)\,|N\rangle$$

2 particle correlation function L

$$L(1,2;3,4) = G(2,4)G(1,3) - G_2(1,2;3,4)$$

or alternatively

$$\frac{\delta G(1,2)}{\delta U(3,4)} = L(1,4;2,3)$$

Susceptibility χ

$$\chi(1,2) = (-i)L(1,2;1^+,2^+)$$

Dyson equation

$$G^{-1}(1,2) = G_0^{-1}(1,2) + iv(1,2')G(2',2'^+)\delta(1,2) - U(1,2) - iv(1,2')L(1,2';3',2'^+)G^{-1}(3',2)$$
$$v_H(1,2) = -iv(1,2')G(2',2'^+)$$
$$\Sigma_{XC}(1,2) = iv(1,2')L(1,2',3',2'^+)G^{-1}(3',2)$$

Derivatives & BSE

Apply splitting $\Sigma_1 = U + v_H$ and $\Sigma_2 = \Sigma_{XC}$

$$\frac{\delta G_H^{-1}(1,2)}{\delta U(3,4)} = -\delta(1-3)\delta(2-4) + iv(1,2')L(2',4,2',3)\delta(1-2)$$
$$\frac{\delta G_H(1,2)}{\delta U(3,4)} = G_H(1,3)G_H(4,2) - iG_H(1,1')G_H(1',2)v(1',2')L(2',4,2',3) = L_H^S(1,4;2,3)$$

C.2 SC equations in frequency space

We will adopt the convention that:

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{i\omega t}dt$$
$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\omega)e^{-i\omega t}d\omega$$

We then have for example:

$$G_H(1,2) = G_H(x_1,x_2|t_1-t_2) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{G}_H(x_1,x_2|\omega)e^{-i\omega(t_1-t_2)}d\omega$$
$$\varepsilon^{-1}(1,2) = \varepsilon^{-1}(x_1,x_2|t_1-t_2) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{\varepsilon}^{-1}(x_1,x_2|\omega)e^{-i\omega(t_1-t_2)}d\omega$$

The hats being redundant we will continue writing without them for the sake of simplicity.

SFC equations approximation 1

$$
\begin{aligned}
W(x_1,x_2|\omega) &= v(x_1,x_2) + M(x_1,x',\omega)v(x',x_2) \\
S_0^W(x_1,x_2|\omega) &= \frac{-i}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega_1}{2\pi}G_H(x_1,x_1'|\omega_1+\omega)W(x_1',x_1|\omega_1)G_H(x_1',x_2|\omega) \\
S^W(x_1,x_2|\omega) &= S_0^W(x_1,x_2|\omega) \\
M(x_1,x_2|\omega) &= \frac{-i}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega_1}{2\pi}G_H(x_1,x_1'|\omega_1+\omega)W(x_1',x_2|\omega)G_H(x_1',x_1|\omega_1)
\end{aligned}
$$

SFC equations approximation 2

The first two equations do not change. As for the remainder one gets:

$$
\begin{aligned}
S^W(x_1,x_2|\omega) &= S_0^W(x_1,x_2|\omega) + S_0^W(x_1,x_1'|\omega)S^W(x_1',x_2|\omega) \\
M(x_1,x_2|\omega) &= \frac{-i}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega_1}{2\pi}G_H(x_1,x_1'|\omega_1+\omega)W(x_1',x_2|\omega)G_H(x_1',x_1|\omega_1) + \\
&\quad + \frac{i}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega_1}{2\pi}G_H(x_1,x_1'|\omega_1+\omega)W(x_1',x_2|\omega)G_H(x_1',x_2'|\omega_1)S^W(x_2',x_2|\omega_1)
\end{aligned}
$$

And finally

SFC equations approximation 3

$$
\begin{aligned}
S^W(x_1,x_2|\omega) &= S_0^W(x_1,x_2|\omega) + S_0^W(x_1,x_1'|\omega)S^W(x_1',x_2|\omega) + \\
&\quad + \frac{i}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega_1}{2\pi}G_H(x_1,x_1'|\omega_1+\omega)S^W(x_1',x_2'|\omega_1+\omega)W(x_2',x_1|\omega_1)G_H(x_2',x_2|\omega) \\
M(x_1,x_2|\omega) &= \frac{-i}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega_1}{2\pi}G_H(x_1,x_1'|\omega_1+\omega)W(x_1',x_2|\omega)G_H(x_1',x_1|\omega_1) + \\
&\quad + \frac{i}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega_1}{2\pi}G_H(x_1,x_1'|\omega_1+\omega)W(x_1',x_2|\omega)G_H(x_1',x_2'|\omega_1)S^W(x_2',x_2|\omega_1) + \\
&\quad + \frac{i}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega_1}{2\pi}G_H(x_1,x_1'|\omega_1+\omega)S^W(x_1',x_2'|\omega_1+\omega)W(x_2',x_2|\omega)G_H(x_2',x_1|\omega_1)
\end{aligned}
$$

C.3 BSE using L^{-1} in frequency space

We start by writing everything in frequency space.

$$L_H^S(1,3,2,3^+) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty} d\omega_1 d\omega_2 L_H^S(x_1,x_2,x_3|\omega_1,\omega_2)e^{-i\omega_2(t_1-t_3)}e^{-i\omega_1(t_1-t_2)} =$$
$$= \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty} d\omega_1 d\omega_2 G_H(x_1,x'|\omega_1+\omega_2)\varepsilon^{-1}(x',x_3|\omega_2)G_H(x',x_2|\omega_1)e^{-i\omega_2(t_1-t_3)}e^{-i\omega_1(t_1-t_2)}$$

The two possible couplings of indices $1 \to 3$ and $1 \to 2$, that enter the actual equation, result in the disappearance of the corresponding exponential term in the equation.

