CEA-R-5856(E)

ISSN 0429 - 3460

COMMISSARIAT À L'ÉNERGIE ATOMIQUE

DIRECTION DES SCIENCES DE LA MATIÈRE

ABSORPTION OPTIQUE DES SEMI-CONDUCTEURS ET ISOLANTS : CALCUL AB-INITIO DES EFFETS À PLUSIEURS CORPS.

par

Stefan ALBRECHT

DIRECTION DES SCIENCES DE LA MATIÈRE

DÉPARTEMENT DE RECHERCHE SUR L'ÉTAT CONDENSÉ, LES ATOMES ET LES MOLÉCULES

LABORATOIRE DES SOLIDES IRRADIÉS

CEA / Saclay

RAPPORT

DIRECTION DE L'INFORMATION SCIENTIFIQUE ET TECHNIQUE

CEA/SACLAY 91191 GIF-SUR-YVETTE CEDEX FRANCE

RAPPORT CEA-R-5856(E)

1999

ABSORPTION OPTIQUE DES SEMI-CONDUCTEURS ET ISOLANTS : CALCUL *AB INITIO* DES EFFETS À PLUSIEURS CORPS

Résumé – Nous présentons une méthode pour inclure des effets de la self-énergie et des effets excitoniques dans des calculs *ab initio* des spectres d'absorption. Ces calculs sont effectués en utilisant l'état de l'art en matière de pseudopotentiels avec une base d'ondes planes.
Nous calculons d'abord l'état fondamental en utilisant la théorie de la fonctionnelle de densité (DFT) et l'approximation de la densité locale (LDA). Nous déterminons ensuite le spectre de quasiparticules physiques en corrigeant le potentiel d'échange-corrélation du calcul DFT-LDA avec la self-énergie, pour laquelle nous employons l'approximation *GW* de Hedin. L'interaction électron–trou est traitée par une équation effective de Schrödinger à deux particules, que nous avons dérivée des équations couplées intégrales de Hedin en passant par l'équation fondamentale de Bethe-Salpeter.
Le kernel d'interaction comporte une interaction écrantée Coulombienne et des effets d'échange électron–trou, qui représentent la structure microscopique du système et qui sont aussi appelés les effets de champ local. Nous obtenons les états propres excitoniques par diagonalisation. Cela permet une analyse détaillée des propriétés optiques. En appliquant des propriétés de symétrie nous réduisons la dimension de la matrice Hamiltonienne à deux particules et nous pouvons minimiser le coût du calcul.
Nous appliquons cette approche au silicium, au diamant, à l'oxyde de lithium et au tétramère de sodium. Un bon accord avec l'expérience est obtenu pour les spectres d'absorption optique de Si et du diamant, la constante diélectrique statique du diamant et pour le seuil d'absorption optique de Li_2O, qui est déterminé par un état excitonique lié. Nous discutons plusieurs approximations de la méthode et nous montrons le mélange très fort des transitions de particules indépendantes dans un état excité de l'agrégat Na_4.
L'influence du calcul de l'état fondamental sur des spectres d'absorption optique est examinée en considérant particulièrement la création des pseudopotentiels et nous discutons le choix de l'échantillonnage de la zone de Brillouin pour des calculs spectroscopiques.

OPTICAL ABSORPTION SPECTRA OF SEMICONDUCTORS AND INSULATORS: *AB INITIO* CALCULATION OF MANY-BODY EFFECTS

Summary – A method for the inclusion of self-energy and excitonic effects in first-principle calculations of absorption spectra, within the state-of-the-art plane wave pseudopotential approach, is presented.
Starting from a ground state calculation, using density functional theory (DFT) in the local density approximation (LDA), we correct the exchange-correlation potential of DFT-LDA with the self-energy applying Hedin's *GW* approximation to obtain the physical quasiparticles states. The electron–hole interaction is treated solving an effective two-particle equation, which we derive from Hedin's coupled integral equations, leading to the fundamental Bethe-Salpeter equation in an intermediate step.
The interaction kernel contains the screened electron–hole Coulomb interaction and the electron–hole exchange effects, which reflect the microscopic structure of the system and are thus also called local-field effects. We obtain the excitonic eigenstates through diagonalization. This allows us a detailed analysis of the optical properties. The application of symmetry properties enables us to reduce the size of the two-particle Hamiltonian matrix, thus minimizing the computational effort. We apply our method to silicon, diamond, lithium oxide and the sodium tetramer. Good agreement with experiment is obtained for the absorption spectra of Si and diamond, the static dielectric constant of diamond, and for the onset of optical absorption of Li_2O due to discrete bound excitons. We discuss various approximations of our method and show the strong mixing of independent-particle transitions to a bound excitonic state in the Na_4 cluster.
The influence of ground state calculations on optical spectra is investigated under particular consideration of the pseudopotential generation and we discuss the use of different Brillouin zone point sampling schemes for spectral calculations.

ÉCOLE POLYTECHNIQUE

Thèse présentée pour obtenir le grade de

DOCTEUR DE L'ÉCOLE POLYTECHNIQUE

Spécialité : **Physique du solide**

par

Stefan ALBRECHT

OPTICAL ABSORPTION SPECTRA OF SEMICONDUCTORS AND INSULATORS: *AB INITIO* CALCULATION OF MANY-BODY EFFECTS

Soutenue le 8 janvier 1999 devant le jury composé de :

Monsieur	Rex Godby	Rapporteur
Monsieur	Gilles Zérah	Rapporteur
Monsieur	Rodolfo Del Sole	Examinateur
Monsieur	François Ducastelle	Examinateur
Monsieur	Michel Lannoo	Examinateur
Madame	Lucia Reining	Examinateur
Monsieur	Angel Rubio	Examinateur

- Rapport CEA-R-5856(E) -

CEA Saclay
Direction des Sciences de la Matière
Département de Recherche sur l'Etat Condensé, les Atomes et les Molécules
Laboratoire des Solides Irradiés

ASBORPTION OPTIQUE DES SEMI-CONDUCTEURS ET ISOLANTS : CALCUL AB-INITIO DES EFFETS A PLUSIEURS CORPS.

par

Stefan ALBRECHT

- Juin1999 -

Acknowledgments

First of all I would like to thank my advisor, Lucia Reining, for supporting and encouraging me, while giving me the freedom to work as I have preferred. I am very grateful for the care with which she has read the thesis. Without the help and the insight of Giovanni Onida and Rodolfo Del Sole in Rome this work would not have been possible and I feel very happy having worked along with them. I am particularly indebted to Michel Lannoo in Lille, who accepted to be the supervisor of my thesis. He let me free to pursue how I wanted, though always being interested and open for discussion.

Many people in our lab SESI have done their very best to make research and life as good as possible. I would just like to mention the ones I have had most contact with. Madeleine Meyer and Marc Hayoun for correcting my French and getting me out of various problems with our workstations. Like Joseph Morillo, who has also shared his interest in good wine with me. Beat Schmidt and Stéphane Berry for always being helpful and open minded. Martin Gander from the CMAP for many exciting discussions about matrices, essential help with formatting and that kind of stuff, and much entertainment. Alexandra Charaya for leaving me our Mac to type the thesis.

Among those who have already left, there are Pierre-Antoine Vieillefond, Sophie Meillon and Emmanuel Frély, who has explained to me the peculiarities of the French administration. Several visitors have contributed to the variety. In particular with Jens Jensen and Geoff Beach I have had many talks and much fun inside and outside the lab. Many thanks to Valerio Olevano in Rome for providing the nonlocal commutator code.

I consider myself privileged having had the opportunity to discuss with researchers interested in my work. Especially I have appreciated the contacts with Friedhelm Bechstedt in Jena, Rex Godby in York and Angel Rubio in Valladolid, who have inspired my work and encouraged me with motivating questions.

I would like to express my gratitude towards the Rapporteurs for assuming the work of reporting and towards the Examinateurs for attending the jury, in particular towards François Ducastelle for chairing it.

Financial support for living and travel expenses from the École Polytechnique is gratefully acknowledged.

Special thanks are for my family for ongoing support on many occasions.

Contents

Chapter 1

Introduction

The spectroscopic properties of solids have been very important in the development of solid state theory [1]. The interaction of light with matter gives rise to many different phenomena and we can learn about the behavior of particles in the solid as well as about their interactions.

In an inverse photoemission measurement, for example, an incoming electron is absorbed in the solid and consequently a photon emitted. Ideally the electron is initially far away from the sample and quickly enough absorbed that we only measure the intrinsic properties of the electron in the solid. This is important since a single excited electron in a polarizable medium – the solid – has different characteristics than a bare electron in vacuum and is thus denoted a *quasiparticle*. The opposite process occurs in a photoemission experiment, where an incoming photon is absorbed and an electron is emitted above the continuum energy level leaving the solid into the vacuum and creating a hole, *i.e.*, a missing electron.

Both processes together describe the energetic levels of the one-particle quasiparticles in the solid, referred to as the band structure. It reveals a wealth of information about electronic and spectroscopic properties. The principal characteristic, for example, of insulators and semiconductors compared to metals is the band gap: In the ground state at zero temperature all energy bands up to a certain level are completely occupied and the first unoccupied state is situated in an energy band separated by an energy gap from the highest occupied level. Metals, on the contrary, have even in the ground state partially filled bands allowing, *e.g.*, the easy flow of electric current.

In order to obtain the electronic structure of a system we follow the state-of-the-art approach of modern *ab initio* calculations. We start from a ground state calculation, using density functional theory (DFT) in the local density approximation (LDA) [2, 3], and obtain the physical quasiparticle (QP) states in a second step. Here we correct the DFT-LDA eigenvalues by using the true electron self-energy operator Σ, which appears at the place of the DFT-LDA exchange-correlation potential. A good approximation for Σ allowing to determine the QP band structure can be obtained within Hedin's *GW* approach [4]. In this scheme, DFT-LDA results

can be used as the starting point in a first-order perturbative approach [5, 6]. This method will be discussed in detail in the next chapter.

The main interest of this thesis are optical absorption spectra of insulators and semiconductors. An electron absorbs the incoming photon and is – in a non-metallic material – energetically put above the band gap from a valence band to a conduction band. However, contrary to a photoemission experiment, the electron stays in the solid and interacts more or less strongly with the hole left behind in the valence band, thus creating an *exciton*. As an immediate consequence, for example, bound exciton states occur within the gap reducing the optical absorption onset and possibly considerably influencing the shape of the absorption spectrum.

A first starting point to describe optical absorption is traditionally nevertheless the one-electron band theory, where we neglect the electron–hole interaction and consider only independent, direct transitions from the filled valence to the empty conduction bands at zero temperature. Although qualitative agreement with experiment was obtained for several materials [7, 8, 9] the calculated spectra are typically too poor to be used as a reference for the interpretation and prediction of experimental results. Indeed, besides the reduction of the optical absorption onset, the spectral lineshape above the continuous–absorption edge is in most cases largely influenced by excitonic effects. Spectral weight is typically shifted to lower energies and can considerably deform the spectrum. These effects are in competition with the local-field effects reflecting the microscopic spatial variation of the material. The latter are not included in the simple independent-particle absorption spectrum either and tend to shift absorption spectra to higher energies. Pioneering work in this context was done by Hanke and Sham, who investigated these two effects at the examples of diamond [10] and silicon [11, 12]. They used a tight-binding calculation to model the electronic structure of the materials.

In this work we solve the problem of including excitonic and local-field effects in optical absorption spectra in an *ab initio* approach, thus not relying on empirical parameters nor restricting the validity of the method to specific optimizations with respect to the material. Since first-principle band structure calculations have become a reliable tool in solid state theory in the 80's [5, 6], the next important step to address two-particle interactions has remained a major challenge.

The polarizability, from which the absorption spectrum can be calculated, is given by the two-particle Green's function [13]. The equation of motion for this two-particle Green's function S is the Bethe-Salpeter equation, which can be derived using Hedin's coupled integral equations [4] and iterating a second time after the first, GW level. It is an integral equation of the form

$$S = S_0 + S_0 \Xi S, \tag{1.1}$$

where S_0 is the non-interacting electron–hole pair, including, however, all band structure and many-particle effects of the one-particle quasiparticles. Ξ is the interaction kernel and includes the exchange effects, reflecting the microscopic structure of the

material, therefore also called the local-field effects, and the screened electron–hole interaction, creating the excitonic effects. In this work we solve for S not by inverting the Eq. (1.1), but by decomposing S in a spectral representation, thus avoiding an inversion, or the mathematical equivalent, for each absorption frequency and – more important – obtaining the excitonic eigenvalues and eigenstates through diagonalization.

In the limit of a large Wannier exciton near the fundamental threshold, Eq. (1.1) can be reduced to the Mott-Wannier model in the effective mass approximation [13]. The electron–hole screening is then given by the long-range dielectric constant. This model is basically valid for direct transitions near the absorption onset in the case of parabolic bands at the band extrema.

Other possible features of band structures are much more difficult to model. The typical example is the E_1 peak of silicon, where nearly parallel conduction and valence bands around the gap lead to expressed continuum-excitonic effects, largely deforming the spectrum calculated in the independent-particle approximation and making simple model assumptions quite useless [12]. The *ab initio* calculation of the absorption spectrum of silicon, being representative for the group IV, III-V and II-VI semiconductors, is thus one of the main topics of this thesis.

In the first part of this work, we will introduce the fundamental theoretical concepts underlying the optical absorption calculations in Chapter 2. These include density functional theory in the Kohn-Sham formalism, using the local density approximation, and the quasiparticle formalism, which allows us to determine the quasiparticle energy levels in the GW approximation. In Chapter 3 we derive in detail the calculation of the macroscopic dielectric function in the independent-particle picture and, of course, consider the inclusion of exchange and/or excitonic effects. A feasible *ab initio* approach is presented, solving the Bethe-Salpeter equation by a spectral decomposition, thus obtaining full information about the excitonic states. Several ways to minimize the computational cost through application of group theory or careful approximations are discussed.

Chapters 4 and 5 apply the theory to silicon and diamond and discuss in detail the various aspects of the calculations at the one-particle and the two-particle level. Special consideration is given to the ground state calculations, notably the pseudopotential method, which is introduced in Appendix A. Here we investigate the construction and the use of soft pseudopotentials and the Kleinman-Bylander separation on the example of lithium oxide and present also results of excitonic and optical calculations of this highly ionic material. In particular, the onset of optical absorption is identified as the lowest excitonic eigenvalue. Appendix B shows the strong mixing of independent-particle transitions to a bound excitonic eigenstate at the example of a small cluster, the sodium tetramer. Chapter 6 finally gives general conclusions of the main topics addressed in this work and briefly summarizes the thesis.

Chapter 2

Fundamental Theories

In this chapter we summarize the underlying theoretical framework of this thesis. We present on introduction to density functional theory and address the problem of quasiparticle states and energies in the so-called *GW* approximation, which will provide us with the band structure.

Since we are interested in electronic states, we limit the discussion to fermions. Furthermore, only the case of zero temperature is considered.

2.1 Key Concepts and Formulas

This section defines essential formulas and concepts needed in the following [14]. We abbreviate the spin and space coordinates by $\mathbf{x}$ and use $x \equiv (\mathbf{x},t) \equiv (\mathbf{r},\sigma,t)$. Atomic units ($\hbar = \frac{e^2}{4\pi\varepsilon_0} = m_e = 1$) are used throughout.

We consider N electrons moving in an external potential w interacting with each other via the Coulomb interaction $v(\mathbf{r},\mathbf{r}') = 1/|\mathbf{r}-\mathbf{r}'|$. Here w is assumed to be time independent. For, *e.g.*, valence electrons in a solid, $w = w(\mathbf{r})$ can be the potential of the ion cores. If relativistic effects can be neglected, the time evolution of the system is determined by the Schrödinger equation,

$$H(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N)\Psi(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N,t) = i\frac{\partial\Psi(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N,t)}{\partial t}. \tag{2.1}$$

The *Hamiltonian* H is given by

$$H(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N) = \sum_{i=1}^{N}\left(-\frac{1}{2}\nabla_{\mathbf{r}_i}^2 + w(\mathbf{x}_i)\right) + \frac{1}{2}\sum_{i\neq j}^{N} v(\mathbf{r}_i,\mathbf{r}_j). \tag{2.2}$$

The stationary states of the system are determined from the time-independent Schrödinger equation,

$$H(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N)\Psi(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N) = E\Psi(\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N), \tag{2.3}$$

and the wavefunction is antisymmetric under the interchange of space and spin variables.

In quantum field theory, the concept of *second quantization* permits the explicit inclusion of the statistics (Bose-Einstein or Fermi-Dirac) and allows to describe creation and annihilation processes in a direct and natural way. The second quantized Hamiltonian $\hat{H}$ reads as

$$\hat{H} = \hat{H}_0 + \hat{V} = \hat{T} + \hat{W} + \hat{V}, \tag{2.4}$$

where

$$\hat{T} = \int d\mathbf{x}\, \psi_H^\dagger(x) \left[-\frac{1}{2}\nabla_{\mathbf{r}}^2\right] \psi_H(x), \tag{2.5}$$

$$\hat{W} = \int d\mathbf{x}\, \psi_H^\dagger(x) w(\mathbf{x}) \psi_H(x), \tag{2.6}$$

$$\hat{V} = \frac{1}{2} \int d\mathbf{x} d\mathbf{x}'\, \psi_H^\dagger(x) \psi_H^\dagger(x') v(\mathbf{r}, \mathbf{r}') \psi_H(x') \psi_H(x). \tag{2.7}$$

The creation and annihilation *field operators* $\psi_H^\dagger(x)$ and $\psi_H(x)$ are used in the Heisenberg representation with time-independent state vectors and time-dependent operators:

$$\psi_H(\mathbf{x}, t) = e^{i\hat{H}t} \psi(\mathbf{x}) e^{-i\hat{H}t}, \tag{2.8a}$$

$$\psi_H^\dagger(\mathbf{x}, t) = e^{-i\hat{H}t} \psi^\dagger(\mathbf{x}) e^{i\hat{H}t}. \tag{2.8b}$$

The defining fermionic anti-commutation relations are

$$\left\{\psi_H(x), \psi_H^\dagger(x')\right\} = \delta(x - x'), \tag{2.9a}$$

$$\{\psi_H(x), \psi_H(x')\} = \left\{\psi_H^\dagger(x), \psi_H^\dagger(x')\right\} = 0, \tag{2.9b}$$

where the anti-commutator $\{A, B\}$ denotes $AB + BA$.

2.2 Density Functional Theory

The electrons in the solid under consideration form an interacting many-particle system. Due to the enormous amount of degrees of freedom – typically of the order of 10^{23} – a direct solution of the describing Schrödinger equation is infeasible. One way to proceed could be to develop models for the Hamiltonian and to apply simplifying computational schemes.

However, instead of considering the many-particle wavefunction,

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N), \tag{2.10}$$

as a solution of the Schrödinger equation (2.3), a very promising approach is to start from the density,

$$n(\mathbf{r}) = N \int d\mathbf{x}_2 ... \int d\mathbf{x}_N \; \Psi^*(\mathbf{r}, \mathbf{x}_2, ..., \mathbf{x}_N)\Psi(\mathbf{r}, \mathbf{x}_2, ..., \mathbf{x}_N), \tag{2.11}$$

as the basic quantity to describe the system. The number of degrees of freedom is thus reduced to three and it is possible to find an exact solution. Moreover, we have direct access to a measurable physical quantity.

Early attempts to base a formulation on the density date back to the work of Thomas and Fermi [15, 16]. In 1964 *density functional theory* (DFT) was introduced by Hohenberg and Kohn [2]. Although it was formulated in principle only for the ground state density of a system, many modern electronic structure calculations start from one of its many powerful computational implications.

Usually one applies the scheme of Kohn and Sham [3] to obtain the ground state properties. This method delivers one-particle wavefunctions, but now of a corresponding non-interacting system. It is a good starting point for more sophisticated and accurate calculations of electronic quasiparticle states (see Section 2.3). In the following we will only present the basic definitions and theorems of DFT together with a short description of the employed computer code. Proofs of the theorems and more detailed technical information can be found, for example, in [17], [18, Chap. 2] and [19, Chap. 1], respectively.

2.2.1 Hohenberg-Kohn Theorem

The original paper by Hohenberg and Kohn [2] is the pioneering work of density functional theory. The two important findings are:

1. The ground state energy $E = \langle \Psi_N | \hat{H} | \Psi_N \rangle$ of a system of N interacting particles with the Hamiltonian $\hat{H}$ of Eq. (2.4) in an external potential $w(\mathbf{r})$ is a *unique* functional of the density $n(\mathbf{r})$: $E = E[n]$. It can be written in the form

$$E[n] = F[n] + \int d\mathbf{r} \; w(\mathbf{r})n(\mathbf{r}), \tag{2.12}$$

   where the universal density functional F depends only on n, but not on w. The functional F contains the complete information about all intrinsic properties of the many-body system, regardless of its actual environment described by the external potential w.

2. The energy functional E has its minimum, the ground state energy E_0, at the physical ground state density $n_0(\mathbf{r})$ of the system: $E_0 = E[n_0]$.

In the proof of the theorem, Hohenberg and Kohn derive a bijective relation between the external potential $w(\mathbf{r})$ and the ground state density $n_0(\mathbf{r})$. Two potentials

w which differ by more than just a trivial constant must lead to different ground state densities. Since the ground state density determines the external potential (to within a constant), the Hamiltonian is completely described, which in turn defines the many-body wavefunction. All *ground state properties* can be derived from the wavefunction of the ground state and are thus functionals of the density. The DFT contains no approximations and is an *exact* theory.

The original proof assumes a non-degenerate ground state and a w -representable particle density, *i.e.*, it is the ground state density of a many-electron system subject to some local external potential w. However, the theory can be generalized [17] to degenerate systems and N-representable densities, *i.e.*, densities being the expectation values of the density operator,

$$n(\mathbf{r}) = \langle \Psi_N | \, \hat{\rho}(\mathbf{r}) \, | \Psi_N \rangle = \langle \Psi_N | \, \psi^\dagger(\mathbf{r}) \psi(\mathbf{r}) \, | \Psi_N \rangle \,, \tag{2.13}$$

of some anti-symmetric wavefunction $|\Psi_N\rangle$.

The Hohenberg-Kohn theorem is very satisfying from the theoretical point of view, but it is not immediately applicable. The density functional of the energy is unknown and one has to resort to approximations. It turns out that the large kinetic energy has to be treated as precisely as possible, in contrast to the Thomas-Fermi approach, where the kinetic energy is expressed by an approximated functional of non-interacting electrons. Kohn and Sham presented a scheme [3], which allows an – in principle – exact determination of the ground state energy and restricts inevitable approximations to some smaller remaining part, namely the exchange-correlation potential.

2.2.2 Kohn-Sham Equations

In order to develop a practical scheme for the DFT, Kohn and Sham represent the system of interacting particles by an artificial one of *non-interacting* particles, which has by construction the same ground state energy and ground state charge density [3]. The ground state itself of this artificial system is found by self-consistently solving a corresponding set of effective Schrödinger equations.

Kohn and Sham start by writing the energy functional of Eq. (2.12) in the form

$$E\,[n] = T_0\,[n] + \int d\mathbf{r}\; n(\mathbf{r}) w(\mathbf{r}) + \frac{1}{2} \int d\mathbf{r} \int d\mathbf{r}'\; \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + E_{\mathrm{xc}}\,[n] \,, \tag{2.14}$$

where $T_0\,[n]$ is the functional of the kinetic energy of non-interacting electrons. The functional $E_{\mathrm{xc}}\,[n]$ contains the *exchange-correlation* contributions as well as corrections to the kinetic term due to the electron–electron interaction. The comparably small exchange-correlation functional is the only unknown part and has to be approximated. In the next section we will discuss the most widely used local density approximation. The distinct advantage of the Kohn-Sham ansatz is the correct

treatment of the large kinetic energy of the non-interacting system, which is close to the one of the real interacting system.

The essential step is to find a relation between the real interacting system and a fictitious non-interacting many-particle system, described by the set of wavefunctions $\{\varphi_i\}$, under the constraint that the latter has exactly the same density,

$$n(\mathbf{r}) = \sum_i^N |\varphi_i(\mathbf{r})|^2 , \tag{2.15}$$

as the former.

We minimize the total energy functional Eq. (2.14) with respect to the density. Preservation of particle number and the orthonormality condition of the wavefunctions leads to Lagrange parameters ϵ_{ij} in the functional

$$\Omega[\{\varphi_i\}] = E[\{\varphi_i\}] - \sum_i^N \sum_j^N \epsilon_{ij} \int d\mathbf{r}\ \varphi_i^*(\mathbf{r})\varphi_j(\mathbf{r}), \tag{2.16}$$

which is minimized with respect to $\varphi_i^*(\mathbf{r})$, in equivalence to a minimization with respect to the density $n(\mathbf{r})$, as can be seen from Eq. (2.15).

A unitary transformation by diagonalization leads to the *Kohn-Sham (KS) equations*

$$\hat{h}_{\mathrm{KS}}^{\mathrm{eff}} \varphi_i = \left[-\frac{1}{2} \nabla^2 + \hat{v}_{\mathrm{eff}} \right] \varphi_i = \epsilon_i \varphi_i, \tag{2.17}$$

with

$$v_{\mathrm{eff}}(\mathbf{r}) \equiv w(\mathbf{r}) + \int d\mathbf{r}'\ \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + V_{\mathrm{xc}}(\mathbf{r}), \tag{2.18}$$

$$V_{\mathrm{xc}}(\mathbf{r}) \equiv \frac{\delta}{\delta n(\mathbf{r})} E_{\mathrm{xc}}[n]. \tag{2.19}$$

Due to construction, the effective potential $v_{\mathrm{eff}}(\mathbf{r})$ of the non-interacting system gives the correct density of the interacting system. The Kohn-Sham equations (2.17) are Schrödinger-like and must be solved self-consistently, since the density Eq. (2.15), which enters in each one-particle equation, depends explicitly on all wavefunctions $\varphi_i(\mathbf{r})$.

Once solved, the system of Eqs. (2.17) would give an exact solution for the ground state density of the interacting system, if the exact exchange correlation potential V_{xc} were known. Unfortunately, this is not the case, and one has to find approximations to V_{xc}.

2.2.3 Local Density Approximation

The simplest and most-widely used approximation for the exchange-correlation functional,

$$E_{\mathrm{xc}}[n] = \int d\mathbf{r}\ n(\mathbf{r}) \epsilon_{\mathrm{xc}}([n], \mathbf{r}), \tag{2.20}$$

is the *local density approximation* (LDA). We assume that the well-known results of the homogeneous electron gas can be locally applied:

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[n] = \int d\mathbf{r}\; n(\mathbf{r}) \epsilon_{\mathrm{xc}}^{\mathrm{hom}}(n(\mathbf{r}))\,. \tag{2.21}$$

The function

$$\epsilon_{\mathrm{xc}}^{\mathrm{hom}}(n(\mathbf{r})) = \epsilon_{\mathrm{xc}}^{\mathrm{hom}}(n(\mathbf{r})) + \epsilon_{\mathrm{xc}}^{\mathrm{hom}}(n(\mathbf{r})) \tag{2.22}$$

is obtained from quantum-mechanical calculations of the interacting homogeneous electron gas. The easiest approximation is given by the Hartree-Fock method, which does not include any correlation and considers only the exchange energy of the homogeneous electron gas [20],

$$\epsilon_{\mathrm{xc}}^{\mathrm{hom}}(n(\mathbf{r})) = -\frac{3}{4}\left[\frac{3n(\mathbf{r})}{\pi}\right]^{1/3}. \tag{2.23}$$

The best nowadays available correlation energies have been calculated with quantum-Monte-Carlo methods by Ceperley and Alder [21]. Their results have been parametrized by, *e.g.*, Perdew and Zunger [22] and also allow the treatment of spin-polarized systems (local spin density approximation, LSD). In this work, however, we will only consider spin-paired electrons. The correlation energy is then given by

$$\epsilon_{\mathrm{xc}}^{\mathrm{hom}}(n(\mathbf{r})) = \gamma/\left(1 + \beta_1\sqrt{r_s} + \beta_2 r_s\right), \tag{2.24}$$

for $r_s \geq 1$, where

$$\frac{1}{n(\mathbf{r})} = \frac{4}{3}\pi r_s^3, \tag{2.25}$$

and

$$\gamma = -0.14230, \quad \beta_1 = 1.05290, \quad \beta_2 = 0.3334. \tag{2.26}$$

Since we are neglecting nonlocal effects, one might expect that the LDA is only valid for systems with slowly varying densities. Yet, even for highly ionic systems like, *e.g.*, an Li_2O crystal, very good results are obtained [23]. In general, quantities derived by comparing total energies, like ground state geometries, phonon frequencies, and moments of the density, are very well reproduced [24].

A natural extension of the LDA would be to add terms containing gradients of the electron density, as already suggested by Hohenberg and Kohn [2]. However, important sum rules are exactly satisfied in the LDA, while they are easily violated in more sophisticated approximations. A systematic improvement of the LDA is not as straightforward as it might appear at first sight [24].

2.2.4 Band Gap Problem

It is crucial to emphasize that the Kohn-Sham eigenvalues do not have any immediate physical relevance and are obtained in a mathematical scheme. Although the relative values of occupied KS-LDA-bands in semiconductors and insulators are often in not too bad agreement with experimental values, the band gaps are far too small by about 50% up to 100% [5, 6, 25, 26, 27, 28, 29] . The *band gap problem* is not an intrinsic feature of DFT nor of the LDA, but rather of the KS-formalism [30, 31].

The highest occupied DFT-KS-eigenvalue $E^{(N)}_{N,\mathrm{DFT}}$ obtained with the exact exchange-correlation potential in an (N)-electron system equals the true ionization potential of the system [32, 33]. For infinite systems, this is the chemical potential μ. The exact minimum band gap is defined by

$$\begin{aligned} E_{\mathrm{gap}} &\equiv E^{(N+1)}_{\mathrm{tot}} - E^{(N)}_{\mathrm{tot}} - \left(E^{(N)}_{\mathrm{tot}} - E^{(N-1)}_{\mathrm{tot}}\right) \\ &= E^{(N+1)}_{N+1,\mathrm{DFT}} - E^{(N)}_{N,\mathrm{DFT}} \end{aligned} \tag{2.27}$$

It is the difference in the chemical potentials of an $(N+1)$-electron system and an (N)-electron system. In the KS-formalism, however, one calculates

$$E_{\mathrm{gap,KS-DFT}} = E^{(N)}_{N+1,\mathrm{DFT}} - E^{(N)}_{N,\mathrm{DFT}}. \tag{2.28}$$

In fact,

$$E_{\mathrm{gap}} = E_{\mathrm{gap,KS-DFT}} + \Delta, \tag{2.29}$$

where

$$\Delta = V^{(N+1)}_{\mathrm{xc}}(\mathbf{r}) - V^{(N)}_{\mathrm{xc}}(\mathbf{r}). \tag{2.30}$$

The last equality follows from the observation that a discontinuity of order one obtained by adding an electron can only come from the exchange-correlation potential, which is not necessarily analytical in N, in contrast to the external and the Hartree potential. The non-interacting KS-electrons move in a DFT effective potential changed by the addition of one electron.

One might assume that Δ would be small and that the local density approximation would be the main cause of the KS-LDA band gap error. However, in the framework of many-particle theory it is possible to obtain the exact DFT exchange correlation potential in a closed set of equations [6, 30, 34]. There is evidence in the cases of Si, GaAs and AlAs that the discontinuity Δ is responsible for about 80% of the total error [6]. In section 2.3 we will develop the quasiparticle formalism, which gives access to realistic particle states and energies.

2.2.5 *Ab Initio* Implementation

In this thesis we use a plane wave Car-Parrinello code [35] for the ground state calculations. It starts from first principles, *i.e.*, no models are applied to the Hamiltonian in Eq. (2.14) besides two approximations: the local density approximation

(see section 2.2.3) and the pseudopotential method, which only includes the valence electrons in the total energy minimization scheme and is described in Appendix A.

Plane Wave Basis

The simplest representation of Bloch wavefunctions in periodic boundary conditions is given by plane waves,

$$\varphi_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{\Omega_0}} e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{\mathbf{G}} c_{n\mathbf{k}}(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}} = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}), \tag{2.31}$$

where $u_{n\mathbf{k}}(\mathbf{r})$ is periodic. The system is represented by a supercell of volume Ω_0 with a basis of atoms at positions $\boldsymbol{\tau}_i$, which are translated according to the group properties of the crystal by the lattice vectors $\mathbf{R}_i$. The vectors $\mathbf{k}$ lie within the first Brillouin zone (BZ), while the $\mathbf{G}$ are reciprocal lattice vectors.

Some distinct advantages of a plane wave basis compared to a localized one are:

- They form a complete and orthogonal basis set, in which most expressions of the Hamiltonian in Eq. (2.14) are analytic.
- The reciprocal space formalism is directly applicable and Fast Fourier Transformations (FFT) are applied without further complications.
- *Exact and smooth* convergence of the total energy calculations can be achieved by gradually increasing the *cutoff energy* E_{cutoff} to include additional plane waves,

  $$\frac{1}{2} |\mathbf{k} + \mathbf{G}|^2 \leq E_{\text{cutoff}}. \tag{2.32}$$

  The number of plane waves N_{PW} scales like

  $$N_{\text{PW}} \propto \Omega_0 \left(E_{\text{cutoff}}\right)^{3/2}. \tag{2.33}$$

- For open systems with some "vacuum", like clusters, surfaces or vacancies, the basis set is independent of the atomic positions and therefore does not influence the electronic orbitals.

Localized basis sets (see, *e.g.*, Ref. [36]) on the other hand, are considerably smaller for systems where highly localized electronic wavefunctions have to be considered in the chemical binding and consequently as valence electrons in the pseudopotential formalism (see Appendix A). Also "vacuum" can be treated more efficiently. Especially for the more involved quasiparticle calculations (see section 2.3) this can be of a great advantage [37].

However, in this work we prefer a plane wave basis, since the above mentioned advantages are especially relevant when developing new methods – in our case the inclusion of excitonic effects in optical spectra. The disadvantages are not relevant for the systems – primarily bulk silicon and diamond – treated here.

Special k Points and Symmetry Properties

All sums over the Brillouin zone are integrals in the limit of an infinite crystal. Thus, the density Eq. (2.15),

$$n(\mathbf{r}) = \frac{1}{N_\mathbf{k}} \sum_n^{\text{occ}} \sum_\mathbf{k}^{\text{BZ}} |\varphi_{n\mathbf{k}}(\mathbf{r})|^2 , \tag{2.34}$$

is indeed

$$n(\mathbf{r}) = \frac{\Omega_0}{(2\pi)^3} \int_{\text{BZ}} d\mathbf{k} \sum_n^{\text{occ}} |\varphi_{n\mathbf{k}}(\mathbf{r})|^2 . \tag{2.35}$$

For practical purposes the integral is, of course, performed as a summation and has in addition to be restricted to a carefully chosen set of so-called *special* $\mathbf{k}$ points that give a good average of the integral over the BZ with volume $V_{\text{BZ}} = (2\pi)^3 / \Omega_0$. Often used methods have been developed by Chadi and Cohen [38] or Monkhorst and Pack [39]. The volume of the crystal is then $\Omega = N_\mathbf{k} \Omega_0$, with $N_\mathbf{k}$ being the number of $\mathbf{k}$ points in the sampling. The above used relation between a sum and an integral over the Brillouin zone,

$$\frac{1}{(2\pi)^3} \int_{\text{BZ}} d\mathbf{k}\, f(\mathbf{k}) = \frac{1}{\Omega} \sum_\mathbf{k}^{\text{BZ}} f(\mathbf{k}), \tag{2.36}$$

is often very useful.

The symmetry properties of the crystal are reflected by the Hamiltonian and therefore also by the wavefunctions [40]. We apply a Seitz operator $\{\mathsf{p} \mid \mathbf{w}\}$ of the *space group* as

$$\{\mathsf{p} \mid \mathbf{w}\}\, \mathbf{r} \equiv w\mathbf{pr} = w(\mathbf{pr}) = \mathbf{pr} + \mathbf{w}, \tag{2.37}$$

where p is a *point group* operation and $\mathbf{w}$ is a translation vector, which is a Bravais lattice vector $\mathbf{R}$ or some fraction of it. By definition, the point group is the subspace with $\mathbf{w} = \mathbf{0}$. The inverse is given by

$$\{\mathsf{p} \mid \mathbf{w}\}^{-1} = \{\mathsf{p}^{-1} \mid -\mathsf{p}^{-1}\mathbf{w}\} . \tag{2.38}$$

Then the operation on a Bloch function transforms as

$$\{\mathsf{p} \mid \mathbf{w}\}\, \varphi_{n\mathbf{k}}(\mathbf{r}) = \varphi_{n\mathbf{k}}(\mathsf{p}^{-1}\mathbf{r} - \mathsf{p}^{-1}\mathbf{w}) = \varphi_{n\mathsf{p}\mathbf{k}}(\mathbf{r}), \tag{2.39}$$

while the corresponding energies are simply related by

$$\epsilon_{n\mathbf{k}} = \epsilon_{n\mathsf{p}\mathbf{k}}. \tag{2.40}$$

For non-degenerate wavefunctions this is true up to a phase factor $e^{i\theta}$, which can be chosen to be one. However, in the case of *degeneracy*, the space group operation leads to a linear combination,

$$\{\mathsf{p} \mid \mathbf{w}\}\, \varphi_{n\mathbf{k}}(\mathbf{r}) = \sum_i^{\text{deg}} \alpha_i \varphi_{i\mathsf{p}\mathbf{k}}(\mathbf{r}), \tag{2.41}$$

of all degenerate bands i with wavevector $\mathbf{pk}$. The α_i are in general complex.

In any case we can reduce the $\mathbf{k}$ points in the full BZ to a corresponding *irreducible* part (IBZ) by taking a set of only those points $\mathbf{k}_{\mathrm{IBZ}}$, which cannot be related to each other by any operation of the type $\mathbf{pk}_{\mathrm{IBZ}} + \mathbf{G}_0$, with $\mathbf{G}_0$ being a reciprocal lattice vector. The set of all mutually inequivalent $\mathbf{k}$ points, expanded from a chosen $\mathbf{k}_{\mathrm{IBZ}}$ point in the IBZ into the full BZ, is called a *star*.

In general, by carefully and consistently choosing the phases of the expanded wavefunctions, we can reduce the sum of a complex valued function f over the full BZ to a sum over only the IBZ and the corresponding inequivalent point group operations D, depending on $\mathbf{k}_{\mathrm{IBZ}}$:

$$\sum_{\mathbf{k}}^{\mathrm{BZ}} f(\mathbf{k}) \rightarrow \sum_{\mathbf{k}_{\mathrm{IBZ}}} \sum_{\mathsf{D}_{\mathbf{k}_{\mathrm{IBZ}}}}^{\mathrm{inequ.}} f(\mathsf{D}_{\mathbf{k}_{\mathrm{IBZ}}} \mathbf{k}_{\mathrm{IBZ}}). \tag{2.42}$$

In the DFT-LDA calculations as well as in the other programs used in this thesis, we only include rotations R in the point group operations and the utilized space groups are symmorphic: $\{\mathsf{R} \mid \mathbf{0}\}$, *i.e.*, they only include the rotations, and no translations $\mathbf{w}$ are applied. However, we explicitly use *time reversal* to get the inversion,

$$\varphi_{n-\mathbf{k}}(\mathbf{r}) = \varphi^*_{n\mathbf{k}}(\mathbf{r}), \tag{2.43}$$

which is a property of quantum mechanics and thus independent of crystal symmetry.

Steepest Descent Method

The method of Car and Parrinello [35] allows realistic *ab initio* molecular dynamics calculations. In this thesis we do not move the ions, but only minimize the electronic configuration to find the ground state. The Car-Parrinello code allows to do so without diagonalizing the full Hamiltonian of the Kohn-Sham scheme Eqs. (2.17). This would include the occupied as well as the unoccupied states up to the number of plane waves N_{PW} chosen by the cutoff energy in Eq. (2.32), if we do not choose an iterative and often more cumbersome diagonalization scheme, which returns only the needed lower energetic eigenstates.

Instead, in the *"steepest descent"* procedure [41] we introduce a fictitious time variable t as a parameter in order to label different configurations in the space of one-particle wavefunctions $\{\varphi_i\}$. The Kohn-Sham energy functional $E[n]$ is, from Eq. (2.15), expressed as $E[\{\varphi_i\}]$ and represents a potential energy surface, on which the electrons move with a fictitious kinetic electronic energy: the φ_i are interpreted as fictitious classical coordinates. The ground state is then obtained by minimizing the total energy of the system whose evolution is described by,

$$\dot{\varphi}_i(\mathbf{r}, t) = -\frac{\delta E}{\delta \varphi^*_i(\mathbf{r}, t)} + \sum_j^{\mathrm{occ}} \epsilon_{ij} \varphi_j(\mathbf{r}, t)$$

$$= -\hat{h}_{\mathrm{KS}}^{\mathrm{eff}}\varphi_i(\mathbf{r},t) + \sum_j^{\mathrm{occ}} \epsilon_{ij}\varphi_j(\mathbf{r},t), \tag{2.44}$$

where $\hat{h}_{\mathrm{KS}}^{\mathrm{eff}}$ is the Kohn-Sham Hamilton operator in Eqs. (2.17). A constraint is included to guarantee orthonormal one-particle wavefunctions.

If $\dot{\varphi}_i = 0$, *i.e.*, in the equilibrium case, we obtain

$$\hat{h}_{\mathrm{KS}}^{\mathrm{eff}}\varphi_i(\mathbf{r},t) = \sum_j^{\mathrm{occ}} \epsilon_{ij}\varphi_j(\mathbf{r},t), \tag{2.45}$$

which are, apart from a unitary transformation, just the Kohn-Sham equations (2.17). Thus, the KS-eigenvalues and wavefunctions can easily be obtained after diagonalization of ϵ_{ij}, which is a small matrix with respect to the number of basis functions used, since only the occupied states are calculated. Experience has shown that for fixed ionic configuration the energy surface E has only one minimum.

In practice, the equations (2.44) have to be numerically integrated. The steepest descent algorithm for a first order minimization reads

$$\bar{\varphi}_i(\mathbf{r},t+\triangle t) = \varphi_i(\mathbf{r},t) - \triangle t\ \hat{h}_{\mathrm{KS}}^{\mathrm{eff}}\varphi_i(\mathbf{r},t) + O(\triangle t)^2. \tag{2.46}$$

The initial guess, $\varphi_i(\mathbf{r}, t = 0)$, must be non-orthogonal to the ground state for the method to work correctly. The time step $\triangle t$ fixes the time scale and governs the convergence of the minimization scheme. The orthogonalization method can be quite arbitrarily chosen and a simple Gram-Schmidt procedure is sufficient. Self-consistency is automatically built-in, as the density and the potential are updated at each time step.

The actual, dynamical Car-Parrinello method starts from a Lagrangian including a fictitious kinetic electronic energy as well as the kinetic ionic energy and leads to a system of second order differential equations of motion. This permits efficient numeric procedures and, most important, ensures that on average the electrons stay close to the Born-Oppenheimer energy surface given by the current ionic configuration. Thus, the system does not heat up due to the fictitious kinetic electronic energy necessary for the minimization scheme.

2.2.6 Dielectric Function

The response of the charge distribution in the system to an external perturbation is measured by the polarizability [42]. Here we will emphasize the formulation in the density-functional approach (see section 2.2). From the DFT polarizability we obtain the inverse DFT dielectric function, which is an important ingredient in the quasiparticle and excitonic calculations performed in the later course of this thesis.

Definitions

In a perturbation approach [43, 44] the *independent-particle polarization* χ^0 relates a change in the electron density δn to a change in the total effective potential δV_{tot},

$$\delta V_{\text{tot}}(x) = \delta w(x) + \delta V_{\text{scr}}(x), \tag{2.47}$$

by

$$\delta n(x) = \int dx'\ \chi^0(x,x')\delta V_{\text{tot}}(x'). \tag{2.48}$$

The external potential w is written including a time-dependence in this section in order to be general. Here we have used

$$\delta V_{\text{scr}}(\mathbf{x},t) = \int d\mathbf{x}'\ \frac{\delta n(\mathbf{x}',t)}{|\mathbf{x}-\mathbf{x}'|} + \int dx'\ K_{\text{xc}}(x,x')\delta n(x'), \tag{2.49}$$

with

$$K_{\text{xc}}(x,x') \equiv \frac{\delta^2 E_{\text{xc}}[n]}{\delta n(x)\delta n(x')}. \tag{2.50}$$

The first term in Eq. (2.49) gives a change in the Hartree potential and the second one a change in the exchange-correlation potential.

In contrast, the *full polarizability* χ is defined via a change in the external potential w only,

$$\delta n(x) = \int dx'\ \chi(x,x')\delta w(x'). \tag{2.51}$$

The two polarizabilities are connected in the DFT framework by

$$\chi = \left[1 - \chi^0(v + K_{\text{xc}})\right]^{-1}\chi^0, \tag{2.52}$$

where we use the notation $v(x,x') \equiv v(\mathbf{r},\mathbf{r}')\delta(t-t')$. Matrix notation is used, so a matrix inversion is required in Eq. (2.52).

We want an explicit formulation for the polarizability. For a time-independent Hamiltonian like in Eq. (2.2) the polarizability depends only on the time difference $\tau = t - t'$. Then it is convenient to Fourier-transform χ into energy-space according to

$$\chi(\mathbf{x},\mathbf{x}',\omega) = \int d\tau\ e^{i\omega\tau}\chi(\mathbf{x},\mathbf{x}',\tau). \tag{2.53}$$

Application of first-order perturbation theory to the Kohn-Sham equations (2.17) yields the standard result for the independent-particle polarization χ^0 in the Adler-Wiser formulation [45, 46],

$$\chi^0(\mathbf{x},\mathbf{x}',\omega) = 2\sum_{i,j}(f_i - f_j)\,\frac{\varphi_i(\mathbf{x})\varphi_j^*(\mathbf{x})\ \varphi_i^*(\mathbf{x}')\varphi_j(\mathbf{x}')}{\epsilon_i - \epsilon_j - \omega - i\eta}, \tag{2.54}$$

with f_i being the occupation numbers $(0, 1)$ and η a positive infinitesimal. Spin has been included in the summation.

Finally, the *inverse dielectric function* measures the screening in the system and is defined by

$$\varepsilon^{-1}(x, x') \equiv \frac{\delta V_{\text{tot}}(x)}{\delta w(x')}. \tag{2.55}$$

In practice we need the inverse dielectric function in momentum space, which is related by a Fourier transform to real space,

$$\varepsilon^{-1}_{\sigma\sigma'}(\mathbf{r}, \mathbf{r}', \omega) = \frac{1}{(2\pi)^3} \int_{\text{BZ}} d\mathbf{q} \sum_{\mathbf{G},\mathbf{G}'} e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} \, \varepsilon^{-1}_{\substack{\mathbf{G}\mathbf{G}' \\ \sigma\sigma'}}(\mathbf{q}, \omega) \, e^{-i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'}. \tag{2.56}$$

The integration runs over the Brillouin zone. The inhomogeneity of the charge distribution, which gives rise to the *local-field effects,* is included via the reciprocal lattice vectors $\mathbf{G}$ and $\mathbf{G}'$.

For later use we also give the explicit Fourier transformation of the independent particle polarization χ^0,

$$\chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega) = 2 \sum_{i,j} (f_i - f_j) \frac{\left\langle j \left| e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} \right| i \right\rangle \left\langle i \left| e^{i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'} \right| j \right\rangle}{\epsilon_i - \epsilon_j - \omega - i\eta}. \tag{2.57}$$

The inverse dielectric function is connected to the polarizability by

$$\varepsilon^{-1}(x, x') = \delta(x - x') + \int dx_1 \, v(x, x_1)\chi(x_1, x'). \tag{2.58}$$

Care must be taken of the origin of the induced screening. If the probe is assumed to be a test particle, it is only affected by the electrostatic Hartree term in Eq. (2.49). Then the widely used *random phase approximation* (RPA) response function is obtained, setting the exchange-correlation contributions to χ to zero,

$$\varepsilon^{-1}_{\text{RPA}} = 1 + v(1 - \chi^0 v)^{-1}\chi^0, \tag{2.59a}$$

$$\varepsilon_{\text{RPA}} = 1 - v\chi^0. \tag{2.59b}$$

For later use we note that $(1 - \chi^0 v)^{-1}\chi^0 = \chi^0(1 - v\chi^0)^{-1}$.

Because of the screening, the electrons do not directly interact via the Coulomb interaction v, but via a *dynamically screened interaction* W, given by

$$W(\mathbf{x}, t, \mathbf{x}', t') \equiv \int d\mathbf{x}_1 dt_1 \, v(\mathbf{x}, t, \mathbf{x}_1, t_1)\varepsilon^{-1}(\mathbf{x}_1, t_1, \mathbf{x}', t'). \tag{2.60}$$

This interaction function W gives the electronic screening in the GW and excitonic calculations (see sections 2.3.2 and 3.1.3).

Plasmon-Pole Model

The computational effort to obtain the full dynamical inverse dielectric function is considerable and often not necessary. It is sometimes possible to make some model assumptions about the frequency dependence of the elements of $\varepsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega)$.

Realistic calculations of the response function show that $\operatorname{Im}\varepsilon^{-1}$ is generally a peaked function in ω. Well-defined peaks correspond to a plasmon excitation of the system associated with an absorption resonance. These excitations are fixed at frequencies $\tilde{\omega}$ in the model [5, 47]. We include only one pole for each matrix element. More plasmon frequencies would increase the accuracy of the model, but also the computational cost.

In the plasmon-pole model the needed *time-ordered* or advanced inverse dielectric function is described by, with ω, Ω^2 and $\tilde{\omega}$ being real,

$$\varepsilon_T^{-1} = 1 + \frac{\Omega^2}{(\omega + i\eta - \tilde{\omega})(\omega - i\eta + \tilde{\omega})} = 1 + \frac{\Omega^2}{\omega^2 - (\tilde{\omega} - i\eta)^2}, \tag{2.61a}$$

$$\operatorname{Re}\varepsilon_T^{-1} = 1 + \frac{\Omega^2}{\omega^2 - \tilde{\omega}^2}, \tag{2.61b}$$

$$\operatorname{Im}\varepsilon_T^{-1} = -\frac{\pi\Omega^2}{2\omega}\left[\delta(\omega - \tilde{\omega}) + \delta(\omega + \tilde{\omega})\right]. \tag{2.61c}$$

In contrast, we note that the *causal* or retarded inverse dielectric function is given by,

$$\varepsilon_R^{-1} = 1 + \frac{\Omega^2}{(\omega + i\eta - \tilde{\omega})(\omega + i\eta + \tilde{\omega})} = 1 + \frac{\Omega^2}{(\omega + i\eta)^2 - \tilde{\omega}^2}, \tag{2.62a}$$

$$\operatorname{Re}\varepsilon_T^{-1} = 1 + \frac{\Omega^2}{\omega^2 - \tilde{\omega}^2}, \tag{2.62b}$$

$$\operatorname{Im}\varepsilon_T^{-1} = -\frac{\pi\Omega^2}{2\omega}\left[\delta(\omega - \tilde{\omega}) - \delta(\omega + \tilde{\omega})\right], \tag{2.62c}$$

where the imaginary part is even instead of odd in ω.

Generally, the time-ordered or advanced dielectric matrix is related to the retarded or causal one by a changed sign of the imaginary part for negative frequencies. For positive frequencies, they are the same.

Here we use the plasmon-pole model of Godby and Needs [47], where the full frequency dependent inverse dielectric matrix elements are fitted at two imaginary frequencies in the energy range important for further calculations. Less relevant higher energies, which may degrade the model, are not considered. Away from the real axis we have

$$\varepsilon_T^{-1} = \varepsilon_R^{-1} = 1 + \frac{\Omega^2}{\omega^2 - \tilde{\omega}^2}, \tag{2.63}$$

which is a well-behaved function of the now complex ω.

Therefore, the plasmon-pole model is defined by, omitting spin dependence,

$$\varepsilon^{-1}_{\mathbf{GG'}}(\mathbf{q},\omega) = \delta_{\mathbf{GG'}} + \frac{\Omega^2_{\mathbf{GG'}}(\mathbf{q})}{\omega^2 - \tilde{\omega}^2_{\mathbf{GG'}}(\mathbf{q})}. \tag{2.64}$$

Two finite frequencies at $\omega_1 = 0$ and $\omega_2 = iE_0$ (with $E_0 \approx \frac{1}{2}$ a.u.) determine the two fit parameters.

This model has been checked at several finite frequencies for very different systems, from bulk silicon to a sodium tetramer [48]: the fit turned out to be very good.

2.3 Quasiparticle Formulation

The KS-DFT scheme in connection with the LDA provides a highly successful way to calculate the ground state properties of a long range of materials, from molecules over clusters to metals, semiconductors and insulators. For electronic structure calculations, however, one often goes beyond the DFT-LDA in order to obtain the "true" levels of the electrons in the solid, which are referred to as the *quasiparticle* (QP) energies. They are given by the energies needed to add or subtract an electron to or from the system and are obtained by spectroscopies like, for example, inverse photoemission and photoemission experiments.

The Kohn-Sham eigenvalues have often failed to deliver physically meaningful gaps. The case of germanium, *e.g.*, is quite dramatic: inclusion of relativistic corrections in the DFT-LDA calculations makes the band gap (experimental value: 0.744 eV) to vanish and leads to a metal [49]. The discrepancies in the details of the band dispersion are typically smaller, but also material dependent: the bandwidth of Ge is well represented by the LDA eigenvalues [50, 51], but for diamond, it is underestimated as compared to X-ray spectra [52].

In the following we will lay out the quasiparticle approach. Further information and more details can be found in many books on many-particle physics, like [14] or [53], and, in particular for solid state applications, in Hedin and Lundqvist [54].

2.3.1 Green's Functions

Like for density functional theory (section 2.2) the basic idea behind the Green's function is the observation that, in general, we do not need to know the detailed behavior of each particle in the system. Theoretically, a real (bare) particle, in our case an electron or a hole, in an interacting system perturbs the particles in its vicinity and moves through the system together with a cloud of surrounding particles. This dressed particle is referred to as a *quasiparticle*. The screening can largely change the properties and lead to, for example, an effective mass different from that of the bare particle and to a finite lifetime. Experiments tell us in fact

about these quasiparticles, as measurements introduce such elementary excitations into the system.

The one-particle Green's function describes the propagation of one quasiparticle through the system and contains important physical information such as the ground state energy, the momentum distribution, and – among other equilibrium properties – the energy and lifetime of quasiparticles. The two-particle Green's function contains information about the coupled behavior of two quasiparticles moving through the system. It has the full information of the one-particle Green's function and directly gives energies and lifetimes of collective excitations, as well as electrical conductivity and other non-equilibrium properties. Even higher-order Green's functions include the information of the lower-order ones in the same manner and contain in addition information about more complicated properties of the system.

Definition

The time-ordered *single-particle* Green's function of the system at zero temperature is defined as

$$G(x,x') \equiv -i \left\langle \Psi_N \left| T\left[\psi_H(x)\psi_H^\dagger(x')\right] \right| \Psi_N \right\rangle , \tag{2.65}$$

where $|\Psi_N\rangle$ is the normalized Heisenberg (time-independent) ground state vector of the interacting N-electron system (see section 2.2.1), satisfying the Schrödinger equation Eq. (2.3),

$$\hat{H}\,|\Psi_N\rangle = E\,|\Psi_N\rangle , \tag{2.66}$$

and E is the ground state energy of the system. As in section 2.1 we abbreviate $x \equiv (\mathbf{x},t) \equiv (\mathbf{r},\sigma,t)$. T is the Wick time-ordering operator [55], which orders an arbitrary number of operators from the right to the left in ascending time order and adds a factor $(-1)^P$, where P is the number of interchanges of fermion operators from the originally given order. Explicitly,

$$T\left[\psi_H(\mathbf{x},t)\psi_H^\dagger(\mathbf{x}',t')\right] = \begin{cases} \psi_H(\mathbf{x},t)\psi_H^\dagger(\mathbf{x}',t'), & if \quad t > t' \\ -\psi_H^\dagger(\mathbf{x}',t')\psi_H(\mathbf{x},t), & if \quad t < t' \end{cases} . \tag{2.67}$$

Relation to Observables

The operator $\psi_H^\dagger(\mathbf{r}',\sigma',t')$, acting on the N-particle ground state, creates an electron with spin σ' at the point $\mathbf{r}'$ and at the time t'. From the definition of G in Eq. (2.65) we see that for $t > t'$, the Green's function is the probability amplitude that we find that particle in the now $(N+1)$-electron system later at the time t at the point $\mathbf{r}$ and with spin σ, without any other modification of the ground state $|\Psi_N\rangle$. Correspondingly, for $t < t'$, G describes the propagation of a hole in an $(N-1)$-electron system. The Green's function is thus a propagator.

With the help of the Green's function we can calculate:

1. the ground state expectation value of any single-particle operator,

2. the ground state energy of the system,

3. certain excitation energies of the system.

The first point is readily obtained considering the one-particle operator

$$\hat{A} = \int d\mathbf{r} \sum_{\sigma\sigma'} \psi_H^\dagger(\mathbf{r},\sigma',t) A_{\sigma'\sigma}(\mathbf{r}) \psi_H(\mathbf{r},\sigma,t), \tag{2.68}$$

which we take to be local in space coordinates for ease. The ground state expectation value of this operator is given by

$$\begin{aligned} \left\langle \hat{A} \right\rangle &= \left\langle \Psi_N \left| \hat{A} \right| \Psi_N \right\rangle = -i \int d\mathbf{r} \lim_{t'\to t^+} \lim_{\mathbf{r}'\to\mathbf{r}} \left[\sum_{\sigma\sigma'} A_{\sigma'\sigma}(\mathbf{r}) G_{\sigma\sigma'}(\mathbf{x}t, \mathbf{x}'t') \right] \\ &= -i \int d\mathbf{r} \lim_{\mathbf{r}'\to\mathbf{r}} \mathrm{Tr}\left[A(\mathbf{r}) G(\mathbf{r}t, \mathbf{r}'t') \right], \end{aligned} \tag{2.69}$$

where t^+ denotes a time infinitesimally later than t. The limit has to be taken after the action of $A_{\sigma'\sigma}(\mathbf{r})$ in order to correctly treat differential operators.

The second point was first proven by Galitskii and Migdal [56] . They derived the following expression for the ground state energy

$$\begin{aligned} E &= \left\langle \hat{T} + \hat{W} + \hat{V} \right\rangle \\ &= -\frac{i}{2} \int d\mathbf{r} \lim_{t'\to t^+} \lim_{\mathbf{r}'\to\mathbf{r}} \left(i\frac{\partial}{\partial t} - \frac{\nabla_{\mathbf{r}}^2}{2} + w(\mathbf{x}) \right) \mathrm{Tr}\left[G(\mathbf{x}t, \mathbf{x}'t') \right]. \end{aligned} \tag{2.70}$$

The third point is the main interest here. Following Ref. [54] we rewrite Eq. (2.65),

$$\begin{aligned} & G(\mathbf{x},t,\mathbf{x}',t') = \\ = & -i \left\langle \Psi_N \left| \psi(\mathbf{x}) \exp\left[-i\left(H - E_N^{(N)} \right)(t-t') \right] \psi^\dagger(\mathbf{x}') \right| \Psi_N \right\rangle \Theta(t-t') \\ & +i \left\langle \Psi_N \left| \psi^\dagger(\mathbf{x}') \exp\left[-i\left(H - E_N^{(N)} \right)(t-t') \right] \psi(\mathbf{x}) \right| \Psi_N \right\rangle \Theta(t'-t), \end{aligned} \tag{2.71}$$

where we have used Eq. (2.8) and the notation of Eq. (2.27) together with

$$\Theta(t) = 1, \qquad t > 0, \tag{2.72a}$$

$$= 0, \qquad t < 0. \tag{2.72b}$$

Next we introduce the complete set of the eigenstates of H for the $(N+1)$ and $(N-1)$ particle system. The quantum labels of theses states are numbered by s

and we set $\tau = t - t'$ for a time-independent Hamiltonian. The limiting energy for injection of electrons respective holes is the chemical potential μ. Thus we obtain

$$\begin{aligned} G(\mathbf{x},\mathbf{x}',\tau) &= -i\sum_s f_s(\mathbf{x}) f_s^*(\mathbf{x}')\, e^{-i\epsilon_s} \\ &\quad \times \left[\Theta(\tau)\Theta(\epsilon_s-\mu) - \Theta(-\tau)\Theta(\mu-\epsilon_s)\right], \end{aligned} \tag{2.73}$$

with

$$\begin{cases} f_s(\mathbf{x}) = \langle \Psi_N \left|\psi(\mathbf{x})\right| \Psi_{N+1,s}\rangle, \\ \epsilon_s = E_s^{(N+1)} - E_N^{(N)}, \end{cases} \quad \text{for } \epsilon_s \geq \mu, \tag{2.74a}$$

$$\begin{cases} f_s(\mathbf{x}) = \langle \Psi_{N-1,s} \left|\psi(\mathbf{x})\right| \Psi_N\rangle, \\ \epsilon_s = E_N^{(N)} - E_s^{(N-1)}, \end{cases} \quad \text{for } \epsilon_s < \mu. \tag{2.74b}$$

The energies ϵ_s are differences of total energies and describe quasiparticle excitation energies. A Fourier transform of Eq. (2.73) to energy space leads to

$$\begin{aligned} G(\mathbf{x},\mathbf{x}',\omega) &= \int_{-\infty}^{\infty} d\tau\; G(\mathbf{x},\mathbf{x}',\tau)\, e^{i\omega\tau} \\ &= \sum_s f_s(\mathbf{x}) f_s^*(\mathbf{x}') \left\{ -i\Theta(\epsilon_s-\mu) \int_0^{\infty} d\tau\; e^{i(\omega-\epsilon_s)\tau} \right. \\ &\quad \left. + i\Theta(\mu-\epsilon_s) \int_{-\infty}^{0} d\tau\; e^{i(\omega-\epsilon_s)\tau} \right\}. \end{aligned} \tag{2.75}$$

Finally we obtain

$$G(\mathbf{x},\mathbf{x}',\omega) = \sum_s \frac{f_s(\mathbf{x}) f_s^*(\mathbf{x}')}{\omega - \epsilon_s + i\eta\,\mathrm{sgn}(\epsilon_s - \mu)}, \tag{2.76}$$

where η is a positive infinitesimal.

The amplitudes $f_s(\mathbf{x})$ are not normalized to unity and are not linearly independent. However, they fulfill the completeness relation,

$$\sum_s f_s(\mathbf{x}) f_s^*(\mathbf{x}') = \left\langle \Psi_N \left| \psi(\mathbf{x})\psi^\dagger(\mathbf{x}') + \psi^\dagger(\mathbf{x}')\psi(\mathbf{x}) \right| \Psi_N \right\rangle = \delta(\mathbf{x}-\mathbf{x}'). \tag{2.77}$$

2.3.2 Self-Energy Concept

The difference in energy between a bare particle and the corresponding quasiparticle is obtained from the *self-energy* operator. The bare particle interacts with the many-particle system and drags a cloud of surrounding particles around it while moving through the system. Thus the particle in some way interacts with itself via the many-particle system, and changes its own energy.

Equation of Motion for the Green's Function

From the equation of motion for the field operators in the Heisenberg representation,

$$i\frac{\partial \psi_H(t)}{\partial t} = \left[\psi_H(t), \hat{H}\right], \tag{2.78}$$

we obtain the equation of motion for the Green's function

$$\left[i\frac{\partial}{\partial t} + \frac{\nabla_{\mathbf{r}}^2}{2} - w(\mathbf{x})\right] G(\mathbf{x}t, \mathbf{x}'t') =$$

$$\delta(\mathbf{x}-\mathbf{x}')\delta(t-t') - i\int d\mathbf{x}_1\, v(\mathbf{r}, \mathbf{r}_1) G_2(\mathbf{x}t, \mathbf{x}_1 t, \mathbf{x}_1 t^+, \mathbf{x}'t'). \tag{2.79}$$

The *two-particle* Green's function G_2 is defined as

$$G_2(x_1, x_2, x_3, x_4) \equiv -\left\langle \Psi_N \left| T\left[\psi_H(x_1)\psi_H(x_2)\psi_H^\dagger(x_3)\psi_H^\dagger(x_4)\right] \right| \Psi_N \right\rangle. \tag{2.80}$$

The physical interpretation of G_2 is similar to the one of G. The two-particle Green's function describes the propagation of two particles being added or removed from the system. The full knowledge of G_2 completely determines the one-particle Green's function G. In the same way the equation of motion of the two-particle Green's function involves the three-particle Green's function, which contains full information about the two-particle Green's function and so on.

In practice, such a chain would have to be terminated by decoupling correlations of higher order. Instead one introduces a nonlocal, time- (or energy-) dependent, non-Hermitian quantity, called the *self-energy* operator Σ, which takes into account the interaction between the particle and the rest of the system. In comparison with Eq. (2.79) we implicitly define Σ by

$$\left[i\frac{\partial}{\partial t} - \hat{h}_0(x)\right] G(x, x') \equiv \delta(x-x') + \int dx_1\, \Sigma(x, x_1) G(x_1, x'). \tag{2.81}$$

Here we also define

$$\hat{h}_0(x) \equiv -\frac{\nabla_{\mathbf{r}}^2}{2} + w(x) + V_{\mathrm{H}}(x), \tag{2.82}$$

with the Hartree potential V_{H}, which can be expressed, using many-body wavefunctions and operators, as

$$V_{\mathrm{H}}(\mathbf{x}, t) \equiv \int d\mathbf{x}'\, v(\mathbf{r}, \mathbf{r}') \left\langle \Psi_N \left| \psi_H^\dagger(\mathbf{x}', t)\psi_H(\mathbf{x}', t) \right| \Psi_N \right\rangle. \tag{2.83}$$

Effective Eigenvalue Equation

If the external potential is time-independent, we obtain in energy-space for Eq. (2.81)

$$\left[\omega - \hat{h}_0(\mathbf{x})\right] G(\mathbf{x}, \mathbf{x}', \omega) = \delta(\mathbf{x}-\mathbf{x}') + \int d\mathbf{x}_1\, \Sigma(\mathbf{x}, \mathbf{x}_1, \omega) G(\mathbf{x}_1, \mathbf{x}', \omega). \tag{2.84}$$

Expressing the Green's function G in terms of energy-dependent wavefunctions $\phi_i(\mathbf{x},\omega)$, which are assumed to be orthonormal and complete in each ω-subspace,

$$G(\mathbf{x},\mathbf{x}',\omega) = \sum_i \frac{\phi_i(\mathbf{x},\omega)\phi_i^*(\mathbf{x}',\omega)}{\omega - E_i(\omega)}, \tag{2.85}$$

we can fulfill Eq. (2.84) by further demanding the following *eigenvalue equations* for the complex energies $E_i(\omega)$ and wavefunctions $\phi_i(\mathbf{x},\omega)$,

$$\hat{h}_0(\mathbf{x})\phi_i(\mathbf{x},\omega) + \int d\mathbf{x}_1\ \Sigma(\mathbf{x},\mathbf{x}_1,\omega)\phi_i(\mathbf{x}_1,\omega) = E_i(\omega)\phi_i(\mathbf{x},\omega). \tag{2.86}$$

The structure of the equations (2.86) is the same as the one of the Kohn-Sham equations (2.17). Indeed, substitution of the self-energy Σ by the exchange-correlation potential V_{xc} in Eqs. (2.86) leads back to Eqs. (2.17). Setting $\Sigma = 0$ results in the elementary Hartree theory. Since the self-energy is a non-Hermitian operator, the eigenvalues $E_i(\omega)$ are in general complex and the imaginary part of $E_i(\omega)$ gives the lifetime of the quasiparticle.

For well-defined excitations, given by sharp peaks at energies $E_i(\omega) = \omega \equiv \epsilon_i$, the lifetime is quasi infinite and the resulting spectrum is real and independent of ω,

$$G(\mathbf{x},\mathbf{x}',\omega) = \sum_i \frac{\phi_i(\mathbf{x})\phi_i^*(\mathbf{x}')}{\omega - \epsilon_i + i\eta\,\mathrm{sgn}(\epsilon_i - \mu)}. \tag{2.87}$$

In the limiting case of independent particles, this form of the spectrum is exact and we use it in this work, since we are only considering sharp quasiparticle states around the Fermi level. Depending on the chemical potential μ, the poles are moved away from the real axis by a positive infinitesimal η.

We note the resemblance to Eq. (2.76), although the amplitudes $f_s(\mathbf{x})$ are of a much more complicated nature than the single-particle QP wavefunctions $\phi_i(\mathbf{x})$.

Spectral Weight Function

In general, the excitation spectrum is given by the *spectral weight function* A, which is related to the Green's function by

$$G(\mathbf{x},\mathbf{x}',\omega) = \int d\omega' \frac{A(\mathbf{x},\mathbf{x}',\omega)}{\omega - \omega' + i\eta\,\mathrm{sgn}(\omega' - \mu)}, \tag{2.88}$$

or also

$$A(\omega) = \frac{1}{\pi}\left|\mathrm{Im}\,G(\omega)\right|. \tag{2.89}$$

The spectral weight function A can always be chosen real except for some problems involving magnetic fields.

For sharp quasiparticle peaks like in Eq. (2.87), A is simply

$$A(\mathbf{x},\mathbf{x}',\omega) = \sum_i \phi_i(\mathbf{x})\phi_i^*(\mathbf{x}')\delta(\omega-\epsilon_i). \tag{2.90}$$

We can model a distinct quasiparticle state in the spectral weight function A by a Lorentzian,

$$A(\mathbf{x},\mathbf{x}',\omega) = \sum_i \phi_i(\mathbf{x})\phi_i^*(\mathbf{x}')\frac{\Gamma_i/\pi}{\Gamma_i^2+(\omega-\epsilon_i)^2}, \tag{2.91}$$

where ϵ_i is the real part of the complex quasiparticle energy $E_i(\omega)$ in Eq. (2.86). Γ_i is the imaginary part, corresponding to a particle decaying like $e^{-\Gamma_i t}$ in time space.

Even when the peak in the spectral weight function A is evident, not all of the spectral weight is contained in this peak. The proportion of the spectral weight under the quasiparticle peak is known as the *renormalization constant* Z. It is a measure of the strength of correlation effects, and it is related to the energy derivative of the matrix element of the self-energy operator by

$$Z_i = \left(1-\left(\frac{\partial\langle\phi_i|\,\Sigma(\omega)\,|\phi_i\rangle}{\partial\omega}\right)\Big|_{\omega=\varepsilon_i}\right)^{-1}. \tag{2.92}$$

A value near unity corresponds to a small associated screening cloud of the bare particle: the quasiparticle is a well-defined particle-like excitation with a long lifetime.

Hedin's Coupled Integral Equations

We have found that the self-energy is a functional of the one-particle Green's function, *i.e.*, $\Sigma = \Sigma[G]$ and that a knowledge of Σ determines G. However, we do not have a practical scheme in order to solve for the self-energy, starting from a reasonable guess for G. Hedin introduced a set of integral equations which fully determines Σ and G [4]. They cannot be exactly solved, but give a good starting point for various approximations for the self-energy.

Taking the functional derivative with respect to an artificial change δw in the external potential w, which is set to zero again at the end, Hedin derived a set of coupled equations [54],

$$W(x,x') = v(x,x') + \int dx_1 dx_2\, W(x,x_1)P(x_1,x_2)v(x_2,x'), \tag{2.93a}$$

$$P(x,x') = -i\int dx_1 dx_2\, G(x',x_1)G(x_2,x')\Gamma(x_1,x_2;x), \tag{2.93b}$$

$$\Sigma(x,x') = i\int dx_1 dx_2\, W(x^+,x_1)G(x,x_2)\Gamma(x_2,x';x_1), \tag{2.93c}$$

$$\Gamma(x,x';x'') = \delta(x-x')\delta(x-x'') +$$

$$\int dx_1 dx_2 dx_3 dx_4 \, \frac{\delta\Sigma(x,x')}{\delta G(x_1,x_2)} G(x_1,x_3)G(x_4,x_2)\Gamma(x_3,x_4;x''), \tag{2.93d}$$

where $x^+ \equiv (\mathbf{x}, t^+)$.

In the Green's function formalism one often uses the irreducible polarization propagators P and P_0 instead of the polarizabilities χ and χ^0 in Eqs. (2.51) and (2.48). The propagator P is defined by a functional derivative,

$$P(x,x') = i \int dx_1 dx_2 \; G(x',x_1)G(x_2,x') \frac{\delta G^{-1}(x_1,x_2)}{\delta V_{\text{tot}}(x)}. \tag{2.94}$$

P is related to ε by

$$\varepsilon(x,x') = \delta(x-x') - \int dx_1 \; v(x,x_1)P(x_1,x'). \tag{2.95}$$

We remind the relation between χ and ε^{-1} in Eq. (2.58). Thus P and χ are generally not equal. We only have $P_0 = \chi^0$.

From the equation of motion for the Green's function Eq. (2.79), we get a functional derivative expression for the self-energy Σ,

$$\Sigma(x,x') = -i \int dx_1 dx_2 \; v(x^+,x_1)G(x,x_2) \frac{\delta G^{-1}(x_2,x')}{\delta w(x_1)}. \tag{2.96}$$

The *vertex function* Γ is defined by

$$\Gamma(x,x';x'') \equiv -\frac{\delta G^{-1}(x,x')}{\delta V_{\text{tot}}(x'')} = \delta(x-x')\delta(x-x'') + \frac{\delta\Sigma(x,x')}{\delta V_{\text{tot}}(x'')}. \tag{2.97}$$

From Eq. (2.93b) we see that the structure of the polarization propagator P is $P = GG\Gamma$, thus describing the creation of an electron and a hole, which are independent, if $\Gamma = \delta$.

GW Approximation

Hedin proposed an iterative solution of the equations, starting with $\Sigma = 0$, *i.e.*, the Hartree approximation. This gives a simple expression for the vertex Γ,

$$\Gamma(x,x';x'') = \delta(x-x')\delta(x-x''). \tag{2.98}$$

The other functional expressions of Eqs. (2.93) are then in first iteration,

$$\Sigma(x,x') = iG(x,x')W(x^+,x'), \tag{2.99a}$$

$$P_0(x,x') = -iG(x,x')G(x',x). \tag{2.99b}$$

This is the so-called *GW approximation*, where vertex corrections are neglected. It is on the same level of approximation as the random-phase approximation (RPA) for the inverse dielectric matrix in Eq. (2.59a). Note that Eq. (2.99b) for P_0 in connection with Eq. (2.87) also gives the Adler-Wiser formulation for the independent polarizability χ^0 in Eq. (2.54).

In a second iteration we would have to calculate the derivation $\delta\Sigma/\delta G$ and obtain terms of the order W^2 in Σ. However, the expressions quickly become very complicated and are usually not considered. Since screening in solids is relatively strong, which leads to a dielectric matrix significantly different from one, a first-order expansion of Σ in powers of W is already quite good, especially compared to an expansion in powers of the unscreened bare Coulomb interaction v.

Calculation of *GW* Corrections

In the DFT-LDA calculations we separate the core from the valence electron contribution in the context of the pseudopotential formalism (see Appendix A). In the same spirit, we split the Green's function,

$$G = G_c + G_v, \tag{2.100}$$

and the polarizability,

$$P = P_c + P_v, \tag{2.101}$$

leading to, from Eqs. (2.58) and (2.60) [54],

$$\Sigma = iG_cW + iG_vW_vP_cW_v + iG_vW_v. \tag{2.102}$$

The first two terms are generally small and have already been approximately incorporated on the DFT level [5]. Σ will only refer to the last term, describing the valence-valence electron interaction.

Since W and G implicitly depend on Σ, it is essential to make a good initial approximation for the *GW* calculation in order to avoid further iterations according to Eqs. (2.93). In fact, the Kohn-Sham wavefunctions have been shown to match very closely the exact quasiparticle wavefunctions for many materials, like diamond, Si, Ge or LiCl [5]:

$$\varphi_i^{\mathrm{KS}}(\mathbf{r}) \approx 0.999\ \phi_i^{\mathrm{QP}}(\mathbf{r}). \tag{2.103}$$

Therefore, the φ_i^{KS} are used to calculate the *ab initio* dielectric matrix as well as to evaluate the desired matrix elements for Σ.

In particular, we approximate G in Eq. (2.87) and χ^0 in Eq. (2.59a) or (2.99b) together with Eq. (2.54), using the KS-LDA wavefunctions and eigenvalues, and evaluate the self-energy corrections in first order perturbation theory,

$$\epsilon_i^{\mathrm{QP}} = \epsilon_i^{\mathrm{KS}} + \left\langle \varphi_i^{\mathrm{KS}} \middle| \Sigma(\varepsilon_i^{\mathrm{QP}}) - V_{\mathrm{xc}}^{\mathrm{LDA}} \middle| \varphi_i^{\mathrm{KS}} \right\rangle + O\left(\left(\Sigma(\epsilon_i^{\mathrm{KS}}) - V_{\mathrm{xc}}^{\mathrm{LDA}} \right)^2 \right), \tag{2.104}$$

where Σ must be computed self-consistently at the energy $\varepsilon_i^{\mathrm{QP}}$. This is achieved by a first order Taylor expansion around ϵ_i^{KS}, since $\langle\Sigma(\omega)\rangle$ is almost linear in energy around a QP peak,

$$\left\langle \Sigma(\epsilon_i^{\mathrm{QP}}) \right\rangle = \left\langle \Sigma\left(\epsilon_i^{\mathrm{KS}}\right) \right\rangle + \left(\epsilon_i^{\mathrm{QP}} - \epsilon_i^{\mathrm{KS}}\right) \left\langle \left. \frac{\partial \Sigma(\omega)}{\partial \omega} \right|_{\omega=\epsilon_i^{\mathrm{KS}}} \right\rangle + O\left(\left(\epsilon_i^{\mathrm{QP}} - \epsilon_i^{\mathrm{KS}}\right)^2\right). \tag{2.105}$$

Thus, the correction for the KS-LDA energies is

$$\epsilon_i^{\mathrm{QP}} - \epsilon_i^{\mathrm{KS}} = \frac{\left\langle \Sigma\left(\epsilon_i^{\mathrm{KS}}\right) \right\rangle - \left\langle V_{\mathrm{xc}}^{\mathrm{LDA}} \right\rangle}{1 - \left\langle \left. \frac{\partial \Sigma(\omega)}{\partial \omega} \right|_{\omega=\epsilon_i^{\mathrm{KS}}} \right\rangle}. \tag{2.106}$$

In the denominator we recognize the renormalization constant Z_i from Eq. (2.92).

The derivation in Eq. (2.106) is done numerically,

$$\left. \frac{\partial \Sigma(\epsilon_i)}{\partial \omega} \right|_{\omega=\epsilon_i} \cong \frac{\Sigma(\epsilon_i + \Delta\omega) - \Sigma(\epsilon_i)}{\Delta\omega}, \tag{2.107}$$

with $\Delta\omega \approx 0.5$ eV.

It turns out that already this first step in the solution of Hedin's equations delivers results in excellent agreement with experiment. Limited updating of only the energy denominators in G and W has been explicitly shown to alter the quasiparticle energies in semiconductors by less than 0.1 eV [25, 26].

Chapter 3

Optical Absorption Spectra

For band structures excellent agreement with experimental data has been obtained for a wide range of materials by applying self-energy corrections – usually evaluated in the *GW* approximation (see Section 2.3.2) – to the Kohn-Sham electronic structure Eqs. (2.17) [5, 6, 25, 27].

However, spectroscopic properties involving two-particle excitations are often only poorly described at this one-particle level. The main example is absorption spectroscopy, where a simultaneously created electron–hole pair interacts more or less strongly. As a consequence, in addition to bound exciton states which occur within the gap, the spectral lineshape above the continuous–absorption edge is distorted.

The reported qualitative agreement with experiment of many computed Kohn–Sham LDA absorption spectra, obtained from one-electron transitions between KS states [7, 8], is indeed due to a partial cancellation between two principal errors: namely, the compensation of the large KS-LDA underestimation of the valence–conduction band gap (see Section 2.2.4), with an overestimation of the absorption onset induced by calculating the dielectric function entirely within the one-particle picture.

The situation often worsens when only the first error is corrected by replacing the KS eigenvalues with the realistic QP energies, as has been shown, *e.g.*, for the case of a sodium tetramer [48] as well as for bulk silicon [57]. On the other hand, going beyond the one-particle picture through inclusion of local-field and/or exchange-correlation effects from Eq. (2.50) within DFT-LDA in the calculation of the absorption spectrum does generally not remove the observed discrepancy [9]. In fact, most of the residual error stems from the neglect of the electron–hole interaction.

In this chapter we will first explain, how the macroscopic dielectric function can be obtained from the dielectric function and consider the independent-quasiparticle approximation as well as the inclusion of the local-field effects.

Then we will develop a new method to include excitonic effects in optical absorption spectra from first principles. We will rederive the Bethe-Salpeter equation,

which is the starting point of our approach, using a second iteration of Hedin's equations after the *GW* approximation. The excitonic problem is solved using an effective two-particle Schrödinger equation in second quantized form.

The excitonic wavefunctions are represented in Bloch space with a basis in terms of $\mathbf{k}$ vectors in the Brillouin zone and valence and conductions bands. We will apply group theory in order to reduce the excitonic Hamiltonian from one large matrix to several smaller ones, thus making the problem feasible for an *ab initio* approach.

We will only consider spin-paired electrons and can therefore omit the spin indices, if not otherwise stated. The appropriate prefactors are included in the formulas in order to account for spin.

3.1 Macroscopic Dielectric Function

The polarization of an inhomogeneous medium induces fluctuations on the scale of the atoms in the material and gives rise to internal *microscopic* fields. While the *macroscopic* field $\mathbf{E}$ varies with frequency ω and has the Fourier component of the limit of long wavelengths of the internal field, the microscopic fields vary with the same frequency ω, but with different wavevector components $\mathbf{q}+\mathbf{G}$, where $\mathbf{q}$ lies within the first Brillouin zone and $\mathbf{G}$ is a reciprocal lattice vector. These microscopic fluctuations are at the origin of the *local-field effects* and reflect the spatial variation of real materials.

According to Adler [45] and Wiser [46], the *macroscopic dielectric function* governing the optical properties of a crystal can be directly related to the inverse of the microscopic dielectric matrix Eq. (2.56),

$$\varepsilon_{\mathrm{M}}(\omega) \equiv \lim_{\mathbf{q}\to 0} \left. \frac{1}{\varepsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega)} \right|_{\mathbf{G}=\mathbf{G}'=0} . \tag{3.1}$$

Only cubic systems are investigated. So we do not have to care about the direction in which the limit of the small $\mathbf{q}$-vector has to be taken. In practice, however, we choose a very small $\mathbf{q}$ in some direction away from high-symmetry regions, as these may bias the result due to a finite basis set.

The optical absorption spectrum is then given by the imaginary part $\varepsilon_2(\omega)$ of $\varepsilon_{\mathrm{M}}(\omega)$. The dielectric constant ε_0 is the value of $\varepsilon_{\mathrm{M}}(\omega)$ at $\omega = 0$.

3.1.1 Independent-Particle Approximation

The dielectric function in the independent-particle approximation uses the polarization for independent-particles χ^0 in Eq. (2.59a) and neglects local-field effects, *i.e.*, Eq. (3.1) is approximated by $\varepsilon_{\mathrm{M}}(\omega) = \lim_{\mathbf{q}\to 0} \varepsilon_{00}(\mathbf{q},\omega)$. We obtain from Eqs. (2.57) and (2.59b)

$$\varepsilon_{\mathrm{M}}(\omega) = 1 + 2\frac{4\pi}{\Omega} \lim_{\mathbf{q}\to 0} \frac{1}{q^2} \sum_{\mathbf{k},c,v} \left\{ \frac{|\langle c\mathbf{k}+\mathbf{q}\,|e^{i\mathbf{q}\cdot\mathbf{r}}|\,v\mathbf{k}\rangle|^2}{\epsilon_{c\mathbf{k}+\mathbf{q}} - \epsilon_{v\mathbf{k}} - (\omega + i\eta)} \right.$$

$$+\frac{|\langle c\mathbf{k}-\mathbf{q}|e^{-i\mathbf{q}\cdot\mathbf{r}}|v\mathbf{k}\rangle|^2}{\epsilon_{c\mathbf{k}-\mathbf{q}}-\epsilon_{v\mathbf{k}}+(\omega+i\eta)}\Bigg\}, \tag{3.2}$$

where we have used the Fourier transformation of the Coulomb potential,

$$\frac{1}{|\mathbf{r}|}=\frac{4\pi}{\Omega}\sum_{\mathbf{q},\mathbf{G}}\frac{e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}}{|\mathbf{q}+\mathbf{G}|^2}, \tag{3.3}$$

and the factor two denotes the spin summation. Eq. (3.2) is the response to a longitudinal field for $\mathbf{q}\to 0$ in the *length gauge*.

Since the anti-resonant (second) term does not contribute to the absorption process, we obtain for the imaginary part ε_2, relevant for absorption,

$$\varepsilon_2(\omega)=2\frac{4\pi}{\Omega}\lim_{\mathbf{q}\to 0}\frac{1}{q^2}\sum_{\mathbf{k},c,v}\mathrm{Im}\,\frac{|\langle c\mathbf{k}+\mathbf{q}|e^{i\mathbf{q}\cdot\mathbf{r}}|v\mathbf{k}\rangle|^2}{\epsilon_{c\mathbf{k}+\mathbf{q}}-\epsilon_{v\mathbf{k}}-\omega-i\eta}. \tag{3.4}$$

Eq. (3.4) could be evaluated in $\mathbf{k}\cdot\mathbf{p}$ – perturbation theory in order to obtain the wavevector $|c\mathbf{k}+\mathbf{q}\rangle$ [18, Chap. 4][58, Vol. 1, p. 76]. However, we would like to clarify some aspects by rederiving it using the golden rule. Following Bassani and Pastori Parravicini [59, Chap. 5], we transform the kinetic energy of a system of N electrons in the presence of an electromagnetic field according to

$$\sum_{i=1}^{N}\frac{\mathbf{p}_i^2}{2}\to\sum_{i=1}^{N}\frac{1}{2}\left(\mathbf{p}_i-\frac{\mathbf{A}(\mathbf{r}_i,t)}{c}\right)^2, \tag{3.5}$$

where $\mathbf{A}$ is the vector potential of the electromagnetic field used in the *Coulomb* or *velocity gauge* $\nabla\cdot\mathbf{A}=0$ in the absence of sources $\phi=0$. Neglecting non-linear effects, we find the interaction Hamiltonian of electrons in a radiation field,

$$H_{\mathrm{eR}}^{\mathrm{local}}=-\frac{1}{c}\sum_i\mathbf{A}(\mathbf{r}_i,t)\cdot\mathbf{p}_i. \tag{3.6}$$

However, this well-known result is only correct for local potentials. For a *nonlocal Hamiltonian* H, as is the case in the pseudopotential formalism (see Appendix A), we must be more careful. First, we note that the commutator with the coordinate operator is given by [60]

$$\mathbf{v}_i=\frac{d}{dt}\mathbf{r}_i=i\,[H,\mathbf{r}_i]=\mathbf{p}_i+i\,[V_{\mathrm{nl}},\mathbf{r}_i], \tag{3.7}$$

where V_{nl} indicates the nonlocal part of the Hamiltonian.

By transforming the nonlocal potential into a momentum-dependent form, making the same replacements $\mathbf{p}\to\mathbf{p}-\mathbf{A}/c$ like in the above standard case and retaining only linear terms in first-order perturbation, one finally obtains the correct interaction Hamiltonian in a radiation field,

$$H_{\mathrm{eR}}^{\mathrm{nonlocal}}=-i\left[H,\sum_i\mathbf{r}_i\right]\cdot\frac{\mathbf{A}}{c}=-\sum_i\mathbf{v}_i\cdot\frac{\mathbf{A}}{c}. \tag{3.8}$$

The index i will be dropped from now on, since we will only deal with one-electron quantities. The wavefunctions and energies refer thus to one-particle quasiparticle states.

This term is then treated in time-dependent perturbation theory. The probability per unit time that a perturbation of the form $\mathfrak{L}e^{\mp i\omega t}$ (where the time dependence is completely contained in $e^{\mp i\omega t}$) induces a transition from the initial state $|i\rangle$ with energy ϵ_i to the final state $|f\rangle$ with energy ϵ_f is

$$P_{i\to f} = 2\pi \left|\langle f|\, \mathfrak{L}\, |i\rangle\right|^2 \delta\left(\epsilon_f - \epsilon_i \mp \omega\right). \tag{3.9}$$

The $-$ sign is for absorption, the $+$ sign for emission of a quantum ω. In the independent particle picture $(\epsilon_f - \epsilon_i)$ is just the energy difference between two one-particle quasiparticle levels. Writing the vector potential as

$$\mathbf{A}(\mathbf{r},t) = A_0\hat{\mathbf{e}}e^{i(\mathbf{q}\cdot\mathbf{r}-\omega t)} + c.c., \tag{3.10}$$

where $\hat{\mathbf{e}}$ is the polarization vector in the direction of the electric field and $\mathbf{q}$ the wave vector of the radiation, we obtain the transition rate for an absorption from $|i\rangle$ to $|f\rangle$,

$$P_{|i\rangle\to|f\rangle} = 2\pi\left(\frac{A_0}{c}\right)^2 \delta_{\sigma_i,\sigma_f} \left|\hat{\mathbf{e}}\cdot\mathbf{D}_{fi}\right|^2 \delta\left(\epsilon_f - \epsilon_i - \omega\right). \tag{3.11}$$

The δ-functions indicate energy and spin conservation. The latter follows from the fact that H_{eR} does not depend on spin.

Since the momentum of the incoming light quantum vanishes, we can only consider direct transitions between two Bloch states with $|i\rangle = |v\mathbf{k}\rangle$ and $|f\rangle = |c\mathbf{k}\rangle$. Therefore we neglect the possibility of momentum decay through three-particle effects like lattice vibrations. Undoubtedly, phonons are responsible for broadening spectral features. However, these lifetime effects are approximately taken into account by a small imaginary constant $i\eta$ added to the absorption energy ω in Eq. (3.2).

The electric dipole transition matrix elements are given by, using Eqs. (3.7) and (3.8),

$$\mathbf{D}_{cv}(\mathbf{k}) = \langle c\mathbf{k}\,|\mathbf{v}|\,v\mathbf{k}\rangle\,. \tag{3.12}$$

The determination of the absorption spectrum requires the knowledge of the absorption coefficient. The macroscopic physical quantities thereby are the complex dielectric constant, $\varepsilon = \varepsilon_1 + i\varepsilon_2$, and the complex refraction index, $N = n + ik$, where n is the ordinary refraction index and k the extinction coefficient. The optical constants ε and N are related by $\varepsilon = N^2$.

From the *absorption coefficient* α, which is given by [59, Chap. 5]

$$\alpha = \frac{2k\omega}{c} = \frac{\omega}{nc}\varepsilon_2, \tag{3.13}$$

we obtain the imaginary part of the dielectric constant, ε_2. α is by definition the energy absorbed per unit time and volume, W, divided by the energy flux,

$$\alpha(\omega) = \frac{\omega W(\omega)}{u\,(c/n)}, \tag{3.14}$$

where c/n is the speed of light in the solid and the average energy density u of the radiation field, Eq. (3.10), is

$$u = \frac{n^2 A_0^2 \omega^2}{2\pi c^2}. \tag{3.15}$$

The absorbed energy $W(\omega)$ is obtained by summing over all states in Eq. (3.11). Finally, we get the imaginary part of the dielectric constant

$$\varepsilon_2(\omega) = 2\frac{4\pi^2}{\Omega}\frac{1}{\omega^2}\sum_{\mathbf{k},c,v} |\hat{\mathbf{e}}\cdot\mathbf{D}_{cv}(\mathbf{k})|^2\,\delta\left(\epsilon_{c\mathbf{k}} - \epsilon_{v\mathbf{k}} - \omega\right). \tag{3.16}$$

In order to understand the relation with Eq. (3.4) and to prove gauge invariance we rewrite Eq. (3.7) slightly differently as [57]

$$v_\alpha = \lim_{q\to 0}\left[H, e^{iqr_\alpha}\right]/q, \tag{3.17}$$

where v_α and r_α ($\alpha = x, y, z$) are the Cartesian components of $\mathbf{v}$ and $\mathbf{r}$, respectively. Placing Eq. (3.17) in Eq. (3.16) and noting that all wavefunctions and energies are quasiparticle quantities, we obtain

$$\varepsilon_2(\omega) = 2\frac{4\pi^2}{\Omega}\lim_{\mathbf{q}\to 0}\frac{1}{q^2}\sum_{\mathbf{k},c,v}\left|\left\langle c\mathbf{k}+\mathbf{q}\left|e^{i\mathbf{q}\cdot\mathbf{r}}\right|v\mathbf{k}\right\rangle\right|^2\delta\left(\epsilon_{c\mathbf{k}+\mathbf{q}} - \epsilon_{v\mathbf{k}} - \omega\right), \tag{3.18}$$

which is just the imaginary part of Eq. (3.4), remembering the rule

$$\lim_{\eta\to 0}\frac{1}{x \pm i\eta} = P\frac{1}{x} \mp i\pi\delta(x). \tag{3.19}$$

We can further simplify, noting that in the GW approximation the quasiparticle (GW) wave functions are nearly identical to the KS-LDA ones (see Section 2.3.2), and applying once again Eqs. (3.7) and (3.17),

$$\lim_{q\to 0}\left\langle c\mathbf{k}+\mathbf{q}\left|e^{iq\hat{\mathbf{e}}\cdot\mathbf{r}}\right|v\mathbf{k}\right\rangle/q = \frac{\hat{\mathbf{e}}\cdot\left\langle c\mathbf{k}\left|\frac{\nabla}{i} + i\left[V_{\mathrm{nl}},\mathbf{r}\right]\right|v\mathbf{k}\right\rangle}{\epsilon^{\mathrm{KS}}_{c\mathbf{k}} - \epsilon^{\mathrm{KS}}_{v\mathbf{k}}}. \tag{3.20}$$

The Kohn-Sham eigenvalues appear in the denominator in Eq. (3.20) since the explicit application of the operators demands for the use of the appropriate Hamiltonian, which is in this case the Kohn-Sham one from Eqs. (2.17). The use of the quasiparticle Hamiltonian of Eq. (2.86) for the commutator in Eq. (3.17) would be

considerable more cumbersome in the evaluation of the matrix element in Eq. (3.20), since the self-energy operator Σ is nonlocal and dynamical.

This gives us the final formula in the *RPA formulation*, neglecting local-field effects,

$$\varepsilon_{\mathrm{M}}(\omega) = 1 + 2\frac{4\pi}{\Omega}\sum_{\mathbf{k},c,v}\left|\frac{\hat{\mathbf{e}}\cdot\left\langle c\mathbf{k}\left|\frac{\nabla}{i}+i\left[V_{\mathrm{nl}},\mathbf{r}\right]\right|v\mathbf{k}\right\rangle}{\epsilon^{\mathrm{KS}}_{c\mathbf{k}}-\epsilon^{\mathrm{KS}}_{v\mathbf{k}}}\right|^2\left\{\frac{1}{\epsilon^{\mathrm{QP}}_{c\mathbf{k}}-\epsilon^{\mathrm{QP}}_{v\mathbf{k}}-(\omega+i\eta)}\right.$$
$$\left.+\frac{1}{\epsilon^{\mathrm{QP}}_{c\mathbf{k}}-\epsilon^{\mathrm{QP}}_{v\mathbf{k}}+(\omega+i\eta)}\right\}. \quad (3.21)$$

In this thesis we omit the nonlocal commutator $i\left[V_{\mathrm{nl}},\mathbf{r}\right]$ (see also Appendix A) in the calculation of silicon in Chapter 4. Its contribution depends on the nonlocality of the pseudopotential of the studied system and tends to decrease the absorption spectrum for about 10% and more in certain cases, in particular strongly bonded group-IV materials, *e.g.*, SiC [42, 61] or diamond, where we include it (see Chapter 5).

Joint Density of States

If we take the matrix elements in Eq. (3.16) to be constant, we find that the absorption spectrum in the independent-particle picture is proportional to $1/\omega^2$ and to the quantity [59]

$$J_{cv}(\omega) = \frac{2}{(2\pi)^3}\int_{\mathrm{BZ}} d\mathbf{k}\sum_{c,v}\delta\left(\epsilon^{\mathrm{QP}}_{c\mathbf{k}}-\epsilon^{\mathrm{QP}}_{v\mathbf{k}}-\omega\right), \quad (3.22)$$

which is called joint density of states, because it gives the density of pairs of states: one state is an occupied valence and the other one an empty conduction band state, separated by an energy ω. Using the Eqs. (2.36) and (3.19), we obtain

$$J_{cv}(\omega) = \frac{2}{\pi\Omega}\,\mathrm{Im}\sum_{\mathbf{k},c,v}\frac{1}{\epsilon^{\mathrm{QP}}_{c\mathbf{k}}-\epsilon^{\mathrm{QP}}_{v\mathbf{k}}-\omega-i\eta}. \quad (3.23)$$

The function $J_{cv}(\omega)$ gives us hints about the absorption spectrum and shows strong variations in the neighborhood of particular transition energies $(\epsilon_{c\mathbf{k}}-\epsilon_{v\mathbf{k}})$, which are called critical point energies [59].

3.1.2 Local-Field Effects

The inclusion of local-field and/or exchange-correlation effects within the DFT-LDA in the calculation of the absorption spectrum is, in principle, straightforward. From Eq. (2.52) we know the polarizability. Placing the Fourier transformation of the inverse dielectric function Eq. (2.58) in Eq. (3.1) directly yields the macroscopic

dielectric function for each frequency ω. The resulting formulas may be found, *e.g.*, in Ref. [9].

For semiconductors and insulators the exchange-correlation effects within the DFT-LDA, given by K_{xc} from Eq. (2.50), have been shown to give only small contributions to the absorption spectrum [9] and we will not consider this possibility here.

Here we would like to derive a different formulation for the macroscopic dielectric function, which enables us to include later on also excitonic effects without much complications.

Given a matrix of the form

$$\mathsf{M} = \begin{pmatrix} m_{00} & \mathbf{m}_1^T \\ \mathbf{m}_2 & \mathsf{m} \end{pmatrix}, \tag{3.24}$$

with m_{00} being a c-number, we get its inverse as

$$\mathsf{M}^{-1} = \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \mathsf{m}^{-1} \end{pmatrix} + \frac{1}{(m_{00} - \mathbf{m}_1^T \mathsf{m}^{-1} \mathbf{m}_2)} \begin{pmatrix} 1 & -\mathbf{m}_1^T \mathsf{m}^{-1} \\ -\mathsf{m}^{-1}\mathbf{m}_2 & \mathsf{m}^{-1}\mathbf{m}_2\mathbf{m}_1^T\mathsf{m}^{-1} \end{pmatrix}. \tag{3.25}$$

Now we use the expression for the dielectric function Eq. (2.95), $\varepsilon = 1 - vP$, and write the macroscopic dielectric function in terms of Fourier components as, using Eq. (3.25) and $\varepsilon_{\text{M}} = 1/\varepsilon_{00}^{-1}(0)$,

$$\varepsilon_{\text{M}}(\omega) = \lim_{\mathbf{q}\to 0} \left[\varepsilon_{00}(\mathbf{q},\omega) - \boldsymbol{\varepsilon}_1^T(\mathbf{q},\omega) \epsilon^{-1}(\mathbf{q},\omega) \boldsymbol{\varepsilon}_2(\mathbf{q},\omega) \right]. \tag{3.26}$$

The quantities ε_{00}, $\boldsymbol{\varepsilon}_1^T$, $\boldsymbol{\varepsilon}_2$ and ϵ are given as

$$\begin{aligned}
\varepsilon_{00}(\mathbf{q},\omega) &= 1 - v_0(\mathbf{q}) P_{00}(\mathbf{q},\omega), && && (3.27a)\\
\left[\boldsymbol{\varepsilon}_1^T(\mathbf{q},\omega)\right]_{\mathbf{G}} &= v_0(\mathbf{q}) P_{0\mathbf{G}}(\mathbf{q},\omega), && \mathbf{G} \neq \mathbf{0}, && (3.27b)\\
\left[\boldsymbol{\varepsilon}_2(\mathbf{q},\omega)\right]_{\mathbf{G}'} &= v_{\mathbf{G}'}(\mathbf{q}) P_{\mathbf{G}'0}(\mathbf{q},\omega), && \mathbf{G}' \neq \mathbf{0}, && (3.27c)\\
\left[\epsilon(\mathbf{q},\omega)\right]_{\mathbf{G}\mathbf{G}'} &= v_{\mathbf{G}}(\mathbf{q}) P_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega), && \mathbf{G},\mathbf{G}' \neq \mathbf{0}. && (3.27d)
\end{aligned}$$

We obtain

$$\begin{aligned}
\varepsilon_{\text{M}}(\omega) = 1 - \lim_{\mathbf{q}\to 0} \Big[& v_0(\mathbf{q}) P_{00}(\mathbf{q},\omega) \\
& - \sum_{\mathbf{G},\mathbf{G}'\neq\mathbf{0}} v_0(\mathbf{q}) P_{0\mathbf{G}}(\mathbf{q},\omega) \epsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) v_{\mathbf{G}'}(\mathbf{q}) P_{\mathbf{G}'0}(\mathbf{q},\omega) \Big].
\end{aligned} \tag{3.28}$$

In order to simplify this expression, we want a matrix expression involving all $\mathbf{G}$ and $\mathbf{G}'$, not only $\mathbf{G}, \mathbf{G}' \neq \mathbf{0}$. Therefore we define

$$\bar{\varepsilon}(\mathbf{q},\omega)_{\mathbf{G}\mathbf{G}'} \equiv \delta_{\mathbf{G},\mathbf{G}'} - \bar{v}_{\mathbf{G}}(\mathbf{q}) P_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega), \tag{3.29}$$

with

$$\bar{v}_{\mathbf{G}}(\mathbf{q}) \equiv \begin{cases} 0, & \mathbf{G} = 0 \\ v_{\mathbf{G}}(\mathbf{q}), & \mathbf{G} \neq 0 \end{cases}, \tag{3.30}$$

where the dash indicates that the $\mathbf{G} = \mathbf{0}$ contribution is to be left out in the bare Coulomb interaction. In matrix notation we have

$$\bar{\varepsilon}_{\mathbf{G}\mathbf{G}'} = \begin{pmatrix} 1 & \mathbf{0}^T \\ \varepsilon_2 & \epsilon \end{pmatrix}. \tag{3.31}$$

Then the inverse of $\bar{\varepsilon}$ is given by, using Eq. (3.25),

$$\bar{\varepsilon}^{-1}_{\mathbf{G}\mathbf{G}'} = \begin{pmatrix} 1 & \mathbf{0}^T \\ -\epsilon^{-1}\varepsilon_2 & \epsilon^{-1} \end{pmatrix}, \tag{3.32}$$

which can be easily checked. Thus we can write Eq. (3.28) as

$$\varepsilon_{\mathrm{M}}(\omega) = 1 - \lim_{\mathbf{q}\to 0} \left[v_0(\mathbf{q}) \bar{P}_{00}(\mathbf{q},\omega) \right], \tag{3.33}$$

where we have defined

$$\bar{P}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) \equiv P_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) + \sum_{\mathbf{K},\mathbf{K}'} P_{\mathbf{G}\mathbf{K}}(\mathbf{q},\omega) \bar{\varepsilon}^{-1}_{\mathbf{K}\mathbf{K}'}(\mathbf{q},\omega) \bar{v}_{\mathbf{K}'}(\mathbf{q}) P_{\mathbf{K}'\mathbf{G}'}(\mathbf{q},\omega). \tag{3.34}$$

The replacement of v by $\bar{v}$ in the second term in Eq. (3.34) allows us to extend the summation over all $\mathbf{K}, \mathbf{K}'$ including $\mathbf{0}$. This is the wanted matrix expression

$$\bar{P} = P + P\bar{\varepsilon}^{-1}\bar{v}P. \tag{3.35}$$

Together with $\bar{\varepsilon}^{-1} = (1 - \bar{v}P)^{-1}$, Eq. (3.29), we finally arrive at the general expression

$$\bar{P} = \left(P^{-1} - \bar{v}\right)^{-1}, \tag{3.36}$$

or in form of a Dyson equation,

$$\bar{P} = P + P\bar{v}\bar{P}. \tag{3.37}$$

Until now, we have not specified P. In the terminology of excitonic calculations the choice $P = P_0$ (we remind that $P_0 = \chi^0$, the independent-particle polarizability) in Eq. (3.37) gives the so-called *exchange* contributions, which is hence the same as including the local-field effects in the RPA. They are explicitly given by the discrepancy between $\varepsilon_{\mathrm{M}}(\omega)$ and $\lim_{\mathbf{q}\to 0} \epsilon_{00}(\mathbf{q},\omega)$ in Eq. (3.26).

3.1.3 Electron–Hole Attraction

Excitonic effects can be included in the macroscopic dielectric function, Eq. (3.33), by taking the polarization function P in Eq. (3.37) not in the RPA approximation, but going one step further and also considering the electron–hole interaction. This can be achieved by an iteration of Hedin's equations (2.93) beyond the *GW* approximation.

Second Iteration of Hedin's Equations

We have to go again through equation Eq. (2.93d), now using $\Sigma = iGW$, Eq. (2.99a). Doing so yields an integral equation for Γ,

$$\begin{aligned}\Gamma(x,x';x'') &= \delta(x-x')\delta(x-x'') \\ &\quad +iW(x^+,x')\int dx_3dx_4\; G(x,x_3)G(x_4,x')\Gamma(x_3,x_4;x''). \quad (3.38)\end{aligned}$$

Here, we have used $\delta\Sigma/\delta G = iW$, neglecting the term $iG\frac{\delta W}{\delta G}$, which contains information about the change in screening due to the excitation and is expected to be small [12, 62, 63].

We can transform Eq. (3.38) to an integral equation for a generalized $\mathbb{P}$, defined as,

$$\mathbb{P}(x'';x_1,x_2) \equiv -i\int dxdx'\;\Gamma(x,x';x'')G(x_1,x)G(x',x_2), \quad (3.39)$$

by multiplying with $-i\,G(x_1,x)G(x',x_2)$ and integrating over $dxdx'$:

$$\begin{aligned}\mathbb{P}(x'';x_1,x_2) &= -iG(x_1,x'')\,G(x'',x_2) \\ &\quad +i\int dxdx'\;\mathbb{P}(x'';x,x')\,W(x^+,x')\,G(x_1,x)G(x',x_2). \quad (3.40)\end{aligned}$$

The polarization P is directly related to $\mathbb{P}$,

$$P(x'',x_2) = \mathbb{P}(x'';x_2,x_2). \quad (3.41)$$

Thus one could possibly find a solution for P by solving for the three-point function $\mathbb{P}$ in Eq. (3.40), of which we only need the diagonal part. However, this approach is not followed in this thesis.

Indeed, like in the previous section on local-field effects, we prefer a matrix expression for Eq. (3.40). Therefore we understand P as a four-point polarization propagator $P(x_1,x_2;x_3,x_4)$ from which $P(x,x')$ is obtained as $P(x,x;x',x')$, which can be done without ambiguity in the given context. In the same sense the bare Coulomb interaction $v(x_1,x_2;x_3,x_4)$ is constructed as

$$v(x_1,x_2) \to v(x_1,x_3)\delta(x_1-x_2)\delta(x_3-x_4), \quad (3.42)$$

and the dynamically screened interaction $W(x_1,x_2;x_3,x_4)$ as

$$W(x_1,x_2) \to W(x_1,x_2)\delta(x_1-x_3)\delta(x_2-x_4) \quad (3.43)$$

Then we get for Eq. (3.40) the integral equation

$$\begin{aligned}P(x_1,x_2;x_3,x_4) &= P_0(x_1,x_2;x_3,x_4) - \int dx_5dx_6dx_7dx_8\; P(x_1,x_2;x_5,x_6) \\ &\quad \times W(x_5,x_6;x_7,x_8)P_0(x_7,x_8;x_3,x_4), \quad (3.44)\end{aligned}$$

which can be written in matrix form as

$$P = P_0 - PWP_0, \tag{3.45}$$

or

$$P = P_0 \left(1 + WP_0\right)^{-1}. \tag{3.46}$$

This expression for P corresponds to an evaluation of the polarization in the time-dependent Hartree-Fock approximation, but with a screened instead of a bare Coulomb interaction. All diagrams in a ladder approximation for the polarization in addition to the non-interacting RPA bubble are summed over [10].

Placing the Eq. (3.46) for P into Eq. (3.36) for the now four-point $\bar{P}$, leads to the *Bethe-Salpeter equation,* Eq. (1.1),

$$\bar{P} = P_0 + P_0(\bar{v} - W)\bar{P}. \tag{3.47}$$

We remind that the exchange contribution $\bar{v}$ is used without the long range term of vanishing wave vector.

Since the total spin of the ground state in semiconductors and insulators equals zero and H_{eR} in Eqs. (3.6) or (3.8) does not depend on spin, only singlet transverse excitons with spin zero can be induced by the radiation field, neglecting the coupling of spin and magnetic field. Then, when spin is not explicitly treated, $\bar{v}$ gets a factor of two for singlet excitons.

We limit ourselves to static screening in W, since dynamical effects in the electron–hole screening and in the one particle Green's function tend to cancel each other [64], which suggests to neglect both of them. Especially, when the plasma frequencies of the investigated systems are much larger than the excitonic binding energies and also than the investigated absorption frequencies, this approximation is adequate.