In equilibrium, L has the same frequency structure i.e. only two frequencies.

$$L(1,3,2,3^+) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty} d\omega_1 d\omega_2 L(x_1,x_2,x_3|\omega_1,\omega_2)e^{-i\omega_2(t_1-t_3)}e^{-i\omega_1(t_1-t_2)}$$

After some mathematical manipulations we obtain the first equation derived in the previous sub-section in frequency space:

$$\begin{aligned}
&L(x_1,x_2,x_3|\omega_1,\omega_2) = L_H^S(x_1,x_2,x_3|\omega_1,\omega_2)+\\
&+\int_{-\infty}^{\infty} L_H^S(x_1,x_1',x_3|\omega_1,\omega_2)iv(x_1',x_2')L(x_1',x_2,x_2'|\omega_1,\omega_3)\frac{d\omega_3}{2\pi}+\\
&-2\delta(\omega_2)\int_{-\infty}^{\infty} G_H(x_1,x_1'|\omega_1)iv(x_1',x_2')L(x_1',x_3',x_2'|\omega_1+\omega_4,\omega_3)\\
&G_H^{-1}(x_3',x_4'|\omega_1+\omega_4)L(x_4',x_2,x_2'|\omega_1,\omega_4)\frac{d\omega_3}{2\pi}\frac{d\omega_4}{2\pi}
\end{aligned}$$

The integration in the non-trivial term cannot be rewritten in a Dyson-like manner i.e. $A(x,x')B(x',x'')C(x'',x''')$. Therefore we cannot find a simple solution.
However, we can compare this equation to the ones obtained applying approximations directly to the derivative $\frac{\delta L}{\delta U}$ and the standard BSE.

In a similar manner we obtain the linearized version of a new Bethe-Salpeter equation in frequency space.

$$\begin{aligned}
&L(x_1,x_2,x_3|\omega_1,\omega_2) = L_H^S(x_1,x_2,x_3|\omega_1,\omega_2)+\\
&+\int_{-\infty}^{\infty} L_H^S(x_1,x_1',x_3|\omega_1,\omega_2)iv(x_1',x_2')L(x_1',x_2,x_2'|\omega_1,\omega_3)\frac{d\omega_3}{2\pi} - \int_{-\infty}^{\infty}\frac{d\omega_3}{2\pi}\frac{d\omega_4}{2\pi}\delta(\omega_2)\times\\
&\times[\, G_H(x_1,x_1'|\omega_1)iv(x_1',x_2')G_H(x_1',x_3'|\omega_1+\omega_3+\omega_4)\varepsilon^{-1}(x_3',x_2'|\omega_3)L(x_3',x_2,x_2'|\omega_1,\omega_4)-\\
&-G_H(x_1,x_1'|\omega_1)iv(x_1',x_2')L(x_1',x_3',x_2'|\omega_1+\omega_4,\omega_3)\varepsilon^{-1}(x_3',x_2'|\omega_4)G_H(x_3',x_2|\omega_1)\,]
\end{aligned}$$

The first non-trivial term can be integrated in ω_3. This will give the density matrix $\rho(x_1',x_3')$. This equation can then be written as $L = \hat{O}L + L_H^S$. To solve it we can first find the self-energy term via integration over ω_2 and then substitute it to get the response function χ. It is essentially solved in the same way as the trivial part.

As for the second term it contains L integrated over two frequencies which not so clear to solve for, because it doesn't have a Dyson-like character.

Nevertheless, let us solve our equation in the simplest form, neglecting the terms multiplied by the delta-function:

$$L(x_1,x_2,x_3|\omega_1,\omega_2) = L_H^S(x_1,x_2,x_3|\omega_1,\omega_2)+ \\ + \int_{-\infty}^{\infty} L_H^S(x_1,x_1',x_3|\omega_1,\omega_2) iv(x_1',x_2') L(x_1',x_2,x_2'|\omega_1,\omega_3)\frac{d\omega_3}{2\pi}$$

Having in mind $-iv\chi = \Sigma G$ we multiply our equation by $iv(x_1,x_3)$ and integrate in over x_3 and ω_2. This gives us, upon introducing

$$S^W(x_1,x_2|\omega_1) = \int_{-\infty}^{\infty} iv(x_1,x_3) L(x_1,x_2,x_3|\omega_1,\omega_2)\frac{d\omega_2}{2\pi}dx_3$$
$$S_0^W(x_1,x_2|\omega_1) = \int_{-\infty}^{\infty} iv(x_1,x_3) L_H^S(x_1,x_2,x_3|\omega_1,\omega_2)\frac{d\omega_2}{2\pi}dx_3$$

the following equations:

$$S^W(x_1,x_2|\omega_1) = S_0^W(x_1,x_2|\omega_1) + S_0^W(x_1,x_1'|\omega_1)S^W(x_1',x_2|\omega_1)$$
$$L(x_1,x_2,x_3|\omega_1,\omega_2) = L_H^S(x_1,x_2,x_3|\omega_1,\omega_2) + L_H^S(x_1,x_1',x_3|\omega_1,\omega_2)S^W(x_1',x_2|\omega_1)$$

This is just the resumation of insertions of self-energy insertions $G_H W$. To get from the second equation χ we should multiply it by $-i$ and integrate over ω_1. We will then get:

$$\chi(x_1,x_3|\omega) = \int_{-\infty}^{\infty}\frac{d\omega_1}{2\pi} L_H^S(x_1,x_1,x_3|\omega_1,\omega) + \int_{-\infty}^{\infty}\frac{d\omega_1}{2\pi} L_H^S(x_1,x_1',x_3|\omega_1,\omega) S^W(x_1',x_2|\omega_1)$$

Using the Dyson equation $G = G_H + G_H S^W$ we get $S^W = G_H^{-1}G - 1$. And this gives us:

$$\chi(x_1,x_3|\omega) = \int_{-\infty}^{\infty} G_H(x_1,x'|\omega_1+\omega)\varepsilon^{-1}(x',x_3|\omega)G(x',x_1|\omega_1)\frac{d\omega_1}{2\pi}$$

Integration over ω_1 can be performed explicitly, giving poles at $\omega = \varepsilon_k^e - \varepsilon_k^h$ with one of the energies being the Hartree energy, and the other one the self-energy corrected energy. Remembering that $\varepsilon^{-1} = 1 + v\chi$ we can see from an iterative point of view that these will be the only poles we will have.

One can actually solve this equation, obtaining:

$$\chi = \frac{GG_H}{1 - vGG_H}$$

This is basically an RPA-type approximation for $\chi_0 = GG_H$.

Appendix D

D.1 LiF. Convergence

Abinit

Before doing any studies on the satellites we need to know what are the parameters required to obtain converged results. Among these are the cut-offs **ecutwfn**, **ecuteps** and **ecutsigx** which account for the number of plane-waves that are used to describe the wave-functions, the susceptibility matrix χ and the self-energy corrections. Some of them can be different for the different parts of the calculation i.e. the number of plane waves in the ground-state and the susceptibility calculations.

In the following we have used a 40 Hartree cut-off for the plane-wave basis set in the ground-state calculation. We then performed the convergence with respect to the value of the G_0W_0 correction to the gap varying the size of the plane-wave basis in the susceptibility and self-energy calculations.

ecutwfn(Ha)	G_0W_0(eV) gap correction
25	4.302
30	4.312
35	4.315
40	4.316

The next step was to converge the results with respect to the size of the susceptibility matrix which is governed by ecuteps.

ecuteps(Ha)	G_0W_0(eV) gap correction
12	4.474
14	4.480
16	4.481
18	4.480

And finally the size of the basis set used to calculate the exchange part of the self-energy (ecutsigx).

ecutsigx(Ha)	G_0W_0(eV) gap correction
40	4.478
44	4.480
48	4.482
52	4.482

In both the plasmon-pole approximation calculation of the self-energy and the contour-deformation approach, one must also converge the results with respect to the number of empty states. Using the density obtained from the ground-state calculation wave-functions for 400 excited states were produced. A convergence study was then performed to get the number of bands required to obtain a G_0W_0 gap.

nbands	G_0W_0(eV)
225	4.475
250	4.480
275	4.484
300	4.487

If we need the band-gap with 0.1 eV precision one sees that the values ecutwfn=30.0 Ha, ecuteps=12.0 Ha, ecutseigx=44Ha together with 275 bands (valence and conduction) are sufficient.

It is worth noting that for spectral function calculations one can reduce the number of bands in the calculation of the susceptibility, while keeping the number of bands in the self-energy fixed (see Fig. D.1).

If we limit ourselves to a plasmon-pole calculation we can stop at this stage and check whether the results depend strongly on the number of k-points chosen to describe the Brillouin zone. However, if we want to use the contour-deformation approach we must also converge with respect of the number of real and imaginary frequencies used in the susceptibility and the maximal real frequency used.

nfreqim	G_0W_0(eV) gap correction
6	4.486
7	4.495
8	4.501
9	4.506

In fact it turns out that even a small number of real frequencies (nfreqre=20) and a conservative freqremax=1 Ha is enough to obtain a converged result for the G_0W_0 correction.

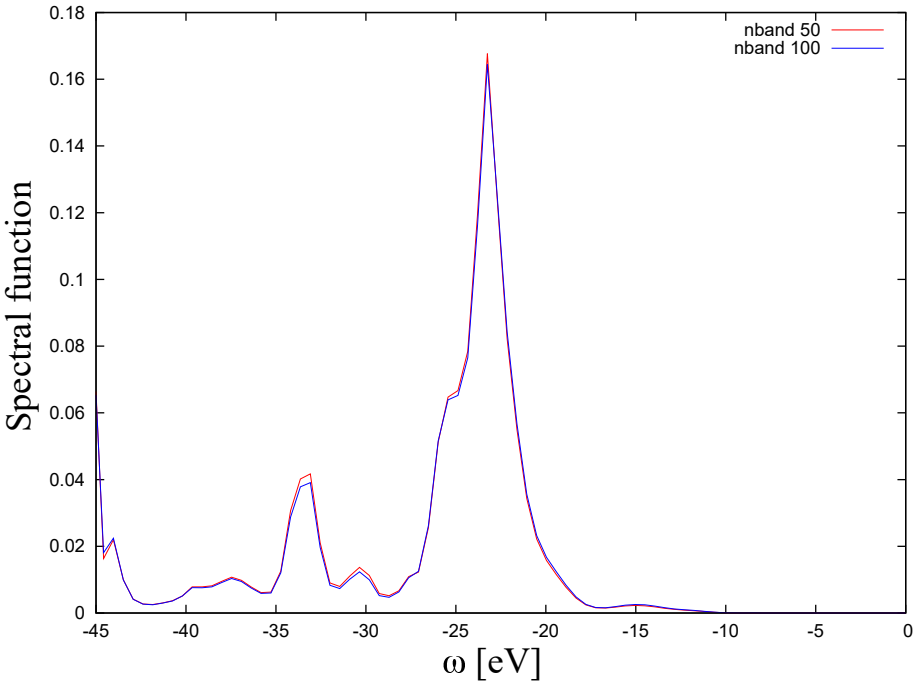
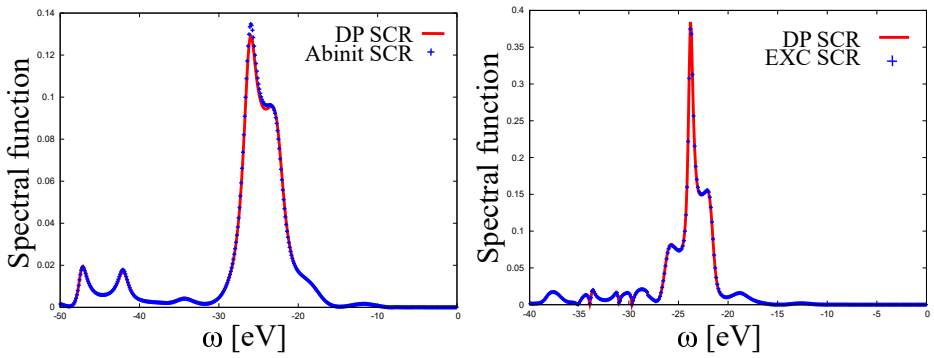


Figure D.1 – Spectral function of LiF for the lowest valence band

Cross-checking

Now that we have on one hand the LDA wave functions and band-structure, and on the other an RPA screening one could proceed to the calculation of the screening accounting for the excitonic effects. However, before making any further studies we compare results obtained with the previously existing codes. This is possible in the RPA approximation for the screening, which is the one implemented in ABINIT. These results are compared to the implementation in the DP-code and the new one, I have implemented, in EXC. As shown in figure Fig. D.2 the spectral functions in the GW approximation to the self-energy match perfectly.