Interaction Kernel Ξ

The connection to the function S in Eq. (1.1) is given by

$$\bar{P}(x_1, x_1'; x_2, x_2') = -iS(x_1, x_1'; x_2, x_2'), \tag{3.48}$$

where the Bethe-Salpeter equation is traditionally written as [12],

$$\begin{aligned} S(x_1, x_1'; x_2, x_2') \;=\; & S_0(x_1, x_1'; x_2, x_2') + S_0(x_1, x_1'; x_3, x_3') \\ & \times \Xi(x_3, x_3'; x_4, x_4') S(x_4, x_4'; x_2, x_2'). \end{aligned} \tag{3.49}$$

Repeated arguments are integrated over. $S(x_1, x_1'; x_2, x_2')$ is indeed the part of the two-particle Green's function Eq. (2.80) which excludes the disconnected term

$$-G(x_1, x_1')G(x_2, x_2'). \tag{3.50}$$

The kernel Ξ can also be found as the functional derivative of the self-energy Σ from Eq. (2.81) including the Hartree potential V_{H} from Eq. (2.83) with respect to the single-particle Green's function [62]:

$$\Xi(x_1, x_1'; x_2, x_2') \equiv \frac{\delta\left(V_{\mathrm{H}}(x_1) + \Sigma(x_2, x_2')\right)}{\delta G(x_1, x_1')}. \tag{3.51}$$

With $\Sigma = iGW$ we obtain then

$$\begin{aligned} \Xi(x_1, x_1'; x_2, x_2') &= -i\delta(x_1, x_1')\delta(x_2, x_2')v(x_1, x_2) \\ &\quad +i\delta(x_1, x_2)\delta(x_1', x_2')W(x_1, x_1'). \end{aligned} \tag{3.52}$$

Like above in Eq. (3.38), we have neglected a term $G\frac{\delta W}{\delta G}$.

Macroscopic Dielectric Function

Finally, the macroscopic dielectric function is given by placing the Fourier transformation of $\bar{P}(\mathbf{r}, \mathbf{r}, \mathbf{r}', \mathbf{r}'; \omega)$ into Eq. (3.33),

$$\varepsilon_{\mathrm{M}}(\omega) = 1 - \lim_{\mathbf{q}\to 0}\left[v_0(\mathbf{q}) \int d\mathbf{r} \int d\mathbf{r}'\, e^{-i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')}\bar{P}(\mathbf{r}, \mathbf{r}, \mathbf{r}', \mathbf{r}'; \omega)\right]. \tag{3.53}$$

The dash indicates once more the omission of the long range term of vanishing wave vector in the exchange contribution.

3.2 Excitonic Effects: Bloch Space

In order to solve Eq. (3.49), we have to invert a four-point function for each absorption frequency. In an *ab initio* plane wave calculation (see Appendix A) such a procedure is clearly prohibitive, when plane waves are chosen as straightforward basis functions.

Instead, the physical picture of interacting electron–hole pairs suggests to use a basis of LDA Bloch functions, $\varphi_{n\mathbf{k}}(\mathbf{r})$, from Eq. (2.31) expecting that only a limited number of electron–hole pairs will contribute to each excitation [48, 65, 66]. We use a spectral representation and transform the inversion into an effective eigenvalue problem, which is solved by diagonalization.

3.2.1 Second Quantization of Bethe-Salpeter Equation

The KS-LDA Bloch functions (numbered by only one label n including spin) are orthonormal and complete and we can transform the polarization P, omitting the frequency dependence, as

$$\bar{P}(\mathbf{r}_1, \mathbf{r}_1; \mathbf{r}_2', \mathbf{r}_2') = \sum_{n_1 \ldots n_4} \varphi_{n_1}^*(\mathbf{r}_1)\varphi_{n_2}(\mathbf{r}_1')\varphi_{n_3}(\mathbf{r}_2)\varphi_{n_4}^*(\mathbf{r}_2')\bar{P}_{(n_1,n_2),(n_3,n_4)}, \tag{3.54}$$

with

$$\bar{P}_{(n_1,n_2),(n_3,n_4)} = P^0_{(n_1,n_2),(n_3,n_4)} - iP^0_{(n_1,n_2),(n_5,n_6)}\bar{\Xi}_{(n_5,n_6),(n_7,n_8)}\bar{P}_{(n_7,n_8),(n_3,n_4)}. \quad (3.55)$$

The independent-particle polarization P_0 is given by

$$P^0_{(n_1,n_2),(n_3,n_4)} = \int d\mathbf{r}_1 ... \int d\mathbf{r}'_2 \; P_0(\mathbf{r}_1,\mathbf{r}'_1;\mathbf{r}_2,\mathbf{r}'_2)\varphi_{n_1}(\mathbf{r}_1)\varphi^*_{n_2}(\mathbf{r}'_1)\varphi^*_{n_3}(\mathbf{r}_2)\varphi_{n_4}(\mathbf{r}'_2), \quad (3.56)$$

with

$$P_0(\mathbf{r}_1,\mathbf{r}'_1;\mathbf{r}_2,\mathbf{r}'_2) = \sum_{m_1...m_4} \frac{\varphi^*_{m_1}(\mathbf{r}_1)\varphi_{m_2}(\mathbf{r}'_1)\varphi_{m_3}(\mathbf{r}_2)\varphi^*_{m_4}(\mathbf{r}'_2)}{\epsilon_{m_2} - \epsilon_{m_1} - \omega - i\eta} \quad (3.57)$$

$$\times \; (f_{m_2} - f_{m_1})\,\delta_{m_1,m_3}\delta_{m_2,m_4}, \quad (3.58)$$

with f_i being the occupation numbers $(0,1)$ of the state i and η a positive infinitesimal. Thus,

$$P^0_{(n_1,n_2),(n_3,n_4)} = \frac{(f_{n_2} - f_{n_1})\,\delta_{n_1,n_3}\delta_{n_2,n_4}}{\epsilon_{n_2} - \epsilon_{n_1} - \omega - i\eta}. \quad (3.59)$$

Now we define

$$\bar{\Pi} \equiv \left[1 + iP_0\bar{\Xi}\right]^{-1}. \quad (3.60)$$

Then Eq. (3.47) becomes

$$\bar{P} = \bar{\Pi}P_0. \quad (3.61)$$

Using the relation

$$\mathsf{M}_{\alpha\beta} = \left[\delta_{ij} - \frac{\beta_{ij}}{c_i}\right]^{-1}_{\alpha\beta} = \left[\delta_{ij}c_i - \beta_{ij}\right]^{-1}_{\alpha\beta} c_\beta, \quad (3.62)$$

we can rewrite $\bar{\Pi}$ as

$$\begin{aligned}
&\bar{\Pi}_{(n_1,n_2),(n_3,n_4)} = \\
&= \left[\delta_{m_1,m_3}\delta_{m_2,m_4} + i\frac{f_{m_2} - f_{m_1}}{\epsilon_{m_2} - \epsilon_{m_1} - \omega - i\eta}\bar{\Xi}_{(m_1,m_2),(m_3,m_4)}\right]^{-1}_{(n_1,n_2),(n_3,n_4)} \\
&= \left[(\epsilon_{m_2} - \epsilon_{m_1} - \omega)\,\delta_{m_1,m_3}\delta_{m_2,m_4}\right. \\
&\quad \left. + i\,(f_{m_2} - f_{m_1})\,\bar{\Xi}_{(m_1,m_2),(m_3,m_4)}\right]^{-1}_{(n_1,n_2),(n_3,n_4)} (\epsilon_{n_4} - \epsilon_{n_3} - \omega)\,. \quad (3.63)
\end{aligned}$$

We define the *excitonic Hamiltonian* as

$$\bar{H}^{\text{exc}}_{(n_1,n_2),(n_3,n_4)} \equiv (\epsilon_{n_2} - \epsilon_{n_1})\,\delta_{n_1,n_3}\delta_{n_2,n_4} + i\,(f_{n_2} - f_{n_1})\,\bar{\Xi}_{(n_1,n_2),(n_3,n_4)}. \quad (3.64)$$

Thus,

$$\begin{aligned}
&\bar{\Pi}_{(n_1,n_2),(n_3,n_4)} = \\
&= \left[\bar{H}^{\text{exc}}_{(m_1,m_2),(m_3,m_4)} - \omega\delta_{m_1,m_3}\delta_{m_2,m_4}\right]^{-1}_{(n_1,n_2),(n_3,n_4)} (\epsilon_{n_4} - \epsilon_{n_3} - \omega)\,,
\end{aligned} \quad (3.65)$$

and

$$\bar{P}_{(n_1,n_2),(n_3,n_4)} = \left[\bar{H}^{\text{exc}} - I\,\omega\right]^{-1}_{(n_1,n_2),(n_3,n_4)} (f_{n_4} - f_{n_3})\,. \tag{3.66}$$

I is the identity operator $I = \delta_{m_1,m_3}\delta_{m_2,m_4}$.

We have to invert the matrix in Eq. (3.66) for each absorption frequency ω. This is computationally very expensive, even in a basis of Bloch functions. The important trick is to apply a spectral representation for the inversion, using the relation

$$\left[\bar{H}^{\text{exc}} - I\,\omega\right]^{-1} = \sum_{\lambda,\lambda'} \frac{|\lambda\rangle\, \mathsf{S}^{-1}_{\lambda,\lambda'}\, \langle\lambda'|}{E_\lambda - \omega}, \tag{3.67}$$

which holds for a system of eigenvectors and eigenvalues of a general non-Hermitian matrix defined by

$$\bar{H}^{\text{exc}}\,|\lambda\rangle = E_\lambda\,|\lambda\rangle\,. \tag{3.68}$$

$\mathsf{S}_{\lambda,\lambda'}$ is the overlap matrix of the (in general non-orthogonal eigenstates) of $\bar{H}^{\text{exc}}$,

$$\mathsf{S}_{\lambda,\lambda'} \equiv \langle\lambda\,|\lambda'\rangle\,. \tag{3.69}$$

The eigenvalues E_λ must nevertheless be real being the optical transition energies of a real physical system. There is no dispersion of momentum nor energy included in the Hamiltonian matrix. Thus our excitons have an infinite lifetime.

More explicitly, we define the *effective two-particle Schrödinger equation* for the excitonic system by

$$\bar{H}^{\text{exc}}_{(n_1,n_2),(n_3,n_4)} A^{(n_3,n_4)}_\lambda = E_\lambda A^{(n_1,n_2)}_\lambda, \tag{3.70}$$

which we solve by diagonalization. The macroscopic dielectric function can then relatively easily be constructed using the excitonic eigenvalues and eigenvectors, as explained in the next section.

The explicit knowledge of the coupling of the various two-particle channels, given by the coefficients $A^{(n_1,n_2)}_\lambda$ of the state $|\lambda\rangle$ in our Bloch basis, allows to identify the character of each transition.

The one-particle transition energies on the diagonal are taken to be the quasiparticle energies, obtained in a *GW* calculation (see Section 2.3.2). This corresponds to an updating of the energy denominators of the independent-particle polarization P_0 in Eqs. (3.60) and (3.61), using quasiparticle instead of Kohn-Sham LDA eigenvalues. The screened interaction W is taken from the preceding *GW* calculations and again we use the very close resemblance of the quasiparticle wavefunctions with LDA-KS ones for the excitonic eigenstates.

Consequently, the *exciton amplitude* is expressed by

$$\Psi_\lambda(\mathbf{r}_1,\mathbf{r}_2) = \sum_{n_1,n_2} A^{(n_1,n_2)}_\lambda \varphi^*_{n_1}(\mathbf{r}_1)\varphi_{n_2}(\mathbf{r}_2). \tag{3.71}$$

Taking only the resonant term of $\bar{H}^{\text{exc}}$, which is later on defined in Eq. (3.78), the amplitude is naturally given in terms of electron and hole coordinates,

$$\Psi_\lambda(\mathbf{r}_\text{h}, \mathbf{r}_\text{e}) = \sum_{v,c,\mathbf{k}} A_\lambda^{(vc\mathbf{k})} \varphi^*_{v\mathbf{k}}(\mathbf{r}_\text{h}) \varphi_{c\mathbf{k}}(\mathbf{r}_\text{e}). \tag{3.72}$$

Furthermore, we are interested in the excited electron–hole pair charge density. We calculate the expectation value of the hole charge density operator in the excited state Eq. (3.72) and obtain

$$\begin{aligned} n_\text{h}^\lambda(\mathbf{r}_\text{h}) &= \langle \Psi_\lambda(\mathbf{r}_\text{h}, \mathbf{r}_\text{e}) | - \delta(\mathbf{r}_1 - \mathbf{r}_\text{h}) | \Psi_\lambda(\mathbf{r}_\text{h}, \mathbf{r}_\text{e}) \rangle \\ &= -\sum_{v,c,\mathbf{k}} \sum_{v',c',\mathbf{k}'} \left(A_\lambda^{(vc\mathbf{k})} \right)^* A_\lambda^{(v'c'\mathbf{k}')} \int\int d\mathbf{r}_1 d\mathbf{r}_2 \varphi_{v\mathbf{k}}(\mathbf{r}_1) \varphi^*_{c\mathbf{k}}(\mathbf{r}_2) \delta(\mathbf{r}_1 - \mathbf{r}_\text{h}) \\ &\quad \times \varphi^*_{v'\mathbf{k}'}(\mathbf{r}_1) \varphi_{c'\mathbf{k}'}(\mathbf{r}_2) \\ &= -\sum_{v,c,\mathbf{k}} \sum_{v'} \left(A_\lambda^{(vc\mathbf{k})} \right)^* A_\lambda^{(v'c\mathbf{k})} \varphi_{v\mathbf{k}}(\mathbf{r}_\text{h}) \varphi^*_{v'\mathbf{k}}(\mathbf{r}_\text{h}). \end{aligned} \tag{3.73}$$

Similarly for the excited electron, we get

$$\begin{aligned} n_\text{e}^\lambda(\mathbf{r}_\text{e}) &= \langle \Psi_\lambda(\mathbf{r}_\text{h}, \mathbf{r}_\text{e}) | \delta(\mathbf{r}_2 - \mathbf{r}_\text{e}) | \Psi_\lambda(\mathbf{r}_\text{h}, \mathbf{r}_\text{e}) \rangle \\ &= \sum_{v,c,\mathbf{k}} \sum_{c'} \left(A_\lambda^{(vc\mathbf{k})} \right)^* A_\lambda^{(vc'\mathbf{k})} \varphi^*_{c\mathbf{k}}(\mathbf{r}_\text{e}) \varphi_{c'\mathbf{k}}(\mathbf{r}_\text{e}). \end{aligned} \tag{3.74}$$

These charge densities are related to the total charge density of the excited system $n^*_\lambda(\mathbf{r})$ and the ground state charge density $n(\mathbf{r})$ by

$$n^*_\lambda(\mathbf{r}) - n(\mathbf{r}) = n_\text{e}^\lambda(\mathbf{r}_\text{e}) + n_\text{h}^\lambda(\mathbf{r}_\text{h}). \tag{3.75}$$

This description of negative hole and positive electron contributions to the excited density is very useful to investigate bound excitonic states in solids [23] and clusters [65, 67]. In Appendix B we discuss as an example the sodium tetramer.

3.2.2 Absorption Spectrum

Structure of Excitonic Hamiltonian

We have to examine the structure of $\bar{H}^{\text{exc}}_{(n_1,n_2),(n_3,n_4)}$ more carefully. As in the case of independent particles (see Section 3.1.1) we only take direct transitions for the electron–hole pairs in the optical absorption, neglecting phonons, and therefore have excitons with zero momentum.

Omitting the spin and wavevector dependences, we obtain for the excitonic Hamiltonian, Eq. (3.64),

$$\bar{H}^{\text{exc}}_{(n_1,n_2),(n_3,n_4)} = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}, \tag{3.76}$$

with

$$A = \begin{array}{c|cc} (n_1,n_2)\backslash(n_3,n_4) & \{v',c'\} & \{c',v'\} \\ \hline \{v,c\} & H^{\mathrm{exc,res}}_{(v,c),(v',c')} & -i\bar{\Xi}_{(v,c),(c',v')} \\ \{c,v\} & i\left[\bar{\Xi}_{(v,c),(c',v')}\right]^* & -\left[\bar{H}^{\mathrm{exc,res}}_{(v,c),(v',c')}\right]^* \end{array}, \tag{3.77a}$$

$$B = \begin{array}{c|cc} (n_1,n_2)\backslash(n_3,n_4) & \{v',\tilde{v}'\} & \{c',\tilde{c}'\} \\ \hline \{v,c\} & -i\bar{\Xi}_{(v,c),(v',\tilde{v}')} & -i\bar{\Xi}_{(v,c),(c',\tilde{c}')} \\ \{c,v\} & i\bar{\Xi}_{(c,v),(v',\tilde{v}')} & i\bar{\Xi}_{(c,v),(c',\tilde{c}')} \end{array}, \tag{3.77b}$$

$$D = \begin{array}{c|cc} (n_1,n_2)\backslash(n_3,n_4) & \{v',\tilde{v}'\} & \{c',\tilde{c}'\} \\ \hline \{v,\tilde{v}\} & (\epsilon_{\tilde{v}}-\epsilon_v)\,\delta_{v,v'}\delta_{\tilde{v},\tilde{v}'} & 0 \\ \{c,\tilde{c}\} & 0 & (\epsilon_{\tilde{c}}-\epsilon_c)\,\delta_{c,c'}\delta_{\tilde{c},\tilde{c}'} \end{array}. \tag{3.77c}$$

The *resonant* part of excitonic Hamiltonian is defined as

$$\bar{H}^{\mathrm{exc,res}}_{(v,c),(v',c')} \equiv (\epsilon_c-\epsilon_v)\,\delta_{v,v'}\delta_{c,c'} - i\bar{\Xi}_{(v,c),(v',c')}. \tag{3.78}$$

It corresponds to transitions at positive absorption frequencies ω.

If we sum over spin, we obtain from Eq. (3.52),

$$\begin{aligned} \bar{H}^{\mathrm{exc,res}}_{(vc\mathbf{k}),(v'c'\mathbf{k}')} &= (\epsilon_{c\mathbf{k}}-\epsilon_{v\mathbf{k}})\,\delta_{v,v'}\delta_{c,c'}\delta_{\mathbf{k},\mathbf{k}'} \\ &\quad + \int d\mathbf{r}\int d\mathbf{r}'\;\varphi^*_{c\mathbf{k}}(\mathbf{r})\varphi_{v\mathbf{k}}(\mathbf{r})\frac{2}{|\mathbf{r}-\mathbf{r}'|}\varphi_{c'\mathbf{k}'}(\mathbf{r}')\varphi^*_{v'\mathbf{k}'}(\mathbf{r}') \\ &\quad - \int d\mathbf{r}\int d\mathbf{r}'\;\varphi^*_{c\mathbf{k}}(\mathbf{r})\varphi_{c'\mathbf{k}'}(\mathbf{r}) \\ &\qquad \times \int d\mathbf{r}_1\;\varepsilon^{-1}(\mathbf{r},\mathbf{r}_1)\frac{1}{|\mathbf{r}_1-\mathbf{r}'|}\;\varphi_{v\mathbf{k}}(\mathbf{r}')\varphi^*_{v'\mathbf{k}'}(\mathbf{r}'). \end{aligned} \tag{3.79}$$

In Eq. (3.66) we have a factor $(f_{n_4}-f_{n_3})$ for the polarization $\bar{P}$. Thus, only the first column of

$$\mathsf{M}^{-1}_{(n_1,n_2),(n_3,n_4)} = \left[\bar{H}^{\mathrm{exc}} - I\,\omega\right]^{-1}_{(n_1,n_2),(n_3,n_4)}, \tag{3.80}$$

i.e., with $(n_3,n_4)=\{c',v'\}$ and $\{v',c'\}$, contributes to the absorption process. We are calculating $\bar{H}^{\mathrm{exc}}$ and want to know, which parts in Eq. (3.76) are significant. We split M and M^{-1} like above in

$$\mathsf{M} = \begin{pmatrix} A_\omega & B \\ 0 & D_\omega \end{pmatrix}, \tag{3.81a}$$

$$\mathsf{M}^{-1} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}. \tag{3.81b}$$

($A_\omega = A - I\,\omega$, $D_\omega = D - I\,\omega$.) From $\mathsf{M}\mathsf{M}^{-1} = I$, we have the condition

$$\begin{pmatrix} \alpha A_\omega & \alpha B + \beta D_\omega \\ \gamma A_\omega & \beta B + \delta D_\omega \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.82}$$

Since only the first column $\binom{\alpha}{\gamma}$ of M^{-1} contributes and, from the above equation, $\gamma \equiv 0$ for all cases, we only need $\alpha = A_\omega^{-1}$. Therefore only the part A, Eq. (3.77a), is relevant for the macroscopic dielectric function, a result which is easy to understand by simply looking on the structure of $\bar{H}^{\text{exc}}$. Only pairs containing one filled and one empty Bloch state contribute.

From now on we will restrict ourselves to this part A, Eq. (3.77a), when referring to $\bar{H}^{\text{exc}}$. It is in general non-Hermitian and can be further separated into two blocks on the diagonal with the quasiparticle transition energies and the interaction kernel Ξ and two off-diagonal coupling blocks with contributions only form the interaction kernel,

$$\bar{H}^{\text{exc}} = \begin{pmatrix} \bar{H}^{\text{exc,res}} & \bar{H}^{\text{coupling}} \\ -\left[\bar{H}^{\text{coupling}}\right]^* & -\left[\bar{H}^{\text{exc,res}}\right]^* \end{pmatrix}. \tag{3.83}$$

The *resonant part* is Hermitian,

$$\left(\bar{H}^{\text{exc,res}}\right)^* = \left(\bar{H}^{\text{exc,res}}\right)^T, \tag{3.84}$$

while the *coupling part* alone is symmetric,

$$\bar{H}^{\text{coupling}} = \left(\bar{H}^{\text{coupling}}\right)^T. \tag{3.85}$$

We denote the part in the lower right *anti-resonant*. It contains negative transition energies, thus only contributing to negative frequencies. It gives the same contributions as the resonant part does for $\omega > 0$ for $\omega < 0$, if we neglect the interaction elements in the coupling parts, which are generally much smaller. Indeed, in very good approximation we can restrict ourselves to the resonant part for the absorption spectrum, as it will be shown in the further course of this thesis.

Macroscopic Dielectric Function

From the effective excitonic Schrödinger equation, Eq. (3.70), the polarization Eq. (3.66) is given by

$$\bar{P}_{(n_1,n_2),(n_3,n_4)}(\omega) = \sum_{\lambda,\lambda'} \frac{A_\lambda^{(n_1,n_2)} \mathsf{S}^{-1}_{\lambda,\lambda'} A_{\lambda'}^{*(n_3,n_4)}}{E_\lambda - \omega} \left(f_{n_4} - f_{n_3}\right). \tag{3.86}$$

Placing this form of $\bar{P}$ into Eq. (3.53), and transforming from Bloch space back to real space, we obtain for the macroscopic dielectric function,

$$\varepsilon_{\text{M}}(\omega) = 1 - \lim_{\mathbf{q}\to 0} \left\{ v_0(\mathbf{q}) \sum_{\lambda,\lambda'} \left[\sum_{n_1,n_2} \left\langle n_1 \left| e^{-i\mathbf{q}\cdot\mathbf{r}} \right| n_2 \right\rangle \frac{A_\lambda^{(n_1,n_2)}}{E_\lambda - \omega - i\eta} \right.\right.

$$\times \mathsf{S}^{-1}_{\lambda,\lambda'} \sum_{n_3,n_4} \left\langle n_4 \left| e^{i\mathbf{q}\cdot\mathbf{r}'} \right| n_3 \right\rangle A^{*(n_3,n_4)}_{\lambda'} \left(f_{n_4} - f_{n_3}\right) \Bigg] \Bigg\}. \tag{3.87}$$

We have added a small imaginary constant $i\eta$ to the frequency ω in order to smooth the curve by a Lorentzian shape and to shift the poles away from the real axis.

If we consider only the Hermitian resonant part of $\bar{H}^{\text{exc}}$, we obtain the simpler formula for the imaginary part,

$$\varepsilon_2(\omega) = 2 \lim_{\mathbf{q}\to 0} v_0(\mathbf{q}) \sum_{\lambda} \text{Im} \frac{\left| \sum_{v,c,\mathbf{k}} \langle v\mathbf{k} - \mathbf{q} \left| e^{-i\mathbf{q}\cdot\mathbf{r}} \right| c\mathbf{k} \rangle A^{(vc\mathbf{k})}_{\lambda} \right|^2}{E_\lambda - \omega - i\eta}. \tag{3.88}$$

The factor 2 considers summation over spin. The exciton wavefunction coefficients $A^{(vc\mathbf{k})}_{\lambda}$ in Eq. (3.88) mix the independent-particle transitions in Eq. (3.4).

Eqs. (3.70) and (3.88) constitute a set of equations which has been frequently used in the non-*ab initio* framework [59, Chap. 6][62]. Here, it appears as a particular approximation to the more general formula Eq. (3.87), with well-defined *ab initio* ingredients, which are consistent with the *GW* approach.

Indeed, we can also include the anti-resonant part of the excitonic Hamiltonian without explicitly solving for the full matrix in the evaluation of the macroscopic dielectric function. Neglecting the coupling parts in the full Hamiltonian, Eq. (3.83), we get two block-diagonal terms, the resonant with solutions λ and the anti-resonant part with solutions $\bar{\lambda}$. From $\bar{H}^{\text{res}} = -\left[\bar{H}^{\text{anti-res}}\right]^*$, we see $E_{\bar{\lambda}} = -E_\lambda$ and $A_{\bar{\lambda}} = A^*_\lambda$. Thus, the macroscopic dielectric function, including the anti-resonant part, but excluding the coupling parts of $\bar{H}^{\text{exc}}$, is

$$\varepsilon_{\text{M}}(\omega) = 1 + 2 \lim_{\mathbf{q}\to 0} v_0(\mathbf{q}) \sum_{\lambda} \Bigg\{ \left| \sum_{v,c,\mathbf{k}} \langle v\mathbf{k} - \mathbf{q} \left| e^{-i\mathbf{q}\cdot\mathbf{r}} \right| c\mathbf{k} \rangle A^{(vc\mathbf{k})}_{\lambda} \right|^2 \times \left(\frac{1}{E_\lambda - (\omega + i\eta)} + \frac{1}{E_\lambda + (\omega + i\eta)} \right) \Bigg\}. \tag{3.89}$$

The energy shift η has also been retained for negative transition energies in order to account for finite life-time effects.

We can obtain the real part of ε_{M} by making a Kramers-Kronig transformation,

$$\varepsilon_1(\omega) = 1 + \frac{1}{\pi} \mathfrak{P} \int_{-\infty}^{+\infty} d\zeta \, \frac{\varepsilon_2(\omega)}{\zeta - \omega}, \tag{3.90}$$

which gives, using $\varepsilon_2(\omega) = -\varepsilon_2(-\omega)$ and an infinitesimal energy shift η for the absorption frequency,

$$\varepsilon_1(\omega) = 1 + \frac{2}{\pi} \int_{0}^{+\infty} d\zeta \, \frac{\zeta \varepsilon_2(\omega)}{\zeta^2 - \omega^2 - 2i\eta\omega}. \tag{3.91}$$

In the above equation for $\varepsilon_1(\omega)$ we have considered the negative frequencies. Thus starting from $\varepsilon_2(\omega)$, obtained from Eq. (3.88), – only taking the resonant term –, we include the anti-resonant term in $\varepsilon_1(\omega)$, though neglecting the coupling terms of the excitonic Hamiltonian.

3.2.3 Formulation in Momentum Space

The computational implementation has profited from the *GW* code of Rex W. Godby and R. J. Needs [6, 47] and the exciton code of Giovanni Onida *et al.* [48] for clusters, where only the Γ point has been programmed for the $\mathbf{k}$ point sampling, since finite systems do not have a dispersion. The original program written by Stefan Albrecht *et al.* to find the optical absorption onset of lithium oxide (see Appendix A) has served as a basis for the developments made in this thesis.

Spin is included in the summation.

Resonant Part

We will first consider the resonant part and split the excitonic Hamiltonian Eq. (3.79) into

$$\bar{H}^{\text{exc,res}} = H^{\text{diag,res}} + \bar{H}^{\text{exch,res}} + H^{\text{Coul,res}}. \tag{3.92}$$

Diagonal Term The diagonal part is trivial,

$$H^{\text{diag,res}}_{(vc\mathbf{k}),(v'c'\mathbf{k}')} = (\epsilon_{c\mathbf{k}} - \epsilon_{v\mathbf{k}})\, \delta_{v,v'} \delta_{c,c'} \delta_{\mathbf{k},\mathbf{k}'}. \tag{3.93}$$

As already discussed above, we use the quasiparticle energies from a *GW* calculation as the independent-particle transition energies.

Exchange Term The unscreened short-range exchange term is readily evaluated. Using the Eqs. (2.31) and (3.3), we obtain

$$\bar{H}^{\text{exch,res}}_{(vc\mathbf{k}),(v'c'\mathbf{k}')} = 2\frac{4\pi}{\Omega} \sum_{\mathbf{G}\neq\mathbf{0}} \frac{1}{|\mathbf{G}|^2} \left\langle c\mathbf{k} \left| e^{i\mathbf{G}\cdot\mathbf{r}} \right| v\mathbf{k} \right\rangle \left\langle v'\mathbf{k}' \left| e^{-i\mathbf{G}\cdot\mathbf{r}} \right| c'\mathbf{k}' \right\rangle. \tag{3.94}$$

The volume of the crystal is $\Omega = N_{\mathbf{k}}\Omega_0$, with $N_{\mathbf{k}}$ being the number of $\mathbf{k}$ points in the sampling over the Brillouin zone.

The inclusion of the nonanalytic term for $\mathbf{G} = \mathbf{0}$ in Eq. (3.94) would account for the transverse-longitudinal splitting [59, Chap. 6], which is not relevant for absorption spectra. It is nevertheless observable for bound states which occur within the gap [23] or discrete states of finite systems like clusters and has to be evaluated in the limit of vanishing wavevector according to Eq. (3.20). This considers the dependence of the transverse-longitudinal splitting on the direction of the electric dipole moment Eq. (3.12) and the excitonic wavevector $\mathbf{k}_{\text{exciton}} = \mathbf{k}_c - \mathbf{k}_v$ relative to each other.

For $\mathbf{G} \neq \mathbf{0}$ the matrix elements are evaluated applying FFT routines,

$$\tilde{\rho}_{\substack{n_1\mathbf{k}' \\ n_2\mathbf{k}}}(\mathbf{G}) \equiv \left\langle n_1\mathbf{k}' \left| e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} \right| n_2\mathbf{k} \right\rangle = \int d\mathbf{r}\; u^*_{n_1\mathbf{k}'}(\mathbf{r}) u_{n_2\mathbf{k}}(\mathbf{r})\; e^{-i\mathbf{G}\cdot\mathbf{r}}, \tag{3.95}$$

with $\mathbf{k}' = \mathbf{k} - \mathbf{q}$.

Coulomb Term The screened Coulomb term is obtained using Eq. (2.56) for the inverse dielectric function in momentum space,

$$H^{\text{Coul,res}}_{(vc\mathbf{k}),(v'c'\mathbf{k}')} = -\frac{4\pi}{\Omega} \sum_{\mathbf{G},\mathbf{G}'} \frac{\tilde{\varepsilon}^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q})}{|\mathbf{q}+\mathbf{G}|\,|\mathbf{q}+\mathbf{G}'|} \left\langle c\mathbf{k} \left| e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} \right| c'\mathbf{k}' \right\rangle \times \left\langle v'\mathbf{k}' \left| e^{-i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'} \right| v\mathbf{k} \right\rangle \cdot \delta_{\mathbf{q},\mathbf{k}-\mathbf{k}'}, \tag{3.96}$$

where we have defined the *symmetrized inverse dielectric matrix* $\tilde{\varepsilon}^{-1}$ by

$$\tilde{\varepsilon}^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) \equiv \varepsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) \frac{|\mathbf{q}+\mathbf{G}'|}{|\mathbf{q}+\mathbf{G}|}. \tag{3.97}$$

Furthermore, we have used Eq. (2.36) to transform the integral over the Brillouin zone into a sum.

The case of $\mathbf{q} = \mathbf{0}$ in Eq. (3.96) has to be considered separately. We have to retain the integration in Eq. (2.56) and examine the limit of $\mathbf{q} \to 0$. From Eqs. (2.57), (3.3) and (3.20) we can estimate for cubic systems the behavior of $\tilde{\varepsilon}^{-1}$ in the RPA formulation Eq. (2.59a),

$$\tilde{\varepsilon}^{-1}_{00}(\mathbf{q} \to 0) \propto 1 + \frac{q^2}{q^2} \propto O(1), \tag{3.98a}$$

$$\tilde{\varepsilon}^{-1}_{0\mathbf{G}'}(\mathbf{q} \to 0) \propto \frac{-\mathbf{q}}{|\mathbf{q}|\,|\mathbf{q}+\mathbf{G}'|} \propto O\left(\frac{-\mathbf{q}}{q}\right), \tag{3.98b}$$

$$\tilde{\varepsilon}^{-1}_{\mathbf{G}0}(\mathbf{q} \to 0) \propto \frac{\mathbf{q}}{|\mathbf{q}+\mathbf{G}|\,|\mathbf{q}|} \propto O\left(\frac{\mathbf{q}}{q}\right), \tag{3.98c}$$

$$\tilde{\varepsilon}^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q} \to 0) \propto \delta_{\mathbf{G},\mathbf{G}'} + 1 \propto O(1). \tag{3.98d}$$

Together with the limiting values of the matrix elements and the denominator of $H^{\text{Coul,res}}$ in Eq. (3.96) we find that the head ($\mathbf{G} = \mathbf{0}$, $\mathbf{G}' = \mathbf{0}$) contributes only notably to ($c = c'$, $v = v'$). This part is analytically integrated out in the manner [5]

$$F = \frac{1}{\Omega} \sum_{\mathbf{q}} \frac{f(\mathbf{q})}{q^2} = f(\mathbf{q}=\mathbf{0}) I_{\text{sz}} + \frac{1}{\Omega} \sum_{\mathbf{q}\neq\mathbf{0}} \frac{f(\mathbf{q})}{q^2}. \tag{3.99}$$

I_{sz} is the integral of $1/q^2$ over a chosen volume of size $V_{\text{BZ}}/N_{\mathbf{k}}$ around $\mathbf{q} = \mathbf{0}$. For the easiest assumption of a small sphere with radius h, we have

$$I_{\text{sz}} = \frac{1}{(2\pi)^3} \int_{V_{\text{BZ}}/N_{\mathbf{k}}} d\mathbf{q}\; \frac{1}{q^2} = \frac{4\pi h}{(2\pi)^3} = \frac{1}{2\pi^2} \sqrt[3]{\frac{6\pi^2}{\Omega}} = \frac{0.197}{\sqrt[3]{\Omega}}. \tag{3.100}$$

The correct consideration of the shape of the $\mathbf{k}$ space for the $\mathbf{q} = \mathbf{0}$ contribution due to the special $\mathbf{k}$ point sampling scheme of Monkhorst and Pack [39] in the integration leads to a factor of 0.188 instead of 0.197.

All other cases for the head as well as the wings ($\mathbf{G} = \mathbf{0}$, $\mathbf{G}' \neq \mathbf{0}$ or $\mathbf{G} \neq \mathbf{0}$, $\mathbf{G}' = \mathbf{0}$) either average to zero or vanish in the limit of a large $\mathbf{k}$ point sampling set ($N_\mathbf{k} \to \infty$). In particular, the case ($c \neq c'$, $v \neq v'$) for the head, which gives – if at all – a very small contribution, is set to zero, since a correct treatment for degenerate bands demands for rather involved second order perturbation theory.

Coupling Part

The coupling part of the excitonic Hamiltonian Eq. (3.79) is split like

$$\bar{H}^{\text{coupling}} = \bar{H}^{\text{exch,coupling}} + H^{\text{Coul,coupling}}. \tag{3.101}$$

Exchange Term The unscreened short-range exchange term is, from Eq. (3.94),

$$\bar{H}^{\text{exch,coupling}}_{(vc\mathbf{k}),(c'v'\mathbf{k}')} = 2\frac{4\pi}{\Omega} \sum_{\mathbf{G}\neq\mathbf{0}} \frac{1}{|\mathbf{G}|^2} \left\langle c\mathbf{k} \left| e^{i\mathbf{G}\cdot\mathbf{r}} \right| v\mathbf{k} \right\rangle \left\langle c'\mathbf{k}' \left| e^{-i\mathbf{G}\cdot\mathbf{r}} \right| v'\mathbf{k}' \right\rangle. \tag{3.102}$$

The inclusion of the $\mathbf{G} = \mathbf{0}$ term in Eq. (3.102) would have the same implications as for the resonant part.

Coulomb Term The screened Coulomb term gives, from Eq. (3.96),

$$\begin{aligned} H^{\text{Coul,coupling}}_{(vc\mathbf{k}),(c'v'\mathbf{k}')} = -\frac{4\pi}{\Omega} \sum_{\mathbf{G},\mathbf{G}'} \frac{\tilde{\varepsilon}^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q})}{|\mathbf{q}+\mathbf{G}|\,|\mathbf{q}+\mathbf{G}'|} \left\langle c\mathbf{k} \left| e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} \right| v'\mathbf{k}' \right\rangle \\ \times \left\langle c'\mathbf{k}' \left| e^{-i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'} \right| v\mathbf{k} \right\rangle \cdot \delta_{\mathbf{q},\mathbf{k}-\mathbf{k}'}. \end{aligned} \tag{3.103}$$

There is no contribution to the case of $\mathbf{q} = \mathbf{0}$ in Eq. (3.103) for the head and the wings. All terms either average to zero or vanish in the limit of a large $\mathbf{k}$ point sampling set, similar to what has been said for Eq. (3.96).

3.2.4 Reduction of Exciton Hamiltonian

The Hamiltonian matrix for the effective eigenvalue equation Eq. (3.70) can be quite large, up to some ten thousands of elements for materials with a large dispersion like silicon. We can reduce the dimension of the matrix by applying symmetry operations of the crystal (see also Section 2.2.5). Instead of solving one large matrix, we treat several small ones.

In general, we consider only those operations which form an Abelian subgroup of the point group, and which altogether allow to reconstruct the whole zone from a

corresponding reduced part. In the case of the diamond group, we find it convenient to use the 180^0 rotations R around the x and the y axis, respectively. These two operations allow us to break the equation $H_{\mathbf{k}\mathbf{k}'}A^{\mathbf{k}'} = EA^{\mathbf{k}}$ (band indices have been suppressed, and repeated indices are summed over) into four equations to be used for points $\mathbf{k}_i$ in a reduced zone only. These equations are of the form $h_{\mathbf{k}_i\mathbf{k}'_i}a^{\mathbf{k}'i} = Ea^{\mathbf{k}_i}$, where $h_{\mathbf{k}_i\mathbf{k}'_i} = H_{\mathbf{k}_i\mathbf{k}'_i} \pm H_{\mathbf{k}_i\mathsf{R}\mathbf{k}'_i}$. The $A^{\mathbf{k}}$ are then reconstructed from the reduced eigenvectors $a^{\mathbf{k}_i}$. Moreover, we apply time reversal and Hermiticity in order to accelerate the calculation of the matrix elements.

In the following we will first examine the symmetry properties of the Hamiltonian H in more detail and then show explicitly, how to construct the small Hamiltonian matrices h, and how to get original $A^{\mathbf{k}}$ from the reduced eigenvectors $a^{\mathbf{k}_i}$. Finally we will consider several approximations in order to speed up the calculation and/or to further reduce H.

Symmetry Properties of the Hamiltonian

The Hamiltonian in momentum space has as essential ingredients the diagonal part, the matrix elements $\tilde{\rho}$ and the inverse dielectric matrix $\tilde{\varepsilon}^{-1}$. We will consider each part separately in the following to find a relation of the form $H_{\mathbf{p}\mathbf{k}\mathbf{p}\mathbf{k}'} = H_{\mathbf{k}\mathbf{k}'}$.

Diagonal Part This part is trivial as $\epsilon_{n\mathbf{k}} = \epsilon_{n\mathbf{p}\mathbf{k}}$ from Eq. (2.40).

Matrix elements The application of symmetry operations to the matrix elements is complicated for degenerate bands and due to the fact that the operations used to expand a $\mathbf{k}$ point from the irreducible Brillouin zone into the full Brillouin to a full star are not uniquely defined (see also Eq. (2.39) and thereafter). We refer to Eq. (3.95).

In order to derive the rules for matrix elements we take $\mathbf{k} = \mathsf{D}\mathbf{k}_{\mathrm{IBZ}}$ and $\mathbf{k}' = \mathsf{D}'\mathbf{k}'_{\mathrm{IBZ}}$, with D and D′ being rotations of the point group. Then we get, applying Eq. (2.39),

$$\begin{aligned}\underset{n\mathbf{k}}{\tilde{\rho}_{n'\mathbf{k}'}}(\mathbf{G}) &= \int d\mathbf{r}\ \varphi^*_{n'\mathbf{k}'}(\mathbf{r})\ e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}\ \varphi_{n\mathbf{k}}(\mathbf{r}) \\ &= \int d\mathbf{r}\ \varphi^*_{n'\mathbf{k}'_{\mathrm{IBZ}}}(\mathsf{D}'^{-1}\mathbf{r})\ e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}\ \varphi_{n\mathbf{k}_{\mathrm{IBZ}}}(\mathsf{D}^{-1}\mathbf{r}). \qquad (3.104)\end{aligned}$$

The phases are deliberately chosen as indicated, $\varphi_{n\mathsf{D}\mathbf{k}_{\mathrm{IBZ}}}(\mathbf{r}) = \varphi_{n\mathbf{k}_{\mathrm{IBZ}}}(\mathsf{D}^{-1}\mathbf{r})$.

We want to find a rule to go from $\underset{n\mathbf{k}}{\tilde{\rho}_{n'\mathsf{D}'\mathbf{k}'_{\mathrm{IBZ}}}}$ to $\underset{n\mathsf{D}'^{-1}\mathbf{k}}{\tilde{\rho}_{n'\mathbf{k}'_{\mathrm{IBZ}}}}$, thus we transform $\mathbf{r} \to \mathsf{D}'\mathbf{r}$ in the above equation. In the case of degenerate bands, which we consider now, this leads to a linear combination of bands for the degenerate wavefunction as given by Eq. (2.41).

For ease we assume here that n at $\mathbf{k}$ is degenerate, while n' at $\mathbf{k}'$ is not. Furthermore, we demand $\mathsf{D} \neq \mathsf{D}'$. Then we obtain from the last expression in Eq. (3.104),

$$\begin{aligned} \underset{n\mathbf{k}}{\tilde{\rho}_{n'\mathsf{D}'\mathbf{k}'_{\mathrm{IBZ}}}}(\mathbf{G}) &= \int d\mathsf{D}'\mathbf{r}\, \varphi^*_{n'\mathbf{k}'_{\mathrm{IBZ}}}(\mathbf{r})\, e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathsf{D}'\mathbf{r}}\, \tilde{\varphi}_{n\mathbf{k}_{\mathrm{IBZ}}}(\underbrace{\mathsf{D}^{-1}\mathsf{D}'}_{\equiv\mathsf{T}^{-1}}\mathbf{r}) \\ &\neq \underset{n\underbrace{\mathsf{D}'^{-1}\mathbf{k}}_{\equiv\tilde{\mathbf{k}}}}{\tilde{\rho}_{n'\mathbf{k}'_{\mathrm{IBZ}}}}(\mathsf{D}'^{-1}\mathbf{G}), \end{aligned} \tag{3.105}$$

with $\tilde{\mathbf{k}} \equiv \mathsf{D}'^{-1}\mathbf{k} = \mathsf{D}'^{-1}\mathsf{D}\mathbf{k}_{\mathrm{IBZ}} = \mathsf{T}\mathbf{k}_{\mathrm{IBZ}}$. The new wavefunction $\tilde{\varphi}_{n\mathbf{k}_{\mathrm{IBZ}}}$ is the result of the linear combination of degenerate bands, Eq. (2.41). We have furthermore used

$$(\mathbf{q}+\mathbf{G})\cdot\mathsf{D}'\mathbf{r} = \mathsf{D}'^{-1}(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}. \tag{3.106}$$

It is essential to note that the inequality in Eq. (3.105) is due to the fact that $\underset{n\mathsf{D}'^{-1}\mathbf{k}}{\tilde{\rho}_{n'\mathbf{k}'_{\mathrm{IBZ}}}}(\mathsf{D}'^{-1}\mathbf{G})$ is indeed consistently taken as,

$$\underset{n\mathsf{D}'^{-1}\mathbf{k}}{\tilde{\rho}_{n'\mathbf{k}'_{\mathrm{IBZ}}}}(\mathsf{D}'^{-1}\mathbf{G}) = \int d\mathbf{r}\, \varphi^*_{n'\mathbf{k}'_{\mathrm{IBZ}}}(\mathbf{r})\, e^{-i\mathsf{D}'^{-1}(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}\, \varphi_{n\mathbf{k}_{\mathrm{IBZ}}}(\mathsf{T}^{-1}\mathbf{r}). \tag{3.107}$$

Evidently, we must be very careful when applying symmetry operations to $\tilde{\rho}$ and we will consider several cases in the following. Immediately we see from the construction of the $\tilde{\rho}$ that we have in the special case $\mathsf{D} = \mathsf{D}'$, and also for degenerate bands,

$$\underset{n\mathsf{D}\mathbf{k}_{\mathrm{IBZ}}}{\tilde{\rho}_{n'\mathsf{D}\mathbf{k}'_{\mathrm{IBZ}}}}(\mathbf{G}) = \underset{n\mathbf{k}_{\mathrm{IBZ}}}{\tilde{\rho}_{n'\mathbf{k}'_{\mathrm{IBZ}}}}(\mathsf{D}^{-1}\mathbf{G}). \tag{3.108}$$

This is very useful and easy to apply for $\mathbf{q} = \mathbf{0}$, *i.e.*, $\mathbf{k}'_{\mathrm{IBZ}} = \mathbf{k}_{\mathrm{IBZ}}$.

The freedom of phase of $\underset{n\mathsf{D}'^{-1}\mathbf{k}}{\tilde{\rho}_{n'\mathbf{k}'_{\mathrm{IBZ}}}}$ in the above example, Eq. (3.105), for the case of degenerate bands is restricted when $\mathsf{D}'^{-1} = \mathsf{D}'$, since twice the application of D' on $\mathbf{k}'$ must then lead back to the identical old phase of $\underset{n\mathbf{k}}{\tilde{\rho}_{n'\mathbf{k}'}}$. This argument leads us to the rule for degenerate bands and could be generalized for an Abelian subgroup of the point group.

Still, however, there might be some ambiguity about the rotation T in Eq. (3.105), since it is not necessarily unique. As it will be discussed below, this does not cause any problems in practice, if we expand the IBZ into the full BZ by first applying all rotations R and then the inversion plus R. Then the choice of phase is consistently fixed for the wavefunctions with $\tilde{\mathbf{k}} = \mathsf{T}\mathbf{k}_{\mathrm{IBZ}}$.

Thus, for *non-degenerate* bands in general and all rotations $\mathsf{R}^{-1} = \mathsf{R}$ in particular we obtain the rule

$$\underset{n\mathsf{R}\mathbf{k}}{\tilde{\rho}_{n'\mathsf{R}\mathbf{k}'}}(\mathbf{G}) = \underset{n\mathbf{k}}{\tilde{\rho}_{n'\mathbf{k}'}}(\mathsf{R}^{-1}\mathbf{G}). \tag{3.109}$$

The *inversion* has to be applied with care. The phase is only then consistently chosen, if we use time reversal only in those cases, where the $\mathbf{k}$ points cannot be related by a proper rotation. Therefore, we must first apply all rotations and then the rotations plus the inversion (time reversal) in order to expand the IBZ into the full BZ, as we have already noted above. This insures a correct book-keeping of the rotations D in the relations $\mathbf{k}_{\mathrm{BZ}} = \mathsf{D}\mathbf{k}_{\mathrm{IBZ}}$. We have from Eq. (2.43),

$$\underset{n-\mathbf{k}}{\tilde{\rho}_{n'-\mathbf{k}'}}(\mathbf{G}) = \underset{n\mathbf{k}}{\tilde{\rho}^{*}_{n'\mathbf{k}'}}(-\mathbf{G}). \tag{3.110}$$

In practice we have only a finite set of $\mathbf{q}$ vectors. Forming the difference

$$\mathbf{q} = \mathbf{k} - \mathbf{k}' \tag{3.111}$$

may lead to a vector outside the first Brillouin zone. However it can be translated back by an appropriately chosen reciprocal lattice vector $\mathbf{G}_0$, since

$$\underset{n\mathbf{k}}{\tilde{\rho}_{n'\mathbf{k}-\mathbf{q}}}(\mathbf{G}) = \underset{n\mathbf{k}}{\tilde{\rho}_{n'\mathbf{k}-\mathbf{q}-\mathbf{G}_0}}(\mathbf{G} - \mathbf{G}_0). \tag{3.112}$$

Finally, we would like to consider *non-symmorphic* operations as well, although we are not employing them in the program. For a space group operator $\{\mathsf{p} \mid \mathbf{w}\}$ we obtain for non-degenerate bands, using Eq. (2.39), and following the same reasoning as above,

$$\underset{n\mathsf{p}\mathbf{k}}{\tilde{\rho}_{n'\mathsf{p}\mathbf{k}'}}(\mathbf{G}) = \underset{n\mathbf{k}}{\tilde{\rho}_{n'\mathbf{k}'}}(\mathsf{p}^{-1}\mathbf{G})\; e^{-i(\mathsf{p}\mathbf{q}+\mathbf{G})\cdot\mathbf{w}}. \tag{3.113}$$

Inverse Dielectric Function In real space the dielectric matrix has the space-group symmetry of the crystal [42], as applied to both arguments. The periodicity is, with $\mathbf{R}_i$ being a lattice translation vector,

$$\varepsilon^{-1}(\mathbf{r},\mathbf{r}') = \varepsilon^{-1}(\mathbf{r}+\mathbf{R}_i, \mathbf{r}'+\mathbf{R}_i). \tag{3.114}$$

The space group symmetry for the operator $\{\mathsf{p} \mid \mathbf{w}\}$ is given by

$$\varepsilon^{-1}(\mathbf{r},\mathbf{r}') = \varepsilon^{-1}(\mathsf{p}\mathbf{r}+\mathbf{w}, \mathsf{p}\mathbf{r}'+\mathbf{w}). \tag{3.115}$$

We can readily obtain the properties in reciprocal space by inspecting Eq. (2.56),

$$\varepsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathsf{p}\mathbf{q}) = e^{i(\mathbf{G}'-\mathbf{G})\cdot\mathbf{w}}\varepsilon^{-1}_{\mathsf{p}^{-1}\mathbf{G}\mathsf{p}^{-1}\mathbf{G}'}(\mathbf{q}). \tag{3.116}$$

If we use only the rotations of the point group, we need the inversion from time reversal,

$$\varepsilon^{-1}_{\mathbf{G}\mathbf{G}'}(-\mathbf{q}) = \left[\varepsilon^{-1}_{-\mathbf{G}-\mathbf{G}'}(\mathbf{q})\right]^{*}, \tag{3.117}$$

since $\varepsilon^{-1}(\mathbf{r},\mathbf{r}')$ is real.

The length $|\mathbf{q}+\mathbf{G}|$ changes like

$$|\mathsf{p}\mathbf{q}+\mathbf{G}| = \left|\mathbf{q}+\mathsf{p}^{-1}\mathbf{G}\right|. \tag{3.118}$$

Hamiltonian Matrix From the above relations it is straightforward to obtain the formula for a space group operator $\{\mathsf{p} \mid \mathbf{w}\}$ and non-degenerate bands,

$$H_{\mathbf{pkpk'}} = H_{\mathbf{kk'}}, \tag{3.119}$$

where we have suppressed band indices. Thus we can obtain by symmetry the elements of H in the row $\mathbf{pk}$ from the row $\mathbf{k}$, since from the above Eq. (3.119),

$$H_{\mathbf{pkk'}} = H_{\mathbf{kp^{-1}k'}}. \tag{3.120}$$

It must be noted that in going from $\mathbf{k}$ to $\mathbf{pk}$ we must stay within the Brillouin zone sampling as used in the program. For example, in a face-centered cubic lattice, the X point $\mathrm{X}_1 = (1,0,0)$ (in units of $\frac{2\pi}{a_0}$, with a_0 being the lattice constant) is equivalent to within a reciprocal lattice vector $\mathbf{G}_0$ to the point $\mathrm{X}_2 = (-1,0,0)$. Instead of taking X_1 and X_2 into the $\mathbf{k}$ point sampling with each weight one half, we may find it more convenient to take only one of them, say X_1, with weight one. Then, however, the inversion of X_1 leads to X_2, not being part of the $\mathbf{k}$ point sampling of the BZ. In these and corresponding cases the Eq. (3.119) is not valid.

We only use the rotations of the point group and restrict us furthermore to rotations $\mathsf{R}^{-1} = \mathsf{R}$ in order to account also for degenerate bands,

$$H_{\mathbf{RkRk'}} = H_{\mathbf{kk'}}. \tag{3.121}$$

The inversion is applied as time-reversal as done for the wavefunctions,

$$H_{\mathbf{-k-k'}} = [H_{\mathbf{kk'}}]^*. \tag{3.122}$$

We remind that the resonant part $H^{\mathrm{exc,res}}_{(vc\mathbf{k}),(v'c'\mathbf{k'})}$, Eq. (3.92), alone of the Hamiltonian is Hermitian, thus

$$H^{\mathrm{exc,res}}_{(vc\mathbf{k}),(v'c'\mathbf{k'})} = \left[H^{\mathrm{exc,res}}_{(v'c'\mathbf{k'}),(vc\mathbf{k})}\right]^*, \tag{3.123}$$

while the coupling part $H^{\mathrm{coupling}}_{(vc\mathbf{k}),(c'v'\mathbf{k'})}$, Eq. (3.101), is symmetric,

$$H^{\mathrm{coupling}}_{(vc\mathbf{k}),(c'v'\mathbf{k'})} = H^{\mathrm{coupling}}_{(c'v'\mathbf{k'}),(vc\mathbf{k})}. \tag{3.124}$$

In practice we only apply Hermiticity and time-reversal explicitly. The relation $H_{\mathbf{RkRk'}} = H_{\mathbf{kk'}}$ of Eq. (3.121) is implicitly used in order to reduce the Hamiltonian by symmetry as explained in the next section.

It is interesting to look on the properties of H^{-1}. Since $H^{-1}_{\mathbf{kk'}} H_{\mathbf{k'k_2}} = \delta_{\mathbf{kk_2}} = \delta_{\mathbf{pkpk_2}}$, we have

$$\sum_{\mathsf{D'k'_{IBZ}}} H^{-1}_{\mathbf{kD'k'_{IBZ}}} H_{\mathbf{D'k'_{IBZ}k_2}} = \sum_{\mathsf{D'k'_{IBZ}}} H^{-1}_{\mathbf{kD'k'_{IBZ}}} H_{\mathbf{k'_{IBZ}D'^{-1}k_2}} \stackrel{!}{=} \delta_{\mathbf{D'^{-1}kD'^{-1}k_2}}. \tag{3.125}$$

Thus,

$$H^{-1}_{\mathbf{kD'k'_{IBZ}}} = H^{-1}_{\mathbf{D'^{-1}kk'_{IBZ}}}, \tag{3.126}$$

and we have the same symmetry properties for H^{-1} as for H.