(a) RPA using Abinit and DP

(b) RPA using DP and EXC

Figure D.2 – Spectral functions obtained with screenings generated using Abinit, DP and EXC codes. Lower valence band.

Acknowledgements

There is a number of people whom I wish to thank for their assistance during my PhD.

First and foremost I would like to thank my supervisor Lucia Reining for creating a unique working atmosphere, her constant advice and encouragement, and availability whenever it was necessary. In addition, I would like to thank Matteo Gatti, Francesco Sottile and Claudia Rödl for extremely fruitful discussions on numerous physical topics; France Pochard and Gaelle Bruant for a very warm welcome and strong support with all kinds of administrative work; Nicolas Tancogne-Dejean for interesting physical and non-physical discussions, and the corrections of my French; Andrea Cucca, Christine Giorgetti, Valérie Véniard and all the current and former members of the LSI Theoretical Spectroscopy Group, for creating an exceptional atmosphere. Furthermore I would like to thank Simo Huotari for inviting me to participate in the measurements at the ESRF in Grenoble.

Last, but not least, I would like to thank my parents for their support during my time as a PhD student at Ecole Polytechnique.

Bibliography

[1] P. Bouguer. *Essai d'optique, sur la gradation de la lumière*. chez Claude Jombert, ruë S. Jacques, au coin de la ruë des Mathurins, à l'Image Notre-Dame, 1729.

[2] J. Clerk Maxwell. A dynamical theory of the electromagnetic field. *Philosophical Transactions of the Royal Society of London*, 155:459–512, 1865.

[3] M. Born and R. Oppenheimer. Zur Quantentheorie der Molekeln. *Annalen der Physik*, 389(20):457–484, 1927.

[4] R. Del Sole and E. Fiorino. Macroscopic dielectric tensor at crystal surfaces. *Phys. Rev. B*, 29:4631–4645, Apr 1984.

[5] Philippe Nozières and David Pines. Electron interaction in solids. general formulation. *Phys. Rev.*, 109:741–761, Feb 1958.

[6] H. Ehrenreich and M. H. Cohen. Self-consistent field approach to the many-electron problem. *Phys. Rev.*, 115:786–790, Aug 1959.

[7] Stephen L. Adler. Quantum theory of the dielectric constant in real solids. *Phys. Rev.*, 126:413–420, Apr 1962.

[8] Nathan Wiser. Dielectric constant with local field effects included. *Phys. Rev.*, 129:62–69, Jan 1963.

[9] A. Einstein. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. *Annalen der Physik*, 322(6):132–148, 1905.

[10] S. P. Kowalczyk, F. R. McFeely, L. Ley, R. A. Pollak, and D. A. Shirley. X-ray photoemission studies of the alkali halides. *Phys. Rev. B*, 9:3573–3581, Apr 1974.

[11] Andrea Damascelli. Probing the electronic structure of complex systems by ARPES. *Physica Scripta*, 2004(T109):61, 2004.

[12] Wikipedia.

[13] Christian Rischel, Antoine Rousse, Ingo Uschmann, Pierre-Antoine Albouy, Jean-Paul Geindre, Patrick Audebert, Jean-Claude Gauthier, Eckhart Froster, Jean-Louis Martin, and Andre Antonetti. Femtosecond time-resolved x-ray diffraction from laser-heated organic films. *Nature*, 390(6659):490–492, Dec 1997.

[14] J. Cao, Z. Hao, H. Park, C. Tao, D. Kau, and L. Blaszczyk. Femtosecond electron diffraction for direct measurement of ultrafast atomic motions. *Applied Physics Letters*, 83(5):1044–1046, 2003.

[15] Sebastian Will, Thorsten Best, Ulrich Schneider, Lucia Hackermuller, Dirk-Soren Luhmann, and Immanuel Bloch. Time-resolved observation of coherent multi-body interactions in quantum phase revivals. *Nature*, 465(7295):197–201, May 2010.

[16] E. Papalazarou, J. Faure, J. Mauchain, M. Marsi, A. Taleb-Ibrahimi, I. Reshetnyak, A. van Roekeghem, I. Timrov, N. Vast, B. Arnaud, and L. Perfetti. Coherent phonon coupling to individual bloch states in photoexcited bismuth. *Phys. Rev. Lett.*, 108:256808, Jun 2012.

[17] Ray Egerton. *Physical principles of electron microscopy: an introduction to TEM, SEM, and AEM*. Springer Science & Business Media, 2006.

[18] W. Schülke. *Electron Dynamics by Inelastic X-Ray Scattering*. Oxford Series on Synchrotron Radiation. Oxford University Press, 2007.

[19] J. Verbeeck, D. Van Dyck, and G. Van Tendeloo. Energy-filtered transmission electron microscopy: an overview. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 59(10–11):1529 – 1534, 2004. 17th International Congress on X-Ray Optics and Microanalysis.

[20] F. J. García de Abajo. Optical excitations in electron microscopy. *Rev. Mod. Phys.*, 82:209–275, Feb 2010.

[21] R. DE L. KRONIG. On the theory of dispersion of x-rays. *J. Opt. Soc. Am.*, 12(6):547–556, Jun 1926.

[22] W. Schülke. Off-diagonal response by means of inelastic x-ray scattering. *Solid State Communications*, 43(11):863 – 866, 1982.

[23] Giovanni Onida, Lucia Reining, and Angel Rubio. Electronic excitations: density-functional versus many-body green's-function approaches. *Rev. Mod. Phys.*, 74:601–659, Jun 2002.

[24] L. D. Landau. Über die Bewegung der Elektronen in Kristalgitter. *Phys. Z. Sowjetunion*, pages 644–645, 1933.