Reduction by Symmetry

Until now, we have found that, when degenerate bands are represented in the $\mathbf{k}$ point sampling, only operations from an Abelian subgroup of the space group can be employed for the reduction of the dimension, and in a different manner, only the inversion by time reversal and the Hermiticity are applicable to construct parts of the Hamiltonian by symmetry without explicitly calculating them. In the following we explain the reduction of the dimension of the Hamiltonian matrix by symmetry.

As already mentioned in the introduction to this section we choose the 180^0 rotations around the x and the y axis, respectively, for the reduction of H. They fulfill $\mathsf{R}^{-1} = \mathsf{R}$ and form an Abelian subgroup together with the 180^0 rotation around the z axis and the identity.

Principle Procedure The idea is to split the $\mathbf{k}$ space into the points $\mathbf{k}_i$ in a reduced zone and $\mathsf{R}\mathbf{k}_i$ in the rest. Here we first assume for ease $\mathsf{R}\mathbf{k}_i \neq \mathbf{k}_i$. Later we will drop this restriction. The Schrödinger equation $H_{\mathbf{k}\mathbf{k}'} A_\lambda^{\mathbf{k}'} = E_\lambda A_\lambda^{\mathbf{k}}$ can then be rewritten as

$$i) \quad \sum_{\mathbf{k}_i'} \left[H_{\mathbf{k}_i\mathbf{k}_i'} A_\lambda^{\mathbf{k}_i'} + H_{\mathbf{k}_i\mathsf{R}\mathbf{k}_i'} A_\lambda^{\mathsf{R}\mathbf{k}_i'} \right] = E_\lambda A_\lambda^{\mathbf{k}_i}, \tag{3.127a}$$

$$ii) \quad \sum_{\mathbf{k}_i'} \left[H_{\mathsf{R}\mathbf{k}_i\mathbf{k}_i'} A_\lambda^{\mathbf{k}_i'} + H_{\mathsf{R}\mathbf{k}_i\mathsf{R}\mathbf{k}_i'} A_\lambda^{\mathsf{R}\mathbf{k}_i'} \right] = E_\lambda A_\lambda^{\mathsf{R}\mathbf{k}_i}. \tag{3.127b}$$

With $H_{\mathsf{R}\mathbf{k}\mathsf{R}\mathbf{k}'} = H_{\mathbf{k}\mathbf{k}'}$ and $\mathsf{R}^{-1} = \mathsf{R}$ we can add $i)$ and $ii)$ to obtain

$$h_{\mathbf{k}_i\mathbf{k}_i'} \equiv H_{\mathbf{k}_i\mathbf{k}_i'} + H_{\mathbf{k}_i\mathsf{R}\mathbf{k}_i'}, \tag{3.128a}$$

$$a_\lambda^{\mathbf{k}_i} \equiv A_\lambda^{\mathbf{k}_i} + A_\lambda^{\mathsf{R}\mathbf{k}_i}, \tag{3.128b}$$

$$h_{\mathbf{k}_i\mathbf{k}_i'} a_\lambda^{\mathbf{k}_i'} = E_\lambda a_\lambda^{\mathbf{k}_i}, \tag{3.128c}$$

and subtract $ii)$ from $i)$ to get

$$\tilde{h}_{\mathbf{k}_i\mathbf{k}_i'} \equiv H_{\mathbf{k}_i\mathbf{k}_i'} - H_{\mathbf{k}_i\mathsf{R}\mathbf{k}_i'}, \tag{3.129a}$$

$$\tilde{a}_{\tilde{\lambda}}^{\mathbf{k}_i} \equiv A_{\tilde{\lambda}}^{\mathbf{k}_i} - A_{\tilde{\lambda}}^{\mathsf{R}\mathbf{k}_i}, \tag{3.129b}$$

$$\tilde{h}_{\mathbf{k}_i\mathbf{k}_i'} \tilde{a}_{\tilde{\lambda}}^{\mathbf{k}_i'} = E_{\tilde{\lambda}} \tilde{a}_{\tilde{\lambda}}^{\mathbf{k}_i}. \tag{3.129c}$$

Since normally, $E_\lambda \neq E_{\tilde{\lambda}}$, we can reconstruct the original eigenvectors $A_\lambda^{\mathbf{k}}$ from $a_\lambda^{\mathbf{k}_i}$. For the solutions λ we have $\tilde{a}_{\tilde{\lambda}}^{\mathbf{k}_i} \equiv 0$, and therefore with the correct normalization,

$$A_\lambda^{\mathbf{k}_i} = A_\lambda^{\mathsf{R}\mathbf{k}_i} = \frac{1}{\sqrt{2}} a_\lambda^{\mathbf{k}_i}. \tag{3.130}$$

Correspondingly, for the solutions $\tilde{\lambda}$ we get $a_\lambda^{\mathbf{k}_i} \equiv 0$, thus,

$$A_{\tilde{\lambda}}^{\mathbf{k}_i} = -A_{\tilde{\lambda}}^{\mathsf{R}\mathbf{k}_i} = \frac{1}{\sqrt{2}} a_{\tilde{\lambda}}^{\mathbf{k}_i}. \tag{3.131}$$

Now we might have a $\mathbf{k}$ point set, where there are some $\mathbf{k}_i$ with $\mathsf{R}\mathbf{k}_i = \mathbf{k}_i$. Then we must transform the Hamiltonian into block-diagonal form by an appropriate unitary mapping. We take as example three $\mathbf{k}$ points in the reduced zone with $\mathbf{k}_0 = \mathsf{R}\mathbf{k}_0$, $\mathbf{k}_2 = \mathsf{R}\mathbf{k}_1 \neq \mathbf{k}_1$. The Hamiltonian H looks like

$$H = \begin{pmatrix} H_{00} & H_{01} & H_{02} \\ H_{10} & H_{11} & H_{12} \\ H_{20} & H_{21} & H_{22} \end{pmatrix}, \tag{3.132}$$

with $H_{01} = H_{02}$, $H_{10} = H_{20}$, $H_{12} = H_{21}$ and $H_{11} = H_{22}$. H does not need to be Hermitian in general.

The transformation matrix S is

$$\mathsf{S} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}. \tag{3.133}$$

We have $\mathsf{S} = \mathsf{S}^{-1} = \mathsf{S}^T$ and $\det \mathsf{S} = -1$. The transformed Hamiltonian $H' = \mathsf{S}H\mathsf{S}^{-1}$ can be found as

$$H' = \begin{pmatrix} H_{00} & \sqrt{2}H_{01} & 0 \\ \sqrt{2}H_{10} & H_{11} + H_{12} & 0 \\ 0 & 0 & H_{11} - H_{12} \end{pmatrix}. \tag{3.134}$$

From there we can deduce the general rule how to reduce the Hamiltonian matrix from one large to two small matrices. For the solutions λ we have

$$h_{\mathbf{k}_i\mathbf{k}_i'} = \begin{cases} H_{\mathbf{k}_i\mathbf{k}_i'} + H_{\mathbf{k}_i\mathsf{R}\mathbf{k}_i'}, & \mathbf{k}_i \neq \mathsf{R}\mathbf{k}_i,\ \mathbf{k}_i' \neq \mathsf{R}\mathbf{k}_i' \\ \sqrt{2}H_{\mathbf{k}_i\mathbf{k}_i'}, & \mathbf{k}_i \neq \mathsf{R}\mathbf{k}_i,\ \mathbf{k}_i' = \mathsf{R}\mathbf{k}_i' \\ H_{\mathbf{k}_i\mathbf{k}_i'}, & \mathbf{k}_i = \mathsf{R}\mathbf{k}_i,\ \mathbf{k}_i' = \mathsf{R}\mathbf{k}_i' \end{cases}, \tag{3.135}$$

$$A_\lambda^{\mathbf{k}} : \begin{cases} A_\lambda^{\mathbf{k}_i} = A_\lambda^{\mathsf{R}\mathbf{k}_i} = \frac{1}{\sqrt{2}} a_\lambda^{\mathbf{k}_i}, & \mathbf{k}_i \neq \mathsf{R}\mathbf{k}_i \\ A_\lambda^{\mathbf{k}_i} = a_\lambda^{\mathbf{k}_i}, & \mathbf{k}_i = \mathsf{R}\mathbf{k}_i \end{cases}, \tag{3.136}$$

and for the solutions $\tilde{\lambda}$, we get

$$\tilde{h}_{\mathbf{k}_i\mathbf{k}_i'} = \begin{cases} H_{\mathbf{k}_i\mathbf{k}_i'} - H_{\mathbf{k}_i\mathsf{R}\mathbf{k}_i'}, & \mathbf{k}_i \neq \mathsf{R}\mathbf{k}_i,\ \mathbf{k}_i' \neq \mathsf{R}\mathbf{k}_i' \\ 0, & \text{otherwise} \end{cases}, \tag{3.137}$$

$$A_{\tilde{\lambda}}^{\mathbf{k}} : \begin{cases} A_{\tilde{\lambda}}^{\mathbf{k}_i} = -A_{\tilde{\lambda}}^{\mathsf{R}\mathbf{k}_i} = \frac{1}{\sqrt{2}} a_{\tilde{\lambda}}^{\mathbf{k}_i}, & \mathbf{k}_i \neq \mathsf{R}\mathbf{k}_i \\ A_{\tilde{\lambda}}^{\mathbf{k}_i} = 0, & \mathbf{k}_i = \mathsf{R}\mathbf{k}_i \end{cases}. \tag{3.138}$$

Setting the corresponding rows and columns of $\tilde{h}$ to zero, reduces the dimension of the matrix. We can either reorder the matrix to a smaller dimension or eliminate after the diagonalization the eigenvalues equal to zero, as they stem from exactly those rows and columns set to zero.

If we use special $\mathbf{k}$ points [38, 39], we can reduce the size of the Hamiltonian matrix by precisely two in each dimension, as we always have the case $\mathbf{k}_i \neq \mathsf{R}\mathbf{k}_i$.

Combination of Two Rotations We can successively apply twice the above outlined procedure with two 180^0 rotations R_1 and R_2. Then we can define four volumes in the Brillouin zone, V_{11} being the reduced zone, which are connected by

$$\mathsf{R}_1 V_{11} = V_{21}, \quad \mathsf{R}_1 V_{12} = V_{22}, \tag{3.139a}$$
$$\mathsf{R}_2 V_{11} = V_{12}, \quad \mathsf{R}_2 V_{21} = V_{22}, \tag{3.139b}$$

since $\mathsf{R}_1\mathsf{R}_2 = \mathsf{R}_2\mathsf{R}_1$. First we will apply R_1,

$$h^1_{\mathbf{k}_i\mathbf{k}'_i}(V_{11}, V_{11}) = H_{\mathbf{k}_i\mathbf{k}'_i}(V_{11}, V_{11}) + \mathrm{sgn}(\mathsf{R}_1) H_{\mathbf{k}_i\mathsf{R}_1\mathbf{k}'_i}(V_{11}, V_{21}), \tag{3.140a}$$
$$h^1_{\mathbf{k}_i\mathsf{R}_2\mathbf{k}'_i}(V_{11}, V_{12}) = H_{\mathbf{k}_i\mathsf{R}_2\mathbf{k}'_i}(V_{11}, V_{12}) + \mathrm{sgn}(\mathsf{R}_1) H_{\mathbf{k}_i\mathsf{R}_1\mathsf{R}_2\mathbf{k}'_i}(V_{11}, V_{22}). \tag{3.140b}$$

The corresponding volumes of the $\mathbf{k}$ points are indicated in the brackets. $\mathrm{sgn}(\mathsf{R}_1)$ stands for the sign $+$ or $-$. The respective Eqs. (3.135) or (3.137) have to be used.

Thus we only have to calculate one forth of the Hamiltonian, namely the part $H(V_{11}, V_{11} \ldots V_{22})$. Also for that small part we can still apply time-reversal and Hermiticity to reduce the computational effort.

In the next step we combine the two h^1 with the same $\mathrm{sgn}(\mathsf{R}_1)$ using rotation R_2,

$$h^2_{\mathbf{k}_i\mathbf{k}'_i}(V_{11}, V_{11}) = h^1_{\mathbf{k}_i\mathbf{k}'_i}(V_{11}, V_{11}) + \mathrm{sgn}(\mathsf{R}_2) h^1_{\mathbf{k}_i\mathsf{R}_2\mathbf{k}'_i}(V_{11}, V_{12}). \tag{3.141}$$

This gives us altogether four equations for the h^2. For special $\mathbf{k}$ points all four matrices have one fourth of H in size in each dimension.

One case has to be separately considered: $\mathsf{R}_1\mathbf{k}'_i = \mathsf{R}_2\mathbf{k}'_i \equiv \tilde{\mathbf{k}}$, *i.e.*, $\mathbf{k}'_i = \mathsf{R}_1\mathsf{R}_2\mathbf{k}'_i$. We can include this constellation, which can also happen for special $\mathbf{k}$ points, in our scheme by treating $\tilde{\mathbf{k}}$ for the rotation R_2 like

$$\begin{cases} \mathbf{k}'_i = \mathsf{R}_2\mathbf{k}'_i \text{ with } \mathrm{sgn}(\mathsf{R}_2) = +, & \text{for } \mathrm{sgn}(\mathsf{R}_1) \times \mathrm{sgn}(\mathsf{R}_2) = +, \\ \mathbf{k}'_i = \mathsf{R}_2\mathbf{k}'_i \text{ with } \mathrm{sgn}(\mathsf{R}_2) = -, & \text{for } \mathrm{sgn}(\mathsf{R}_1) \times \mathrm{sgn}(\mathsf{R}_2) = -, \end{cases} \tag{3.142a}$$

in order to avoid double counting.

The operations R_1 and R_2 must commute in order to preserve the relation

$$h^1_{\mathsf{R}_2\mathbf{k}_i\mathsf{R}_2\mathbf{k}'_i} = h^1_{\mathbf{k}_i\mathbf{k}'_i}. \tag{3.143}$$

This can be seen from

$$\begin{aligned} h^1_{\mathsf{R}_2\mathbf{k}_i\mathsf{R}_2\mathbf{k}'_i} &= H_{\mathsf{R}_2\mathbf{k}_i\mathsf{R}_2\mathbf{k}'_i} + H_{\mathsf{R}_2\mathbf{k}_i\mathsf{R}_1\mathsf{R}_2\mathbf{k}'_i} \\ &= H_{\mathsf{R}_2\mathbf{k}_i\mathsf{R}_2\mathbf{k}'_i} + H_{\mathsf{R}_2\mathbf{k}_i\mathsf{R}_2\mathsf{R}_1\mathbf{k}'_i} \\ &= H_{\mathbf{k}_i\mathbf{k}'_i} + H_{\mathbf{k}_i\mathsf{R}_1\mathbf{k}'_i} \\ &= h^1_{\mathbf{k}_i\mathbf{k}'_i}. \end{aligned} \tag{3.144}$$

The eigenvectors are reconstructed going backwards according to Eqs. (3.136) and (3.138). First we apply $V_{11} \xrightarrow{\mathsf{R}_2} V_{12}$, then $V_{11} \xrightarrow{\mathsf{R}_1} V_{21}$ and $V_{12} \xrightarrow{\mathsf{R}_1} V_{22}$ to obtain all $A^{\mathbf{k}}_\lambda$ in the full BZ.

Some Approximations

The calculation has several important parameters: the size of the $\mathbf{k}$ point sampling in the BZ, the number of valence and conduction bands included and the cutoff of the summation over the reciprocal lattice vectors in Eqs. (3.94) and (3.96) for the resonant part and the corresponding Eqs. (3.102) and (3.103) for the coupling part of the Hamiltonian. In particular, the inverse dielectric function can be considerably simplified. Thus there are several possibilities to minimize the computational effort. We will explore them in practice in the next chapter.

Bands Typically, for silicon as standard test material, we need four conduction bands in addition to the four valence bands in the pseudopotential formalism in order to get a converged absorption spectrum [7, 9] . However, we might only need the higher two valence bands and lower two conduction bands, or so, in order to understand already the main feature of the calculation.

A more severe approximation is the *decoupling* of the bands in the Hamiltonian matrix. This neglects the mixing of band pairs in the exciton wave function: that is, an exciton state is associated with valence band v_1 and conduction band c_1, another exciton state with bands v_1 and c_2, and so on. The resonant part of the Hamiltonian looks then like $H^{\text{res}}_{(vc\mathbf{k}),(v'c'\mathbf{k}')}\delta_{v,v'}\delta_{c,c'}$ and its size is only determined by the $\mathbf{k}$ point sampling for each band-pair subspace.

This approximation is reasonable when the bands are energetically well separated, but might be bad in the case of degenerate bands. It avoids band-pair mixing in the exciton wavefunctions and can limit the possibility of constructing strongly bound electron–hole states. It might work well, for example, for bound exciton states below the gap, if the bands do only slightly interact.

The nonlocality of space is represented by the off-diagonal elements of the inverse dielectric function Eq. (2.56). Although this nonlocality is important for the *GW* calculations, even for the quite homogeneous material silicon [5], in case of the excitonic interaction, it is more the mixing of the various single particle states, which creates the effect. Thus, it is well possible to ignore the off-diagonal elements of $\varepsilon^{-1}_{\mathbf{GG}'}(\mathbf{q})$ in certain cases, as will be shown in Chapters 4 and 5. Furthermore, the long-range character of the Coulomb interaction might demand for only a smaller cutoff in the summation over the reciprocal lattice vectors than the short-range contact exchange contribution.

Moreover, also with a model dielectric function, as used in Refs. [68, 69] by Shirley *et al.*, one can obtain quite good results for a range of semiconductors and also insulators.

In the limit of a long range interaction, we can replace the screened electron–hole interaction $W(\mathbf{r},\mathbf{r}')$ with the Coulomb interaction screened by the macroscopic dielectric constant, $-1/\varepsilon_0|\mathbf{r}-\mathbf{r}'|$, like in a Mott-Wannier model. This corresponds to taking $\varepsilon^{-1}_{\mathbf{GG}'}(\mathbf{q}) = 1/\varepsilon_0 \cdot \delta_{\mathbf{GG}'}$, overlooking in this way the reduction of screening at short distances. This approximation overestimates the screening and therefore

underestimates the electron–hole interaction.

Chapter 4

Optical Absorption: Silicon

In this chapter we will apply the theory developed to optical absorption calculations of bulk silicon, which is representative for the group IV, III-V, and II-VI semiconductors. These materials show qualitatively similar optical spectra, with two major structures at 3–5 eV. The first peak (E_1) has been interpreted as a M_1 type critical point transition, and the second peak (E_2) as a M_2 type one [7].

First, from a ground state calculation in a standard plane wave pseudopotential Car-Parrinello approach, we can readily obtain an LDA absorption spectrum. Like early empirical pseudopotential approaches [70, 71] and many other *ab initio* DFT-LDA works, *e.g.* [9], we get an underestimation of the E_1 peak by as much as 50%, reducing it to a weak shoulder of the generally overestimated E_2 peak. Then, after applying the *GW* corrections to the Kohn-Sham eigenvalues we obtain the quasiparticle band structure and the independent-particle (quasiparticle) absorption spectrum Eq. (3.21), which is, like the LDA spectrum, also quite different from the experimental spectrum [72, 73, 74].

In order to go beyond, we include local-field effects in the calculation of the dielectric matrix. Louie *et al.* [75] showed that the resulting spectrum is significantly improved at higher energies (above 15 eV), but not in the region of interest around 4 eV. As also reported by Hanke and Sham [11, 12], who performed a semi-empirical LCAO (linear combination of atomic orbitals) calculation, we find that local-field effects alone transfer oscillator strength to higher energies and hence increase the discrepancy with experiment at lower energies.

Several authors suggested that strong contributions to the E_1 peak could arise from saddle point excitons [12, 76, 77, 78, 79], related to van Hove singularities in the joint density of states. Experimentally, excitonic effects allowed to explain the measured temperature and pressure dependence of the lineshape and the symmetry in wavelength modulation reflectivity spectra [78, 79].

Until recently, the most sophisticated calculation of excitonic effects on the spectral lineshape of silicon was the LCAO calculation of Hanke and Sham [11, 12]. The electron–hole interaction shifted the position of the E_1 peak to lower energies, and almost doubled its intensity, while the oscillator strength of the higher energy peaks

was decreased. The overall agreement with experiment was hence improved, and clear evidence was given for the importance of excitonic effects. However, the final intensity ratio between the E_1 and E_2 peaks was reversed, in disagreement with the experimental spectrum. As pointed out by Wang *et al.* [7], the reliability of semi-empirical approaches is limited. For instance, there are important differences, already at the one-electron level, between the spectra of Refs. [12] and [75].

Here we will examine in detail the influence of the local-field and excitonic effects on the absorption spectrum in our *ab initio* approach demonstrating the performance of the method. We report the reliability of various approximations and interpolation schemes in order to speed up the calculation.

4.1 One-Particle Calculations

4.1.1 Kohn-Sham LDA Results

Ground State Calculations

We use a norm-conserving *pseudopotential* [80], which was generated within the Martins-Troullier scheme [81] (see also Appendix A). The cutoff radii are $r_c = 1.89$ for each orbit, with the local reference potential being the d orbital (spd).

The space group of silicon is the diamond structure O_h^7, with a basis of two atoms in a face-centered cubic (fcc) cell,

$$000; \quad \frac{1}{4}\frac{1}{4}\frac{1}{4}; \quad \text{(fcc)}. \tag{4.1}$$

The coordinates are measured relative to the Bravais lattice in units of the lattice constant a_0. The point group has 24 non-symmorphic rotations together with 24 symmorphic operations.

In reciprocal space the fcc direct lattice transforms to a body-centered (bcc) Bravais lattice with lattice parameter $\frac{2\pi}{a_0}$. The first Brillouin zone (BZ) is shown in Fig. 4.1. Some high symmetry points are given in Tab. 4.1.

We take a cutoff energy Eq. (2.32) of 15 Ry and ten special $\mathbf{k}$ points [38, 39] in the IBZ, which give 256 $\mathbf{k}$ points in the full BZ, for the ground state, which is sufficient for the following electronic structure calculations [42].

The *Kleinman-Bylander* (KB) form of the pseudopotential [82] offers an efficient computational scheme to obtain the DFT Hamiltonian for the Kohn-Sham electronic structure (see also Appendix A). In order to examine its influence on the Kohn-Sham eigenvalues we perform the ground state calculations with and without this separation. In each case we take about 8 values around the experimental lattice constant of $a_0^{\text{exp}} = 10.260$ a.u. [83] and fit the total energies with the Murnaghan

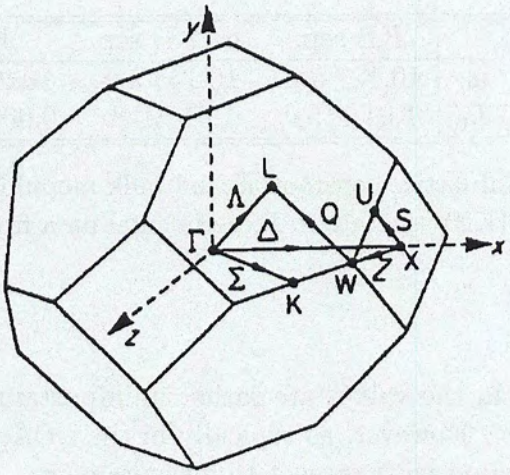


Figure 4.1: First Brillouin zone of the fcc cell. It has the shape of a bcc cell.

	x	y	z
Γ	0	0	0
Δ	.5	0	0
Σ	.5	.5	0
L	.5	.5	.5
K	.75	.75	0
X	1	0	0
U	1	.25	.25
W	1	.50	0

Table 4.1: Some high symmetry points in units of $\frac{2\pi}{a_0}$ along reciprocal Cartesian axes. The points K and U are equivalent.

equation of state [84] to obtain the theoretical equilibrium lattice constant a_0 and the bulk modulus B_0. The results are given in Tab. 4.2.

	KB sep.	no KB sep.	Exp.
a_0	10.127 a.u.	10.145 a.u.	10.260 a.u.
B_0	1.01 Mbar	1.01 Mbar	0.990 Mbar

Table 4.2: Theoretical lattice constant a_0 and bulk modulus B_0 for Si with and without Kleinman-Bylander (KB) separation. Experimental data from Ref. [83].

The differences in the values are rather unimportant concerning the Kleinman-Bylander separation. However, as typically in the LDA, the theoretical lattice constant is underestimated with respect to measurement.

It is interesting to look on the Kohn-Sham eigenvalues. The minimum gap of silicon is indirect and between the highest occupied band at Γ and around 0.85 in direction X. With and without using the KB separation we obtain $E_{\mathrm{Gap}}(\Gamma - 0.85\mathrm{X}) = 0.55$ eV, calculated at the respective theoretical a_0. For the values calculated at the experimental lattice constant we get 0.61 eV. The higher value is due to the stretched lattice. At the experimental lattice constant the differences between energies, obtained with and without using the Kleinman-Bylander separation, at high symmetry points like Γ, X, Δ, or L are largest for the lowest valence band (up to 60 meV) and only up to 20 meV for bands around the Fermi level.

A closer examination of the eigenvalues at the points Γ, X and L in Fig. 4.2 shows that up to about 40 eV above the Fermi level there is virtually no difference induced by the Kleinman-Bylander separation on the energy scale presented. For higher values, the eigenvalues shift and regroup in a somewhat different manner. However, these changes are not significant for the optical absorption properties studied in this work. These are mainly described by the four lowest conduction bands, *i.e.* , typically below 20 eV. Also, even for sums over many conduction bands, like some 100 bands for the independent particle polarization χ^0 in Eq. (2.57), the changes are small and unimportant with respect to the difference to the Fermi level.

In Fig. 4.3 we show the *band structure* for the four valence and first eight conduction bands without using the Kleinman-Bylander separation. Although the influence of the KB separation is small, we will not use it in the following for silicon, unless otherwise stated, in order not to bias our results. The eigenvalues are aligned with Γ'_{25}. For optical properties the nearly parallel bands around the gap especially from the Γ point in the directions of the X and the L point are important. They lead to M_1 type critical point transitions and are responsible for strong continuum excitonic effects [12].

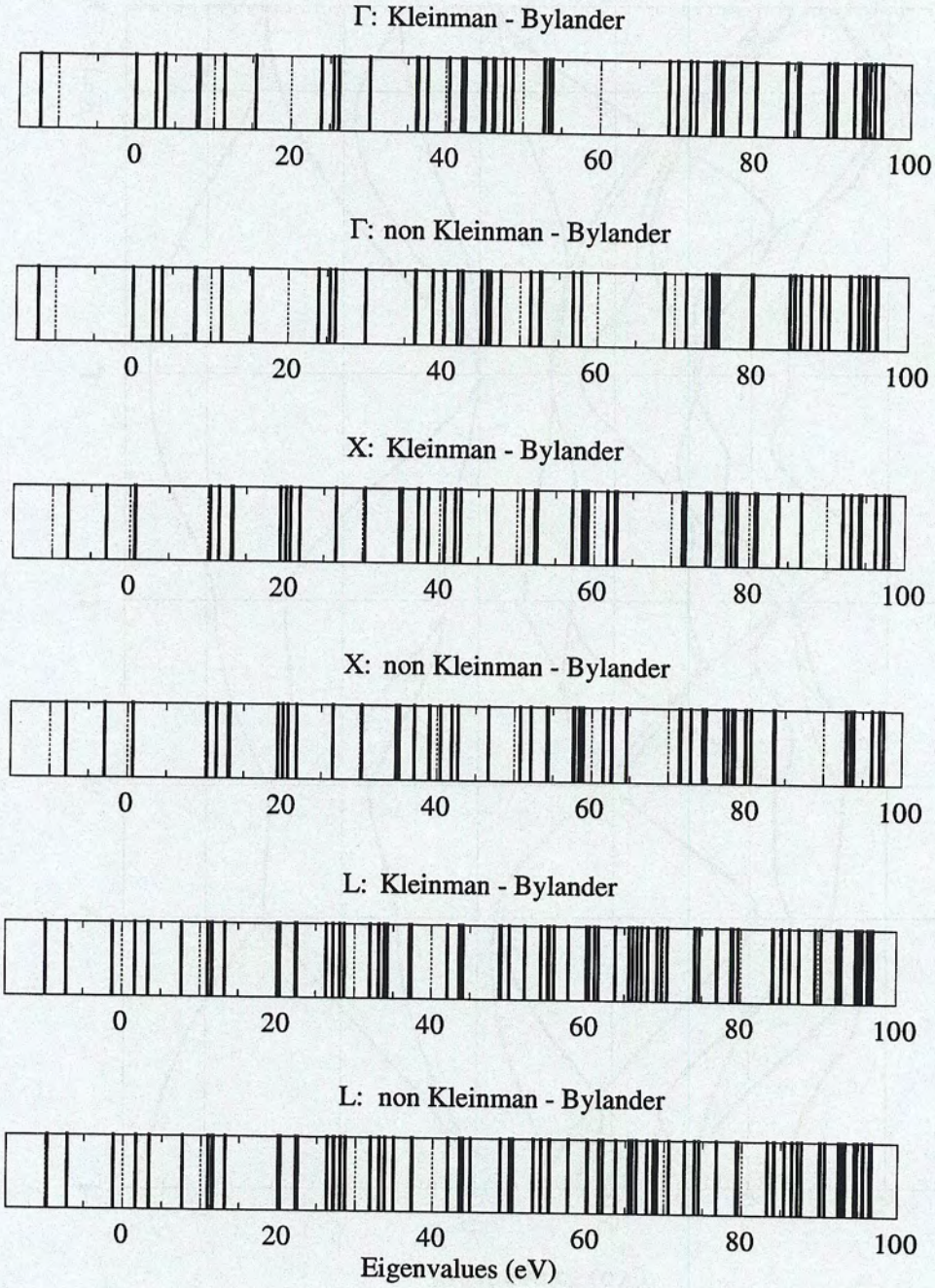


Figure 4.2: Comparison of Kohn-Sham LDA eigenvalues of Si up to 100 eV at the Γ, X and L points with and without Kleinman-Bylander separation. Eigenvalues are aligned with Γ'_{25}.

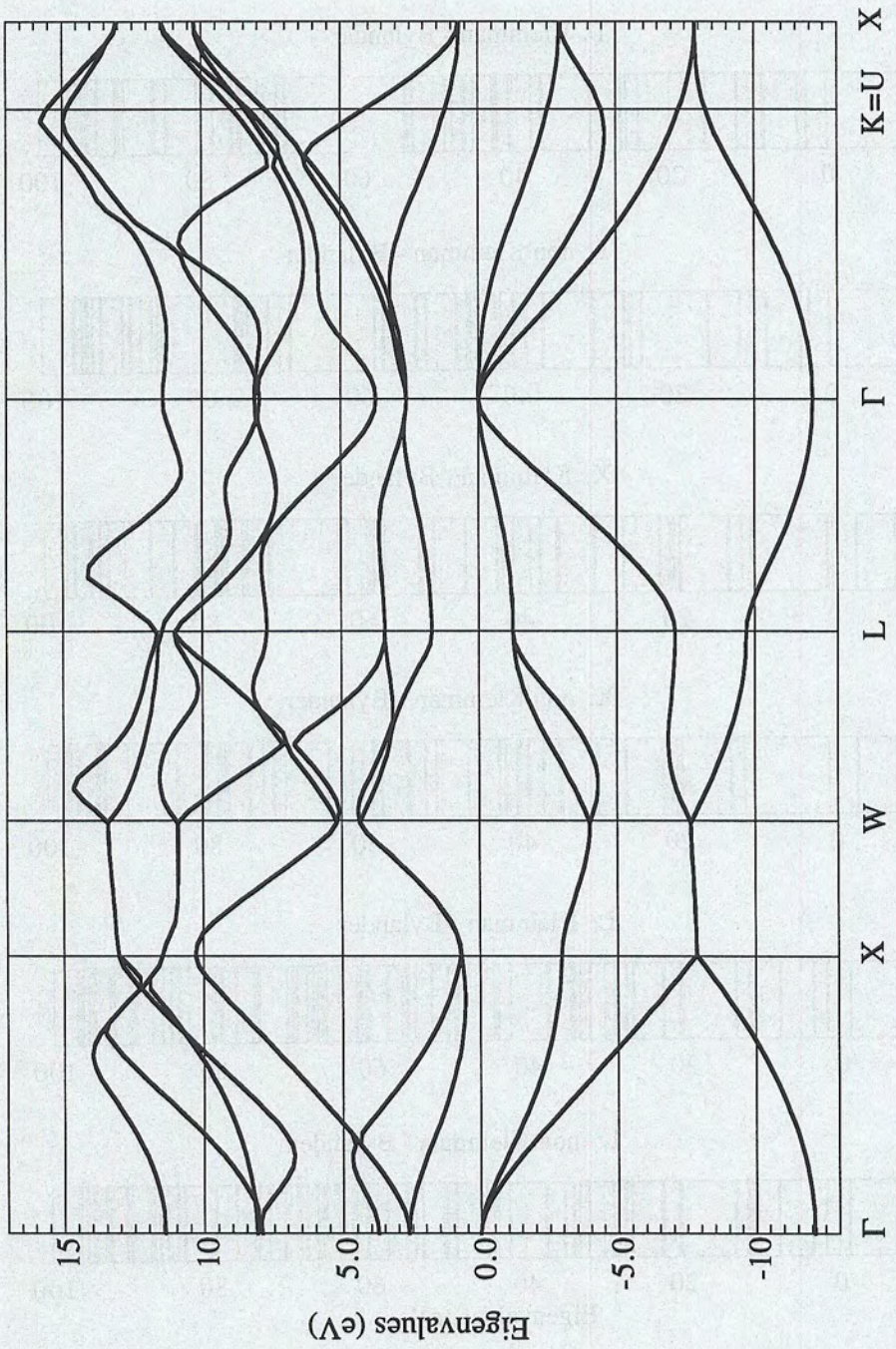


Figure 4.3: Kohn-Sham LDA band structure of Si for four valence and the first eight conduction bands. Eigenvalues are aligned with Γ'_{25}.

LDA Absorption Spectrum

The absorption spectrum in the independent particle approximation with the Kohn-Sham LDA eigenvalues as transition energies is calculated according to Eq. (3.21) with all energies taken to be the Kohn-Sham LDA eigenvalues. Eight bands (four valence and four conduction bands) are enough in the region of interest of excitations around the gap [7, 9]. For a fully converged spectrum we need 408 **k** points in the IBZ, which give 16384 **k** points in the full BZ.

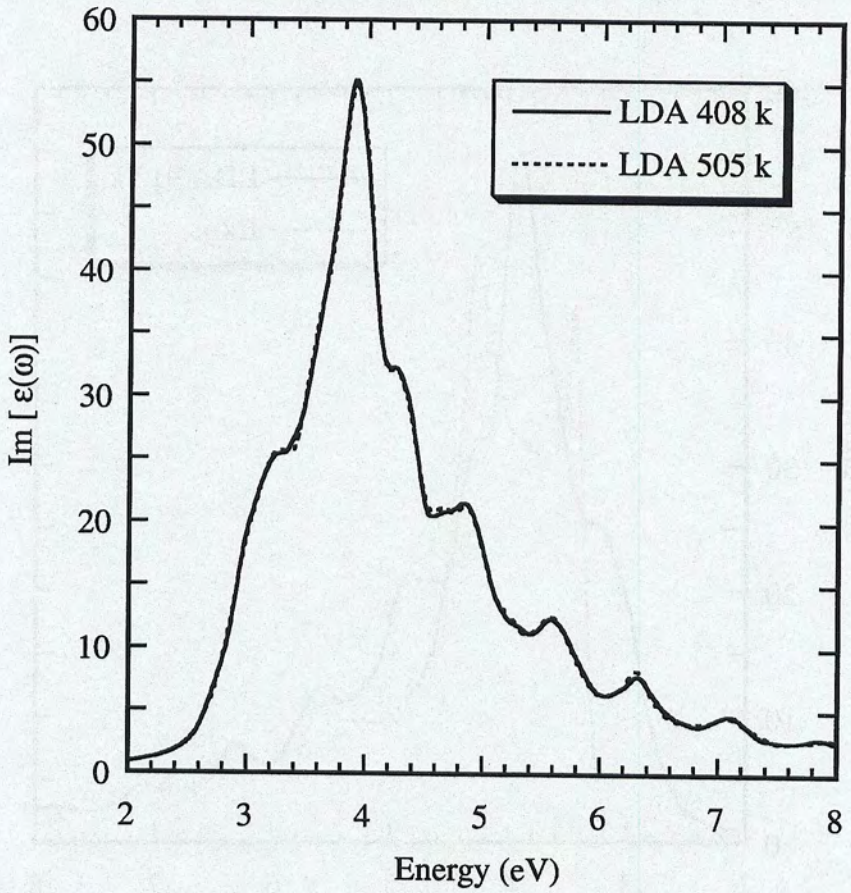


Figure 4.4: LDA absorption spectra of Si with 408 and 505 **k** points in the IBZ. 8 bands are used.

In Fig. 4.4 we show two spectra in order to verify the convergence of the **k** point sampling. The 408 **k** point set is shifted to high symmetry points, including Γ, leading to a set with 505 **k** points in the IBZ, which gives by construction also 16384 **k** points in the full BZ. We take 0.10 eV for the broadening parameter η,

which is also used in all other spectra of Si in the following, if not otherwise stated.

The two curves are very close to each other and we can add them up in Fig. 4.5 to one single curve with 913 **k** points in the IBZ in order to compare with the experimental result [74] measured ellipsometrically at 30 K. Like already mentioned in the introduction to this chapter we get an underestimation of the E_1 peak at 3.2 eV in the theoretical spectrum by as much as 50%. It is only a weak shoulder next to the overestimated E_2 peak around 3.9 eV. Furthermore, the LDA spectrum is shifted towards lower energies by about 0.2 eV with respect to the E_1 peak at 3.4 eV and the E_2 peak at 4.25 eV in measurement.

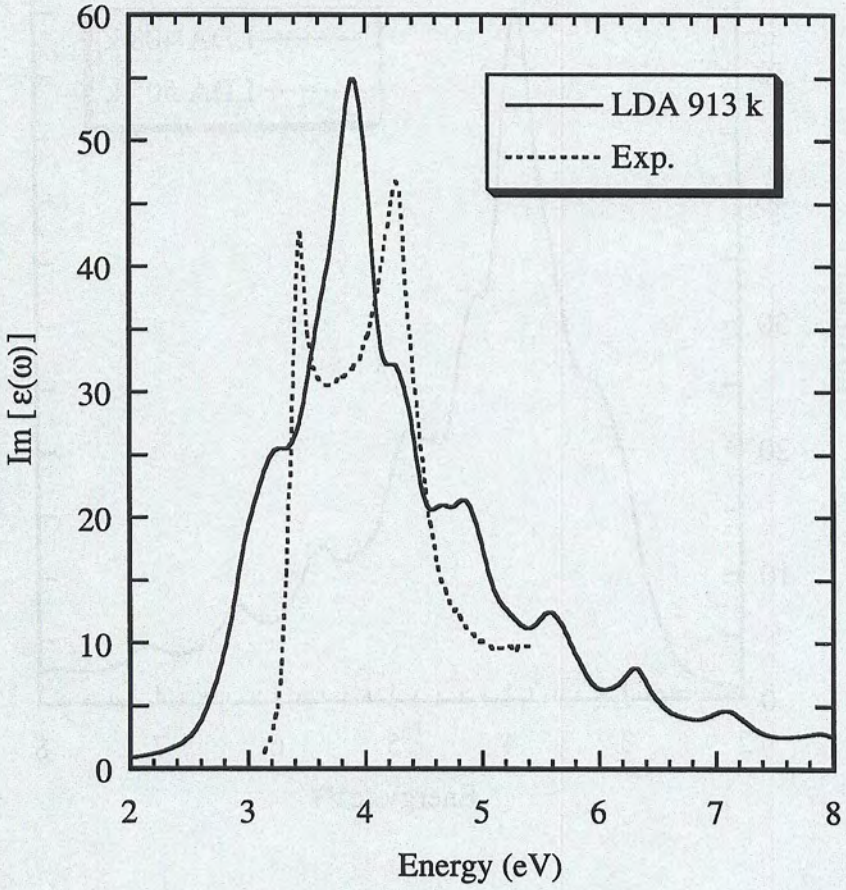


Figure 4.5: LDA absorption spectrum of Si with 913 **k** points in the IBZ. 8 bands are used. Experimental curve from Ref. [74] at 30 K.

The theoretical curve shows three small structures on the shoulder around 3.0 eV, 3.2 eV and 3.7 eV. The first one stems from transitions around the Γ point, with some dispersion in the X and L directions. The second one is also close to Γ, but

with more weight shifted further along the X and also the L direction. The third structure is less concentrated at one point in $\mathbf{k}$ space, but mainly from some region around Δ. The main peak at 3.9 eV has contributions from large regions, primarily around Σ in K direction and Δ in X direction. In all cases, only bands around the gap contribute. This identification is conform with Wang and Klein [7], although some authors have reported slightly different attributions [79, 85, 86].

4.1.2 *GW* Calculations

Quasiparticle Band Structure

The band structure with the physical *quasiparticle energies* is obtained by applying the *GW* corrections to the Kohn-Sham eigenvalues as outlined in Section 2.3.2.

For silicon we take the parameters from Ref. [6] and use 65 bands and 169 wavefunctions for the calculation of the independent particle polarization χ^0, Eq. (2.57), and the *GW* corrections, Eq. (2.106). The polarization χ^0 is calculated with 10 $\mathbf{k}$ points in the IBZ, giving 256 $\mathbf{k}$ points in the full BZ. The head of the dielectric matrix in the long wavelength limit, $\varepsilon_{00}(\mathbf{q} \to 0)$, is 16.14, while the inclusion of local-field effects, $1/\varepsilon^{-1}_{00}(\mathbf{q} \to 0)$, leads to 14.51. This is not a too well converged value for the macroscopic dielectric constant in the LDA-RPA formulation. A typical value from Ref. [42] is 12.2. The inclusion of the nonlocal commutator in the matrix elements, like in Eq. (3.20), would decrease the dielectric constant by about 10% [42].

The element $\varepsilon_{00}(\mathbf{q} \to 0)$ converges much slower than the large rest of the matrix. It must be noted that since it is only one entry in a large matrix used for the *GW* (and also excitonic) calculations we can tolerate the error. Furthermore, the contribution of this element vanishes in full convergence with respect to the $\mathbf{k}$ point sampling.

The *GW* corrections are obtained for the first eight bands. In order to correctly treat high symmetry points in the band structure, we shift the 10 special $\mathbf{k}$ points to a set with 19 high symmetry points in the IBZ, including Γ, which gives by construction also 256 $\mathbf{k}$ points in the full BZ. With these high symmetry points we calculate the quasiparticle energies for the band structure.

The calculated minimum *GW* gap is 1.19 eV, with a correction of 0.64 eV for the LDA value. At Γ, the gap is widened by 0.66 eV, which is representative for other high symmetry points. The experimental gap is 1.17 eV [83]. This is a typical result for silicon [5, 6].

In Fig. 4.6 we show the quasiparticle band structure as a full line for the first eight bands (four valence and four conduction bands) together with the corresponding LDA Kohn-Sham eigenvalues, represented by the dashes. The *GW* corrections have been interpolated with splines from 10 calculated values for each band, which is justified since the dispersion of the band structure is only changed by about 0.2 eV for each band, although with a maximum of 0.5 eV for the lowest valence band.

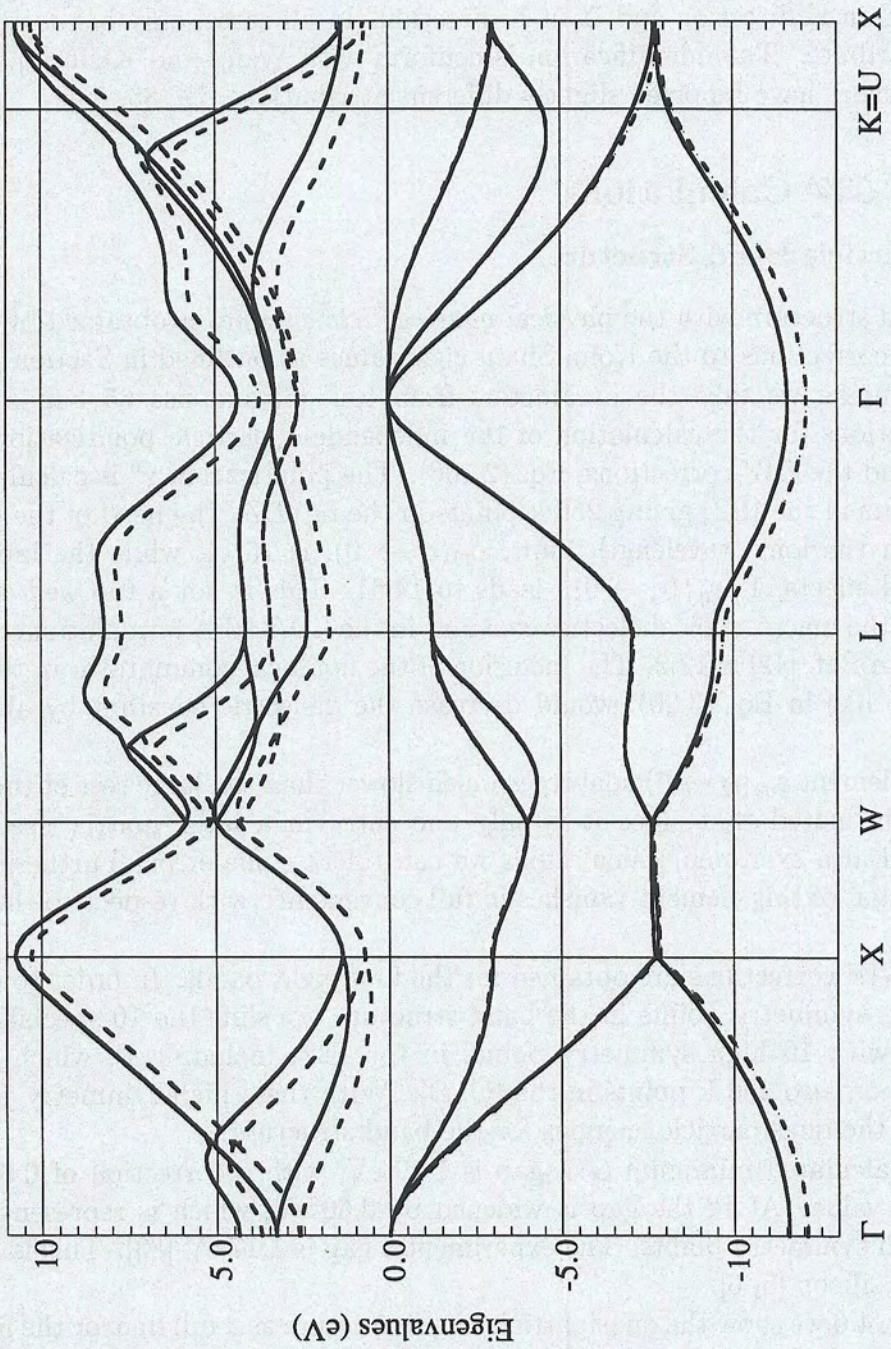


Figure 4.6: *GW* (full lines) and Kohn-Sham LDA (dashed lines) band structures of Si for the first eight bands. Eigenvalues are aligned with Γ'_{25}.

The plasmon pole model defined in Eq. (2.64) can lead to poles in the *GW* calculation for certain band energies. These poles, which correspond to unbroadened resonances, drop out in a full integration over the whole Brillouin zone, as positive and negative contributions cancel. The divergencies occur most likely for high and especially low bands, notably the lowest valence band. Since the *GW* corrections typically alter the band dispersion only slightly, we can correct diverged calculations by inspection. For the lowest valence band, however, we normally have to omit all plasmon frequencies $\tilde{\omega}_{\mathbf{GG'}}(\mathbf{q})$ lying within the valence bandwidth. The induced error can be estimated by repeating the procedure for an non-divergent band, *e.g.*, Γ'_{25}, and turns out to be of the order of some 10 meV for silicon and is thus within the overall error range of about 50 meV for the *GW* corrections in general. The positive corrections for the energies of the lowest valence band have rather to be attributed to the plasmon pole model Eq. (2.64) than to the energy cutoff in the plasmon frequencies.

Quasiparticle Absorption Spectrum

The spectrum using the physical quasiparticle energies, obtained in a *GW* calculation described above, is obtained according to Eq. (3.21). The *GW* corrections are interpolated from values obtained with 10 special **k** points in the IBZ with a tetrahedron interpolation method [87, 88], which works quite well due to the small dispersion of the corrections.

The comparison of the quasiparticle spectrum with experiment in Fig. 4.7 shows about the same discrepancy as the LDA one with experiment, now only with a blueshift instead of a redshift. As for the LDA spectrum we have a weak shoulder instead of the E_1 peak, which shows three small structures at around 3.7 eV, 3.9 eV and 4.4 eV. The main E_2 peak is at 4.6 eV. Their attributions to regions in **k** space are the same as for the LDA spectrum, since the dispersion of the band structure is only slightly changed.

4.1.3 Reduction of k Point Sampling

One of the bottlenecks in the calculations of absorption spectra is the very large number of **k** points necessary in the sampling over the Brillouin zone. Indeed, though in the independent-particle approximation we have been able to include quite a large number of special **k** points, the consideration of many-particle interactions, in particular excitonic effects, in the next section drastically increases the computational effort and one would desire to have a **k** point set as small as possible.

Regular Grid for k Point Sampling

The special **k** point schemes, like [38, 39], have been developed in order to have a maximum accuracy when using a limited number of **k** points for summing a complex

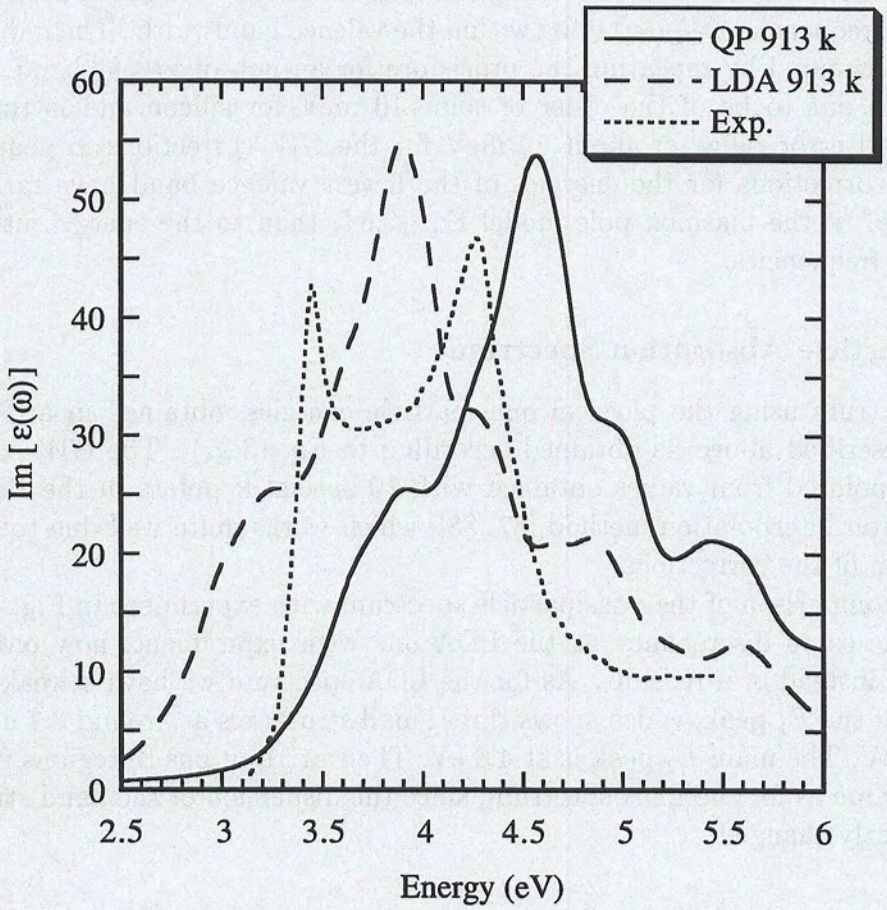


Figure 4.7: QP and LDA absorption spectra of Si with 913 **k** points in the IBZ. 8 bands are used. Experimental curve from Ref. [74] at 30 K.

valued function f over the Brillouin zone, see also Eq. (2.36),

$$\bar{f} = \frac{1}{(2\pi)^3} \int_{\mathrm{BZ}} d\mathbf{k}\, f(\mathbf{k}) \to \frac{1}{\Omega} \sum_{\mathbf{k}}^{\mathrm{BZ}} f(\mathbf{k}). \tag{4.2}$$

Increasing the number of special $\mathbf{k}$ points in our optical calculations enhances the accuracy of our spectra in general. However, the special $\mathbf{k}$ points are not optimized for spectral calculations by construction. The independent-particle absorption spectrum, for example, has the structure, from Eq. (3.2),

$$\varepsilon_{\mathrm{M}}(\omega) \sim \sum_{\mathbf{k},c,v} |f_{vc}(\mathbf{k})|\,. \tag{4.3}$$

Thus, only real valued functions are summed up, and most important, the band structure and dispersion play a crucial role for the result. This is an important notion, since, naturally the special $\mathbf{k}$ point schemes do not consider the material dependent band dispersion when constructing the $\mathbf{k}$ point set. Indeed, the special $\mathbf{k}$ points are regularly dispersed in the Brillouin zone and lie along high-symmetry directions. Thus, they may reflect to some extent local band maxima and minima rather than giving a good average of the band dispersion.

A naive and simple way to construct a $\mathbf{k}$ point set could be a regular grid in the Brillouin zone. We have started from the Cartesian fcc cell and placed a regular grid with 576 $\mathbf{k}$ points along and parallel to the high-symmetry axes. This is, of course, a rather insufficient approach to sample over the Brillouin zone. A better way is to shift the regular grid slightly in a non-symmetric direction, like $(1, 2, 3)$, thus breaking the symmetry of the $\mathbf{k}$ point sampling, but using a set more representative for the bulk volume properties.

In Fig. 4.8 we show two independent-particle spectra using these regular grids. The first one (dotted line) uses a Cartesian high symmetry grid with a nearest distance of $s = 2/11$ (in units of $\frac{2\pi}{a_0}$) between the points, giving 576 $\mathbf{k}$ points in the full BZ. The differences with the converged spectrum (full line) using 408 special $\mathbf{k}$ points in the IBZ are quite drastic. The second regular grid (in the above sense) has been created from a high-symmetry grid, with $s = 2/10$, which has been shifted by $(1, 2, 3) \cdot s/8$ in order to break symmetry. This gives 456 $\mathbf{k}$ points in the full BZ. Using this latter set for the spectrum (dotted-dashed line), we obtain good agreement with the converged full curve. The differences are mainly the exact position of the E_2 peak and its flanks.

However, since this $\mathbf{k}$ point set has no symmetry, we cannot, for example, reduce the size of the excitonic Hamiltonian as discussed in Section 3.2.4. Since we are primarily interested in the low energy part of the spectrum, especially in the positions and shapes of the first peaks, we will use the special $\mathbf{k}$ point scheme, as it gives a better representation of this part of the spectrum (see also the discussion to Fig. 4.12 in Section 4.2.1). It may be noted that Shirley *et al.* used a Brillouin zone

with the shape of a rhombohedron and slightly shifted a regular **k** point grid in an off-symmetric direction for their optical spectra calculations [68].

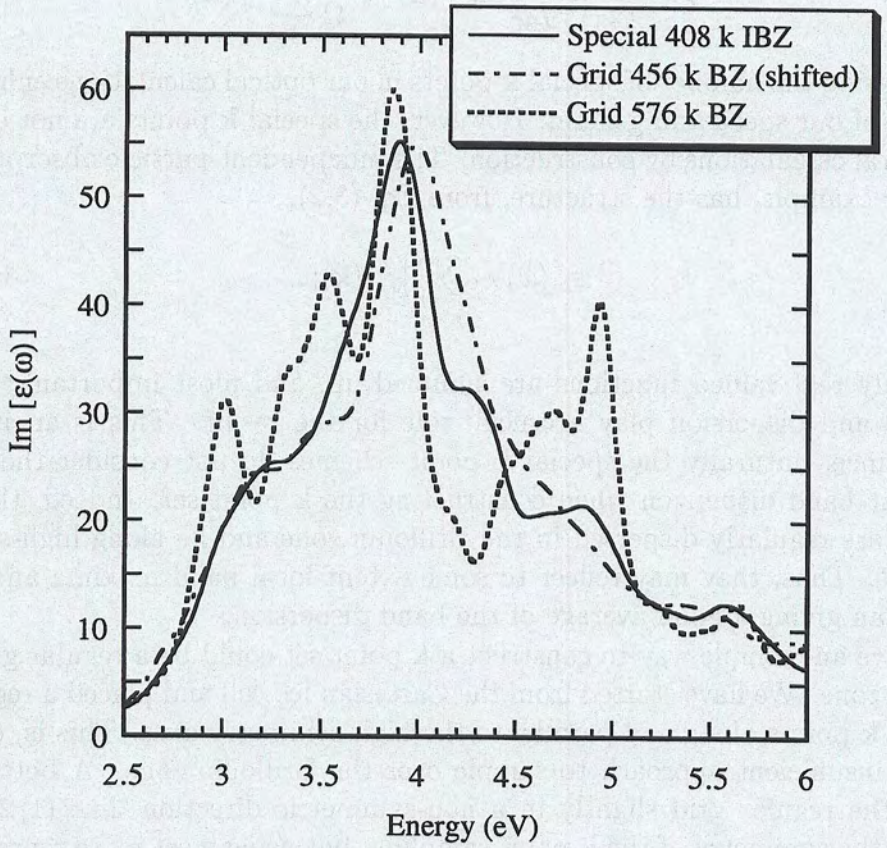


Figure 4.8: LDA absorption spectra of Si. 8 bands are used. Full curve: 408 special **k** points in the IBZ. Dotted-dashed curve: Regular grid of 456 **k** in the full BZ (slightly shifted to break symmetry). Dotted curve: Regular grid of 576 **k** in the full BZ (along high-symmetry directions).

Taylor Expansion in Imaginary Plane

We have also applied a method proposed by Godby *et al.* [89], which exploits the smooth behavior of the dielectric function in the complex plane in order to reduce the basis set needed to represent the dielectric function.

In the work of Godby *et al.* [89] the Taylor expansion scheme proved to be quite efficient in minimizing the number of bands and plane waves needed for the calculation of the dielectric function in a basis of reciprocal lattice vectors. Here,

however, we want to minimize the **k** point sampling in the Brillouin zone used for the absorption spectrum.

Following Ref. [89] we represent ε_M as a Taylor series expansion derived from function values further away from the real axis. We expect a faster convergence of our calculations, since much of the noise resulting from a too small basis will be smoothened out. For a function f, this expansion from a point z_0 is given by

$$f(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(z_0)}{k!}(z - z_0)^k. \tag{4.4}$$

$f^{(k)}$ is the k-th derivation of f with respect to its argument.

In order to get a numerically stable algorithm we take advantage of the analytic nature of ε_M and use Cauchy's expression for the derivative. We evaluate the contour integral numerically with m values lying on a circle of radius r, centered on z_0. Thus we have

$$f^{(k)}(z_0) = \frac{k!}{mr^k} \sum_{j=0}^{m-1} f\left(z_0 + re^{i2\pi j/m}\right) e^{-i2\pi jk/m}. \tag{4.5}$$

The expansion is made for $\varepsilon_M(\omega)$ with ω lying in the complex plane. Some heuristic search shows that a good choice is $z_0 = (0.05,\ 0.5)$ eV, $r = 0.10$ eV, $m = 10$ and $\eta = 0.10$ eV. The z_0 are shifted by steps of 0.1 eV along the real axis. In going much further away from the real axis than one eV for the expansion center we loose too much information about the absorption curve. It is not possible to take many more orders in the Taylor expansion than 10 for a numerically stable procedure. Depending somewhat on $\mathrm{Im}\, z_0$, the radius r can be varied over almost one order of magnitude without influencing the results.

We can best see the difference to a simple Lorentzian broadening, which is of course also included in the Taylor expansion scheme, by using a small basis set like 10 **k** points in the IBZ. In Fig. 4.9 we note that several small wiggles, especially in the higher energy regime, are washed out. However, the main discrepancy to the fully converged spectrum with 913 **k** points in the IBZ remains.

Furthermore, in excitonic calculations, where the size of the **k** point sampling is most crucial, we generally only use the resonant part of the Hamiltonian, Eq. (3.78). In shifting the absorption frequency ω into the complex plane in Eq. (3.88), the anti-resonant part for the negative frequencies together with the coupling terms of the Hamiltonian get more and more important, the larger $\mathrm{Im}\,\omega$ is, which makes the approximation of taking only the resonant part doubtful. Although the expansion is in principle exact, the limited number of Taylor coefficients influences the result depending on the expansion center z_0.

Thus we do not further employ the Taylor expansion method in our calculations. Indeed, the best way to minimize the basis set is to use an optimized one for the specific purpose, maybe along the lines of Ref. [90] or of the rather simple, but effective scheme, using a regular, slightly shifted grid of **k** points in the Brillouin zone as discussed in the previous section, if one does not want to use symmetry.

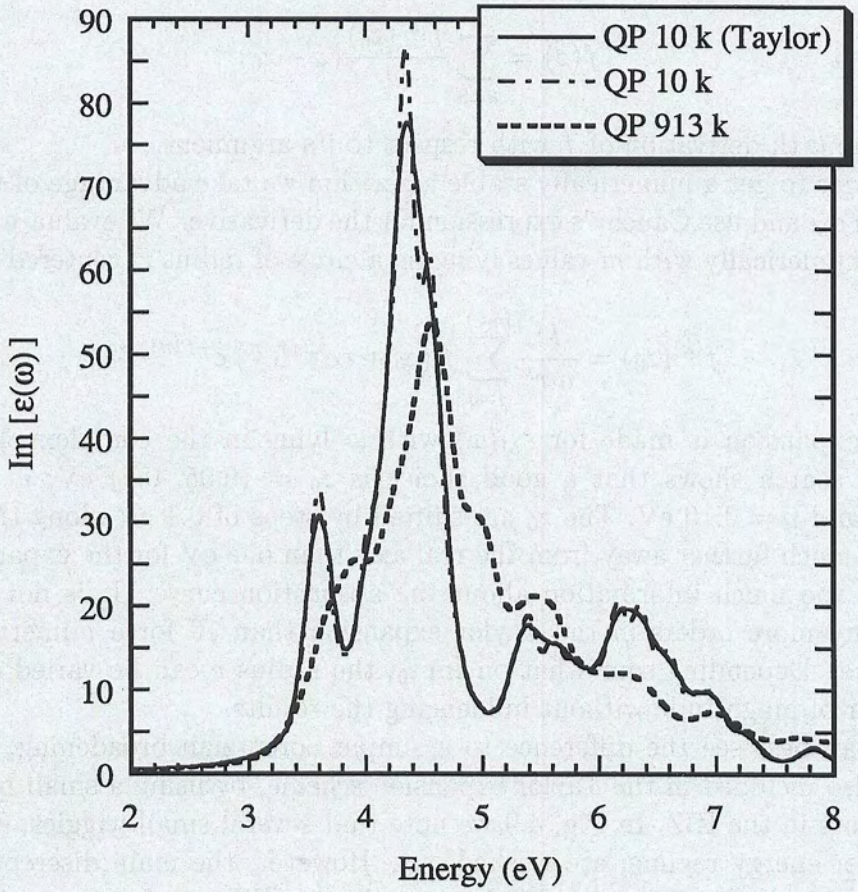


Figure 4.9: QP absorption spectrum of Si with 10 **k** points in the IBZ with Taylor expansion method and with standard Lorentzian broadening. The QP spectrum with 913 **k** points in IBZ is also shown for reference. 8 bands are used. Values for expansion: $z_0 = (0.05, 0.5)$ eV, $r = 0.10$ eV, $m = 10$. $\eta = 0.10$ eV.

4.2 Many-Particle Calculations

The interaction of the one-particle quasiparticle states leads to considerable deformations of the independent-particle absorption spectrum. The non-locality of space introduces the local-field effects, which are baptized exchange effects in the exciton terminology, while the created electron–hole pairs give rise to the excitonic effects.

In the following we will examine our approach, outlined in Section 3.2, to the inclusion of local-field effects in the absorption spectrum and compare it to the standard formulation in a basis of reciprocal lattice vectors and of valence and conduction bands [9].

We present our results for the optical absorption and reflectivity spectra of silicon including excitonic interaction [66]. The effects of degenerate bands on the application of symmetry properties are explained, and some approximations in order to simplify the calculations, already proposed in Section 3.2, are discussed [91].

4.2.1 Local-Field Effects

Comparison with Plane Wave Basis Calculations

The standard approach to the inclusion of local-field effects in the calculation of optical spectra is done by, in principle, inverting the dielectric matrix for each absorption frequency ω [9].

Using the RPA formulation Eq. (2.59b) of the dielectric matrix, we obtain the macroscopic dielectric function after Fourier transforming to momentum space from Eq. (3.1). Indeed, we already calculate $\varepsilon^{-1}_{\mathbf{GG'}}(\mathbf{q},\omega)$ in the *GW* approximation for two frequencies, 0 and iE_0, in the plasmon pole model Eq. (2.64).

The code of Godby and Needs [6, 47] allows us to make several tests and comparisons with our approach, where we describe the Hamiltonian in a basis of Bloch functions. First, we investigate the convergence with respect to the number of wavefunctions and bands of the approach in reciprocal space. Using 6 special $\mathbf{k}$ points in the IBZ, which give 108 $\mathbf{k}$ points in the full BZ, and 27 wavefunctions, we find that already 8 bands (4 valence and 4 conduction bands) give reasonable values for the local-field effects, as can be seen from Tab. 4.3. The small differences between the values of $\left[\varepsilon^{-1}_{00}(0)\right]^{-1}$ and ε_{M} are due to the rounding errors.

We want to know the influence on the local-field effects, if we evaluate the dielectric matrix for imaginary frequencies ω. Taking 27 wavefunctions, 8 bands and using 10 special $\mathbf{k}$ points, we find in Tab. 4.4 that the relative effect of the local field does not notably change in the range of about 1 eV away from the real axis. This is relevant for the Taylor expansion method, introduced in Section 4.1.2, and also for the Lorentzian broadening.

$\omega = 0$, 6 **k** IBZ	5 bnds	6 bnds	8 bnds	65 bnds	65 bnds
$\varepsilon_{00}(0)$	14.23	15.79	17.50	17.73	18.15
$\left[\varepsilon_{00}^{-1}(0)\right]^{-1}$	13.76	14.97	16.26		
$\varepsilon_{\mathrm{M}} = 1 - v\bar{P}$	13.76	14.98	16.25	16.38	16.41
LF effects	3.4%	5.5%	7.6%	8.2%	10.6%

Table 4.3: Standard reciprocal space approach. Results for $\omega = 0$ for several numbers of bands for the head of the dielectric matrix in the long wavelength limit, $\varepsilon_{00}(0)$; the inverse of the head of the inverse dielectric matrix in the long wavelength limit, $\left[\varepsilon_{00}^{-1}(0)\right]^{-1}$; the macroscopic dielectric constant, ε_{M}, calculated according to Eq. (3.33); and the percentage of the local-field (LF) effects on the values for each number of bands. In the last column, we use 169 wavefunctions instead of 27 for the basis sets.

27 wfc, 8 bands, 10 **k** IBZ	$\omega = 0$	$\omega = i0.5$ eV	$\omega = i1.0$ eV
$\varepsilon_{00}(0)$	15.50	15.25	14.56
$\left[\varepsilon_{00}^{-1}(0)\right]^{-1}$	14.31	14.08	13.46
Local-field effects	8.3%	8.3%	8.2%

Table 4.4: Influence on local-field effects taking imaginary frequencies ω. The calculations are done with the standard reciprocal space method.

Now we compare the standard approach in reciprocal space with our method of representing the dielectric matrix in a basis of Bloch functions like in Eq. (3.66). For only a limited number of frequencies, we can perform a direct inversion of the matrix $\left[\bar{H}^{\mathrm{exc}} - I\,\omega\right]^{-1}_{(n_1,n_2),(n_3,n_4)}$, thus avoiding to take care of the non-Hermiticity of $\bar{H}^{\mathrm{exc}}$ in general. The resulting equation is

$$\varepsilon_{\mathrm{M}}(\omega) = 1 - \lim_{\mathbf{q}\to 0} v_0(\mathbf{q}) \sum_{n_1 \ldots n_4} \left\langle n_1 \left| e^{-i\mathbf{q}\cdot\mathbf{r}} \right| n_2 \right\rangle \left\langle n_4 \left| e^{i\mathbf{q}\cdot\mathbf{r}'} \right| n_3 \right\rangle \bar{P}_{(n_1,n_2),(n_3,n_4)}(\omega), \quad (4.6)$$

with $\bar{P}$ from Eq. (3.66). We have to include the full excitonic Hamiltonian $\bar{H}^{\mathrm{exc}}$, Eq. (3.64), including the anti-resonant and coupling parts, for this comparison.

We are using two sets of parameters in Tabs. 4.5 and 4.6. For purely imaginary values of ω we find excellent agreement of the two approaches in Tab. 4.5.

In Tab. 4.6, the differences at $\omega = (3 + i0.05)$ eV are due to a sharply peaked structure in $\varepsilon_1 = \mathrm{Re}\,\varepsilon_{\mathrm{M}}$ around 3 eV causing numerical noise.

The excellent agreement of the two approaches confirms the analytical derivation of the Bloch representation and is a good check of the overall structure and implementation of our code.

27 wfc, 5 bands, 6 **k** IBZ	$\omega = 0$	$\omega = i0.27$ eV	$\omega = i13.6$ eV
ε_M (std. basis)	13.76	13.68	1.87
ε_M (Bloch basis)	13.76	13.68	1.87

Table 4.5: Comparison of standard reciprocal space method with representation in Bloch basis for macroscopic dielectric function ε_M from Eq. (4.6) for ω on imaginary axis.

27 wfc, 8 bands, 6 **k** IBZ	$\omega = (2 + i0.05)$ eV	$\omega = (3 + i0.05)$ eV
ε_M (std. basis)	$23.22 + i0.55$	$-34.48 + i86.90$
ε_M (Bloch basis)	$23.22 + i0.56$	$-31.27 + i86.28$

Table 4.6: Comparison of standard reciprocal space method with representation in Bloch basis for macroscopic dielectric function ε_M from Eq. (4.6) for ω along real axis.

Absorption Spectrum

The diagonalization of the Hamiltonian matrix in order to find the eigenvalues and eigenstates limits the maximal size of the basis set. The Cray C98 of IDRIS–CNRS can treat in the standard job queues complex matrices in the memory with up to about 5000 elements in each of two dimensions. We can reduce the number of **k** points by a factor of four using symmetry properties as outlined in Section 3.2.4. Then the maximal size of the **k** point set with four valence and four conduction bands is about 1250, since we only use the resonant part of $\bar{H}^{\text{exc}}$, Eq. (3.78), for the absorption spectra, if not otherwise stated.

Actually, even for the exchange contribution we do not need 8 bands for a reasonable absorption spectrum. In Fig. 4.10 we show by the full line a spectrum including the local-field effects obtained with 28 special **k** points in the IBZ, which give 864 **k** points in the full BZ, and 8 bands compared to one with only 6 bands (dashed line). The number of reciprocal lattice vectors in the evaluations of the matrix elements of Eq. (3.20) is taken as 169, like in the other spectra shown for silicon, if not otherwise stated. We will test the convergence of this summation in the next section on excitonic effects. Also the sum in Eq. (3.94) over the reciprocal lattice is taken, as in the other spectra, over 169 **G** vectors.

The agreement is quite good and we can restrict ourselves to four valence and two conduction bands in order to compare the effects of the local field with the independent-quasiparticle spectrum and the experimental curve in Fig. 4.11, now using 60 special **k** points in the IBZ, expanding to 2048 **k** points in the full BZ. The findings are interesting. First, at about 5 eV we find an anomalous peak in both theoretical spectra, which reduces to a small shoulder at full convergence, as can be seen from Fig. 4.12. There we compare the independent-quasiparticle spectra for 60 and 913 **k** points in the IBZ. We can see that 60 **k** points are enough for our region of interest up to 5 eV in that graph, but not for higher energies.

More important, the two main peaks E_1 and E_2 are shifted toward higher energies

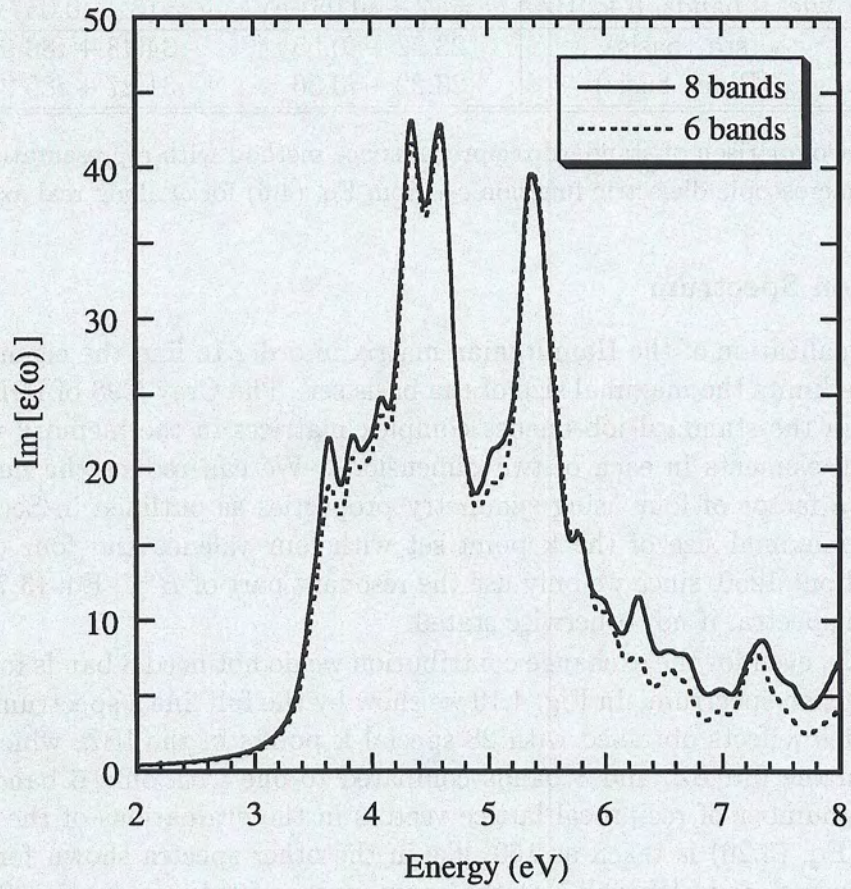


Figure 4.10: Absorption spectra of Si with QP energies and exchange effects included. 28 **k** points in the IBZ are used. 8 bands (full curve) and 6 bands (dashed curve) are used.

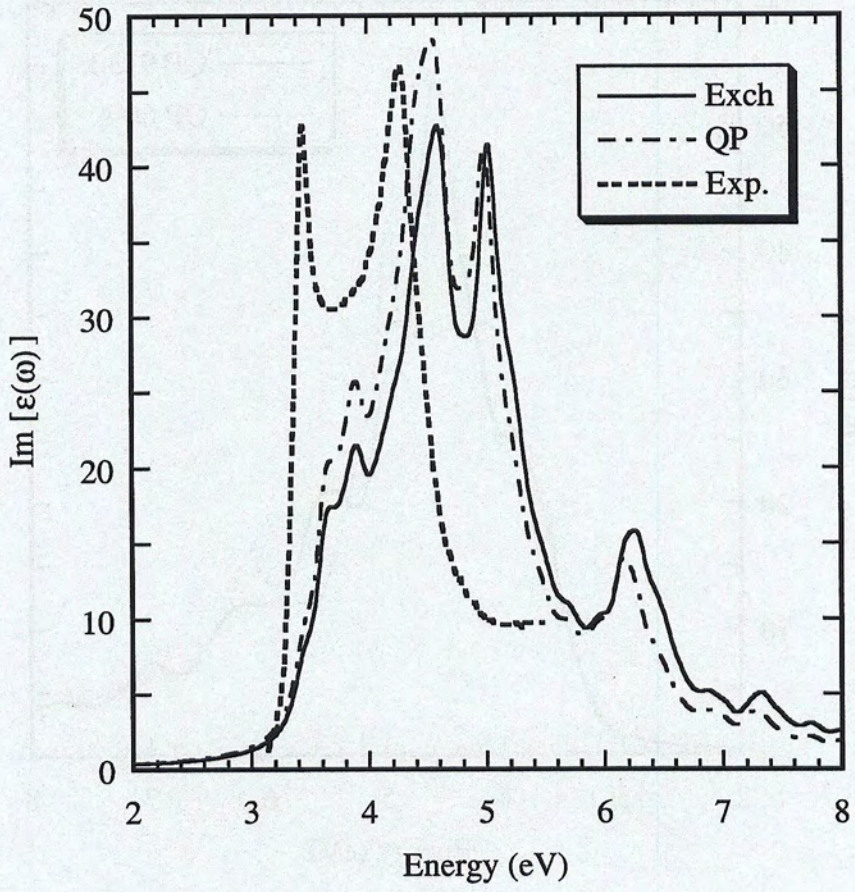


Figure 4.11: Absorption spectra of Si with QP energies (dotted-dashed curve) and exchange effects (full curve) included. 60 **k** points in the IBZ and 6 bands are used. Experimental data (dotted curve) from Ref. [74] at 30 K.

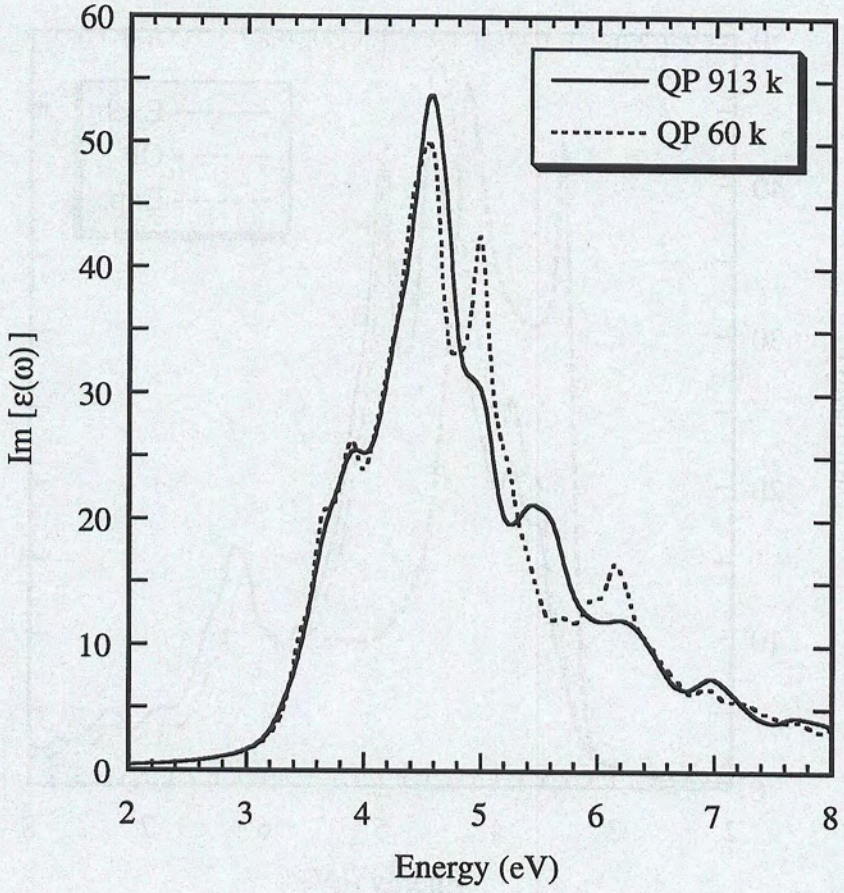


Figure 4.12: QP absorption spectra of Si with 913 (full curve) and 60 (dashed curve) **k** points in the IBZ. 8 bands are used. $\eta = 0.10$ eV.

by about 0.15 eV and both peaks are decreased in height. This increases even further the discrepancy with experiment. Our results are in general agreement with Louie *et al.* [75], using an empirical pseudopotential method and with Hanke and Sham [12], using a linear combination of atomic orbitals method, who also observe a decrease of the spectrum in the region of the two main peaks and an increase for higher energies.

On the contrary, an *ab initio* plane wave pseudopotential calculation of Gavrilenko and Bechstedt [9] shows a decrease of the E_1 peak and an increase of E_2 due to local-field effects. They employ the standard reciprocal space method and use a tetrahedron interpolation scheme [87, 88] also for the one-particle LDA energy levels and matrix elements, not only for the *GW* corrections as we do.

4.2.2 Excitonic Effects

Effects of Degenerate Bands on Symmetry

In Section 3.2.4 we have discussed the correct use of symmetry properties when applying to the excitonic Hamiltonian in order to reduce the dimension of the matrix, making four small matrices from the large Hamiltonian, and to minimize the number of elements which have to be calculated, constructing the others by symmetry according to Eq. (3.121).

Here we want to show the effect on the spectrum using the property $H_{\mathbf{pkpk'}} = H_{\mathbf{kk'}}$, Eq. (3.119) to obtain elements of Hamiltonian matrix when degenerate bands are involved. We take two special $\mathbf{k}$ points in the IBZ ($\mathbf{k}_1 = (0.25, 0.25, 0.25)$ and $\mathbf{k}_2 = (0.25, 0.25, 0.75)$ in Cartesian coordinates), which give 32 $\mathbf{k}$ points in the full BZ, and 5 bands. The only degenerate bands are then 3 and 4 of $\mathbf{k}_2$. We compare the full Hamiltonian matrix obtained from $H_{\mathbf{pkk'}} = H_{\mathbf{kp^{-1}k'}}$ (here p: rotation and inversion by time reversal) and the application of Hermiticity with a full Hamiltonian matrix, where we have calculated all elements explicitly, unless we can obtain them from Hermiticity. Although the gain in computer time is considerable, we cannot use all symmetry operations p.

Indeed, the two matrices – one obtained using $H_{\mathbf{pkk'}} = H_{\mathbf{kp^{-1}k'}}$ and one calculated not using this relation – are only different for elements of $H_{(vc\mathbf{k}),(v'c'\mathbf{k'})}$ ($c = c' = 5$) with $\mathbf{k} = \mathbf{k}_1$ and $\mathbf{k'} = \mathbf{k}_2$, $v' = \{3, 4\}$, looking only on one triangle due to Hermiticity. For $\mathbf{k} = \mathbf{k'} = \mathbf{k}_1$ we observe no differences, since the first 5 bands of $\mathbf{k}_1$ are not degenerate. For $\mathbf{k} = \mathbf{k'} = \mathbf{k}_2$, we also get no differences, since only 180^0 rotations around the axes x, y, z together with the identity and inversion are used to construct the star of $\mathbf{k}_2$ in the full BZ. Thus, the problems with degenerate bands do not arise calculating elements according to $H_{\mathbf{pk}_2\mathbf{k}'_2} = H_{\mathbf{k}_2\mathsf{p}^{-1}\mathbf{k}'_2}$, as has been explained in Section 3.2.4.

In Fig. 4.13 we depict the influence on the spectrum of the wrong application of the symmetry properties to the excitonic Hamiltonian, again using two special $\mathbf{k}$ points in the IBZ, but now 6 bands. Also in this case, the only degenerate bands

are 3 and 4 of $\mathbf{k}_2$. The spectrum, obtained using $H_{\mathbf{pkpk'}} = H_{\mathbf{kk'}}$ (dotted curve), is considerably different from the one obtained using only the Hermiticity (full curve) in order to reduce the number of elements of H to be calculated.

In Fig. 4.13 we also show a spectrum (dotted-dashed curve) taking all 32 $\mathbf{k}$ points of the BZ into the IBZ, thus not applying any symmetry operations (besides the identity) to the expansion of the $\mathbf{k}$ points in the IBZ into the full BZ and also not for the construction of the wavefunctions from the IBZ. The full and the dotted-dashed curve are very close together, the tiny differences probably stemming from the inverse dielectric matrix in the Coulomb interaction term Eq. (3.96). For technical reasons it has been constructed in both cases by application of all symmetry operations to $\tilde{\varepsilon}^{-1}_{\mathbf{GG'}}(\mathbf{q})$. These findings confirm the importance of a consistent use of symmetry properties.

Now we want to check the reduction of the dimension of the excitonic Hamiltonian. In Fig. 4.14 we show a spectrum calculated from the full Hamiltonian, using only the Hermiticity (full curve) in order to reduce the number of elements of H to be calculated and a spectrum obtained with an Hamiltonian split into four smaller matrices as outlined in Section 3.2.4, employing both time reversal and Hermiticity to speed up the calculation (dotted curve). The two curves are virtually indistinguishable.

In addition, we look at the *Kleinman-Bylander* separation. In Fig. 4.15 the dotted curve gives the spectrum using only Hermiticity, but without the Kleinman-Bylander separation, which has been used in this section for all other calculations with 2 $\mathbf{k}$ points in the IBZ. The influence in comparison to the non-Kleinman-Bylander separation is very small and supports our arguments discussing Fig. 4.2.

Some Approximations

The computational cost can be considerably reduced by introducing some approximations as discussed in Section 3.2.4.

Dielectric function First, we will ignore the off-diagonal elements of $\varepsilon^{-1}_{\mathbf{GG'}}(\mathbf{q})$ in the screening in Eq. (3.96). In Fig. 4.16 we depict a calculation with 10 $\mathbf{k}$ points in the IBZ and 8 bands, clearly indicating that it is enough to take the diagonal elements of $\varepsilon^{-1}_{\mathbf{GG'}}(\mathbf{q})$. In fact, this approximation is so good that we will not explicitly mention for each spectrum shown for Si, whether the off-diagonal elements have been neglected or not. Physically, the electron–hole pair does evidently not depend on the microscopic angular distribution of the local field. This might be due to a relative movement of the two-body system around each other, averaging over the angles, and/or to such a large distance $\mathbf{r}_{\mathrm{e-h}}$ between electron and hole that only the macroscopic, long-range dielectric constant ε_0 is relevant.

The latter can be easily ruled out, setting $\varepsilon^{-1}_{\mathbf{GG'}}(\mathbf{q}) = 1/\varepsilon_0 \cdot \delta_{\mathbf{GG'}}$, $\varepsilon_0 = 11.7$, in the screening in Eq. (3.96). Fig. 4.17 shows clearly that the overestimation of screening at short distances destroys most of the excitonic effect. We follow that the

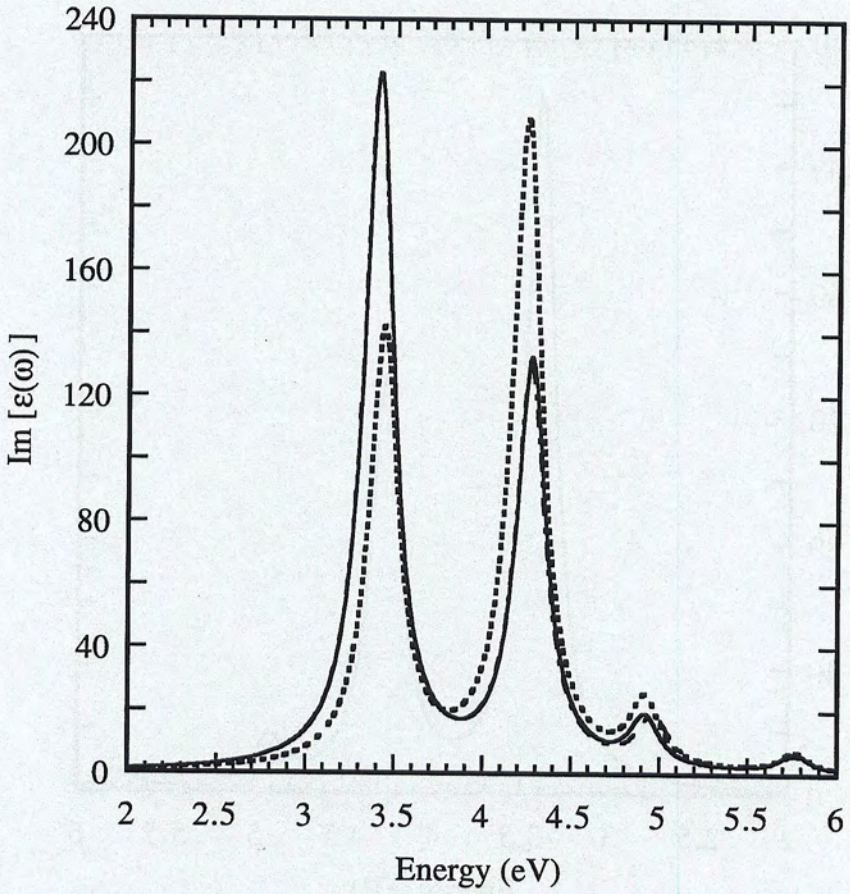


Figure 4.13: Absorption spectra of Si with excitonic and exchange effects included using 2 special $\mathbf{k}$ points in the IBZ and 6 bands. Bands 3 and 4 of $\mathbf{k}_2$ are degenerate. Full curve: only Hermiticity is used to reduce the number of elements of Hamiltonian H to be calculated. Dotted curve: In addition, also relation $H_{\mathbf{pkpk'}} = H_{\mathbf{kk'}}$ is used. Dotted-dashed curve: No symmetry operations are used for $\mathbf{k}$ points and their wavefunctions besides identity (and Hermiticity for H): 32 $\mathbf{k}$ points in the BZ, which is taken as IBZ.

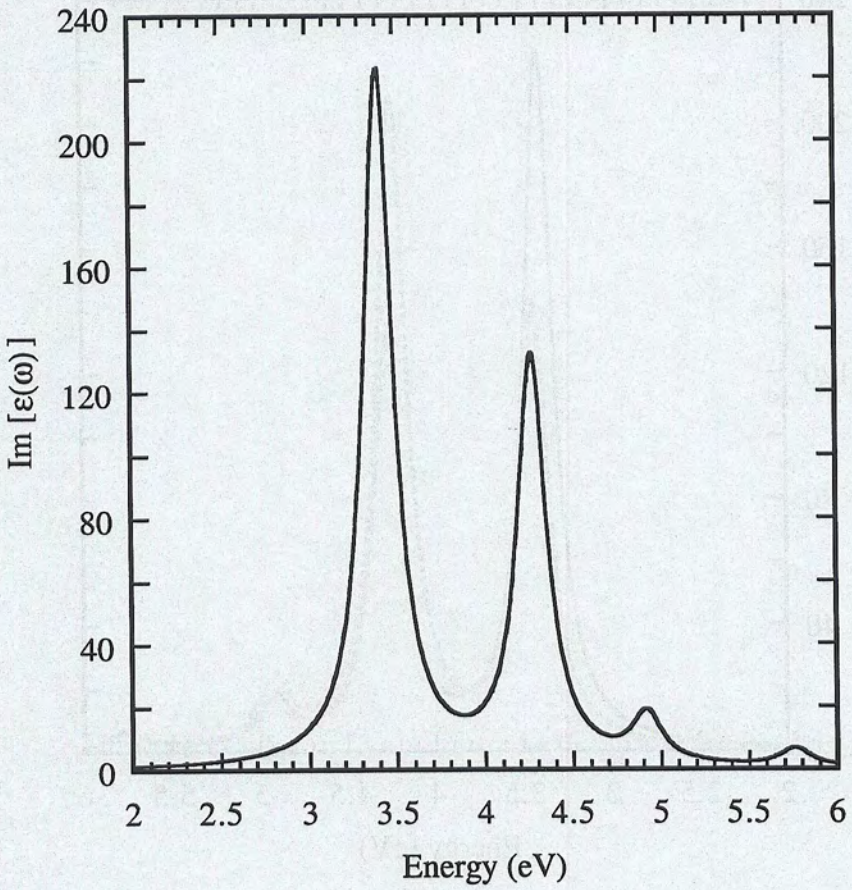


Figure 4.14: Absorption spectra of Si with excitonic and exchange effects included using 2 special $\mathbf{k}$ points in the IBZ and 6 bands. Bands 3 and 4 of $\mathbf{k}_2$ are degenerate. Full curve: Only Hermiticity is used to reduce the number of elements of Hamiltonian H to be calculated. Dotted curve: Like before, but H is split by symmetry into four matrices of one quarter in size in each dimension.

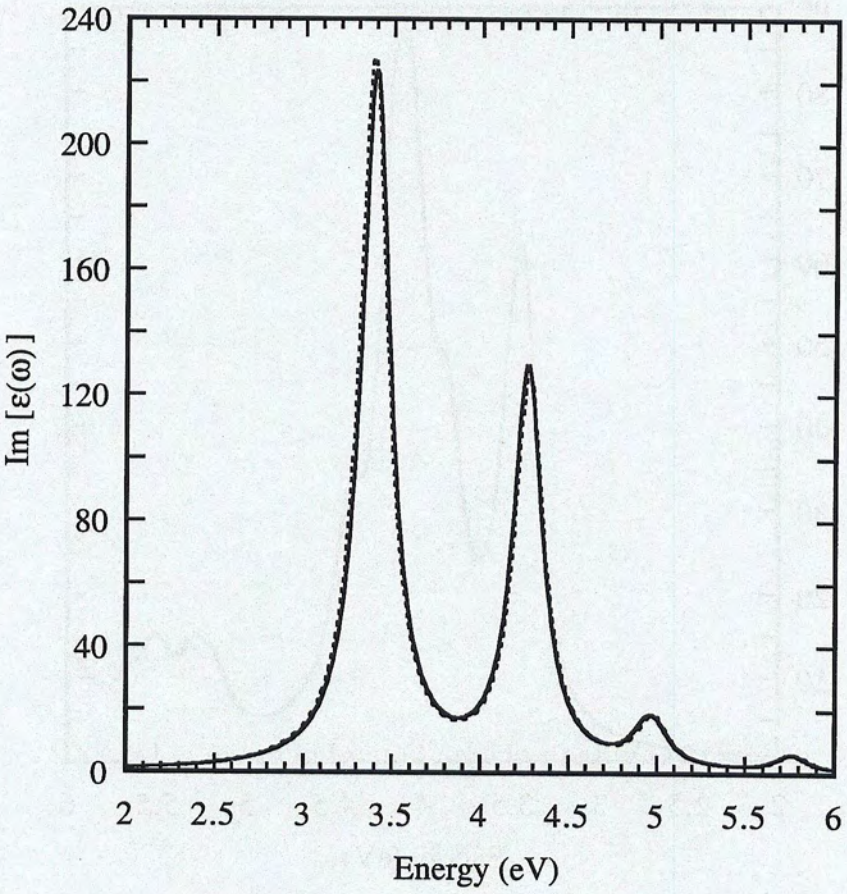


Figure 4.15: Absorption spectra of Si with excitonic and exchange effects included using 32 special $\mathbf{k}$ points in the BZ and 6 bands. Only Hermiticity is used to reduce the number of elements of Hamiltonian H to be calculated. Full curve: Kleinman-Bylander separation used. Dotted curve: Kleinman-Bylander separation not used.

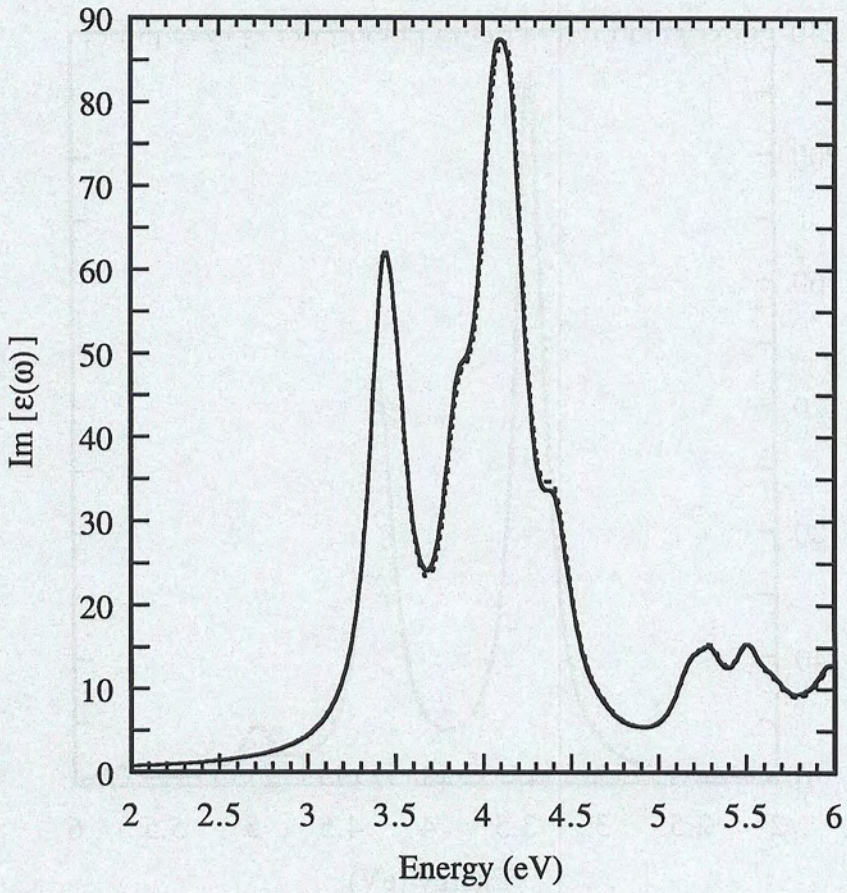


Figure 4.16: Absorption spectra of Si with excitonic and exchange effects included using 10 special $\mathbf{k}$ points in the IBZ and 8 bands. Full curve: Full $\varepsilon^{-1}_{\mathbf{GG'}}(\mathbf{q})$ taken as screening. Dotted curve: Diagonal approximation to $\varepsilon^{-1}_{\mathbf{GG'}}(\mathbf{q})$.

distance $\mathbf{r}_{e-h}$ between electron and hole is too short for a Mott-Wannier model to be applicable. $\mathbf{r}_{e-h}$ does thus not extend over a very large number of lattice constants.

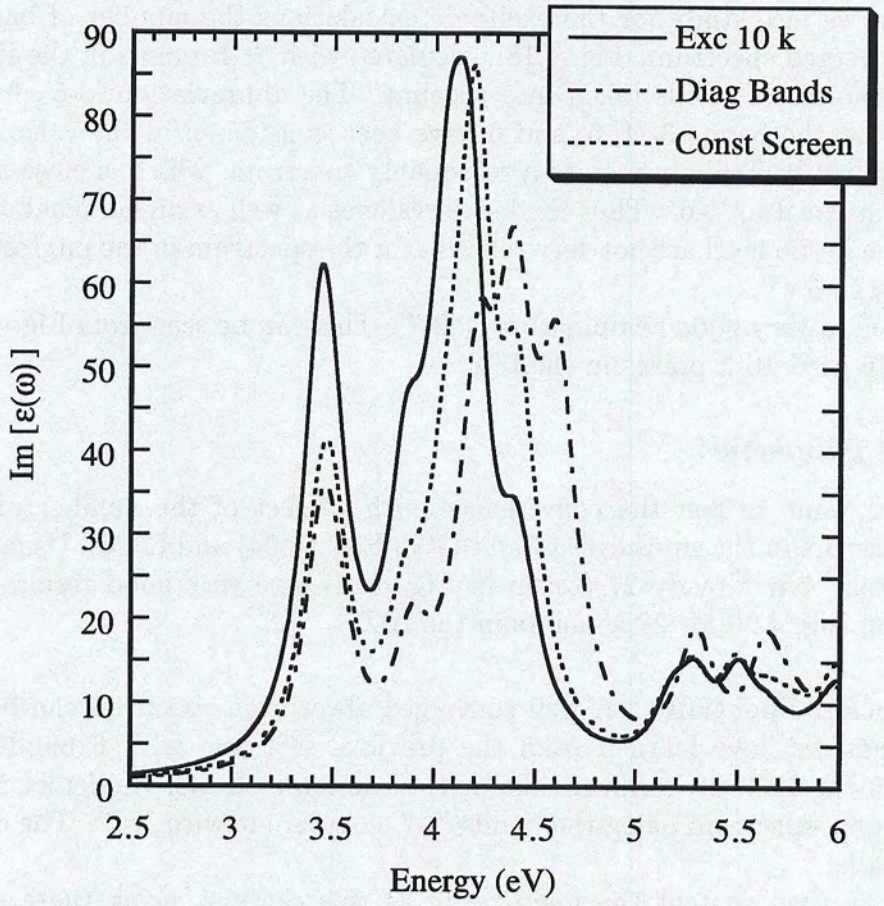


Figure 4.17: Absorption spectra of Si with excitonic and exchange effects included using 10 special $\mathbf{k}$ points in the IBZ and 8 bands. Full curve: Standard calculation with full dielectric matrix and all bands coupled. Dotted-dashed curve: Bands decoupled, but full screening. Dotted curve: Constant screening, $\varepsilon^{-1}_{\mathbf{GG'}}(\mathbf{q}) = 1/\varepsilon_0 \cdot \delta_{\mathbf{GG'}}$, but all bands coupled.

Bands The decoupling of the bands in the Hamiltonian matrix reduces considerably the computational effort, since each pair of valence and conduction bands can be separately considered. The size of each matrix is thus given by the number of $\mathbf{k}$ points in the reduced Brillouin zone alone. However, this approximation is not valid for silicon, as can been seen from Fig. 4.17. The mixing of the bands in addition to

the one of the **k** points is essential for the excitonic effects. The effect on the low energy side of the decoupling of the bands is very close to the effect of a constant screening, discussed in the above paragraph.

Next, we also study for the excitonic calculations the number of bands needed for a converged spectrum. Fig. 4.18, calculated with 28 **k** points in the IBZ, depicts various choices of bands taken into account. The abbreviation '3-6', for example, means that the bands 3, 4, 5, and 6 have been considered in the calculation. Obviously, already '3-6' gives quite a reasonably spectrum, which is close to '1-8' and almost identical to '1-6'. Thus the lowest valence as well as higher conduction bands above the Fermi level are not very relevant for the spectrum in the interesting energy range up to 6 eV.

Indeed, a very good combination is '2-7'. This can be seen from Fig. 4.19, where have only used 10 **k** points in the IBZ.

Optical Properties

First, we want to test the convergence with respect of the number of reciprocal lattice vectors in the sums over **G** and **G'** in Eqs. (3.94) and (3.96). Usually, we take 169 vectors, but already 27 (taken in Fig. 4.18) give very good results, as can be seen from Fig. 4.20 for 28 **k** points in the IBZ.

Absorption Spectrum A well converged absorption spectrum can be obtained, after what we have learned from the previous sections, with 6 bands and 60 **k** points in the IBZ. To confirm this for the excitonic effects, we depict in Fig. 4.21 an excitonic spectrum using the bands '2-7' compare it with '1-6'. The difference is only small.

Now we can present the main result of this chapter, using these parameters. In Fig. 4.22 we compare an RPA calculation with the correct quasiparticle energies (short-dashed curve), Eq. (3.21), a spectrum obtained including the local-field effects (long-dashed curve) and a full calculation including the electron–hole interaction (continuous curve), Eq. (3.88), with experiment (dotted curve) [74] measured ellipsometrically at 30 K. The absolute intensities of the excitonic spectrum agree well with experiment.

The remaining slight overestimate is of the order of what has been predicted to be the contribution of dynamical effects. The inclusion of dynamical effects only in the Green's function, used to construct the polarization function, largely reduces the curve and even increases the discrepancy with experiment at this level of approximation [92]. This reduction is compensated by the inclusion of dynamical effects also in the vertex correction [64].

More important, the peak positions and the relative intensity of the main structures are both in good agreement with experiment. The structure at 3.8 eV is overestimated due to the finite **k** point sampling. It is, however, not completely artificial and has been repeatedly observed in both theoretical [7, 85] and experimental

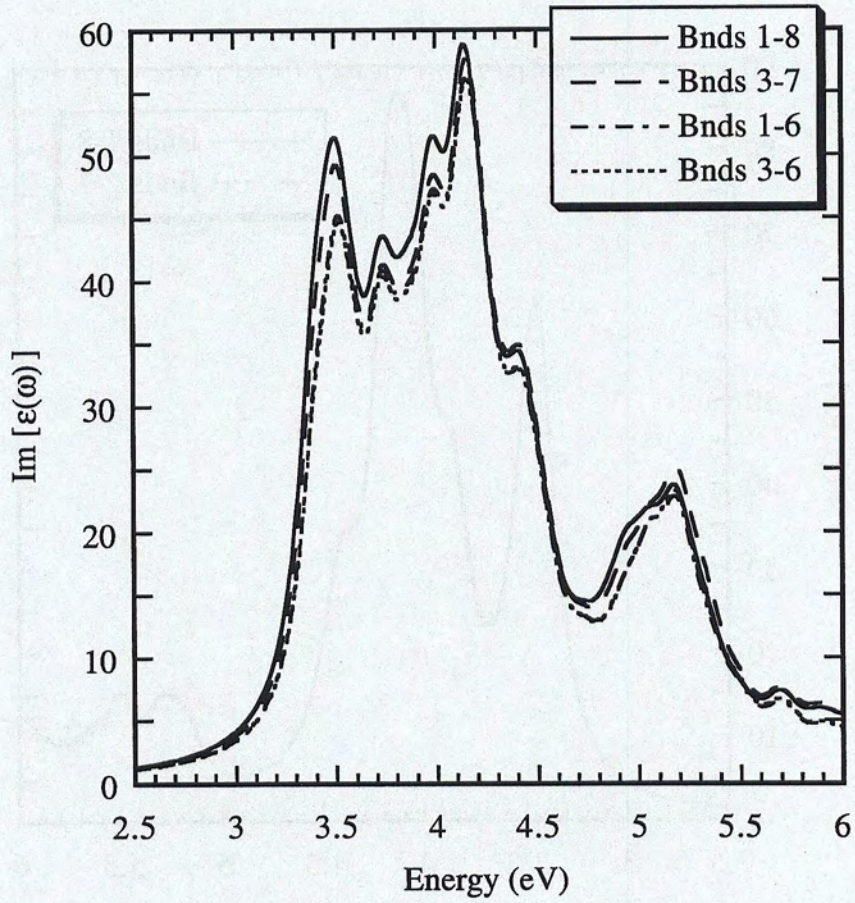


Figure 4.18: Absorption spectra of Si with excitonic and exchange effects included using 28 special **k** points in the IBZ and various combinations of bands. '3-7', *e.g.*, means including the bands 3, 4, 5, 6 and 7. 27 reciprocal lattice vectors have been taken in the respective sums of the Hamiltonian.

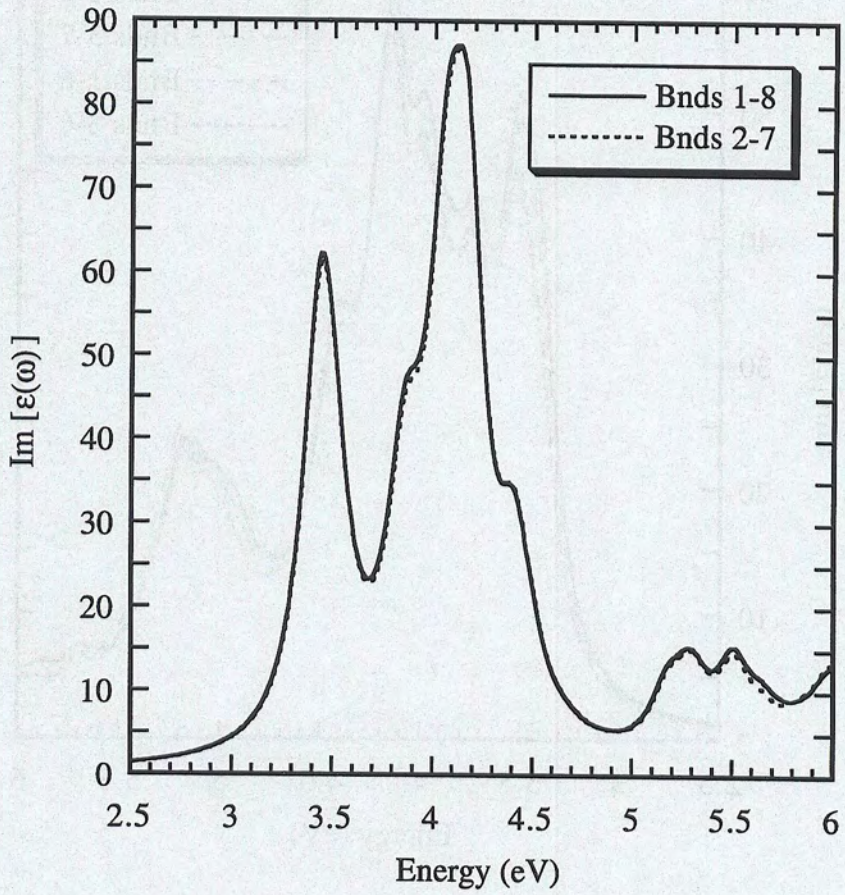


Figure 4.19: Absorption spectra of Si with excitonic and exchange effects included using 10 special **k** points in the IBZ and two combinations of bands. '2-7', *e.g.*, means including the bands 2, 3, 4, 5, 6 and 7.