[25] J. Frenkel. On the transformation of light into heat in solids. i. *Phys. Rev.*, 37:17–44, Jan 1931.

[26] J. Frenkel. On the transformation of light into heat in solids. ii. *Phys. Rev.*, 37:1276–1294, May 1931.

[27] Gregory H. Wannier. The structure of electronic excitation levels in insulating crystals. *Phys. Rev.*, 52:191–197, Aug 1937.

[28] R. J. Elliott. Intensity of optical absorption by excitons. *Phys. Rev.*, 108:1384–1389, Dec 1957.

[29] I. Egri. Electronic interband transitions: Plasmons, frenkel- and wannier-excitons. *Zeitschrift für Physik B Condensed Matter*, 42(2):99–106, 1981.

[30] I. Egri. The internal structure of plasmons. *Zeitschrift für Physik B Condensed Matter*, 53(3):183–189, 1983.

[31] W. Kohn. Nobel lecture: Electronic structure of matter—wave functions and density functionals. *Rev. Mod. Phys.*, 71:1253–1266, Oct 1999.

[32] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, Nov 1965.

[33] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, Nov 1964.

[34] C.-O. Almbladh and U. von Barth. Exact results for the charge and spin densities, exchange-correlation potentials, and density-functional eigenvalues. *Phys. Rev. B*, 31:3231–3244, Mar 1985.

[35] EKU Gross and RM Dreizler. *Density functional theory*, volume 337. Plenum Press ASI series B, 1995.

[36] Stephan Kümmel and Leeor Kronik. Orbital-dependent density functionals: Theory and applications. *Rev. Mod. Phys.*, 80:3–60, Jan 2008.

[37] Eleonora Luppi, Hans-Christian Weissker, Sandro Bottaro, Francesco Sottile, Valérie Veniard, Lucia Reining, and Giovanni Onida. Accuracy of the pseudopotential approximation in *ab initio* theoretical spectroscopies. *Phys. Rev. B*, 78:245124, Dec 2008.

[38] Ricardo Gómez-Abal, Xinzheng Li, Matthias Scheffler, and Claudia Ambrosch-Draxl. Influence of the core-valence interaction and of the pseudopotential approximation on the electron self-energy in semiconductors. *Phys. Rev. Lett.*, 101:106404, Sep 2008.

[39] Erich Runge and E. K. U. Gross. Density-functional theory for time-dependent systems. *Phys. Rev. Lett.*, 52:997–1000, Mar 1984.

[40] M.A.L. Marques, C. Ullrich, F. Nogueira, A. Rubio, K. Burke, and E.K.U. Gross. *Time-Dependent Density Functional Theory*. Springer Berlin Heidelberg, 2006.

[41] A. Zangwill and Paul Soven. Density-functional approach to local-field effects in finite systems: Photoabsorption in the rare gases. *Phys. Rev. A*, 21:1561–1572, May 1980.

[42] Andreas Heßelmann and Andreas Görling. Blindness of the exact density response function to certain types of electronic excitations: Implications for time-dependent density-functional theory. *Phys. Rev. Lett.*, 102:233003, Jun 2009.

[43] Wolfgang Hieringer and Andreas Görling. Failure of time-dependent density functional methods for excitations in spatially separated systems. *Chemical Physics Letters*, 419(4–6):557 – 562, 2006.

[44] C. A. Ullrich. Time-dependent density-functional theory beyond the adiabatic approximation: Insights from a two-electron model system. *The Journal of Chemical Physics*, 125(23):–, 2006.

[45] R. van Leeuwen and E. J. Baerends. Exchange-correlation potential with correct asymptotic behavior. *Phys. Rev. A*, 49:2421–2431, Apr 1994.

[46] G. Vignale, C. A. Ullrich, and S. Conti. Time-dependent density functional theory beyond the adiabatic local density approximation. *Phys. Rev. Lett.*, 79:4878–4881, Dec 1997.

[47] Lucia Reining, Valerio Olevano, Angel Rubio, and Giovanni Onida. Excitonic effects in solids described by time-dependent density-functional theory. *Phys. Rev. Lett.*, 88:066404, Jan 2002.

[48] Francesco Sottile, Valerio Olevano, and Lucia Reining. Parameter-free calculation of response functions in time-dependent density-functional theory. *Phys. Rev. Lett.*, 91:056402, Jul 2003.

[49] I. V. Tokatly and O. Pankratov. Many-body diagrammatic expansion in a kohn-sham basis: Implications for time-dependent density functional theory of excited states. *Phys. Rev. Lett.*, 86:2078–2081, Mar 2001.

[50] I. V. Tokatly, R. Stubner, and O. Pankratov. Many-body diagrammatic expansion for the exchange-correlation kernel in time-dependent density functional theory. *Phys. Rev. B*, 65:113107, Feb 2002.

[51] Silvana Botti, Arno Schindlmayr, Rodolfo Del Sole, and Lucia Reining. Time-dependent density-functional theory for extended systems. *Reports on Progress in Physics*, 70(3):357, 2007.

[52] H. Nyquist. Thermal agitation of electric charge in conductors. *Phys. Rev.*, 32:110–113, Jul 1928.

[53] Herbert B. Callen and Theodore A. Welton. Irreversibility and generalized noise. *Phys. Rev.*, 83:34–40, Jul 1951.

[54] H. Lehmann. Über Eigenschaften von Ausbreitungsfunktionen und Renormierungskonstanten quantisierter Felder. *Il Nuovo Cimento (1943-1954)*, 11(4):342–357, 1954.

[55] N. N. Bogoliubov. Kinetic equations. *Journal of Experimental and Theoretical Physics (in Russian)*, 16(8):691–702, 1946.

[56] Richard P. Feynman and Albert R. Hibbs. *Quantum mechanics and path integrals*. McGraw-Hill, New York, 1965.

[57] Lars Hedin. New method for calculating the one-particle green's function with application to the electron-gas problem. *Phys. Rev.*, 139:A796–A823, Aug 1965.

[58] F Aryasetiawan and O Gunnarsson. The GW method. *Reports on Progress in Physics*, 61(3):237, 1998.

[59] Mark S. Hybertsen and Steven G. Louie. First-principles theory of quasiparticles: Calculation of band gaps in semiconductors and insulators. *Phys. Rev. Lett.*, 55:1418–1421, Sep 1985.

[60] Mark S. Hybertsen and Steven G. Louie. Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies. *Phys. Rev. B*, 34:5390–5413, Oct 1986.

[61] R. W. Godby, M. Schlüter, and L. J. Sham. Trends in self-energy operators and their corresponding exchange-correlation potentials. *Phys. Rev. B*, 36:6497–6500, Oct 1987.

[62] Y. Pavlyukh and W. Hübner. Analytic solution of hedin's equations in zero dimensions. *Journal of Mathematical Physics*, 48(5):–, 2007.

[63] J A Berger, Pina Romaniello, Falk Tandetzky, Bernardo S Mendoza, Christian Brouder, and Lucia Reining. Solution to the many-body problem in one point. *New Journal of Physics*, 16(11):113025, 2014.

[64] Arno Schindlmayr. Analytic evaluation of the electronic self-energy in the GW approximation for two electrons on a sphere. *Phys. Rev. B*, 87:075104, Feb 2013.

[65] G. Strinati, H. J. Mattausch, and W. Hanke. Dynamical aspects of correlation corrections in a covalent crystal. *Phys. Rev. B*, 25:2867–2888, 1982.

[66] F Aryasetiawan and O Gunnarsson. The gw method. *Reports on Progress in Physics*, 61(3):237, 1998.

[67] G. Strinati. Application of the green's functions method to the study of the optical properties of semiconductors. *La Rivista del Nuovo Cimento (1978-1999)*, 11(12):1–86, 1988.

[68] Antoine Georges, Gabriel Kotliar, Werner Krauth, and Marcelo J. Rozenberg. Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. *Rev. Mod. Phys.*, 68:13–125, Jan 1996.

[69] P. Romaniello, S. Guyot, and L. Reining. The self-energy beyond GW: Local and nonlocal vertex corrections. *The Journal of Chemical Physics*, 131(15):–, 2009.

[70] Pina Romaniello, Friedhelm Bechstedt, and Lucia Reining. Beyond the GW approximation: Combining correlation channels. *Phys. Rev. B*, 85:155131, Apr 2012.

[71] Michael Rohlfing and Steven G. Louie. Electron-hole excitations and optical spectra from first principles. *Phys. Rev. B*, 62:4927–4944, Aug 2000.

[72] J R Schrieffer. Inelastic processes in photoemission from adsorbed atoms. *Physica Scripta*, 21(3-4):472, 1980.

[73] E. E. Salpeter and H. A. Bethe. A relativistic equation for bound-state problems. *Phys. Rev.*, 84:1232–1242, Dec 1951.

[74] R. E. Cutkosky. Solutions of a bethe-salpeter equation. *Phys. Rev.*, 96:1135–1141, Nov 1954.

[75] L. J. Sham and T. M. Rice. Many-particle derivation of the effective-mass equation for the wannier exciton. *Phys. Rev.*, 144:708–714, Apr 1966.

[76] W. Hanke and L. J. Sham. Many-particle effects in the optical spectrum of a semiconductor. *Phys. Rev. B*, 21:4656–4673, May 1980.

[77] V. Olevano, L. Reining, and F. Sottile. DP-code, http://www.dp-code.org/.

[78] X Gonze, G Rignanese, M Verstraete, J Betiken, Y Pouillon, R Caracas, F Jollet, M Torrent, G Zerah, M Mikami, P Ghosez, M Veithen, J-Y Raty, V Olevano, F Bruneval, L Reining, R Godby, G Onida, D Hamann, and D Allan. A brief introduction to the ABINIT software package. *Zeitschrift für Kristallographie.(Special issue on Computational Crystallography.)*, 220:558–562, 2005.

[79] A. Fleszar and W. Hanke. Spectral properties of quasiparticles in a semiconductor. *Phys. Rev. B*, 56:10228–10232, Oct 1997.

[80] Valerio Olevano and Lucia Reining. Excitonic effects on the silicon plasmon resonance. *Phys. Rev. Lett.*, 86:5962–5965, Jun 2001.

[81] Matteo Gatti and Francesco Sottile. Exciton dispersion from first principles. *Phys. Rev. B*, 88:155113, Oct 2013.

[82] Hans-Christian Weissker, Jorge Serrano, Simo Huotari, Eleonora Luppi, Marco Cazzaniga, Fabien Bruneval, Francesco Sottile, Giulio Monaco, Valerio Olevano, and Lucia Reining. Dynamic structure factor and dielectric function of silicon for finite momentum transfer: Inelastic x-ray scattering experiments and *ab initio* calculations. *Phys. Rev. B*, 81:085104, Feb 2010.

[83] B. Arnaud, S. Lebègue, and M. Alouani. Excitonic and quasiparticle lifetime effects on silicon electron energy loss spectra from first principles. *Phys. Rev. B*, 71:035308, Jan 2005.

[84] Roger Alan Smith. Planar version of baym-kadanoff theory. *Phys. Rev. A*, 46:4586–4597, Oct 1992.

[85] Lorenzo Sponza. *Damping, satellites and multiple excitations in oxides and nanostructures: efficient theoretical and numerical approaches towards a dynamical many-body theory*. PhD thesis, Ecole Polytechnique, 2013.

[86] Adrian Stan, Pina Romaniello, Santiago Rigamonti, Lucia Reining, and JA Berger. Multiple solutions in many-body theories. *arXiv preprint arXiv:1503.07742*, 2015.

[87] Carsten A Ullrich. *Time-dependent density-functional theory: concepts and applications*. Oxford University Press, 2011.

[88] M. Petersilka, U. J. Gossmann, and E. K. U. Gross. Excitation energies from time-dependent density-functional theory. *Phys. Rev. Lett.*, 76:1212–1215, Feb 1996.

[89] Mark E Casida. Time-dependent density functional response theory for molecules. *Recent advances in density functional methods*, 1:155, 1995.