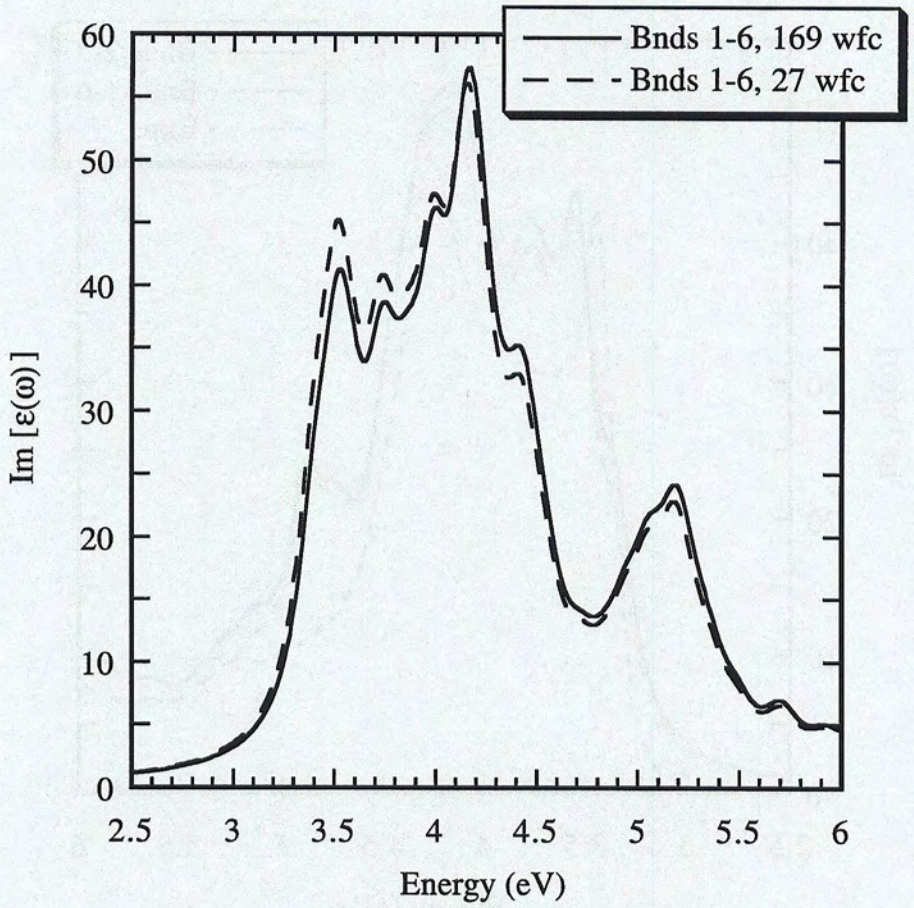


Figure 4.20: Absorption spectra of Si with excitonic and exchange effects included using 28 special **k** points in the IBZ and 6 bands. 169 (full curve) and 27 (dashed curve) reciprocal lattice vectors have been taken in the respective sums of the Hamiltonian.

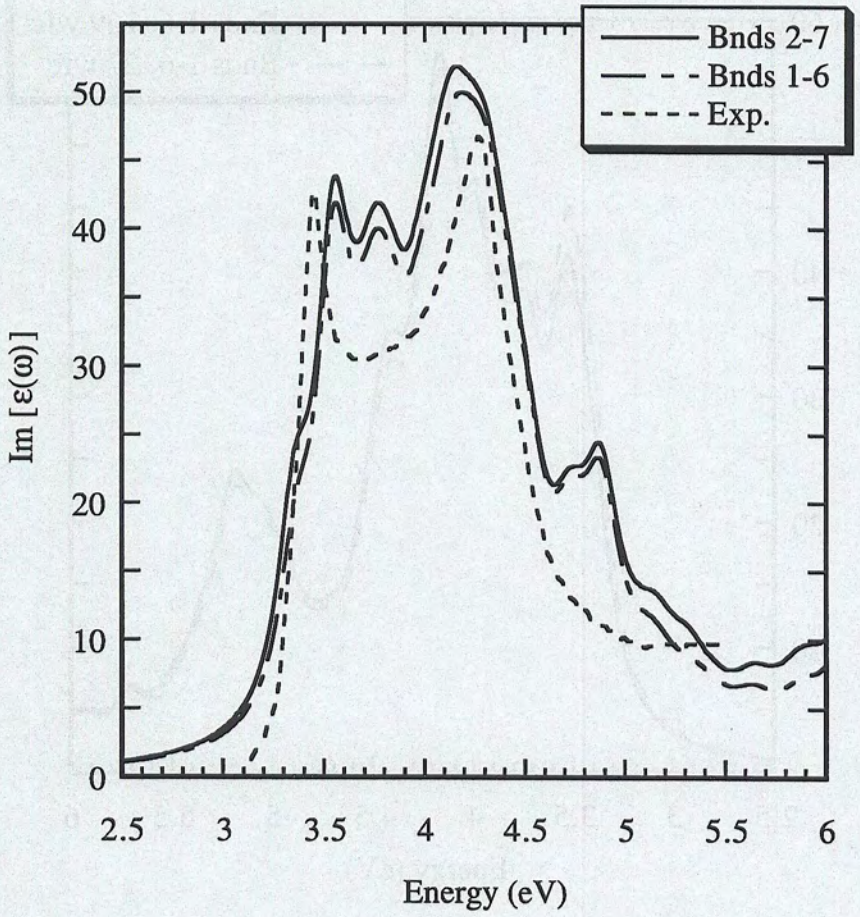


Figure 4.21: Absorption spectra of Si using 60 special **k** points in the IBZ with excitonic and exchange effects included and bands '2-7' (full curve) and bands '1-6' (dotted-dashed curve). Dotted curve: Experiment [74].

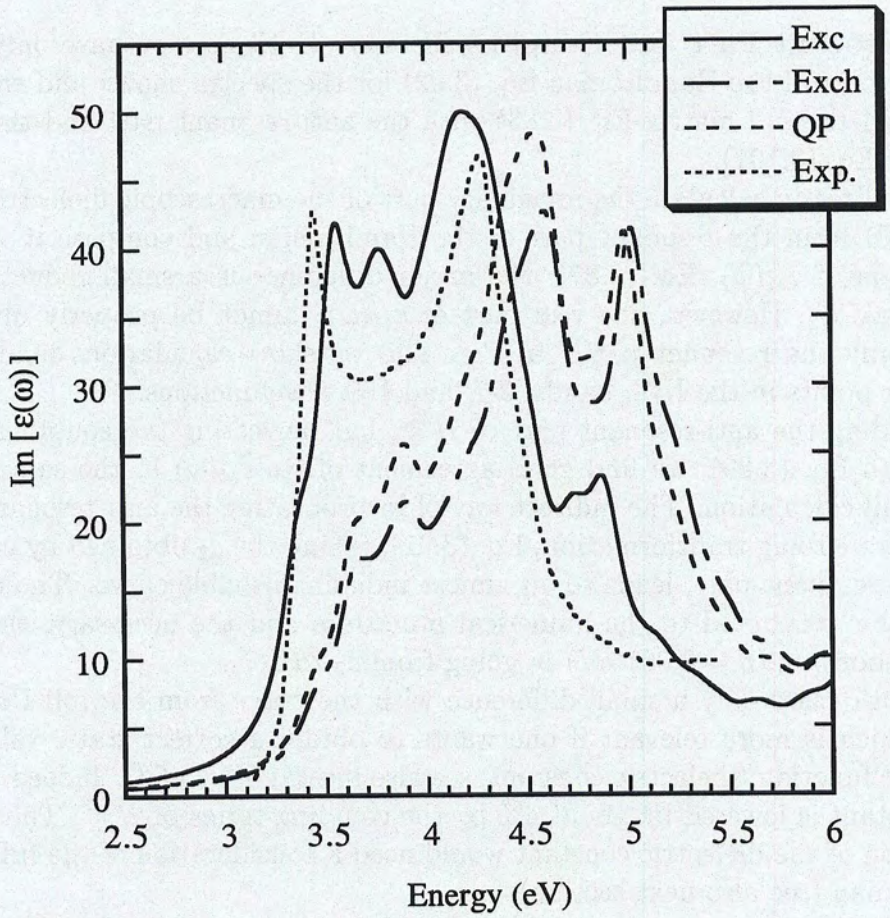


Figure 4.22: Absorption spectra of Si using 60 special **k** points in the IBZ and 6 bands. Full curve: Excitonic and exchange effects included. Dotted-dashed curve: Only exchange effects included. Short-dashed curve: RPA calculation with the quasiparticle energies. Dotted curve: Experiment [74].

work [79, 86, 93] of reflectivity measurements, which we shall examine below.

This result opens the way to first-principle calculations of optical properties with a precision comparable to that typically achieved in ground state calculations. We can thus expect to obtain quantitatively correct theoretical absorption spectra as a reference for the interpretation and even prediction of experimental results.

Anti-Resonant Part and Coupling Terms Until now we have only used the resonant part of the Hamiltonian Eq. (3.92) for the spectra shown and we have not considered the full matrix Eq. (3.83) with the anti-resonant part and the coupling elements Eq. (3.101).

If we directly calculate the imaginary part of the macroscopic dielectric function Eq. (3.88) from the resonant part of the Hamiltonian and compare it with a full calculation of $\varepsilon_M(\omega)$, Eq. (3.87), the major difference is a small reduction of the main peak E_2. However, the real part of $\varepsilon_M(\omega)$ cannot be properly obtained by taking only the resonant part. In Fig. 4.23 we show calculations made with 10 special $\mathbf{k}$ points in the IBZ, bands '2-7' and 169 wavefunctions.

Including the anti-resonant part of H^{exc}, but neglecting the coupling parts according to Eq. (3.89), we find good agreement of Re $\varepsilon_M(\omega)$ in the same Fig. 4.23 with a full calculation. The indirect way of incorporating the anti-resonant term by a Kramers-Kronig transformation, Eq. (3.91), taking the ε_2 obtained by considering only the resonant part, leads to an almost indistinguishable curve. The differences have to be attributed to the numerical procedure and the necessary, though very small smoothing ($\eta = 0.001$ eV) in going from ε_2 to ε_1.

In both cases only a small difference with the curve from the full Hamiltonian rests, which is more relevant if one wants to obtain a correct static value for the dielectric function (dielectric constant, see also inset in Fig. 4.23). Indeed the dielectric constant is lowered by about 4% by the coupling terms of H^{exc}. Thus a precise evaluation of the dielectric constant would need a consideration of the full excitonic Hamiltonian (see also next section).

Obviously, one can with confidence only take the resonant excitonic matrix for the absorption spectra, which is done in this work, if not otherwise mentioned.

Dielectric Constant In the previous section we have seen that for an exact calculation of the dielectric constant we have to consider the full excitonic Hamiltonian H^{exc}.

Nevertheless, we can examine the convergence parameters here and give values using the results obtained so far in this work. We do not consider the coupling parts of H^{exc}, but include the anti-resonant terms in all values given in Tab. 4.7.

Evidently, it is enough to take 60 special $\mathbf{k}$ points in the irreducible Brillouin zone and 12 bands in the summation over the states. However, using only bands

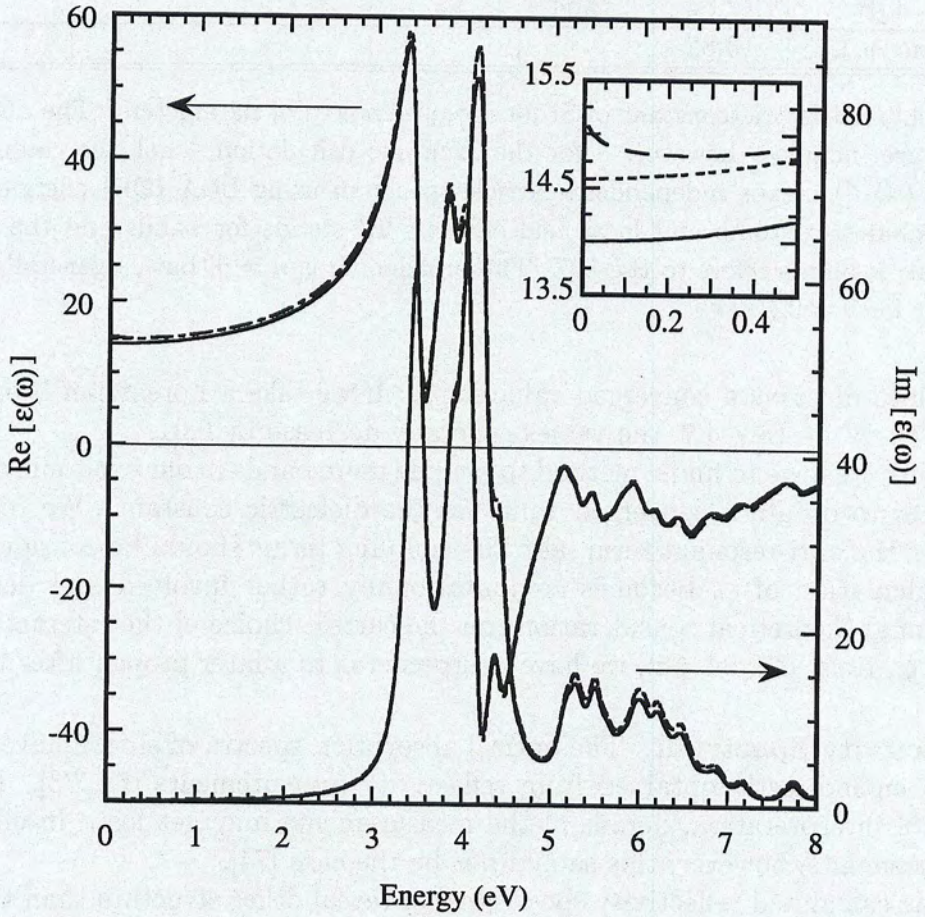


Figure 4.23: Macroscopic dielectric function of Si using 10 special **k** points in the IBZ and bands '2-7'. Real and imaginary parts are depicted in the main figure. The inset shows the real part only. Full curves: Full excitonic matrix including anti-resonant and coupling terms. Dotted curves: Including anti-resonant, but excluding coupling terms (according to Eq. (3.89)). The curves are virtually indistinguishable from the respective full curves on the large scale. Dashed curve: Real part obtained with Kramers-Kronig transformation from imaginary part. Dotted-dashed curve: Only resonant excitonic matrix (imaginary part only).

ε_0 $\eta = 0$	B 2-7 60 **k**	B 2-7 408 **k**	B 1-12 10 **k**	B 1-65 10 **k**	B 1-65 60 **k**	B 1-200 10 **k**
LDA	14.61		16.12	16.14	15.21	16.14
QP	12.64	12.64				
Exc + LF	13.82					

Table 4.7: Dielectric constant of Si for various choices of parameters. The anti-resonant terms are included, however, – for the excitonic calculation – not the coupling terms. 'LDA' ('QP') means independent-particle spectrum using LDA (QP) energies. 'Exc + LF' includes excitonic and local-field effects. 'B' stands for bands and the number of (special) **k** points refers to the IBZ. The broadening is $\eta = 0$, but ε_0 generally decreases by 0.01 for $\eta = 0.10$ eV.

'2-7' does not give a converged value for ε_0. If we take a Lorentzian broadening of $\eta = 0.10$ eV in Tab. 4.7, the values generally decrease by 0.01.

Thus, we have to find a method to include more bands in our excitonic calculation in order to obtain a converged value for the dielectric constant. We remind that besides the anti-resonant term, also the coupling terms should be considered. Since the calculation of ε_0 becomes computationally rather involved and demands for additional theoretical considerations on the correct choice of the interaction kernel, see, *e.g.*, Refs. [42, 94, 95], we have postponed it to a later project after this thesis.

Reflectivity Spectrum The optical absorption spectra of silicon have traditionally been indirectly obtained from reflectivity measurements [72, 73]. Depending on data interpretation, details of the measurements may get lost. In ellipsometry measurements, however, this should not be the case [74].

The calculated reflectivity spectrum can reveal other structure than the absorption curve and it is interesting to compare it with experiment. In Section 3.1.1 we introduced the complex refraction index, $N = n + ik$, which is related to the dielectric function by $\varepsilon = N^2$. Then the *reflectivity* for perpendicular incidence onto a plane surface is given as, by matching both the electric and magnetic fields at the surface [96]:

$$R = \left|\frac{1-N}{1+N}\right|^2 = \frac{(n-1)+k^2}{(n+1)+k^2}. \tag{4.7}$$

Actually, in reflectance-modulation measurements, like wavelength-modulation [79, 86], the *logarithmic derivative reflectivity* is obtained,

$$\frac{1}{R}\frac{\Delta R}{\Delta E}. \tag{4.8}$$

In Fig. 4.24 we depict both the reflectivity and the logarithmic derivative reflectivity spectra obtained with 60 **k** points in the IBZ and bands '2-7'. The Lorentzian broadening is chosen as $\eta = 0.15$ eV in order to smoothen the small wiggles of the

logarithmic derivative reflectivity spectrum. The anti-resonant part of H^{exc} is included, but not the coupling terms. The agreement with experiment is quite good [79, 86, 93]. Even the theoretically overestimated peak at 3.8 eV is visible as a shoulder in measurements [86, 79, 93].

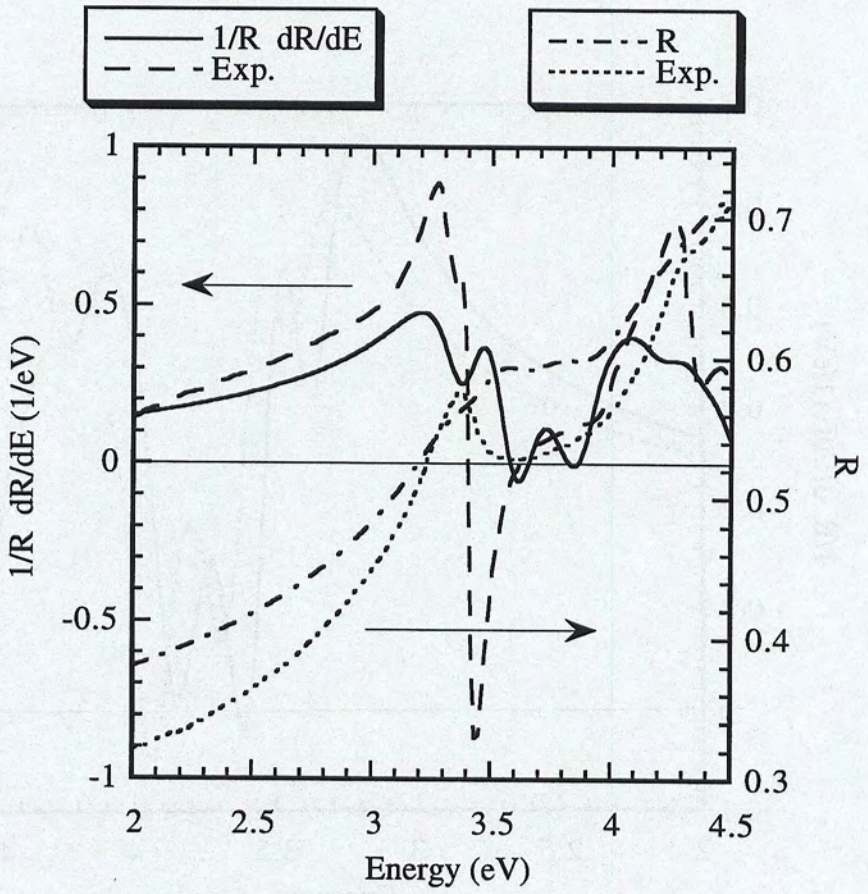


Figure 4.24: Reflectivity (dotted-dashed curve) and logarithmic derivative reflectivity (full curve) spectra of Si with excitonic and exchange effects included using 60 special **k** points in the IBZ and bands '2-7'. Anti-resonant part of excitonic Hamiltonian included, but not coupling terms. $\eta = 0.15$ eV. Experiment from Ref. [86].

The excitonic effects are not easily identified in a logarithmic derivative reflectivity spectrum, when we compare it to a spectrum taking only the exchange effects into account. Although the energetic shifts of the peaks are evident in Fig. 4.25, the shapes of the curves are quite similar to each other. Here we only use the resonant part of H^{exc} for ease.

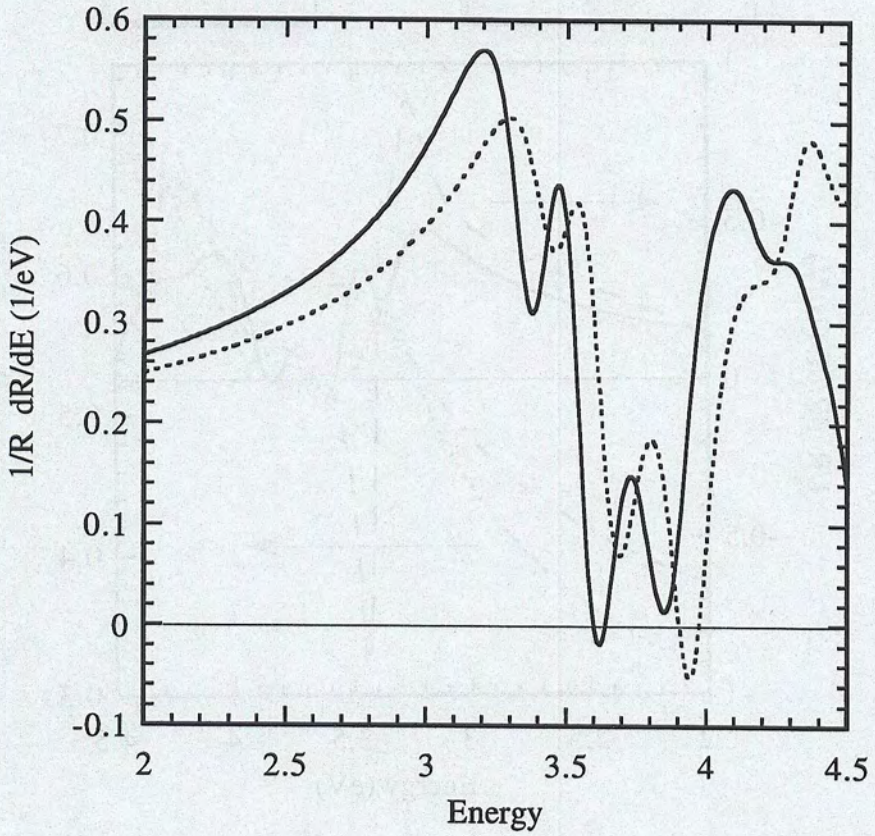


Figure 4.25: Logarithmic derivative reflectivity spectra of Si with excitonic (full curve) and exchange only (dotted curve) effects included using 60 special **k** points in the IBZ. Bands '2-7' for excitonic and '1-6' for exchange effects. $\eta = 0.15$ eV.

Origin of Peaks It is interesting to examine the origin of the two excitonic peaks E_1 and P_1 denoted in Fig. 4.26. If we restrict the energy range of the excitonic spectrum Eq. (3.88) to a certain interval $[E_{\lambda_1}, E_{\lambda_2}]$ and take in the absorption formula the quasiparticle transition energies instead of the excitonic energies,

$$\varepsilon_{\mathrm{M}}^{\mathrm{peak}}(\omega) = 1 + 2 \lim_{\mathbf{q}\to 0} v_0(\mathbf{q}) \sum_{\lambda=\lambda_1}^{\lambda_2} \sum_{v,c,\mathbf{k}} \left\{ \left| \langle v\mathbf{k} | e^{-i\mathbf{q}\cdot\mathbf{r}} | c\mathbf{k} \rangle A_\lambda^{(vc\mathbf{k})} \right|^2 \times \left(\frac{1}{\epsilon^{\mathrm{QP}}_{c\mathbf{k}+\mathbf{q}} - \epsilon^{\mathrm{QP}}_{v\mathbf{k}} - (\omega + i\eta)} + \frac{1}{\epsilon^{\mathrm{QP}}_{c\mathbf{k}+\mathbf{q}} - \epsilon^{\mathrm{QP}}_{v\mathbf{k}} + (\omega + i\eta)} \right) \right\}, \quad (4.9)$$

with the sum over all $\lambda \in \{\lambda \,|E_{\lambda_1} \leq E_\lambda \leq E_{\lambda_2}\}$, we get the contributions to the selected interval, weighted by the excitonic wavefunctions, from the independent-quasiparticle spectrum. The anti-resonant part is considered, though it does not contribute to the imaginary part of $\varepsilon_{\mathrm{M}}(\omega)$ for positive absorption frequencies ω. Fig. 4.26 shows that the small structures around 3.6 eV and 3.9 eV in the independent-quasiparticle spectrum are at the origin of the excitonic peaks E_1 and P_1. These two peaks are shifted by only 0.10 eV - 0.15 eV toward lower energies in contrast to E_2, which is displaced by 0.4 eV.

Joint Density of States In the independent-particle picture the optical transition energies are given by the differences $(\epsilon_{c\mathbf{k}} - \epsilon_{v\mathbf{k}})$ in Eq. (3.23). However, the correct values considering many-body effects are given by the excitonic eigenvalues E_λ of Eq. (3.70). Thus, using these energies in $J_{cv}(\omega)$, we obtain an excitonic density of states,

$$J_{\mathrm{exc}}(\omega) = \frac{2}{\pi\Omega} \operatorname{Im} \sum_\lambda \frac{1}{E_\lambda - \omega - i\eta}, \quad (4.10)$$

which can be directly compared with the independent-quasiparticle one. The differences between the two transition energy densities indicate the shifts of the transition energies due to excitonic and exchange effects.

In Fig. 4.27 we see that the QP transition energies are mostly unaltered and that the excitonic and local-field effects change the energy spectrum by less than 0.1 eV. Thus, the shift of the E_1 peak in the absorption spectrum Fig. 4.22 is not due to a mere reduction of the transitions energies, but stems from a redistribution of spectral weight due to excitonic effects. The electron–hole interaction leads at lower energies to a coherent superposition and to a destructive superposition at energies around about 4 eV. Thus, the formerly largely underestimated E_1 peak in the absorption spectrum is almost doubled in intensity and shifted, like the E_2 peak, to lower energies.

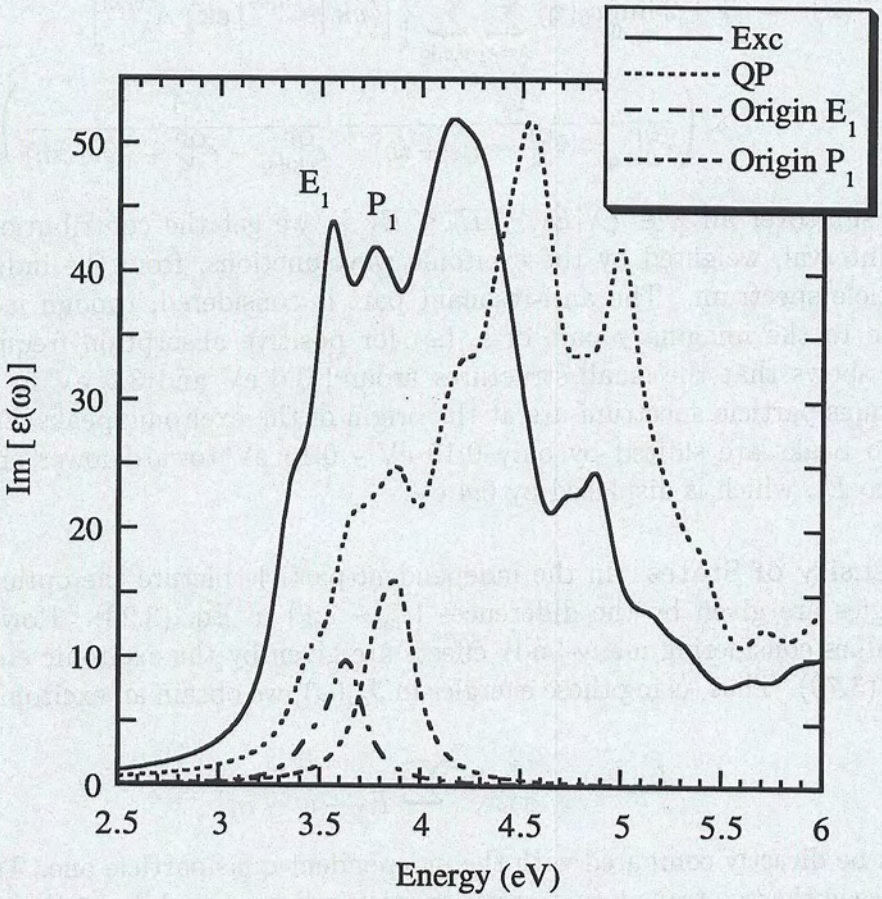


Figure 4.26: Absorption spectra of Si using 60 special **k** points in the IBZ and bands '2-7'. Full curve: Excitonic and exchange effects included. Dotted curve: RPA spectrum with QP energies. Dotted-dashed curve: According to Eq. (4.9) with 3.4 eV$\leq \omega \leq$ 3.6 eV for E_1. Dashed curve: Same, but with 3.65 eV$\leq \omega \leq$ 3.9 eV for P_1.

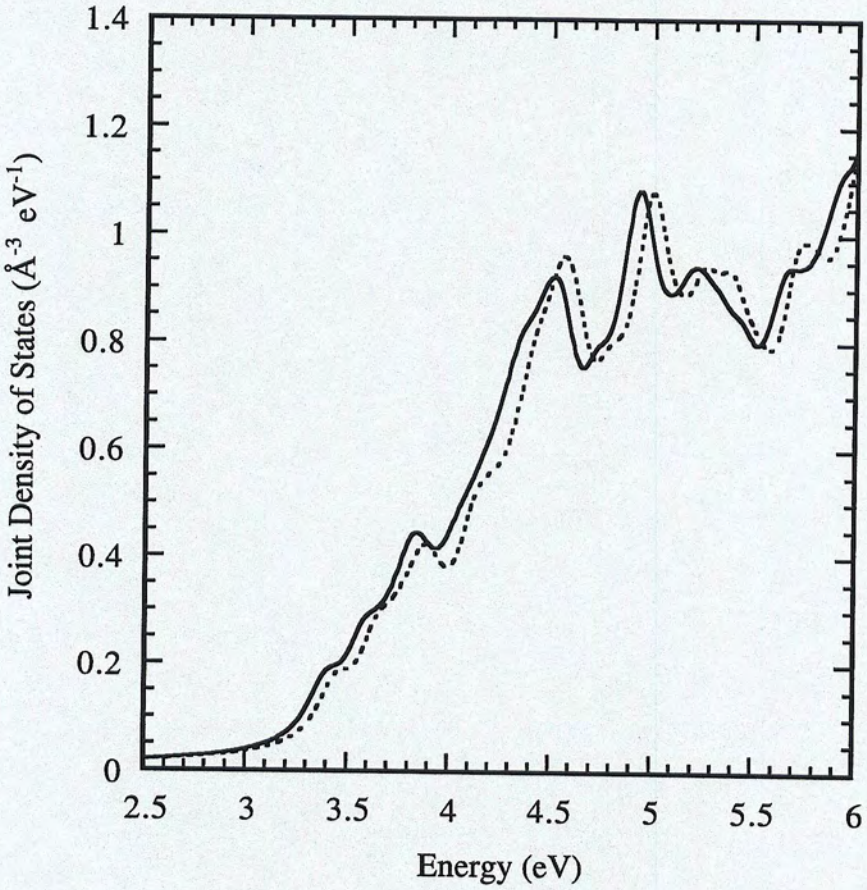


Figure 4.27: Joint density of states of silicon using QP transition energies (dashed curve), Eq. (3.23), compared with excitonic density of states (full curve), Eq. (4.10). 60 **k** points in the IBZ and bands '2-7' are used. $\eta = 0.10$ eV.

Chapter 5

Optical Absorption: Diamond

In the previous chapter we have discussed in detail the optical properties of the prototype semiconductor Si. Although clearly not a metal, the charge distribution of silicon is quite homogeneous [85]. However, diamond, which has the same space group as Si, has a less uniform charge more closely concentrated to the atoms [97]. The band gap is as wide as 5.48 eV, compared to 1.17 eV of Si [83]. Thus, we expect different aspects of the optical absorption process to be important. An application of our theory will allow us to further examine the validity of some approximations and to better understand the importance of excitonic effects in optical spectra.

Starting with a ground state calculation like for silicon, we get the Kohn-Sham band structure and the LDA absorption spectrum in the DFT-LDA framework. The discrepancy with experiment is much smaller than for Si, as the E_1 peak is quite well reproduced. However, the E_2 structure is overestimated, giving rise to a peak with a higher intensity than observed in measurement. The absorption spectrum using the correct quasiparticle energies basically only shifts the spectrum to higher energies without changing the shape significantly [9].

Like other authors before, we include local-field effects in the calculation of the dielectric matrix. However, contrary to Gavrilenko and Bechstedt [9] and Van Vechten and Martin [98], we find that the exchange effects decrease the overestimated E_2 and in addition transfer spectral weight to higher energies, although the overall effect is less significant than for Si.

In the literature, two realistic calculations have shown that the inclusion of excitonic effects is mandatory in order obtain good agreement with experiment. Hanke and Sham performed a tight-binding LCAO calculation [10]. Starting from an independent-quasiparticle absorption spectrum, which underestimates considerably the E_1 and E_2 peaks, they got close agreement with the measured data after inclusion of excitonic effects in an unscreened close contact approximation. Recently, Shirley *et al.* published an *ab initio* calculation of diamond [68], presenting good agreement with experiment.

Here we will report calculations including the *GW* corrections for the transition energies and consider the exchange as also the excitonic effects.

5.1 One-Particle Calculations

5.1.1 Ground State Calculations

We use a norm-conserving *pseudopotential* (PP) generated within the Martins-Troullier (MT) scheme [81] (see also Appendix A). The atomic state configuration is $2s^2p^2$ with cutoff radii of $r_c(s) = 1.50$ a.u. and $r_c(p) = 1.60$ a.u. The local reference potential is the p orbital (sp). The space group is the diamond structure O_h^7 (see Section 4.1.1).

We take a cutoff energy Eq. (2.32) of 50 Ry and ten special $\mathbf{k}$ points [38, 39] in the IBZ for the ground state, which is sufficient for the following electronic structure calculations [42, 81]. The Kleinman-Bylander formulation [82] (see also Appendix A) is used throughout for diamond.

We calculate 7 values around the experimental lattice constant $a_0^{\text{exp}} = 6.741$ a.u. [99] and fit the total energies with the Murnaghan equation of state [84] to obtain the theoretical equilibrium lattice constant a_0 and the bulk modulus B_0 ($B_0^{\text{exp}} = 4.43$ Mbar [100]). The results are $a_0^{\text{theor}} = 6.698$ a.u. and $B_0^{\text{theor}} = 4.47$ Mbar.

The minimum gap of diamond is indirect and between the highest occupied band at Γ and around 0.73 in direction X. At the theoretical a_0, we obtain $E_{\text{Gap}}(\Gamma - 0.73\text{X}) = 4.26$ eV.

In Fig. 5.1 we show the *band structure* for the four valence and first eight conduction bands. The eigenvalues are aligned with Γ'_{25}. In contrast to silicon (see Fig. 4.3) the bands around the gap, especially from the Γ point in the directions of the X and the L point, are not parallel. Therefore, their contributions to M_1 type critical point transitions and strong continuum excitonic effects are not as important as for Si [68].

For the following spectral calculations, we investigate the gap at the X point, around which a large region of the Brillouin zone contributes to the most prominent E_2 peak of the absorption spectrum [101], see also Fig. 5.5 and discussion in Section 5.1.3. The above configuration (MT PP with sp reference at a_0^{theor}) gives a gap at X of 11.29 eV. Taking the experimental lattice constant leads to $E_{\text{Gap}}(\text{X}) = 11.13$ eV, which is about 0.2 eV lower. This result is also obtained when using 28 special $\mathbf{k}$ points instead of ten. The inclusion of the d orbital into the pseudopotential (spd) gives 11.19 eV at a_0^{theor}, being 0.1 eV lower than our reference calculation.

Finally, the influence of an insufficient cutoff energy can further shrink the gap. Using a cutoff energy of 30 Ry instead of 50 Ry for the (spd) configuration, we obtain $E_{\text{Gap}}(\text{X}) = 11.03$ eV, which is about 0.2 eV lower. The same effect can be observed using a relatively hard Bachelet-Hamann-Schlüter (BHS) type PP for diamond [80] with a (spd) configuration and a cutoff energy of 50 Ry. We obtain at a_0^{exp} a gap of 10.78 eV for the X point, which is the value also reported by Hybertsen and Louie in Refs. [5, 42]. In contrast we get 11.01 eV with a cutoff energy of 100 Ry. Again the difference is about 0.2 eV. We note that the difference at the Γ point for this case is only 0.07 eV, so the calculation would be falsely considered to be converged

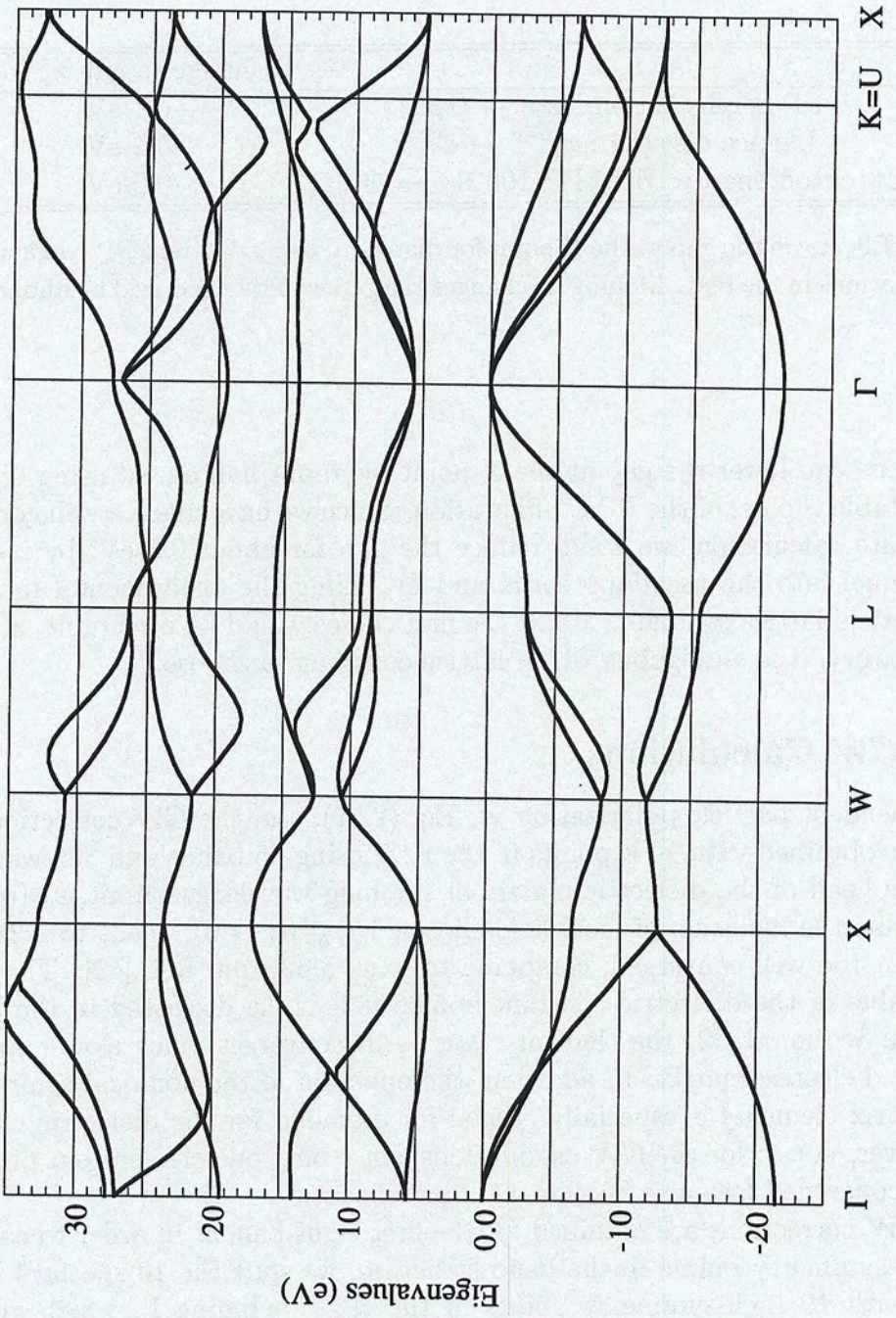


Figure 5.1: Kohn-Sham LDA band structure of diamond for four valence and the first eight conduction bands. Eigenvalues are aligned with Γ'_{25}.

looking only on this point. In Tab. 5.1 our findings are summarized.

	Change in gap at X
PP configuration: $(sp) \rightarrow (spd)$	-0.1 eV
Lattice constant: $a_0^{\text{theor}} \rightarrow a_0^{\text{exp}}$	-0.2 eV
Low cutoff energy: BHS PP 100 Ry $\rightarrow$ 50 Ry	-0.2 eV

Table 5.1: Effects on the gap at the X point for diamond, using different PP configurations. See explanations in the text. Making all changes the gap would be reduced by about 0.5 eV.

Thus one can lower the gap at the X point by about half an eV using different, but acceptable choices of the PP configuration. Since we have made a well converged ground state calculation, we could reduce the gap for about 0.3 eV, by including the d channel into the pseudopotential and by taking the experimental instead of the theoretical lattice constant. While the first choice would be preferable, although not mandatory, it is ambiguous which lattice constant to choose.

5.1.2 *GW* Calculations

The independent particle polarization χ^0, Eq. (2.57), and the *GW* corrections, Eq. (2.106) are obtained with 10 $\mathbf{k}$ points in the IBZ, using 65 bands and 339 wavefunctions. The head of the dielectric matrix in the long wavelength limit, $\varepsilon_{00}(\mathbf{q} \rightarrow 0)$, is 6.71, while the inclusion of local-field effects, $1/\varepsilon_{00}^{-1}(\mathbf{q} \rightarrow 0)$, leads to 6.23. This value is not too well converged, compared to, *e.g.*, 5.62 from Ref. [42]. The experimental value of the dielectric constant is 5.66 [83]. Like discussed in the case of silicon (see Section 4.1.2) the element $\varepsilon_{00}(\mathbf{q} \rightarrow 0)$ converges much slower than the rest of the dielectric matrix. In addition, the omission of the nonlocal commutator in the matrix elements is especially crucial for diamond for the dielectric constant [61], however, not so for the *GW* calculations, since only one element out of a large matrix is concerned (see also Section 4.1.2 on silicon).

The *GW* corrections are obtained for the first eight bands. In order to correctly treat high-symmetry points in the band structure, we shift the 10 special $\mathbf{k}$ points to a set with 19 high-symmetry points in the IBZ, including Γ, which gives by construction also 256 $\mathbf{k}$ points in the full BZ.

The plasmon pole model defined in Eq. (2.64) does not lead to expressed poles in the evaluation of the self-energy operator for diamond and we can correct small divergencies by inspection. The calculated minimum *GW* gap is 5.64 eV, with a correction of 1.38 eV for the LDA value. At Γ, the gap is widened by 1.75 eV, which is typical for other high-symmetry points, although the band dispersion is changed

by up to 0.4 eV. The experimental gap is 5.48 eV [83]. These results are typical for diamond [5, 6].

A calculation with 65 bands and 169 wavefunctions [6] does lead to the same *GW* corrections to within about 50 meV, disregarding the fact that the bands 7, 8 and especially 1 converge much slower. Nevertheless, we use all 339 wavefunctions as stated above.

In Fig. 5.2 we show the quasiparticle band structure as a full line for the first eight bands (four valence and four conduction bands) together with the corresponding LDA Kohn-Sham eigenvalues, represented by the dashes. The *GW* corrections have been interpolated with splines from 10 calculated values for each band.

Although silicon has the same group symmetry as diamond, we do find some differences in the band structures of these two materials. Besides the much smaller gap of Si and the overall larger dispersion in diamond, the important observation for optical properties is the fact that there are no parallel bands around the gap for diamond. Thus a continuum excitonic effect will be, if at all present, only small.

5.1.3 Independent-Particle Absorption Spectrum

Using the physical quasiparticle energies, obtained in a *GW* calculation described above, we calculate the independent-particle absorption spectrum according to Eq. (3.21). The *GW* corrections are interpolated from 10 special $\mathbf{k}$ points in the IBZ with a tetrahedron method [87, 88] for each band.

Like for silicon, 8 bands are enough for the region of interest up to 20 eV. The number of wavefunctions used for the matrix elements in Eq. (3.21) is taken as 411. For a very dense $\mathbf{k}$ point sampling set the optimal choice of the broadening parameter η turns out to be 0.20 eV. In Fig. 5.3 we compare the convergence of several $\mathbf{k}$ points sets, namely 10, 60 and 408 special $\mathbf{k}$ points in the IBZ. The curve with 60 $\mathbf{k}$ points is still quite rough. However, a broader smoothening with $\eta = 0.30$ eV, using only the bands '2-7' leads to a reasonable agreement with the converged curve with 408 $\mathbf{k}$ points (see Fig. 5.4).

The comparison of the quasiparticle spectrum with experiment [101, 102] in Fig. 5.5 shows a much greater discrepancy than the LDA one with experiment. Indeed, while the main peak E_2 is overestimated for both calculated curves, its energy is quite correct for the LDA spectrum, but has a large shift to higher energies for the QP one.

The inclusion of the *nonlocal commutator* into the optical matrix elements Eq. (3.20) improves the result, as can be seen in Fig. 5.5. The curve is lowered by about 15% in the interesting energy range from 7 to 17 eV, while the energies are only very slightly affected. The height of the main peak E_2 is now in good agreement with experiment.

Gavrilenko and Bechstedt find an absorption curve in Ref. [9], using a tetrahedron interpolation scheme, with again a sharp E_2 peak like for silicon. Shirley *et al.* depict an independent quasiparticle spectrum, where the E_2 peak also has a higher

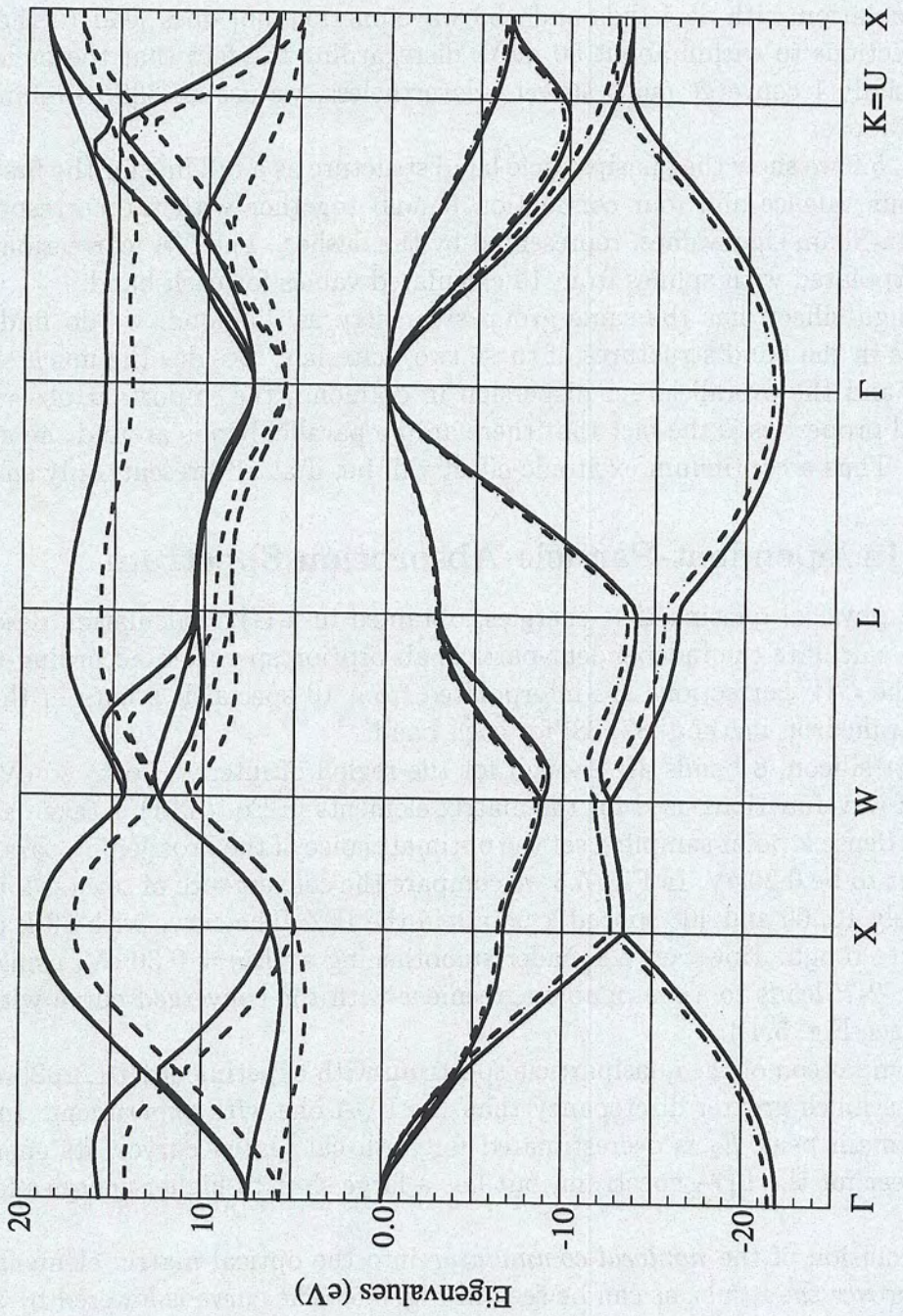


Figure 5.2: *GW* (full lines) and Kohn-Sham LDA (dashed lines) band structures of diamond for the first eight bands. Eigenvalues are aligned with Γ'_{25}.

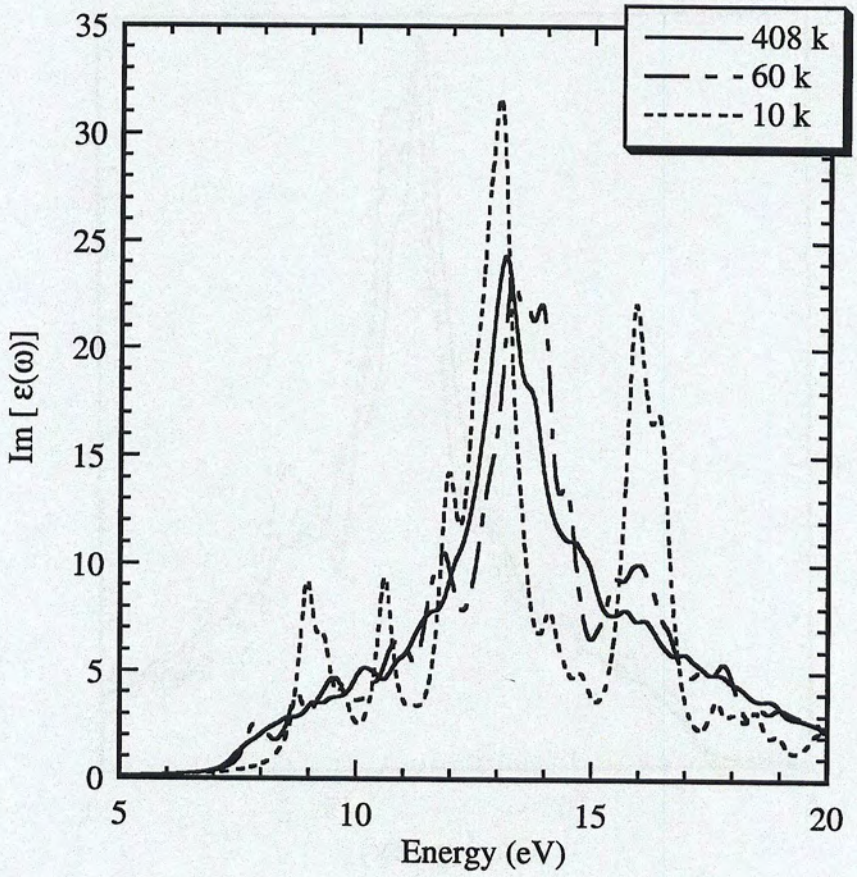


Figure 5.3: QP absorption spectra of diamond with 408 (full curve), 60 (dotted-dashed curve) and 10 (dotted curve) **k** points in the IBZ. 8 bands are used.

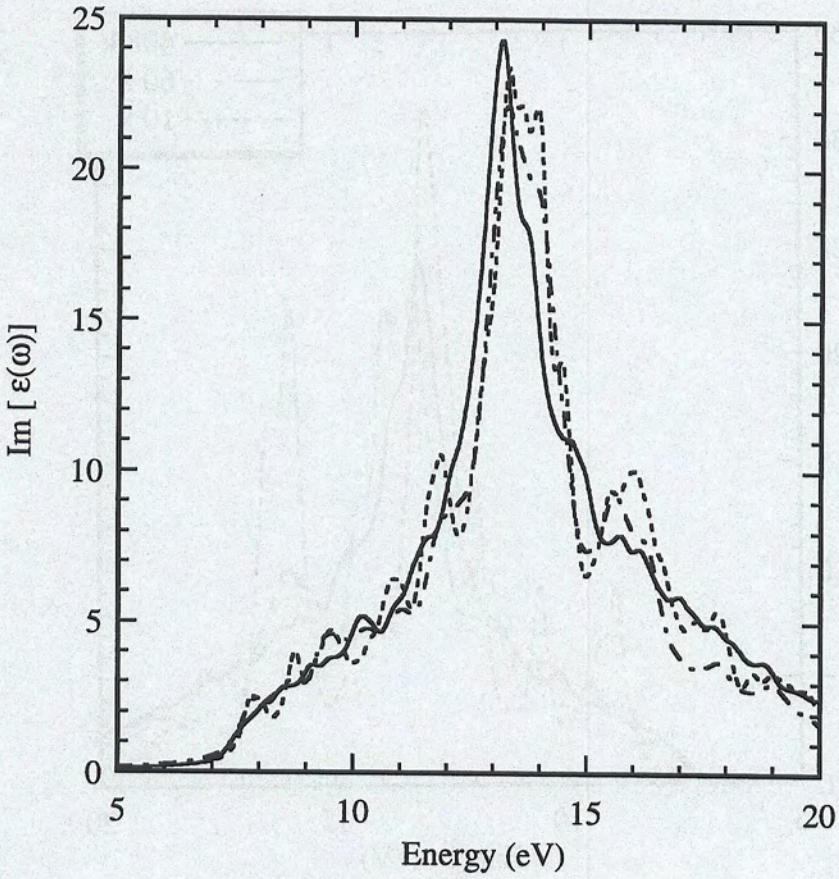


Figure 5.4: QP absorption spectra of diamond using various $\mathbf{k}$ points in the IBZ, broadening η and bands. Full curve: 408 $\mathbf{k}$, bands '1-8', $\eta = 0.20$ eV. Dotted curve: 60 $\mathbf{k}$, bands '1-8', $\eta = 0.20$ eV. Dotted-dashed curve: 60 $\mathbf{k}$, bands '2-7', $\eta = 0.30$ eV. The bands 1 and 8 do not notably contribute in the shown energy range.

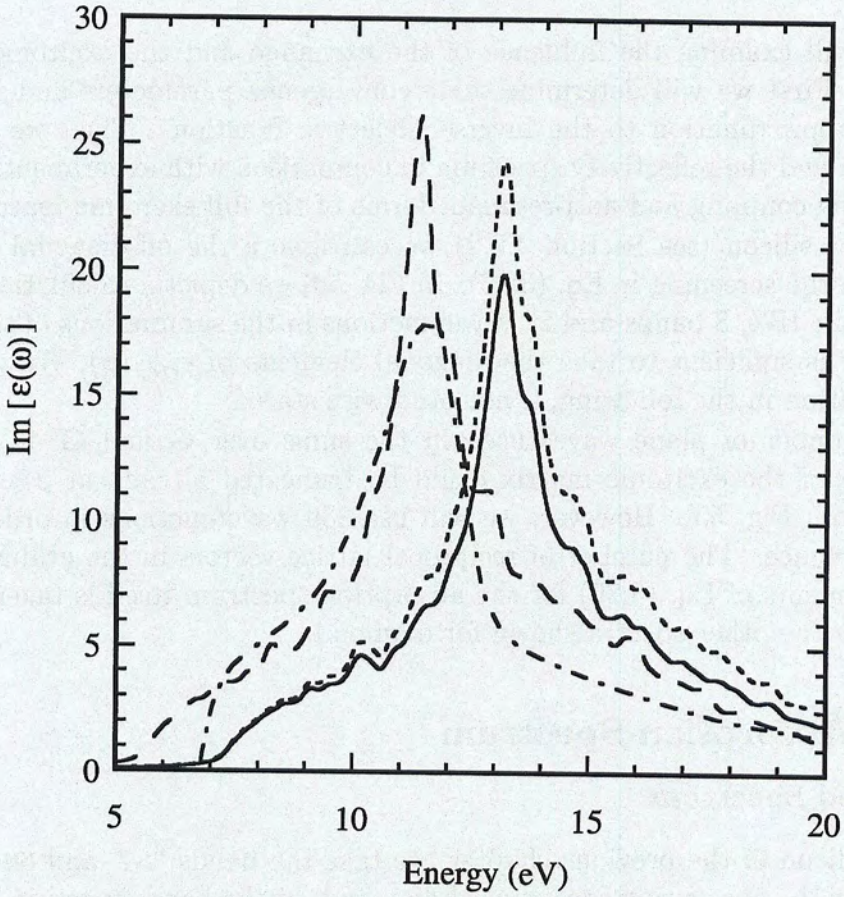


Figure 5.5: QP and LDA absorption spectra of diamond with 408 **k** points in the IBZ. 8 bands are used. Full curve: QP energies, with inclusion of nonlocal commutator in the optical matrix elements Eq. (3.20). Short-dashed curve: Same, but without nonlocal commutator. Long-dashed curve: LDA energies, without nonlocal commutator. Dotted-dashed curve: Experimental curve from Ref. [101].

intensity than in the measured curve [68]. The latter use a slightly shifted regular grid in the Brillouin zone in order to get a better average of the band dispersion in the **k** point sampling.

5.2 Many-Particle Calculations

Here we will examine the influence of the exchange and the excitonic effects for diamond. First we will determine some convergence parameters and validate the diagonal approximation to the inverse dielectric function. Then we discuss the absorption and the reflectivity spectrum in comparison with experiment. Finally we consider the coupling and anti-resonant terms of the full excitonic matrix.

Like for silicon (see Section 4.2.2) we can ignore the off-diagonal elements of $\varepsilon^{-1}_{\mathbf{GG'}}(\mathbf{q})$ in the screening in Eq. (3.96). In Fig. 5.6 we depict a calculation with 10 **k** points in the IBZ, 8 bands and 27 wavefunctions in the summations of the excitonic matrix. It is sufficient to take the diagonal elements of $\varepsilon^{-1}_{\mathbf{GG'}}(\mathbf{q})$. We will use this approximation in the following, if not otherwise stated.

The number of plane waves used in the sums over **G** and **G'** in Eqs. (3.94) and (3.96) of the excitonic matrix could be truncated already at 27, as can also be seen from Fig. 5.6. However, we will use 339 wavefunctions in order to assure full convergence. The number of reciprocal lattice vectors in the evaluation of the matrix elements of Eq. (3.20) for the absorption spectrum itself is taken as high as 411, like in the other spectra shown for diamond.

5.2.1 Absorption Spectrum

Converged Spectrum

Like for silicon in the previous chapter, we take the bands '2-7' and 60 **k** points in the IBZ for the absorption spectrum of diamond, as the band structure has roughly a similar dispersion. In Fig. 5.7 we compare an RPA calculation with the correct quasiparticle energies (short-dashed curve), a spectrum obtained including the local-field effects (long-dashed curve) and a full calculation including the electron–hole interaction (full curve) with experiment (dotted curve) [101, 102].

The already overestimated E_2 peak in the independent-quasiparticle spectrum sharpens due to the *excitonic effects* and rises above the experimental curve. Also its energy is still about 0.7 eV too high, though it is shifted by around 1.5 eV from the peak including the exchange effects only. On the low and the high energy side of the main peak the agreement with experiment has benefitted from a redistribution of spectral weight (compare also with the LDA curve of Fig. 5.5). The E_1 peak does not arise as for Si, which is a direct effect of the non-parallel bands around the gap.

The inclusion of *local-field effects* alone in the calculation of the dielectric matrix decreases the E_2 peak and transfers spectral weight to higher energies. This effect is

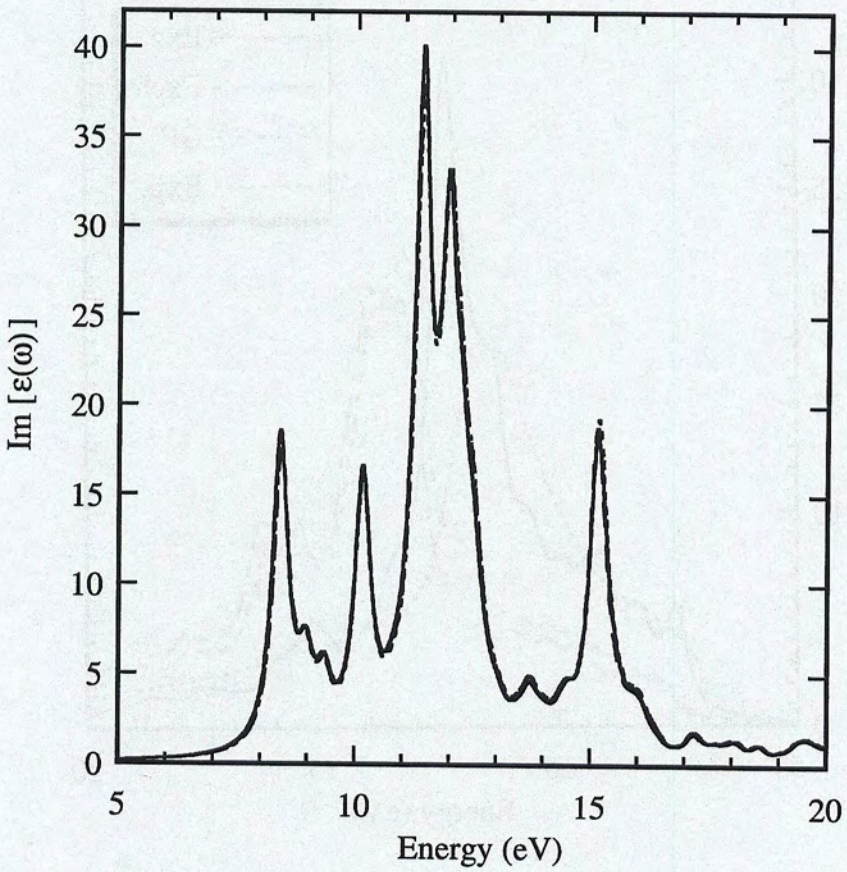


Figure 5.6: Absorption spectra of diamond with excitonic and exchange effects included using 10 special $\mathbf{k}$ points in the IBZ, 8 bands. Full curve: Diagonal approximation to $\varepsilon^{-1}_{\mathbf{GG'}}(\mathbf{q})$ in screening and 339 plane waves in excitonic matrix. Dashed curve: Like before, but 27 plane waves. Dotted curve: Full $\varepsilon^{-1}_{\mathbf{GG'}}(\mathbf{q})$ taken as screening with 27 plane waves.

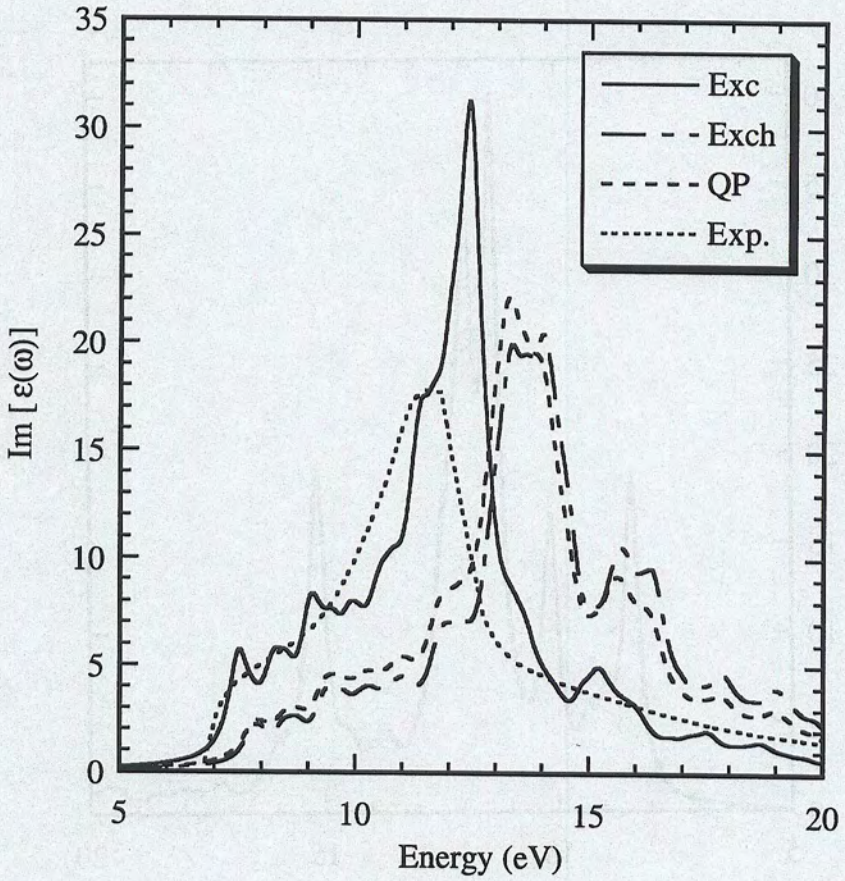


Figure 5.7: Absorption spectra of diamond using 60 special **k** points in the IBZ and bands '2-7'. Full curve: Excitonic and exchange effects included. Dotted-dashed curve: Only exchange effects included. Short-dashed curve: RPA calculation with the quasiparticle energies. Dotted curve: Experiment [101]. See also Fig. 5.8, where the nonlocal commutator is included.

less significant than for Si. As already mentioned in the introduction to this chapter, this is in contrast to Van Vechten and Martin [98] and Gavrilenko and Bechstedt [9] and is similar to the situation we have found for silicon. One may speculate if these discrepancies are due to the employed techniques. Van Vechten and Martin use an empirical pseudopotential method, while Gavrilenko and Bechstedt use the standard reciprocal space method and apply a tetrahedron interpolation scheme [87, 88]. Since we observe the same tendency as already done for silicon in the previous chapter, where we have carefully checked our approach, we believe in the effects reported here.

It is essential to include also the *nonlocal commutator* into the optical matrix elements Eq. (3.20), which can be seen in Fig. 5.8. Like for the independent-quasiparticle spectrum the curve is lowered by 15% in the important energy range from 7 to 17 eV and even more for the main peak E_2, while the energies are only very slightly affected. The height of the main peak E_2 is now in much better agreement with experiment.

In Tab. 5.1 we have explained that different, but valid choices of the PP configuration can shrink the gap at the X point by about 0.3 eV in our case. The direct effect on the RPA spectrum for independent particles is a shift of the E_2 peak (at 10.35 eV in the LDA absorption spectrum in Fig. 5.5) to lower energies, since the region around the X point in the Brillouin zone, besides the region around the Σ point, contributes most to the E_2 peak [101]. Evidently, this shift also translates into the excitonic spectrum and would give a better alignment with measurement in Fig. 5.8. Thus small and often not considered features of the ground state calculation can have noticeable effects on the spectral properties.

Dense k Point Set

In order to understand the influence of a limited $\mathbf{k}$ point set, like 60 $\mathbf{k}$ points in the IBZ as used above, on the excitonic effects, we only use the bands '4-5', but 408 $\mathbf{k}$ points in the IBZ. This set for the Brillouin zone sampling should be dense enough to overcome spurious effects in the Coulomb interaction. In Fig. 5.9 we show a spectrum including the excitonic and exchange effects in comparison with a calculation in the RPA picture for independent quasiparticles for 60 and 408 $\mathbf{k}$ points in the IBZ.

Although there are some differences in the two RPA spectra, which are indeed bigger than in Fig. 5.4, where more bands are included, the excitonic spectra are very close to each other. This confirms the use of 60 $\mathbf{k}$ points in the IBZ for the excitonic calculations. The choice of the broadening is taken from Fig. 5.4 to account for the different roughness of the curves due to the different $\mathbf{k}$ point samplings.

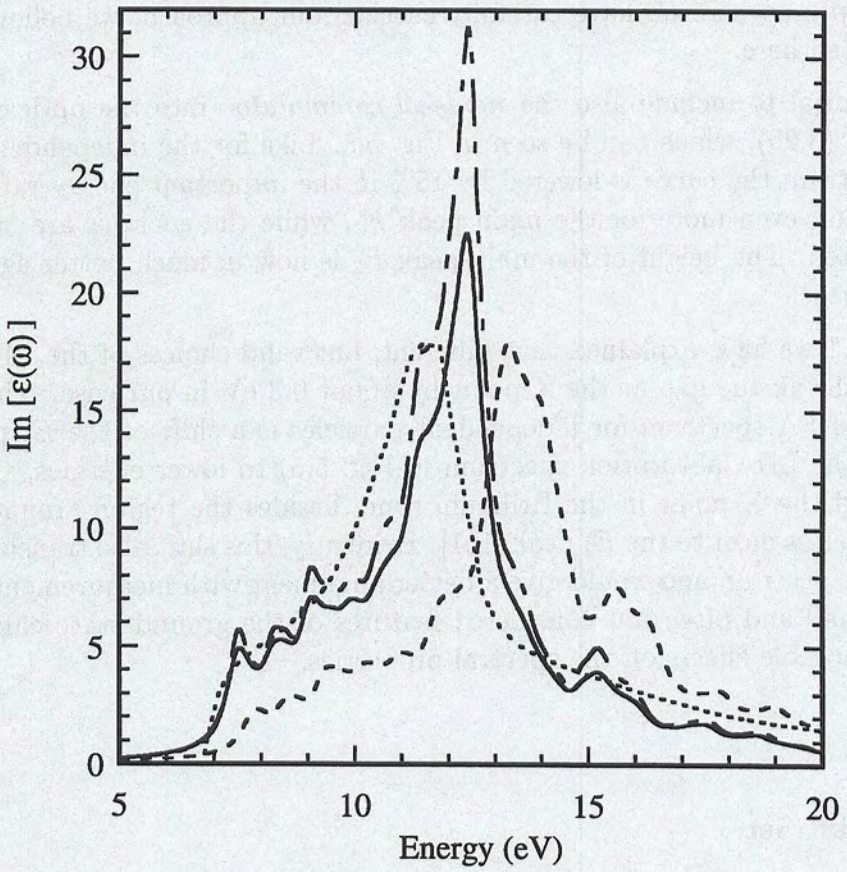


Figure 5.8: Absorption spectra of diamond using 60 special **k** points in the IBZ and bands '2-7'. Full curve: Excitonic and exchange effects with inclusion of nonlocal commutator in the optical matrix elements Eq. (3.20). Dotted-dashed curve: Like before, but nonlocal commutator not included. Dashed curve: RPA calculation with the quasiparticle energies and nonlocal commutator included. Dotted curve: Experiment [101].

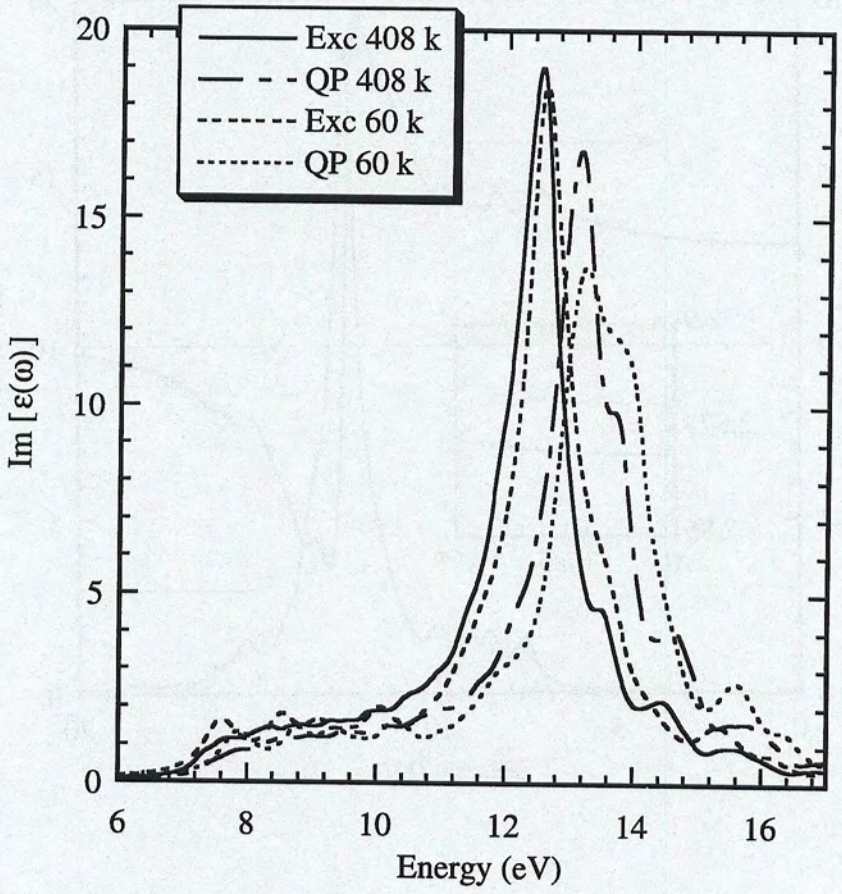


Figure 5.9: Absorption spectra of diamond using bands '4-5' with different **k** point samplings and broadening paramters η. Full curve: Excitonic and exchange effects included, 408 special **k** points in the IBZ, $\eta = 0.20$ eV. Dotted-dashed curve: Like before, but RPA calculation with quasiparticle energies. Long-dashed line: Excitonic and exchange effects included, 60 special **k** points in the IBZ, $\eta = 0.30$ eV. Short-dashed curve: Like before, but RPA calculation with quasiparticle energies.

5.2.2 Anti-Resonant Part and Coupling Terms

We want to consider the full Hamiltonian matrix Eq. (3.83) with the anti-resonant part and the coupling elements.

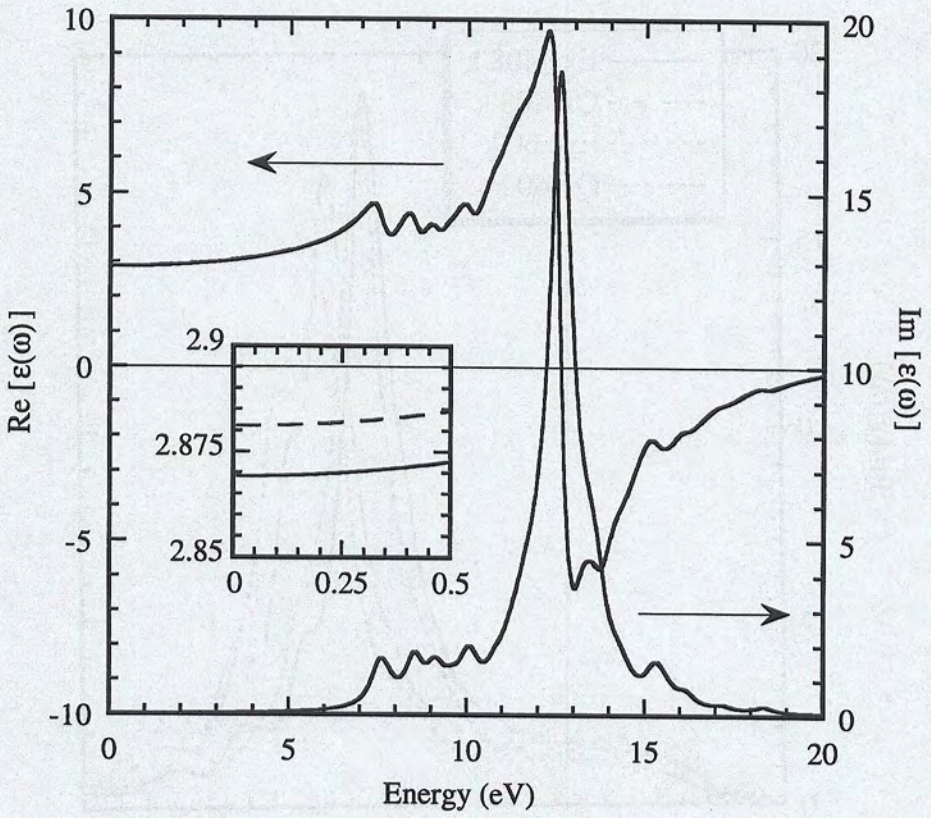


Figure 5.10: Macroscopic dielectric function of diamond using 60 special **k** points in the IBZ and bands '4-5'. Real and imaginary parts are depicted. The inset shows the real part only. Full curves: Full excitonic matrix including anti-resonant and coupling terms. Dashed curves: Including anti-resonant, but excluding coupling terms (according to Eq. (3.89)). Dotted curve: Only resonant excitonic matrix (only for imaginary part).

In Fig. 5.10 we show calculations made with 60 special **k** points in the IBZ, bands '4-5' and 339 wavefunctions for the sums over the **G** vectors in the excitonic Hamiltonian. Like for silicon, the full exciton matrix does give very similar results for Re $\varepsilon_M(\omega)$ compared to the resonant one, if the anti-resonant solutions – though omitting the coupling terms – are properly included. For the imaginary part of $\varepsilon_M(\omega)$, Eq. (3.87), there is hardly any difference visible between the full matrix, omitting the coupling term, but including the anti-resonant part and taking only

the resonant part.

Inclusion of the anti-resonant matrix according to Eq. (3.89) gives very close agreement with the full Hamiltonian for the real as well as the imaginary part of ε_M. Thus, again we can restrict ourselves to the resonant excitonic matrix for absorption spectra.

Also the static value of the macroscopic dielectric function (dielectric constant) is only slightly influenced by the coupling terms of H^{exc}. The exact evaluation of the dielectric constant for diamond does hence not demand a consideration of the full excitonic Hamiltonian.

5.2.3 Dielectric Constant

The dielectric constant of diamond can be calculated in very good approximation including only the anti-resonant term, but not the coupling terms in H^{exc}, as seen in the previous section.

Like for silicon in Tab. 4.7, we examine the convergence parameters and give values using the results obtained so far in this work. We do not consider the coupling parts of H^{exc}, but include the anti-resonant terms in all values given in Tab. 4.7.

ε_0 $\eta = 0$	B 2-7 60 **k**	B 1-12 10 **k**	B 1-65 10 **k**	B 1-65 60 **k**	B 1-200 10 **k**
LDA (no NLC)	6.22	6.69	6.71	6.71	6.71
LDA (with NLC)		5.80			
QP (no NLC)	5.45				
QP (with NLC)	4.78				
Exc + LF (no NLC)	6.14	6.70			
Exc + LF (with NLC)	5.38	5.81			

Table 5.2: Dielectric constant of diamond for various choices of parameters. The anti-resonant terms are included, however, – for the excitonic calculation – not the coupling terms. 'LDA' ('QP') means independent-particle spectrum using LDA (QP) energies. 'Exc + LF' includes excitonic and local-field effects. 'B' stands for bands and the number of (special) **k** points refers to the IBZ. 'NLC' means the nonlocal commutator in the optical matrix elements. The broadening is $\eta = 0$, but ε_0 generally decreases by only 0.01 for $\eta = 0.30$ eV.

Evidently, it is enough to take 10 special **k** points in the irreducible Brillouin zone and 12 bands in the summation over the states. Using only bands '2-7' does not give a converged value for ε_0. If we take a Lorentzian broadening of $\eta = 0.30$ eV in Tab. 5.2, the values generally decrease by 0.01.

The dielectric constant measured at ambient temperature is 5.66 [83]. Our converged value of $\varepsilon_0 = 5.81$, including excitonic and local-field effects, agrees within less than 3%. We note that almost the same number is obtained with a simple RPA calculation using Kohn-Sham LDA eigenvalues, whereas the RPA QP calculation gives a much lower value.

5.2.4 Reflectivity Spectrum

In Fig. 5.11 we depict the reflectivity, Eq. (4.7), obtained with 60 $\mathbf{k}$ points in the IBZ and bands '2-7'. The anti-resonant part of H^{exc} is included for the excitonic spectrum, but not the coupling terms. Also in the RPA spectrum the anti-resonant contributions have been considered. The nonlocal commutator in the optical matrix elements Eq. (3.20) has been included. For the excitonic calculation the agreement with experiment for the reflectivity is quite good for energies up to 12 eV [101, 102], while the RPA spectrum is shifted to higher energies and shows a somewhat different curvature. For higher energies one must use more bands and also a denser $\mathbf{k}$ point set in order to better represent the band structure in these energy transition regions.

5.2.5 Joint Density of States

Like for silicon in Fig. 4.27 we find that the excitonic effects only slightly shift the transition energies of independent quasiparticles, $(\epsilon_{c\mathbf{k}} - \epsilon_{v\mathbf{k}})$, by up to 0.2 eV to the excitonic eigenstates E_λ. This can be seen in Fig. 5.12, where we compare the joint density of states, Eq. (3.23), with the excitonic density of states, Eq. (4.10). Again the changes in the excitonic spectrum, compared to the RPA QP calculation, see Fig. 5.8, is due to a redistribution of spectral weight and not to a simple shift of transition energies.

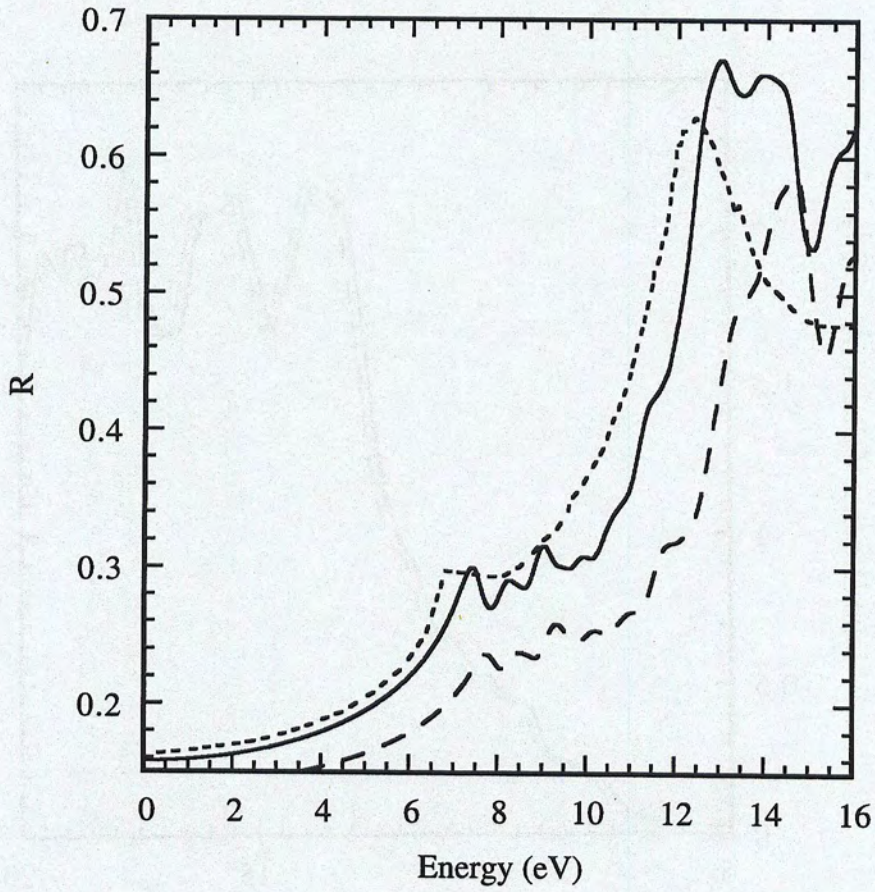


Figure 5.11: Reflectivity spectra of diamond using 60 special $\mathbf{k}$ points in the IBZ and bands '2-7'. $\eta = 0.30$ eV. Anti-resonant parts of spectra included, but not coupling terms of excitonic Hamiltonian. The nonlocal commutator in the optical matrix elements has been included. Full curve: Excitonic and exchange effects included. Long-dashed curve: RPA calculation with quasiparticle energies. Short-dashed curve: Experiment [101].

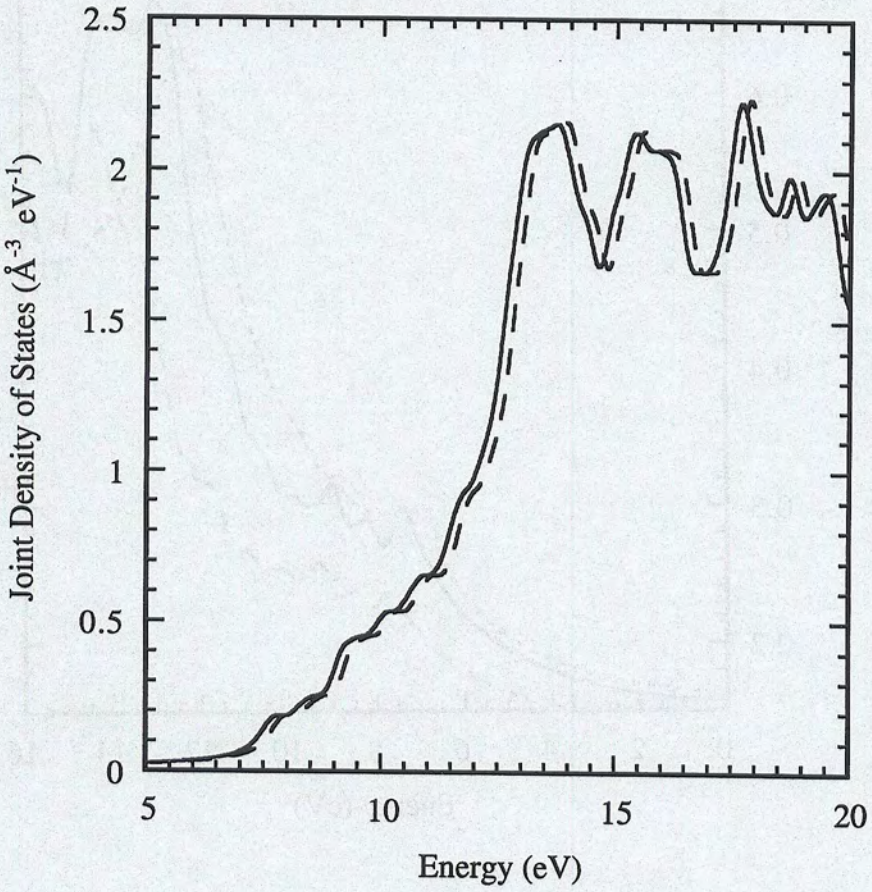


Figure 5.12: Joint density of states of diamond using QP transition energies (dashed curve), Eq. (3.23), compared with excitonic density of states, Eq. (4.10), (full curve). 60 **k** points in the IZB and bands '2-7' are used. $\eta = 0.30$ eV.

Chapter 6

Conclusion

In this thesis we have for the first time presented an *ab initio* approach to include excitonic effects in optical absorption spectra of semiconductors and insulators. We have demonstrated the feasibility of our method at the examples silicon and diamond obtaining results in good agreement with experiment.

We have derived the relevant formulas in a consistent way and developed and tested the computer codes of the excitonic calculations with, in particular, the implementation of the symmetry properties in order to reduce the excitonic Hamiltonian. In addition, the routines to obtain the macroscopic dielectric function in various ways presented in the course of this work have been written: independent-particle picture, inclusion of excitonic and/or local-field effects, consideration of the anti-resonant and coupling terms, Taylor expansion scheme of Section 4.1.3, and others.

Furthermore, several changes in the existing Car-Parrinello code for the ground state calculations and the *GW* programs for the RPA inverse dielectric matrix and the self-energy corrections to the Kohn-Sham LDA energies have been made. Finally, some small, but important routines, like the tetrahedron interpolation scheme, here used for the *GW* corrections of the Kohn-Sham LDA energies in the spectral calculations, have been developed.

6.1 Theoretical Developments

From the theoretical point of view, the observed spectrum is given by the imaginary part of the macroscopic dielectric function, which essentially depends on the polarization being the response of the system to an external perturbation. Schematically this can be written in the form, Eq. (3.33),

$$\varepsilon_{\mathrm{M}} = 1 - v\bar{P}. \tag{6.1}$$

In the absorption process two principal many-body effects can be identified: the excitonic effects, resulting from electron–hole Coulomb interaction and the electron–hole exchange contributions, reflecting the microscopic structure of the system, thus also called local-field effects.

Starting from Hedin's coupled integral equations we have derived an integral equation for the polarization by an iteration beyond the GW approximation, Eq. (3.45),

$$P = P_0 - PWP_0. \tag{6.2}$$

This expression can be used as an input for another integral equation describing the connection between the polarization propagator and – indirectly – the macroscopic dielectric function, Eq. (3.37),

$$\bar{P} = P + P\bar{v}\bar{P}. \tag{6.3}$$

The resulting Bethe-Salpeter equation, Eq. (3.47),

$$\bar{P} = P_0 + P_0(\bar{v} - W)\bar{P}, \tag{6.4}$$

is the fundamental four-point equation governing the optical response of the system, which we have to solve for $\bar{P}$ [67, 103]. The interaction kernel includes the exchange, $\bar{v}$, as well as the excitonic, W, effects. The dash indicates the omission of the long-range term of vanishing wave vector in the bare Coulomb interaction v for the macroscopic dielectric function.

The task of solving for $\bar{P}$ is technically very demanding, since even in a natural basis of Bloch function, given by valence and conduction bands and $\mathbf{k}$ vectors in the Brillouin zone, we have to represent $\bar{P}$ as a matrix of the order of some $10^4 \times 10^4$ elements.

Our ansatz starts with the transformation of Eq. (6.4) to an eigensystem [66], Eq. (3.68),

$$H\,|\lambda\rangle = E_\lambda\,|\lambda\rangle\,, \tag{6.5}$$

with the excitonic, in general non-Hermitian Hamiltonian having the structure, Eq. (3.64),

$$H = \Delta E^{\mathrm{QP}} + (\bar{v} - W)\,. \tag{6.6}$$

Thus we avoid an inversion process for each absorption frequency, since from the spectral decomposition, we can then directly obtain $\bar{P}$ as, Eq. (3.86),

$$\bar{P}(\omega) = \sum_{\lambda,\lambda'} \frac{|\lambda\rangle\,\langle\lambda\,|\lambda'\rangle^{-1}\,\langle\lambda'|}{E_\lambda - \omega}. \tag{6.7}$$

The Hamiltonian H has four principal blocks contributing to the macroscopic dielectric function. The most important block for absorption is the resonant part with positive independent-particle energies ΔE^{QP} on its diagonal, Eq. (3.78). It has positive absorption frequencies and is by itself Hermitian. The second block is the anti-resonant part, being the negative Hermitian complex conjugate of the resonant part, see also Eq. (3.83). It is especially relevant for the static dielectric constant. These two blocks are connected by the coupling parts having only the interaction kernel $(\bar{v} - W)$ as contribution.

An important advantage of our approach is the explicit knowledge of the excitonic eigenstates $|\lambda\rangle$ and eigenvalues E_λ, which allows us a detailed analysis of the optical properties. Knowing the wavefunctions we can, for example, plot the charge densities of the excited electron and hole. This is particularly instructive in the case of a bound excitonic state in a cluster, where strong mixing of independent-particle transitions to the excitonic eigenstate occurs, as has been demonstrated in Appendix B with the sodium tetramer. The analysis in the solid is more complicated, but we can still identify the principal independent-particle transitions to a given energy region, see Fig. 4.26 for silicon.

We can also find the transverse-longitudinal splitting in the excitonic eigenvalues by including the long-range term of vanishing wave vector in the bare Coulomb interaction v, although it does not contribute to the macroscopic dielectric function. In the case of Li_2O we have found a split of about 50 meV for the lowest excitonic eigenvalue being a bound state with a binding energy of 0.8 eV compared to a minimum direct gap of 7.4 eV [23]. This eigenstate has been identified as the onset of optical absorption (see Appendix A).

Parallel to this work, Shirley and co-workers presented an approach in a mixed basis in real and reciprocal space [68, 69], which enables an *ab initio* solution by inversion. Using the Lanczos or Haydock recursion method [104], their ansatz is numerically very efficient. However, they only obtain the spectral curve and therefore do not directly get information about the excitonic eigenstates. Furthermore, the repeated recursion process does not allow to obtain fine details of the spectrum due to limited resolution.

Recently, Rohlfing and Louie have demonstrated the advantages of our method [105]. They have calculated, using a localized basis set, the absorption spectrum for GaAs and determined the bound excitonic states below the fundamental gap, including the transverse-longitudinal splitting, in good agreement with experiment. Furthermore, they have shown that also a highly ionic material, LiF, with strongly bound excitons below the gap can be adequately described. The spectrum including excitonic and exchange effects is dramatically changed, now in good agreement with experiment compared to the calculation in the independent-particle picture.

6.2 Optical Absorption: Silicon

The paradigmatic case for the optical absorption in solids is silicon, being the standard example of a semiconductor. It has served as a test case for our programs and then as the first application of our method to a real material. Indeed, silicon is quite a demanding system considering its large excitonic effects, mixing states lying in broad regions of the Brillouin zone.

The simple RPA picture of independent-particle transitions from the valence to the conduction bands in the absorption process does not lead to a satisfying agreement with experiment. This result has been know since a long time, see, *e.g.*,

Refs. [7, 9]. In this context the spectrum does not improve when using *GW* energies, *i.e.*, the real quasiparticle energies, instead of Kohn-Sham LDA eigenvalues for the one-particle transition energies, as can be seen in Fig. 4.7.

Also the inclusion of local-field effects is not enhancing the agreement with experiment. Rather we have found in Fig. 4.11 that oscillator strength is transferred to higher energies in contrast to the measured curve and agreement is even worsened, as already reported by other workers before, see Refs. [9] or [12].

Only after the inclusion of excitonic effects in the calculation of the absorption curve we have been able to obtain good agreement with experiment in Fig. 4.22. This result for silicon was previously reported in the LCAO tight-binding calculation of Hanke and Sham [11] and has recently been confirmed by Shirley *et al.* [68]. We have seen in Fig. 4.27 that a redistribution of spectral weight due to excitonic effects doubles the formerly largely underestimated E_1 peak in the absorption spectrum. Such modifications of spectra also occur in other semiconductors [68, 105].

At the example of silicon we have investigated the possibilities to approximate the Hamiltonian matrix by neglecting the coupling between the bands or to consider only a constant screening in the electron–hole interaction. These simple assumptions are part of the Mott-Wannier model. In Fig. 4.17 we have shown that both approximations are not able to reproduce the excitonic effects in the E_1 peak. The nearly parallel conduction and valence bands around the gap lead to expressed continuum-excitonic effects and several bands around the gap couple. The spatial extension of the electron–hole pair is obviously not big enough to be accounted for by only taking a macroscopic long-range screening. However, the electron–hole screening can be taken diagonal, thus neglecting the off-diagonal elements of $\varepsilon^{-1}_{\mathbf{GG'}}(\mathbf{q})$, as shown in Fig. 4.16. This approximation has also been applied to diamond (see below and Fig. 5.6).

Neglecting the coupling parts, the Hamiltonian, Eq. (3.83), becomes block-diagonal and the anti-resonant contributions can be easily incorporated, Eq. (3.89). In the real part of the macroscopic dielectric function the anti-resonant part of the excitonic Hamiltonian is essential, while the coupling terms can be neglected for the spectrum in good approximation, see Fig. 4.23. However, the macroscopic dielectric constant, $\varepsilon_{\mathrm{M}}(0)$, is notably changed by the coupling parts: about 4% for a limited $\mathbf{k}$ point sampling, Tab. 4.7. For a converged value we would have to include at least 8 conduction bands. The imaginary part of $\varepsilon_{\mathrm{M}}(\omega)$ is only slightly influenced by both the anti-resonant as well as the couplings terms.

In the reflectivity spectra, both the real and the imaginary part of ε_{M} are included and we have obtained good agreement with experiment in Fig. 4.24. The excitonic effects are by far not as evident in the logarithmic derivative reflectivity spectrum as in the absorption spectrum, see Fig. 4.25.

6.3 Optical Absorption: Diamond

Although diamond has the same space group as silicon, there are two important differences concerning the optical properties: i) diamond has a much larger band gap of 5.48 eV instead of 1.17 eV for silicon [83], which reduces the dielectric constant and thus the screening; ii) diamond does not have parallel or nearly parallel bands around the band gap, see Fig. 5.2. The reduced screening leads to a stronger electron–hole interaction and one would therefore expect more expressed excitonic effects. However, the minimum gap is indirect – like for silicon – and the nonparallel bands around the gap do give a less coherent superposition of the oscillator strengths.

Starting with an RPA calculation and using LDA energies, we have noted in Fig. 5.5 that the inclusion of the self-energy corrections worsens agreement with experiment. The curve using the quasiparticle energies is shifted to higher energies. In both cases the shape of the spectrum is not in too good agreement with experiment. The inclusion of the nonlocal commutator (see also next section) in the optical matrix elements, Eq. (3.20), is found to be crucial for diamond, in particular to reduce the overestimated E_2 peak.

Consideration of the local-field effect deteriorates agreement with experiment. Like for silicon, we have found a decrease of the spectrum on the low energy side and a shift of oscillator strength to higher energies, see Fig. 5.7. Only after including the excitonic effects we come to good agreement with experiment in Fig. 5.8. The shift of the absorption curve is due to, as for silicon, a redistribution of spectral weight and not just to a reduction of the transition energies (Fig. 5.12). The inclusion of the nonlocal commutator is even more important for the excitonic spectrum than for the simple RPA spectrum.

The excitonic absorption spectrum still suffers from a small misalignment of the most prominent E_2 peak, which can, however, be partly attributed to the choices made in the pseudopotential configuration in the ground state calculation (see also next section). Thus small changes in the first step, the DFT-LDA calculation, can be more expressed in the final step, the spectral calculation.

The coupling elements of the full excitonic Hamiltonian, Eq. (3.83), have been shown to be less important for diamond than for silicon in Fig. 5.10, which is attributed to the larger gap, making the diagonal RPA transition elements relatively more important. Again, we have to include the anti-resonant part of H^{exc} into the real part of $\varepsilon_{\mathrm{M}}(\omega)$ for the spectrum, while we can neglect it for the imaginary part, the absorption spectrum. The calculated dielectric constant of 5.81 is in good agreement with experiment, see Tab. 5.2.

Finally, the reflectivity spectrum agrees up to energies of about 12 eV quite well with experiment and excitonic effects are not more expressed than in the absorption spectrum, see Fig. 5.11. For higher energies above 14 eV, more bands and also $\mathbf{k}$ points should be included in the calculation.

6.4 Ground State Calculations and Spectroscopic Properties

The choice of the pseudopotential (see Appendix A) can be important for the optical spectrum and we have investigated this aspect at various points. Indeed, already the first step, namely the ground state calculation, is crucial for the details in the final spectrum. In the case of diamond we have seen that different, but reasonable choices of the pseudopotential (PP) configuration can shift the main peak in the absorption spectrum by up to half an eV, which is of the same order as the *GW* corrections for the Kohn-Sham LDA energies or the excitonic effects for the shift of oscillator strength in the absorption spectrum, see Tab. 5.1.

The ground state calculations are generally speeded up using the separable Kleinman-Bylander (KB) form of the pseudopotential. However, this can influence the minimal Kohn-Sham LDA gap. In Tab. A.3 we have shown at the example of Li_2O that even when the theoretical minimal lattice constant a_0 is only slightly changed by the KB separation, Tab. A.1, the gaps and therefore the position of the peaks in the absorption spectrum can nevertheless be somewhat different according to the chosen PP configuration.

We note that silicon is generally not very sensitive to the PP configuration and consequently we have observed only a small influence of the KB separation. Nevertheless, in order not to bias our results we have not employed the KB separation being silicon the first application of our method.

Furthermore, for spectral calculations we have to be careful about the individual optical matrix elements, since they give the transition probability. For Li_2O we have used a test model system to check the optical matrix elements, taking only the Γ point in the Brillouin zone sampling, and we have compared three PP configurations with and without employing the KB separation, Figs. A.5-A.9. The results support the use of the KB separation and of very soft PPs, though the application of very soft PPs in spectral calculations reduces to some extent the accuracy. We have hence learned from Li_2O that we can apply very soft PPs of the Martins-Troullier type even for highly ionic materials for the ground state and the *GW* as well as the excitonic eigenvalues calculations, but for spectral calculations we might encounter worse results than when using hard PPs.

The nonlocal commutator $i\,[V_{\mathrm{nl}}, \mathbf{r}]$ (see Appendix A) in the optical matrix elements, Eq. (3.20), generally tends to decrease the absorption spectrum for about 15% in certain energy regions of the spectrum. Its contribution depends on the nonlocality of the pseudopotential of the studied system and thus we have included it in diamond (Fig. 5.8).

6.5 k Point Sampling for Spectral Calculations

The size of the excitonic Hamiltonian matrix, which is one of the bottlenecks in the calculations of the absorption spectrum, is given by the number of bands and $\mathbf{k}$ points included in the basis of the electron–hole pairs. Since the minimal number of bands is determined by the energy range we want to investigate, we can effectively only minimize the basis by choosing an optimal set of $\mathbf{k}$ points.

In this thesis we have used the special $\mathbf{k}$ point scheme of Chadi and Cohen [38] for the sampling and developed an efficient method, using symmetry properties of the Hamiltonian, to reduce one big Hamiltonian matrix into several small ones. At the examples of the diamond group materials silicon and diamond, we have obtained four small matrices. The excitonic eigenstates have been reconstructed by symmetry from the eigenvectors of the small matrices. This scheme has proven to deliver results in good agreement with experiment. However, the implementation is rather involved and depending on the symmetry properties of the material.

It would be better to have an approach which is easier to apply. Although not used here, Shirley *et al.* [68] and Rohlfing and Louie [105] have used a relatively small regular $\mathbf{k}$ point grid of around 500 points in the full Brillouin zone and shifted the grid slightly into an off-symmetry direction in order to better map the band dispersion of the bulk as a whole, and not just at high-symmetry points and in high-symmetry directions. The former have used a Brillouin zone with the shape of a rhombohedron, while the latter have employed the usual fcc Brillouin zone and used a shifted grid of Monkhorst and Pack points [39]. In both cases the agreement with experiment has been good.

However, the shifted grid destroys the symmetry of the crystal and thus does not allow the use of symmetry operations like we have done. It would be desirable to develop a $\mathbf{k}$ point sampling scheme, which gives a grid minimal in size and still possessing the full symmetry of the crystal. Indeed, this is not all obvious and it may be best to use an optimized one for the specific purpose, maybe along the lines of Ref. [90].

We note that the $\mathbf{k}$ point grid for the excitonic Hamiltonian matrix must be regular in order to be able to use only a small number of $\mathbf{q}$ vectors for the inverse dielectric matrix, Eq. (2.56), in the evaluation of the Coulomb terms Eqs. (3.96) and (3.103) of the excitonic Hamiltonian.

6.6 Summary

We have developed a method for including excitonic and exchange effects in *ab initio* calculations of optical properties. Starting from Hedin's coupled integral equations, we have rederived the fundamental equation for the absorption process, the Bethe-Salpeter equation, and we have discussed the ingredients of our approach. Our method allows us to study in detail the various two-particle channels contributing

to each excitonic eigenstate.

We have discussed the results of applications to silicon, diamond, lithium oxide and the sodium tetramer. Good agreement with experiment has been obtained for the absorption spectra of Si and diamond, the static dielectric constant of diamond and for the onset of optical absorption of Li_2O due to discrete bound excitons. We have discussed various approximations of our method and have shown the strong mixing of independent-particle transitions to a bound excitonic state in the Na_4 cluster.

Furthermore, the influence of ground state calculations on optical spectra has been investigated under particular consideration of the pseudopotential generation and we have discussed the use of different Brillouin zone point sampling schemes for spectral calculations.

Appendix A

Pseudopotentials

Since the early work of Fermi [106] there has been great interest in replacing the effect exerted by the chemically inert core states on the chemically active valence states by an effective pseudopotential. With the introduction of empirical pseudopotentials realistic band structure calculations of semiconductors and simple metals could be performed. Although experimental data had to be fitted in the construction, the results demonstrated the feasibility of the concept [107, 108]. The development of *ab initio*, *i.e.*, parameter free, and norm-conserving pseudopotentials marked an important step forward [80, 109].

In this appendix we give an overview over the construction and the use of pseudopotentials in the framework of density functional calculations. Some aspects of pseudopotentials will be discussed at the example of lithium oxide.

A.1 Concept of Pseudopotentials

We are considering one isolated atom and estimate that well within the core the potential felt by an electron is roughly given by

$$-\frac{Z_{\text{eff}}}{r}, \tag{A.1}$$

with Z_{eff} being some effective charge due to the positive nucleus screened by other electrons.

The high negative Coulomb energy $\bar{V}$ gives from the Virial theorem [110, Chap. 3.4],

$$\bar{T} = -\frac{1}{2}\bar{V}, \tag{A.2}$$

also a high positive kinetic energy $\bar{T}$ for the core electrons. Since the velocity operator, used in the evaluation of the kinetic energy, is basically measuring the spatial change of the wavefunctions, the core electron wavefunctions will have significant wiggles typical for low lying atomic states. Compared to the core electrons, the

valence electrons have in total a higher energy. Of course, in the core region the valence electrons are far less likely to be found. However, they experience the same Coulomb potential there and must therefore have an even higher kinetic energy. This results in very rapidly changing wavefunctions.

The expressed oscillations of the valence states can also be understood by noting the orthogonality relation between core and valence states, $\phi_{ck}(\mathbf{r})$ and $\phi_{vk}(\mathbf{r})$,

$$\int d\mathbf{r}\, \phi^*_{ck}(\mathbf{r})\phi_{vk}(\mathbf{r}) = 0. \tag{A.3}$$

Within the core region the oscillations have to interlace in order to give a vanishing integral.

Many computational codes use a plane wave basis for the wave functions. The rapid oscillations of the wave functions in real space would result in a very large number of expansion coefficients in reciprocal space, see Eq. (2.31), after Fourier transforming and therefore in extremely expensive calculations. The problem is even worse, as the energy range of the core states is of the order of keV, while the relevant bonding energies are in the range of some eV. So one needs a very high relative precision for total energies, despite the fact that normally only differences are of interest.

Since the core is virtually inert to changes in the chemical environment, we can construct a pseudopotential from atomic wavefunctions following the lines of Phillips and Kleinman [111]. The true wavefunction ϕ is expressed as the sum of a smooth wavefunction φ and a sum over occupied core states φ_c,

$$\phi = \varphi + \sum_c b_c \varphi_c. \tag{A.4}$$

The wavefunction is forced to be orthonormal to the core states $\langle \varphi_c | \phi \rangle = 0$. Solving for b_c yields

$$\phi = \varphi - \sum_c \langle \varphi_c | \varphi \rangle \varphi_c. \tag{A.5}$$

Operating on ϕ with the Hamiltonian $H = T + V_c$ (where V_c is the attractive core potential) gives the correct energy eigenvalue E. Substituting $H\phi = E\phi$ yields after some rearrangement

$$(H + V_R)\varphi = E\varphi, \tag{A.6}$$

where

$$V_R = \sum_c (E - E_c) |\varphi_c\rangle \langle \varphi_c|, \tag{A.7}$$

and

$$H\varphi_c = E_c \varphi_c. \tag{A.8}$$

The potential V_R acts like a short-ranged repulsive potential. The pseudopotential is in general spatially nonlocal and depends on $\mathbf{r}$ and $\mathbf{r}'$.

The Eq. (A.6) is the new wave equation for the pseudowavefunction φ. It is important to note that the eigenvalue E is the true energy corresponding to the true wavefunction ϕ of the system. The potential of Eq. (A.6), $V = V_c + V_R$, is the sum of an attractive long-ranged and a repulsive short-ranged part. Near the core both parts cancel fairly completely and the whole potential is well-behaved and weak over the whole range (*Phillips-Kleinman cancellation theorem*).

If the true wavefunction ϕ is normalized to unity, then φ is only approximately normalized in the Phillips-Kleinman scheme. The condition $\langle\phi\,|\phi\rangle = 1$ with Eq. (A.5) gives

$$1 = \langle\varphi\,|\varphi\rangle - 2\sum_c \langle\varphi_c\,|\varphi\rangle\,\langle\varphi\,|\varphi_c\rangle + \sum_c |\langle\varphi_c\,|\varphi\rangle|^2 . \tag{A.9}$$

Therefore,

$$\langle\varphi\,|\varphi\rangle = 1 + \sum_c |\langle\varphi_c\,|\varphi\rangle|^2 , \tag{A.10}$$

The incorrect distribution of the valence charge between the core and valence region is known as *orthogonality hole* and causes serious problems in self-consistent calculations.

Finally, we would like to mention that the Phillips-Kleinman pseudopotential is non-unique. This can be seen by noting that replacing $\varphi \to \varphi + \sum_c \delta\alpha_c\varphi_c$ still gives the same true wavefunction ϕ in Eq. (A.5).

A.2 Construction of Pseudopotentials

Modern *ab initio* pseudopotential avoid the problems arising from the above described orthogonality hole by choosing a different approach [80, 109, 112]. Here, we first describe a general prescription given by Martins and Troullier [81] to construct norm-conserving pseudopotentials and introduce then the Martins-Troullier method itself. Finally, we discuss the separable pseudopotentials as invented by Kleinman and Bylander [82].

A.2.1 General Description

Pseudopotentials are constructed from isolated atoms, whose electronic structure is obtained in an all-electron (AE) calculation within density-functional theory. In free space we can assume spherical symmetry and decompose each atomic wavefunction (with a principal quantum number n) in a radial and an angular part [81],

$$\phi^{\mathrm{AE}}(\mathbf{r}) = \sum_{l,m} R_l^{\mathrm{AE}}(r) Y_{lm}(\Omega). \tag{A.11}$$

The Y_{lm} are spherical harmonics.

The radial Kohn-Sham equations for spherical screening,

$$\left(-\frac{1}{2}\frac{d^2}{dr^2}+\frac{l(l+1)}{2r^2}+V^{\mathrm{AE}}(r)\right) rR_l^{\mathrm{AE}}(r)=\epsilon_l^{\mathrm{AE}} rR_l^{\mathrm{AE}}(r), \tag{A.12}$$

are solved self-consistently, where $V^{\mathrm{AE}}(r)$ is the self-consistent one-electron potential in the local-density approximation,

$$V^{\mathrm{AE}}(r)=-\frac{Z}{r}+V_{\mathrm{H}}(r)+V_{\mathrm{xc}}^{\mathrm{LDA}}\left(n(r)\right). \tag{A.13}$$

Z is the atomic number and V_{H} the Hartree potential from Eq. (2.83). The solutions of Eq. (A.12) give the all-electron radial wavefunctions $R_l^{\mathrm{AE}}(r)$ and eigenvalues ϵ_l^{AE}.