[90] Rüdiger Bauernschmitt and Reinhart Ahlrichs. Treatment of electronic excitations within the adiabatic approximation of time dependent density functional theory. *Chemical Physics Letters*, 256(4–5):454 – 464, 1996.

[91] Mark E Casida, Christine Jamorski, Kim C Casida, and Dennis R Salahub. Molecular excitation energies to high-lying bound states from time-dependent density-functional response theory: Characterization and correction of the time-dependent local density approximation ionization threshold. *The Journal of chemical physics*, 108(11):4439–4449, 1998.

[92] Cladia Rödl. *Elektonische und exzitonische Anregungen in magnetischen Isolatoren*. PhD thesis, Fredrich-Schiller-Universitat Jena, 2009.

[93] Myrta Gruning, Andrea Marini, and Xavier Gonze. Exciton-plasmon states in nanoscale materials: Breakdown of the tamm-dancoff approximation. *Nano Letters*, 9(8):2820–2824, 2009. PMID: 19637906.

[94] Marie Lopez del Puerto, Murilo L. Tiago, and James R. Chelikowsky. *Ab initio* methods for the optical properties of cdse clusters. *Phys. Rev. B*, 77:045404, Jan 2008.

[95] L. E. Ramos, J. Paier, G. Kresse, and F. Bechstedt. Optical spectra of si nanocrystallites: Bethe-salpeter approach versus time-dependent density-functional theory. *Phys. Rev. B*, 78:195423, Nov 2008.

[96] Murilo L. Tiago, Juan C. Idrobo, Serdar Öğüt, Julius Jellinek, and James R. Chelikowsky. Electronic and optical excitations in ag_n clusters $(n = 1\tilde{}8)$: Comparison of density-functional and many-body theories. *Phys. Rev. B*, 79:155419, Apr 2009.

[97] Zeng-hui Yang and Carsten A. Ullrich. Direct calculation of exciton binding energies with time-dependent density-functional theory. *Phys. Rev. B*, 87:195204, May 2013.

[98] Lorenzo Sponza, Valerie Veniard, Francesco Sottile, Christine Giorgetti, and Lucia Reining. Role of localized electrons in electron-hole interaction: The case of SrTiO3. *Phys. Rev. B*, 87:235102, Jun 2013.

[99] Roland Winkler. *Quasi-Degenerate Perturbation Theory*, volume 191 of *Springer Tracts in Modern Physics*, pages 201–206. Springer Berlin Heidelberg, 2003.

[100] V. Turkowski, A. Leonardo, and C. A. Ullrich. Time-dependent density-functional approach for exciton binding energies. *Phys. Rev. B*, 79:233201, Jun 2009.

[101] Tobias Sander, Emanuele Maggio, and Georg Kresse. Beyond the tamm-dancoff approximation for extended systems using exact diagonalization. *Phys. Rev. B*, 92:045209, Jul 2015.

[102] W. Schülke. Compton scattering of photons from standing wave fields in the bragg case of x-ray diffraction. *Physics Letters A*, 83(9):451 – 454, 1981.

[103] W. Schülke, J. R. Schmitz, H. Schulte-Schrepping, and A. Kaprolat. Dynamic and static structure factor of electrons in Si: Inelastic x-ray scattering results. *Phys. Rev. B*, 52:11721–11732, Oct 1995.

[104] K. Sturm and W. Schülke. Shape of plasmon bands at the brillouin-zone boundary. *Phys. Rev. B*, 46:7193–7195, Sep 1992.

[105] P. Nozières and D. Pines. Electron interaction in solids. characteristic energy loss spectrum. *Phys. Rev.*, 113:1254–1267, Mar 1959.

[106] W. Schülke. Compton scattering of photons from standing wave fields in the bragg case of x-ray diffraction. *Physics Letters A*, 83(9):451 – 454, 1981.

[107] P. Abbamonte, K. D. Finkelstein, M. D. Collins, and S. M. Gruner. Imaging density disturbances in water with a 41.3-attosecond time resolution. *Phys. Rev. Lett.*, 92:237401, Jun 2004.

[108] Peter Abbamonte, Gerard C. L. Wong, David G. Cahill, James P. Reed, Robert H. Coridan, Nathan W. Schmidt, Ghee Hwee Lai, Young Il Joe, and Diego Casa. Ultrafast imaging and the phase problem for inelastic x-ray scattering. *Advanced Materials*, 22(10):1141–1147, 2010.

[109] Peter Abbamonte, Tim Graber, James P. Reed, Serban Smadici, Chen-Lin Yeh, Abhay Shukla, Jean-Pascal Rueff, and Wei Ku. Dynamical reconstruction of the exciton in lif with inelastic x-ray scattering. *Proceedings of the National Academy of Sciences*, 105(34):12159–12163, 2008.

[110] M Ehrnsperger and H Bross. Calculation of the dielectric matrix of si. *Journal of Physics: Condensed Matter*, 9(6):1225, 1997.

[111] W. A. Caliebe, J. A. Soininen, Eric L. Shirley, C.-C. Kao, and K. Hämäläinen. Dynamic structure factor of diamond and LiF measured using inelastic x-ray scattering. *Phys. Rev. Lett.*, 84:3907–3910, Apr 2000.

[112] David C. Langreth. Singularities in the x-ray spectra of metals. *Phys. Rev. B*, 1:471–477, Jan 1970.

[113] F. Aryasetiawan, L. Hedin, and K. Karlsson. Multiple plasmon satellites in Na and Al spectral functions from *Ab Initio* cumulant expansion. *Phys. Rev. Lett.*, 77:2268–2271, Sep 1996.

[114] Matteo Guzzo, Giovanna Lani, Francesco Sottile, Pina Romaniello, Matteo Gatti, Joshua J. Kas, John J. Rehr, Mathieu G. Silly, Fausto Sirotti, and Lucia Reining. Valence electron photoemission spectrum of semiconductors: *Ab Initio* description of multiple satellites. *Phys. Rev. Lett.*, 107:166401, Oct 2011.

[115] Guzzo, M., Kas, J.J., Sottile, F., Silly, M.G., Sirotti, F., Rehr, J.J., and Reining, L. Plasmon satellites in valence-band photoemission spectroscopy. *Eur. Phys. J. B*, 85(9):324, 2012.

[116] Fabio Caruso and Feliciano Giustino. Spectral fingerprints of electron-plasmon coupling. *Phys. Rev. B*, 92:045123, Jul 2015.

[117] Fabio Caruso, Henry Lambert, and Feliciano Giustino. Band structures of plasmonic polarons. *Phys. Rev. Lett.*, 114:146404, Apr 2015.

[118] A. S. Kheifets, V. A. Sashin, M. Vos, E. Weigold, and F. Aryasetiawan. Spectral properties of quasiparticles in silicon: A test of many-body theory. *Phys. Rev. B*, 68:233205, Dec 2003.

[119] M. Vos, A.S. Kheifets, V.A. Sashin, E. Weigold, M. Usuda, and F. Aryasetiawan. Quantitative measurement of the spectral function of aluminum and lithium by electron momentum spectroscopy. *Phys. Rev. B*, 66:155414, Oct 2002.

[120] Johannes Lischner, Derek Vigil-Fowler, and Steven G. Louie. Satellite structures in the spectral functions of the two-dimensional electron gas in semiconductor quantum wells: A GW plus cumulant study. *Phys. Rev. B*, 89:125430, Mar 2014.

[121] Jianqiang Zhou, J. Kas, Lorenzo Sponza, Igor Reshetnyak, Matteo Guzzo, Christine Giorgetti, Matteo Gatti, Francesco Sottile, J.J. Rehr, and Lucia Reining. Dynamical effects in electron spectroscopy. *JCP*, page submitted, 2015.

[122] Michele Casula, Alexey Rubtsov, and Silke Biermann. Dynamical screening effects in correlated materials: Plasmon satellites and spectral weight transfers from a green's function ansatz to extended dynamical mean field theory. *Phys. Rev. B*, 85:035115, Jan 2012.

[123] N. Troullier and José Luís Martins. Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B*, 43:1993–2006, Jan 1991.

[124] J. Shy-Yih Wang and M. Rasolt. Study of the density gradient expansion in the surface energy calculation for metals. *Phys. Rev. B*, 13:5330–5337, Jun 1976.

[125] S. Lebègue, B. Arnaud, M. Alouani, and P. E. Bloechl. Implementation of an all-electron GW approximation based on the projector augmented wave method without plasmon pole approximation: Application to Si, SiC, AlAs, InAs, NaH, and KH. *Phys. Rev. B*, 67:155208, Apr 2003.

[126] Eric L. Shirley, Louis J. Terminello, John E. Klepeis, and Franz J. Himpsel. Detailed theoretical photoelectron angular distributions for LiF(100). *Phys. Rev. B*, 53:10296–10309, Apr 1996.

[127] Neng-Ping Wang, Michael Rohlfing, Peter Krüger, and Johannes Pollmann. Quasiparticle band structure and optical spectrum of LiF(001). *Phys. Rev. B*, 67:115111, Mar 2003.

[128] Santiago Rigamonti. Private communications. 2015.

[129] M. Piacentini, D. W. Lynch, and C. G. Olson. Thermoreflectance of LiF between 12 and 30 ev. *Phys. Rev. B*, 13:5530–5543, Jun 1976.

[130] F. J. Himpsel, L. J. Terminello, D. A. Lapiano-Smith, E. A. Eklund, and J. J. Barton. Band dispersion of localized valence states in LiF(100). *Phys. Rev. Lett.*, 68:3611–3614, Jun 1992.

[131] Alex Zunger and A. J. Freeman. Ground- and excited-state properties of LiF in the local-density formalism. *Phys. Rev. B*, 16:2901–2926, Sep 1977.

[132] A. Barry Kunz. Study of the electronic structure of twelve alkali halide crystals. *Phys. Rev. B*, 26:2056–2069, Aug 1982.

[133] R. W. Godby and R. J. Needs. Metal-insulator transition in Kohn-Sham theory and quasiparticle theory. *Phys. Rev. Lett.*, 62:1169–1172, Mar 1989.

[134] Fabien Bruneval, Nathalie Vast, and Lucia Reining. Effect of self-consistency on quasiparticles in solids. *Phys. Rev. B*, 74:045102, Jul 2006.

[135] Eric K. Chang, Michael Rohlfing, and Steven G. Louie. Excitons and optical properties of α-quartz. *Phys. Rev. Lett.*, 85:2613–2616, Sep 2000.

[136] D. A. Shirley. High-resolution x-ray photoemission spectrum of the valence bands of gold. *Phys. Rev. B*, 5:4709–4714, Jun 1972.

[137] Marisa Scrocco. Satellites in x-ray photoelectron spectroscopy of insulators. ii. multielectron excitations in LiF, NaF, and KF. *Phys. Rev. B*, 32:1306–1310, Jul 1985.

[138] S. P. Kowalczyk, F. R. McFeely, L. Ley, R. A. Pollak, and D. A. Shirley. X-ray photoemission studies of the alkali halides. *Phys. Rev. B*, 9:3573–3581, Apr 1974.

[139] Hendrik J. Monkhorst and James D. Pack. Special points for brillouin-zone integrations. *Phys. Rev. B*, 13:5188–5192, Jun 1976.

[140] V. Hernandez, J. E. Roman, A. Tomas, and V. Vidal. Krylov-schur methods in SLEPc. Technical Report STR-7, Universitat Politècnica de València, 2009. Available at http://www.grycap.upv.es/slepc.

[141] Iurii Timrov, Nathalie Vast, Ralph Gebauer, and Stefano Baroni. Electron energy loss and inelastic x-ray scattering cross sections from time-dependent density-functional perturbation theory. *Phys. Rev. B*, 88:064301, Aug 2013.

[142] Jane K. Cullum. Arnoldi versus nonsymmetric lanczos algorithms for solving nonsymmetric matrix eigenvalue problems. *BIT*, 36:470–493, 1996.