The radial pseudopotential (PP) wavefunction $R_l^{\mathrm{PP}}(r)$ is then constructed according to the following quite general conditions:

1. The pseudowavefunction $R_l^{\mathrm{PP}}(r)$ originating from the solution of the pseudoatom Kohn-Sham equation,

$$\left(-\frac{1}{2}\frac{d^2}{dr^2}+\frac{l(l+1)}{2r^2}+V_l^{\mathrm{PP}}(r)\right) rR_l^{\mathrm{PP}}(r)=\epsilon_l^{\mathrm{PP}} rR_l^{\mathrm{PP}}(r), \tag{A.14}$$

   must be nodeless in order to be smooth and free of unwanted wiggles. Note the l-dependence of the pseudopotential.

2. Beyond a certain cutoff radius $r_c(l)$ characterizing the core radius, the wavefunction $R_l^{\mathrm{PP}}(r)$ must be equal to $R_l^{\mathrm{AE}}(r)$:

$$R_l^{\mathrm{PP}}(r)=R_l^{\mathrm{AE}}(r) \quad \text{for } r>r_c(l), \tag{A.15}$$

   or at least converge rapidly to that value.

3. The charge enclosed within $r_c(l)$ for the two wavefunctions must be equal,

$$\int_0^{r_c(l)} dr\left|R_l^{\mathrm{PP}}(r)\right|^2 r^2=\int_0^{r_c(l)} dr\left|R_l^{\mathrm{AE}}(r)\right|^2 r^2. \tag{A.16}$$

4. Clearly, the valence all-electron and pseudopotential eigenvalues must be equal

$$\epsilon_l^{\mathrm{PP}}=\epsilon_l^{\mathrm{AE}}. \tag{A.17}$$

Pseudopotentials meeting the above conditions are referred to as *norm-conserving* and are relatively well transferable. The latter means that the PP is adequate to describe well the electronic structure also in a changed chemical environment, like in a solid or an excited atomic state, and not only in the atomic state for which

it is has been constructed. The freedom given within the general rules allows the construction of soft pseudopotentials, *i.e.*, having a low cutoff energy Eq. (2.32).

By analytically inverting the effective Schrödinger equation (A.14) we obtain the screened (scr) pseudopotential

$$V_{\mathrm{scr},l}^{\mathrm{PP}}(r) = \epsilon_l - \frac{l(l+1)}{2r^2} + \frac{1}{2rR_l^{\mathrm{PP}}(r)} \frac{d^2}{dr^2} \left[rR_l^{\mathrm{PP}}(r) \right] . \tag{A.18}$$

For a nodeless pseudowavefunction the pseudopotential Eq. (A.18) does not have any singularities, except possibly at the origin. They can be avoided by choosing a pseudowavefunction behaving like r^l near the origin. For a continuous pseudopotential the pseudowavefunction must have continuous derivatives up to and including the second derivative.

The screening from the valence electrons, which is contained in $V_{\mathrm{H}}(r)$ and in $V_{\mathrm{xc}}^{\mathrm{LDA}}(n(r))$, is then removed *(unscreening)*. For that purpose, we split the density in the core density n_c and the part coming from the pseudowavefunctions n_v , $n = n_c + n_v$.

The ionic pseudopotential is now constructed by subtracting the Hartree $V_{\mathrm{Hart}}^{\mathrm{PP}}(r)$ and the exchange-correlation $V_{\mathrm{xc}}^{\mathrm{PP}}(r)$ potentials, calculated from the *valence* pseudowavefunctions, from the screened potential,

$$V_{\mathrm{ion},l}^{\mathrm{PP}}(r) = V_{\mathrm{scr},l}^{\mathrm{PP}}(r) - V_{\mathrm{Hart}}^{\mathrm{PP}}(r) - V_{\mathrm{xc}}^{\mathrm{PP}}(r). \tag{A.19}$$

Here we are assuming the core to be completely inert to changes in the electronic environment *(frozen core approximation)*. The valence electrons, on the other hand, are, of course, strongly influenced by the chemical bonding. In subsequent solid state calculations the removed valence part of the pseudopotential is then again added, now self-consistently in order to describe the new screening. In the following we omit the index ion for convenience.

Due to the construction of the pseudopotential we have an explicit dependence on the components of the angular momentum l, which can be expressed via the projection operators

$$V^{\mathrm{PP}} = \sum_{l,m} |lm\rangle \, V_l^{\mathrm{PP}} \, \langle lm| \equiv \sum_l \hat{P}_l V_l^{\mathrm{PP}}. \tag{A.20}$$

Writing the above in real space representation,

$$V^{\mathrm{PP}}(r, \Omega, \Omega') = \sum_{l,m} |Y_{lm}(\Omega)\rangle \, V_l^{\mathrm{PP}}(r) \, \langle Y_{lm}(\Omega')| \, , \tag{A.21}$$

shows the explicit nonlocality in the angular coordinates. Since the radial component is local, we have a semilocal pseudopotential. It can be further split into a local and

a semilocal part by using a local reference potential $V_{\text{ref}}^{\text{PP}}$. We get

$$\begin{aligned} V^{\text{PP}} &= V_{\text{loc}}^{\text{PP}} + V_{\text{nl}}^{\text{PP}} = \sum_{l=0}^{\infty} |l\rangle \langle l| \left(V_{\text{ref}}^{\text{PP}} + \Delta V_l^{\text{PP}}\right) \\ &= V_{\text{ref}}^{\text{PP}} + \sum_{l=0}^{\infty} |l\rangle \Delta V_l^{\text{PP}} \langle l| . \end{aligned} \tag{A.22}$$

One often chooses $V_{\text{ref}}^{\text{PP}}$ to be equal to the one of the nonlocal components, $V_{\text{ref}}^{\text{PP}} \equiv V_{l_{\text{ref}}}^{\text{PP}}$. Then,

$$\Delta V_l^{\text{PP}} = V_l^{\text{PP}} - V_{l_{\text{ref}}}^{\text{PP}} \tag{A.23}$$

vanishes for that l. The truncation of the sum over l, $l \leq \bar{l}$ is unavoidable and takes high angular momentum contributions, which are unimportant for the chemical bond, to be equal to $V_{l_{\text{ref}}}^{\text{PP}}$.

The maximal l orbital up to which the sum in Eq. (A.22) is developed depends not only on the atom, but also on the electronic environment in which it is placed. In the solid the atomic electronic charge is redistributed and may effectively be in an excited state compared with the atomic ground state. For the construction of excited angular momentum pseudopotential components one rearranges thus the valence charge into excited states to get bound states of the atom, maybe even in an ionic configuration (*e.g.* $l = 2$ for Si: Si [Ne] $3s^1 3p^{0.75} 3d^{0.25}$ [80]).

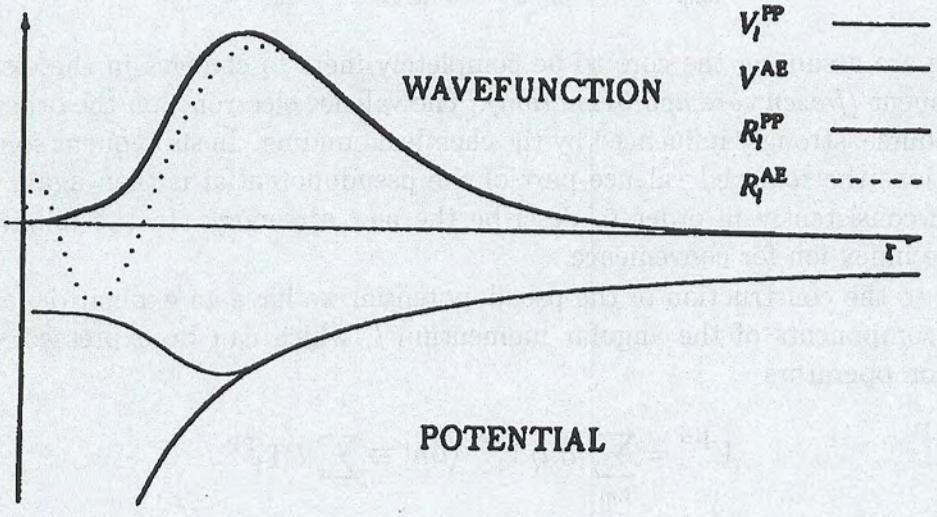


Figure A.1: Schematic representation of the pseudopotential method. The all-electron (AE) potential V^{AE} and the AE wavefunction are modified inside the core radius r_c to the pseudopotential (PP) V_l^{PP} and the PP wavefunction R_l^{PP}.

Fig. A.1 shows schematically the PP- and the AE-radial wavefunctions as well as the pseudopotential and the all-electron potentials for some angular momentum

l. The PP wavefunction has no nodes and is much smoother near the core than the AE one. The PP does not have the Coulomb singularity for $r \to 0$, but rather tends to be flat near $r = 0$. This is especially characteristic for the Martins-Troullier scheme, which is discussed in the next section.

The *transferability* of a pseudopotential is largely influenced by the cutoff radius $r_c(l)$. A high radius reduces the influence of the oscillating core states and permits a smoother pseudopotential. However, the transferability is lowered due to a loss of physically relevant information. In any case one has to assure that $r_c(l)$ is greater than the radius of the outermost node of all AE wavefunctions.

A good first criterion for transferability is given by the logarithmic derivative of the wavefunction $R(r)$. Since the Schrödinger equation is a second order differential equation, the value at a certain radius in combination with its first derivative and the proper normalization condition (see condition 3 above) uniquely defines $R(r)$. From condition 2 we have for any $r \geq r_c$:

$$\frac{1}{R_l^{\mathrm{PP}}(r,\epsilon)} \frac{dR_l^{\mathrm{PP}}(r,\epsilon)}{dr} = \frac{1}{R_l^{\mathrm{AE}}(r,\epsilon)} \frac{dR_l^{\mathrm{AE}}(r,\epsilon)}{dr}. \tag{A.24}$$

For a perfect pseudopotential Eq. (A.24) holds for every value ϵ, not just for ϵ_l (condition 4). However, from a version of Friedel's sum rule [109, 113, 114] for $R(r)$ it follows together with the norm-conserving condition that also in a region around ϵ_l the Eq. (A.24) is closely satisfied.

An additional quality requirement is the check that the AE excited state energies should be equal to the ones obtained from the ionic pseudopotential. A final answer on the quality of a pseudopotential can, of course, only be given by the results of realistic calculations of molecules or condensed matter phases. For each atom the appropriate cutoff radius has to be chosen as a compromise between fast convergence and transferability. In addition, one should avoid overlaps of neighboring core regions.

A.2.2 Martins-Troullier Pseudopotentials

The four conditions for a norm-conserving pseudopotential given in the last section are by far not sufficient to uniquely define a pseudopotential. The remaining freedom can be used to construct smooth pseudopotentials which have a relatively high cutoff radius $r_c(l)$, while still guaranteeing satisfying logarithmic derivatives in Eq. (A.24). Especially for atoms with $1s$, $2p$, $3d$, $4f$ or $5g$ valence electrons highly efficient pseudopotentials are crucial. In the expansion in angular momentum components in Eq. (A.20) the repulsive short-ranged potential V_R of Eq. (A.6) is missing, which stemmed form the condition of the orthogonality of the true wavefunction ϕ with respect to the core states. Therefore, calculations involving atoms of the above type are computationally quite demanding. In this work we are faced with this kind of problem by the diamond and the oxygen atom, which have $2p$ valence electrons.

Martins and Troullier have introduced a procedure based on the method of Kerker [112] to construct fairly soft pseudopotentials [81]:

The radial PP wavefunction is modeled inside the core by

$$R_l^{\mathrm{PP}}(r) = r^l e^{p_l(r^2)}, \tag{A.25}$$

where p_l is a polynomial of order six in r^2 fulfilling the conditions:

1. Norm-conservation of charge within the core radius $r_c(l)$ according to Eq. (A.16).

2. The continuity of the PP wavefunction and its first four derivatives at $r_c(l)$, which in effect imposes the continuity of $V_{\mathrm{scr},l}^{\mathrm{PP}}(r)$, Eq. (A.18), and its first two derivatives at $r_c(l)$.

3. The zero curvature of the screened pseudopotential at the origin. This contributes significantly to the smoothness of the pseudopotential.

Although there are still no absolute criteria for efficient pseudopotentials, these characteristics have been proven to deliver excellent results concerning high transferability in combination with a fast convergence with respect to the basis set size.

A.2.3 Separable Pseudopotentials

We use a basis set of plane waves for the wavefunctions, Eq. (2.31), in the computer codes and thus perform the total energy calculations, at least partly, in momentum space. The semilocal part of the pseudopotential Eq. (A.22) is then essentially described – after Fourier transforming – by

$$\sum_{\mathbf{G},\mathbf{G}'} |\mathbf{k}+\mathbf{G}\rangle\, V_{\mathrm{nl}}^{\mathrm{PP}}(\mathbf{k}+\mathbf{G},\mathbf{k}+\mathbf{G}')\, \langle \mathbf{k}+\mathbf{G}'|\,, \tag{A.26}$$

which scales like N_{PW}^2, where N_{PW} is the number of plane waves Eq. (2.33) and $V_{\mathrm{nl}}^{\mathrm{PP}}$ denotes the nonlocal part of the pseudopotential. This is prohibitively expensive in time and memory use for many large systems like defects or surfaces.

The coupling of the two sums over the reciprocal lattice vectors can be avoided by writing the nonlocal pseudopotential after *Kleinman and Bylander* (KB) [82] in a separable form

$$\Delta V_{\mathrm{nl}}^{\mathrm{KB}} = \sum_{l,m} \frac{\left|\Delta V_l^{\mathrm{PP}} \varphi_{l,m}^{\mathrm{PP}}\right\rangle \left\langle \Delta V_l^{\mathrm{PP}} \varphi_{l,m}^{\mathrm{PP}}\right|}{\left\langle \varphi_{l,m}^{\mathrm{PP}}\right| \Delta V_l^{\mathrm{PP}} \left|\varphi_{l,m}^{\mathrm{PP}}\right\rangle}. \tag{A.27}$$

The eigenfunctions of the atomic Hamiltonian, given in radial form in Eq. (A.14), $\varphi_{l,m}^{\mathrm{PP}}(r) = R_l^{\mathrm{PP}}(r) Y_{lm}(\Omega)$, are obtained with the atomic pseudopotential. Evidently,

the Fourier transform of Eq. (A.27),

$$\sum_{l,m} \frac{\left[\sum_{\mathbf{G}} |\mathbf{k}+\mathbf{G}\rangle \langle \mathbf{k}+\mathbf{G}| \Delta V_l^{\mathrm{PP}} \varphi_{l,m}^{\mathrm{PP}} \rangle\right] \left[\sum_{\mathbf{G}'} \langle \Delta V_l^{\mathrm{PP}} \varphi_{l,m}^{\mathrm{PP}} | \mathbf{k}+\mathbf{G}' \rangle \langle \mathbf{k}+\mathbf{G}'|\right]}{\langle \varphi_{l,m}^{\mathrm{PP}} | \Delta V_l^{\mathrm{PP}} | \varphi_{l,m}^{\mathrm{PP}} \rangle}, \quad \text{(A.28)}$$

scales like N_{PW} due to the separate dependence on $\mathbf{G}$ and $\mathbf{G}'$.

Both nonlocal pseudopotentials deliver the same pseudowavefunctions in Eq. (A.14), but not necessarily the same physical results in a solid state calculation. It is possible that ghost states appear, *e.g.* in GaAs: these are unphysical eigenstates of the KB Hamiltonian which are possible, since it does not obey the Wronskian theorem [115]. The latter implies that atomic eigenfunctions are energetically ordered such that the energies increase with the number of nodes (for a given quantum number l). The unphysically bound ghost states can be avoided using certain prescriptions and care in the pseudopotential construction [116].

Nonlocal Commutator

In Section 3.1.1 we discussed the contribution of the nonlocal pseudopotential to the electric dipole matrix element Eq. (3.12). For the *GW* and excitonic calculations this contribution is naturally very small, since only one element of the – in principle – infinitely large dielectric matrix is affected. The negligible influence on the results in easily verified.

However, the transition probability in absorption spectra is essentially given by the absolute square of just this electric dipole, as seen in Eqs. (3.21) or (3.88). It can thus be important to include these contributions into the code for the absorption spectra in order to obtain quantitatively correct spectra.

We will follow the lines of Refs. [42], [61] and [117]. The code has developed by Valerio Olevano [117] and has then been integrated in our programs for spectral calculations.

The required matrix element is, from Eqs. (2.31) and (3.20),

$$\begin{aligned} & \hat{\mathbf{e}} \cdot \langle v\mathbf{k} | i\,[V_{\mathrm{nl}}, \mathbf{r}] | c\mathbf{k} \rangle = \\ = \; & \hat{\mathbf{e}} \cdot \int d\mathbf{r} \int d\mathbf{r}'\; \varphi_{v\mathbf{k}}^*(\mathbf{r})\; V_{\mathrm{nl}}(\mathbf{r},\mathbf{r}')\; i(\mathbf{r}'-\mathbf{r})\; \varphi_{c\mathbf{k}}(\mathbf{r}') \\ = \; & \hat{\mathbf{e}} \cdot \frac{1}{\Omega_0} \sum_{\mathbf{G},\mathbf{G}'} c_{v\mathbf{k}}^*(\mathbf{G}) c_{c\mathbf{k}}(\mathbf{G}') \int d\mathbf{r} \int d\mathbf{r}'\; e^{-i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}\; V_{\mathrm{nl}}(\mathbf{r},\mathbf{r}')\; i(\mathbf{r}'-\mathbf{r})\; e^{i(\mathbf{k}+\mathbf{G}')\cdot\mathbf{r}'}. \end{aligned} \quad \text{(A.29)}$$

The hat indicates a normalized vector: $\hat{\mathbf{v}} \equiv \mathbf{v}/v$, with $v = |\mathbf{v}|$. For ease of notation we set $\mathbf{K} \equiv \mathbf{k}+\mathbf{G}$, $\mathbf{K}' \equiv \mathbf{k}+\mathbf{G}'$ and define

$$V_{\mathrm{nl}}(\mathbf{K},\mathbf{K}') \equiv \frac{1}{\Omega_0} \int d\mathbf{r} \int d\mathbf{r}'\; e^{-i\mathbf{K}\cdot\mathbf{r}}\; V_{\mathrm{nl}}(\mathbf{r},\mathbf{r}')\; e^{i\mathbf{K}'\cdot\mathbf{r}'}. \quad \text{(A.30)}$$

Then we obtain

$$\begin{aligned} &\hat{\mathbf{e}}\cdot\langle v\mathbf{k}\,|i\,[V_{\rm nl},\mathbf{r}]|\,c\mathbf{k}\rangle = \\ &= \hat{\mathbf{e}}\cdot\sum_{\mathbf{G},\mathbf{G}'} c^*_{v\mathbf{k}}(\mathbf{G})c_{c\mathbf{k}}(\mathbf{G}')\ (\nabla_{\mathbf{K}}+\nabla_{\mathbf{K}'})\ V_{\rm nl}(\mathbf{K},\mathbf{K}'). \end{aligned} \tag{A.31}$$

In the KB formulation the matrix elements $V_{\rm nl}$ become

$$V^{\rm KB}_{\rm nl}(\mathbf{K},\mathbf{K}') = \sum_s e^{-i(\mathbf{K}-\mathbf{K}')\cdot\boldsymbol{\tau}_s}\sum_l P_l(\hat{\mathbf{K}}\cdot\hat{\mathbf{K}}')F_{sl}(K)F_{sl}(K'), \tag{A.32}$$

with, using Eq. (A.23) for $\Delta V_l^{\rm PP}$,

$$F_{sl}(K) = \sqrt{\frac{4\pi}{\Omega_0}(2l+1)}f_{sl}(K), \tag{A.33}$$

$$f_{sl}(K) = \int dr\ r^2\Delta V_l^{\rm PP}(r)j_l(Kr)R_l^{\rm PP}(r). \tag{A.34}$$

For later use, we also give the derivative

$$\frac{\partial f_{sl}(K)}{\partial K} = \int dr\ r^3\Delta V_l^{\rm PP}(r)j_l'(Kr)R_l^{\rm PP}(r). \tag{A.35}$$

The atomic structure factor $e^{-i(\mathbf{K}-\mathbf{K}')\cdot\boldsymbol{\tau}_s}$ reflects the basis of the supercell. The $P_l(x)$ are Legendre polynomials and the $j_l(x)$ are Bessel functions, with the $j_l'(x)$ being the first derivatives of them with respect to their argument. The $f_{sl}(K)$ are called KB form factors.

Finally, the matrix element is

$$\begin{aligned} &\hat{\mathbf{e}}\cdot\langle v\mathbf{k}\,|i\,[V_{\rm nl},\mathbf{r}]|\,c\mathbf{k}\rangle = \\ &= \hat{\mathbf{e}}\cdot\sum_{\mathbf{G},\mathbf{G}'} c^*_{v\mathbf{k}}(\mathbf{G})c_{c\mathbf{k}}(\mathbf{G}')\sum_s e^{-i(\mathbf{K}-\mathbf{K}')\cdot\boldsymbol{\tau}_s}\sum_l\Big\{P_l'(\hat{\mathbf{K}}\cdot\hat{\mathbf{K}}')\mathbf{A}(K,K')F_{sl}(K)F_{sl}(K') \\ &+ P_l(\hat{\mathbf{K}}\cdot\hat{\mathbf{K}}')\left[\mathbf{B}(K,K')+\mathbf{B}(K',K)\right]\Big\}, \end{aligned} \tag{A.36}$$

with

$$\begin{aligned} \mathbf{A}(K,K') &\equiv (\nabla_{\mathbf{K}}+\nabla_{\mathbf{K}'})\frac{\mathbf{K}\cdot\mathbf{K}'}{KK'} \\ &= \frac{1}{KK'}\left[\mathbf{K}\left(1-\frac{\mathbf{K}\cdot\mathbf{K}'}{K^2}\right)+\mathbf{K}'\left(1-\frac{\mathbf{K}\cdot\mathbf{K}'}{K'^2}\right)\right], \end{aligned} \tag{A.37}$$

being the derivative of the argument in the Legendre polynomials, and

$$\begin{aligned} \mathbf{B}(K,K') &\equiv \nabla_{\mathbf{K}}F_{sl}(K)F_{sl}(K') \\ &= \left.\frac{\partial F_{sl}(K)}{\partial K}\right|_K\hat{\mathbf{K}}\ F_{sl}(K'). \end{aligned} \tag{A.38}$$

A.3 Soft Pseudopotentials: Lithium Oxide

In this section we will examine the use of soft pseudopotentials (PP) of the Martins-Troullier (MT) type [81] at the example of an highly ionic crystal, lithium oxide, Li_2O. In an earlier work we have already investigated the properties Li_2O [18].

In the course of this thesis we have revisited lithium oxide under special consideration of the PP configuration, in particular the use of the Kleinman-Bylander separation, on the ground state and also on the optical properties. We will discuss the effects on the optical matrix elements Eq. (3.12), or explicitly Eq. (3.20), using different pseudopotential configurations. To this end, we will compare a very soft MT PP with a harder MT one and with a standard Bachelet-Hamann-Schlüter (BHS) PP [80].

Lithium oxide is a material of technological interest, with possible applications in deuterium-tritium fusion reactors as blanket breeding material [118], and in solid-state batteries [119]. Therefore, many experimental [120] and theoretical [121] studies are devoted to its defects, from the properties of simple point defects [118, 119] to the formation of Li colloids under irradiation [122]. The electronic structure of the material plays an important role in most of the studied processes, and is needed for the interpretation of the results of many of the applied experimental techniques, like optical absorption [123]. On the other hand, only few investigations of the electronic structure of bulk Li_2O itself have been published: like the measurement of the absorption onset [124], or the Hartree-Fock calculations of the occupied valence states [125]. Further theoretical investigations of the electronic properties of lithium oxide have also been suggested by a recent photoemission and electron-loss study [126].

A.3.1 Electronic Structure Calculations

Our procedure starts with a DFT-LDA calculation of the ground state properties and the KS electronic structure. We have carefully tested the very soft, norm-conserving pseudopotentials used in the calculations, and we discuss in particular their influence on the electronic structure. We then determine the QP energies by evaluating self-energy corrections in first order perturbation theory, using Hedin's *GW* approximation. This yields a precise picture of the occupied and the empty bands of Li_2O. Finally, we determine the optical transition energies – including excitonic effects – from a two-particle effective equation derived from Green's function theory Eq. (3.78), which includes the screened electron–hole interaction as well as an unscreened electron–hole exchange term. This calculation allows to analyze in detail the nature of the electronic excitations which contribute to the results [23].

DFT-LDA Results

Pseudopotentials and Ground State Lithium oxide has anti-fluorite structure and can be described by an fcc cell with one oxygen and two lithium atoms. We determine the ground state properties within the DFT-LDA pseudopotential-based Car-Parrinello method [41]. Since plane wave calculations with the traditional Bachelet-Hamann-Schlüter pseudopotentials for the oxygen atom demand for extremely large computational efforts, we use the Martins-Troullier pseudopotential generation scheme, in order to obtain a considerably softer norm-conserving pseudopotential. We have first used, as reference configuration, a relatively smooth MT PP for oxygen which has already been applied and tested in the past [127]. The cutoff radii are 1.45 a.u. for both the s and the p component, and the latter is taken as local reference (sp). This potential (denoted hMTO in the following) has yielded perfect convergence with a plane wave cutoff of 80 Ry, which means about 2100 plane waves for the Li_2O primitive unit cell.

In order to reduce the computational effort we have investigated the properties of a smoothed version (denoted sMTO), which has also already been used in the past [128]. The cutoff radii are 2.30 a.u. for the s component and 1.95 a.u. for the p component. In this case, the required cutoff energy for full convergence is 50 Ry (1000 plane waves). For the lithium atom we choose a pseudopotential with cutoff radii of 1.75 a.u. for both the local s as well as the p component (sp), the latter one created from a slightly excited atomic configuration. Using two special $\mathbf{k}$ points [39], we obtained very good agreement with experiment [129] for the lattice constant and the bulk modulus with both pseudopotential configurations (see Tab. A.1). The influence of the Kleinman-Bylander separation on the lattice constant and the bulk modulus is negligible. Increasing the number of special $\mathbf{k}$ points to ten changes the total energy per atom by less than 1 meV.

	with KB	non KB	Exp.
a_0	4.534 Å	4.525 Å	4.573 Å
B_0	0.90 Mbar	0.91 Mbar	0.90 Mbar

Table A.1: Lattice constant a_0 and bulk modulus B_0 of Li_2O calculated with and without the Kleinman-Bylander separation. Experimental data from Ref. [129]. Both pseudopotential configurations yield the same results with KB separation. Only for the softer oxygen pseudopotential (sMTO) the calculation without the KB separation has been done.

The good agreement of the two pseudopotential configurations for the ground state calculations is particularly remarkable, if we compare the cutoff radii with the nearest atomic distance in the crystal. It is between the Li and the O atom and – for $a_0 = 4.573$ Å (exp. value) – we have $\overline{\mathrm{LiO}} = \frac{\sqrt{3}}{4} a_0 = 1.98$ Å. The other distances are

$\overline{\mathrm{LiLi}} = \frac{1}{2}a_0 = 2.29$ Å and $\overline{\mathrm{OO}} = \frac{1}{\sqrt{2}}a_0 = 3.23$ Å. In Tab. A.2 we compare the Li-O distance with the sums of the various cutoff radii for the hMTO and the sMTO with the lithium PP cutoff radius. We find that there is an overlap of the oxygen sMTO s orbital with the lithium s and p orbitals: the ratio of $[r_c(\mathrm{Li}, s, p) + r_c(\mathrm{O}, s)]/\overline{\mathrm{LiO}}$ for the sMTO is 1.08.

	hMTO	sMTO
$r_c(\mathrm{Li}, s, p) + r_c(\mathrm{O}, s)$	1.69 Å	2.14 Å
$r_c(\mathrm{Li}, s, p) + r_c(\mathrm{O}, p)$	1.69 Å	1.96 Å
Ratio w.r.t. $\overline{\mathrm{LiO}}$	0.85	**1.08**

Table A.2: Comparison of the sums of the cutoff radii of the hMTO and the sMTO with the lithium PP cutoff radius. The ratio of the larger sum for each MTO is given with respect to the nearest atomic distance $\overline{\mathrm{LiO}} = 1.98$ Å for $a_0 = 4.573$ Å (exp. value).

Clearly, this is not supposed to happen. Therefore, one is surprised by the excellent agreement of the two PP configurations with experiment for the ground state calculations. Indeed, as we shall see in the following sections, also the *GW* corrections and even the optical matrix elements are close to each other for the two configurations.

Charge Distribution Due to the norm-conserving pseudopotentials the total charge is correct in each case and the details of the wavefunctions are obviously less important. In Fig. A.2 the total charge is depicted in the (110) plane of the conventional unit cell. Characteristically for a strongly ionic bond, the valence charge is almost entirely concentrated around the oxygen atom.

In addition, the use of large cutoff radii is not unusual in the literature. Hybertsen and Louie [42] applied soft pseudopotentials of the Kerker type [112] for the highly ionic material LiCl. This pseudopotential generation scheme is similar to the one of the Martins-Troullier used by us. In their work, the r_c for Li and Cl add up to 2.24 Å, while the nearest Li-Cl distance is 2.57 Å. The ratio of the sum of r_c to the nearest distance (0.87) is thus comparable in our case, if one takes the harder of the two oxygen pseudopotentials (0.85 in our work). In a recent publication [130] Surh, Chacham and Louie used Martins-Troullier type soft pseudopotentials for electronic structure calculations of an F center in LiCl. In that case the defect states are, of course, strongly localized around the halogen vacancy and it is not difficult to see that the reasoning concerning norm-conservation, charge distribution and detailed knowledge of the wavefunctions is similar to our case.

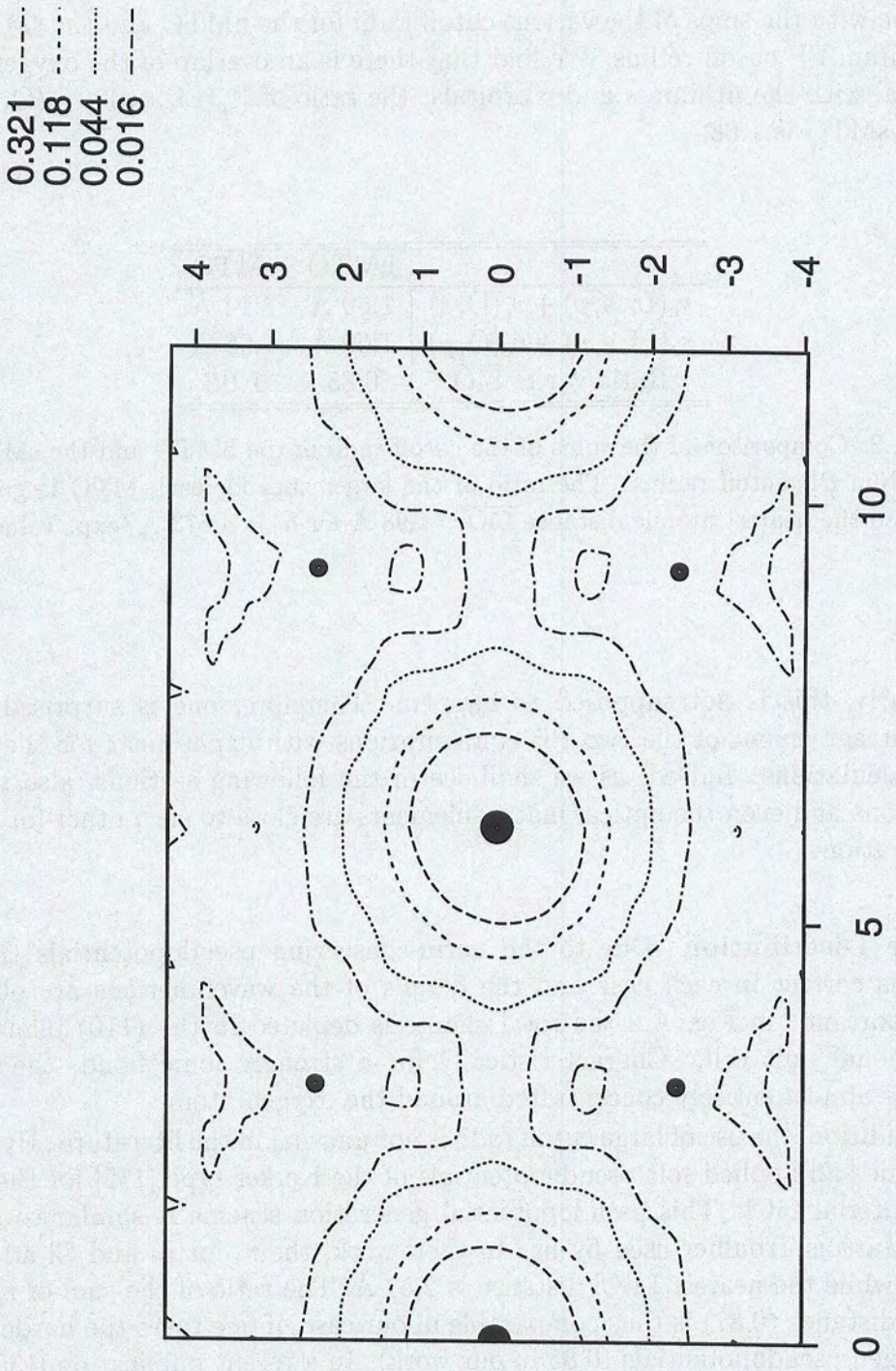


Figure A.2: Total charge density of Li_2O in the (110) plane of the conventional unit cell. Large and small dots indicate the positions of the O and Li atoms, respectively. Distances in a.u. Density in electrons per unit cell.

Kohn-Sham LDA Eigenvalues Although the Kohn-Sham LDA eigenvalues are converged to better than 10 meV for the two pseudopotential configurations, we find that with the low-cutoff one the gap is slightly overestimated (for about 0.15 eV). This error can be tolerated: the minimum LDA gap is direct, at the Γ point, and as large as 5.3 eV without the KB separation and 5.5 eV using it – calculated at the respective theoretical equilibrium lattice constant a_0, see Tab. A.3 –, so that the relative error is small. Nevertheless, it should be noted that a shift of 0.3 eV, like for the Γ point in Tab. A.3, is visible in an absorption spectrum. Also the possible choice of the experimental lattice constant instead of the theoretical one influences the calculated minimal gap.

Gaps	KB, $a_0^{\text{theor. KB}}$	KB, $a_0^{\text{exp.}}$	non-KB, $a_0^{\text{theor. non-KB}}$
Γ (min.)	5.52	5.33	5.24
X	6.62	6.50	6.54
L	10.10	9.89	10.17

Table A.3: Kohn-Sham LDA gaps of Li_2O at some high-symmetry points in the BZ using the sMTO with and without the KB separation at the indicated lattice constant. Values in eV.

The dashed lines in Fig. A.3 show our results for the LDA band structure. The occupied bands are mostly due to the oxygen ($2s$, $2p$) and are very flat. The lowest unoccupied band is also mostly localized on the oxygen atoms and is very flat along the $\overline{\Gamma X}$ direction, consistent with the geometry of the underlying oxygen sublattice.

As expected, the LDA gap underestimates the value of the onset of optical absorption, found experimentally at about 6.6 eV [124].

GW and Excitonic Calculations

Using 10 special $\mathbf{k}$ points the *GW* corrections open the gap at the Γ point by 2.1 eV, yielding a minimum direct gap of 7.4 eV, using the LDA gap of the KB calculation at the experimental lattice constant. Convergence of about 30 meV for the eigenvalues around the gap and for the gap itself is obtained with 941 wavefunctions for the hMTO, and with 531 plane waves for the sMTO. The convergence in the number of $\mathbf{k}$ points is about 0.15 eV. Comparison of the corrections shows that shifts are identical to within 30 meV for both pseudopotentials. This can be expected from a first-order perturbation correction, since the energy correction is an expectation value using very similar wavefunctions. Roughly, the unoccupied bands are shifted down and the occupied ones up by about 1 eV in either case (Fig. A.3). The band dispersion is only slightly modified, up to 0.40 eV.

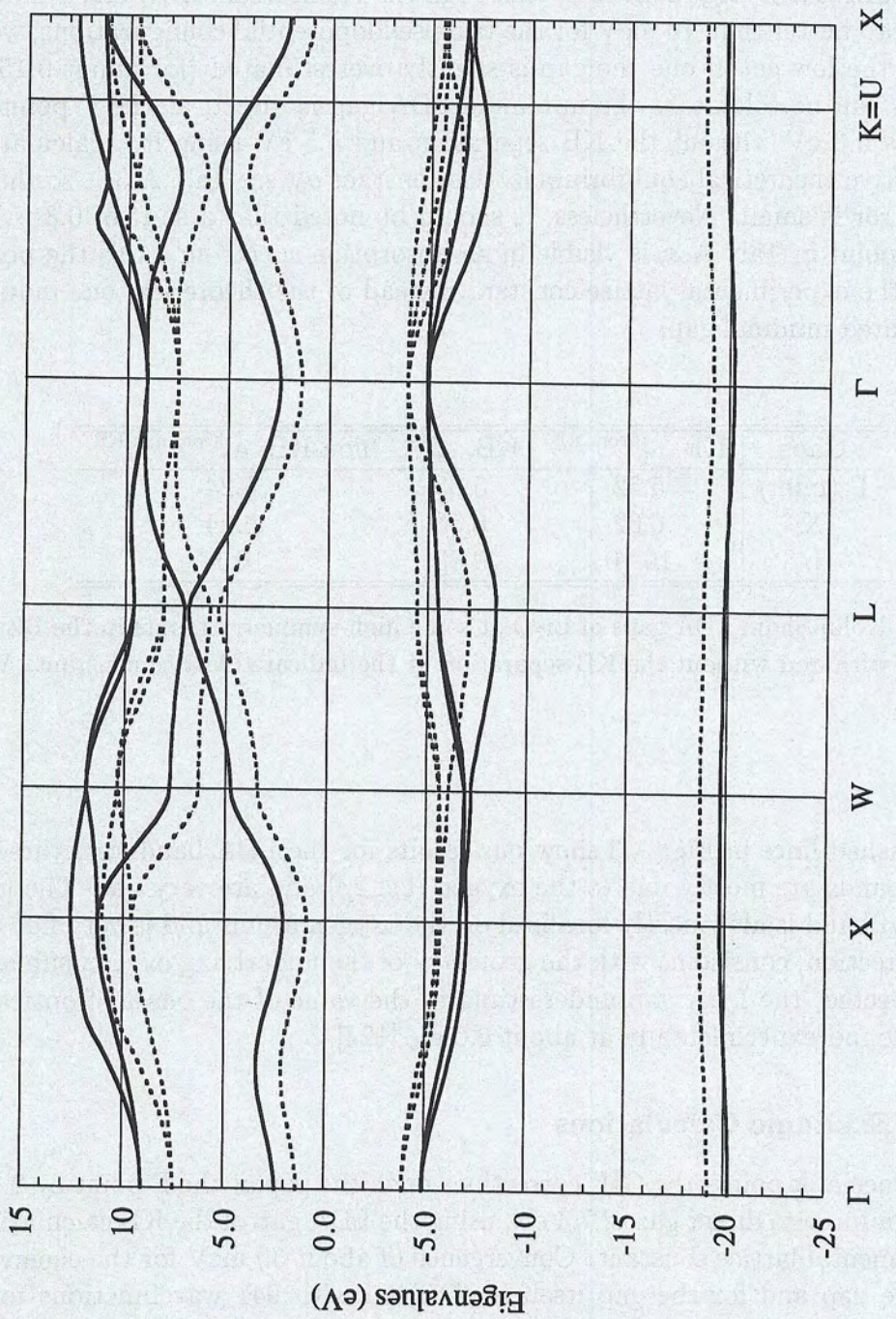


Figure A.3: QP (full lines) and DFT-LDA (dashed lines) band structure of Li_2O along high-symmetry directions for the valence and the first four unoccupied bands. The top of the LDA valence bands at Γ is set at -3.9 eV.

In spite of the well established reliability of *ab initio GW* calculations, we have hence found a QP gap which differs by more than half an eV from the measured "optical gap". The situation would not improve if the calculations were done at the theoretical instead of the experimental lattice constant: the *GW* gap would be even slightly bigger, 7.6 eV.

However, screening in Li_2O is weak: as a by-product of the *GW* calculation, we obtain the macroscopic dielectric constant ϵ_0, which is 3.46 for both pseudopotential configurations. This value overestimates the experimental value of 2.68 [131]. This is partly due to the well-known general overestimation of the LDA dielectric constant [132], and also to the neglect of the nonlocality of the pseudopotential in the evaluation of the $\mathbf{q} \to 0$ limit of the polarization matrix elements [42]. (It must be emphasized that the results of our calculations are not sensitive to the exact value of ϵ_0, see also Section 4.1.2.) This fact, together with the small dispersion of the top valence and bottom conduction bands, suggests that the electron-hole interaction should be large, which would lead to optical transitions at energies well below the QP gap.

We compute the binding energies of the excitons from the two-particle effective equation (3.79), using the sMTO. In order to obtain convergence (within 30 meV), we use 19 high symmetry points in the IBZ, including Γ, which give 256 $\mathbf{k}$ points in the full BZ, 8 bands and about 300 plane waves. The excitonic Hamiltonian has not been reduced by symmetry as outlined in Section 3.2.4.

We find the lowest exciton eigenvalue at 6.6 eV for the two transverse transitions (see Eq. (3.94)) with a small split of about 50 meV to the longitudinal component; this value is of the order of the global precision of our calculations. The dominant contribution comes from the transitions from the threefold degenerate $2p$ bands to the $3s$ band mostly due to the oxygen. The corresponding electric dipole transition Eq. (3.12) is allowed. Thus, this lowest eigenvalue of 6.6 eV corresponds to the onset of the optical absorption. We have hence found agreement with experiment [124].

Concerning the choice of the pseudopotential in the case of the exciton calculation, we have to consider two contributions: the repulsive exchange and the attractive Coulomb part. Since the pseudopotential is norm-conserving, total charges are correct and the long-range Coulomb interaction is adequately described. Only for very small distances we introduce an error due to the finite size of the core region given by the cutoff radius, inside of which the pseudowavefunctions are not correct. The exchange part is in fact short-range. It reflects the explicit shape of the wavefunctions used and will thus be influenced by the choice of the pseudopotential. However, the matrix elements depend only weakly on the cutoff radius, as shown below, and, more important, the Coulomb attraction is as big as +0.95 eV, while the exchange part contributes only -0.09 eV to the total binding energy of about 0.8 eV of the exciton. This is typical for this type of material.

These calculations have been made with wavefunctions generated by using the Kleinman-Bylander separation. The influence on the *GW* and excitonic calculations

is expected, in contrast to the Kohn-Sham LDA gaps (see Tab. A.3), to be very small, as will be further demonstrated in the next section.

A.3.2 Optical Transitions

In this section we examine the influence of the pseudopotential and the Kleinman-Bylander separation on the optical matrix elements. They are essential ingredients in the evaluation of the absorption spectrum, see, for example, Eqs. (3.4) or (3.87). Compared to ground state and *GW* or excitonic calculations the calculation of optical spectra is different in two main aspects.

Firstly and most important, pseudopotentials are generally constructed in order to find the ground state and excited states are not directly accessible for higher energies. The Eq. (A.24) for transferability of the PP is typically only fulfilled for some 15 eV around the Fermi level. Nevertheless, in the *GW* calculations we sum up eigenstates up to 100 eV and more above the Fermi energy in Eqs. (2.54) and (2.106). It is, however, not necessary for a reasonable *GW* correction that each individual state is exact. Only the sums must be correct and we can thus tolerate errors in the energy level and the matrix elements, if they average out on the whole in the sums. This situation is different for the absorption spectrum, as – in particular for the independent-particle approximation Eqs. (3.4) – each transition is considered separately and the optical matrix elements give the probability of the transition and thus the height of the absorption curve. On the other hand, only the first 20 eV above the Fermi level are important in the optical and ultraviolet range and the transferability of the PP is normally good. One can therefore assume the PP description to work.

The second difference with respect to ground state calculations is the crucial importance of the band dispersion. The choice of the pseudopotential influences somewhat the band structure, which is reflected in the spectrum. In order to precisely determine the energy of prominent peaks in the spectrum, we need even for large band gap materials a small absolute error in the band dispersion.

Test System: Γ Point Only

Here we will consider a test system with only Γ as sampling point in order to simplify the calculations. We take an fcc cell for lithium oxide with only the Γ point for the Brillouin zone sampling at the experimental lattice constant of $a_0 = 4.573$ Å. In addition to the sMTO and the hMTO we also use a Bachelet-Hamann-Schlüter pseudopotential [80] for oxygen (BHSO) with cutoff radii of $r_c(s) = 0.48$ a.u. and $r_c(p) = 0.28$ a.u. This pseudopotential is much harder than both MT PP are. The local reference component is p ($\underline{sp}$) and the required cutoff energy for full convergence is 200 Ry (7727 plane waves).

Eigenvalues All PP configurations have been calculated with and without using the Kleinman-Bylander separation. In Tab. A.4 we summarize the gaps of the Kohn-Sham LDA eigenvalues. The gaps of the sMTO and the BHSO are closer to each other for the KB separation than to the hMTO, which a maximal difference of 10% for all PP.

Gap Γ	sMTO	hMTO	BHSO
KB	3.02	2.85	3.06
non KB	2.80	2.79	3.08

Table A.4: Kohn-Sham LDA gaps of Li_2O for various PP configurations using only the Γ point for the BZ sampling with and without the KB separation at the experimental lattice constant. Values in eV.

Similar to silicon in Fig. 4.2 we examine the eigenvalues at the Γ point in Fig. A.4 for the first 100 bands. Up to about 40 eV above the Fermi level there is virtually no difference induced by the Kleinman-Bylander separation on the energy scale presented and the various PP configurations resemble each other very much. For higher values, the eigenvalues shift and regroup again in a somewhat different manner. These changes are unimportant for sums over many conduction bands, like for the independent particle polarization χ^0 in Eq. (2.57), since they are small and unimportant with respect to the difference to the Fermi level.

Optical Matrix Elements We have investigated the optical matrix elements for the first 100 bands by comparing the results of the three pseudopotential configurations (sMTO, hMTO and BHSO) at the Γ point. Only the three highest, degenerate valence bands (bands 2-4) are considered, as they give the major contribution in the excitonic calculations and the absorption spectra. Groups of degenerate valence and conduction bands are summed over. The plotted data is thus,

$$P(c_i^{\min}) = \sum_{v=2}^{4} \sum_{c=c_i^{\min}}^{c_i^{\max}} |\mathbf{M}_{cv}(\Gamma)|^2 , \tag{A.39}$$

where

$$\mathbf{M}_{cv}(\mathbf{k}) = \left\langle c\mathbf{k} \left| \frac{\boldsymbol{\nabla}}{i} \right| v\mathbf{k} \right\rangle , \tag{A.40}$$

from Eq. (3.20), and the sum over the conduction bands c includes in each case i all bands of a degenerate group. The index $c_i^{\min}$ is the number of the lowest band of that group.

Figure A.4: Comparison of Kohn-Sham LDA eigenvalues of Li_2O for the first 100 bands using only the Γ point in the BZ sampling with and without the KB separation for various PP configurations at the experimental lattice constant.

In Fig. A.5 we compare the three PPs using the KB separation, while in Fig. A.6 we do not employ it. In both figures the differences between the PP configurations are generally less than 10%, though in some cases it can be larger.

The KB separation induces only small differences. This can be more explicitly seen in Figs. A.7-A.9, where we directly compare for each PP configuration the influence of the KB separation. Evidently, especially in the region near the Fermi level, which is most important for the exciton calculation and the absorption spectra, the differences between the three configuration are within 10%. This is the typical error margin of experiment and, thus, our results should be compared within this global precision.

We have hence shown that not only standard, already soft Martins-Troullier pseudopotentials are very well applicable for ground state and electronic structure calculations for oxides. One can make even softer pseudopotentials of this type and still get good results. However, differences, which are unimportant for ground state or electronic structure calculations, can be more visible in optical properties.

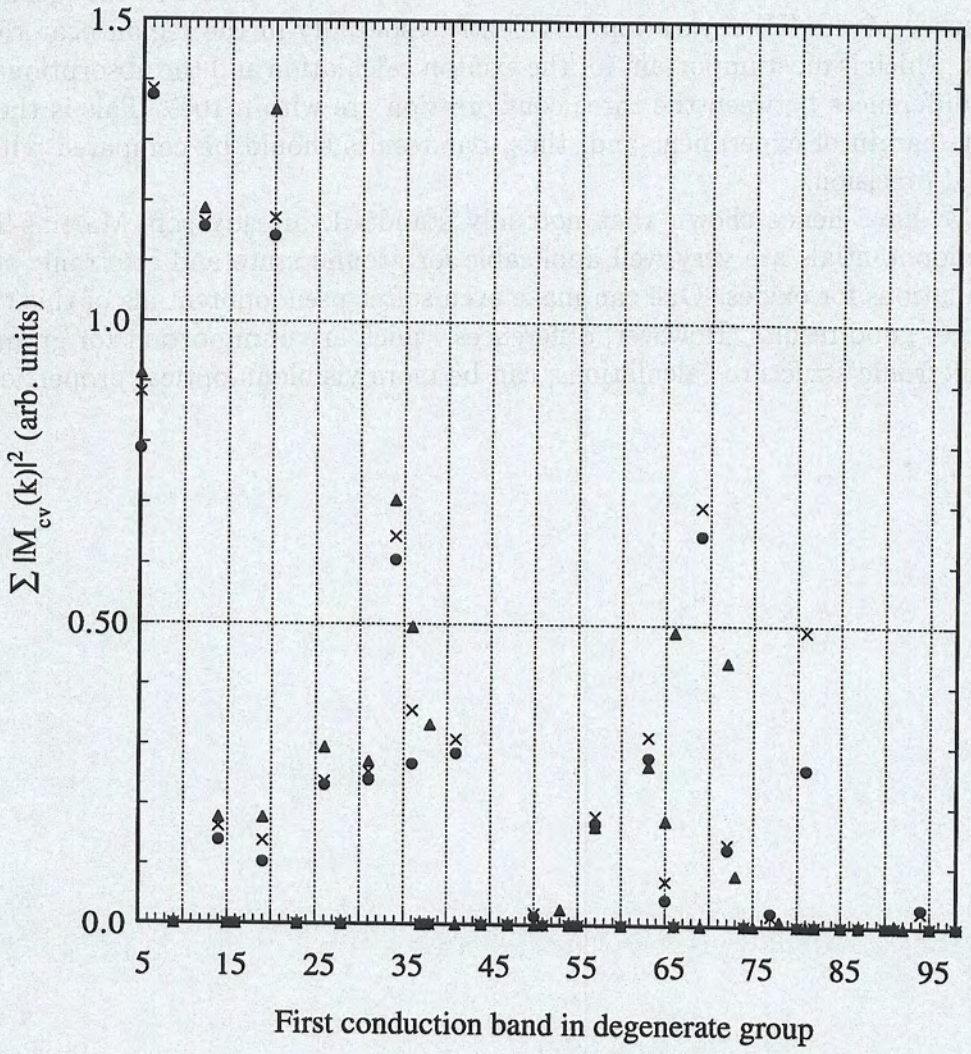


Figure A.5: Comparison of optical matrix elements $|\mathbf{M}_{cv}(\Gamma)|^2$ of Li_2O for the first 100 bands using only the Γ point in the BZ sampling and the KB separation for various PP configurations at the experimental lattice constant according to Eq. (A.39). ▲: sMTO, ×: hMTO, •: BHSO.

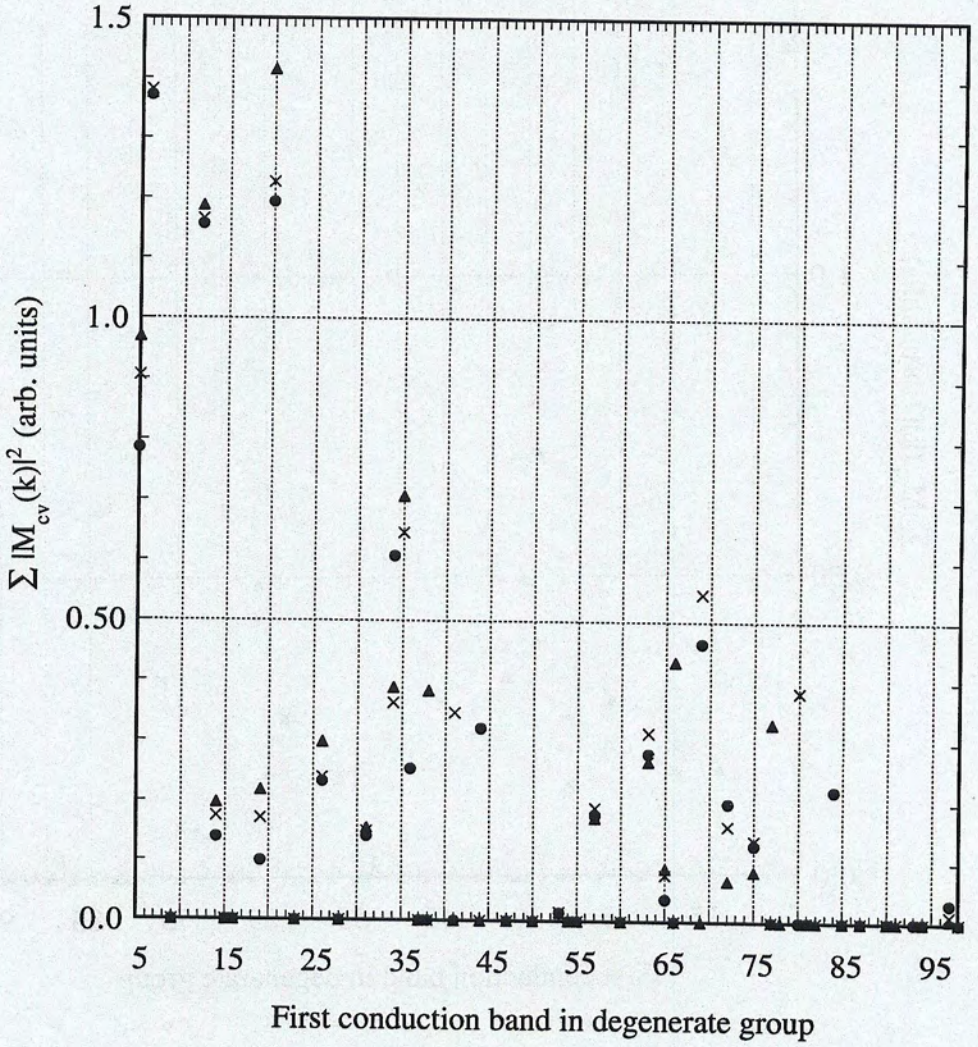


Figure A.6: Like Fig. A.5, but not employing the KB separation.

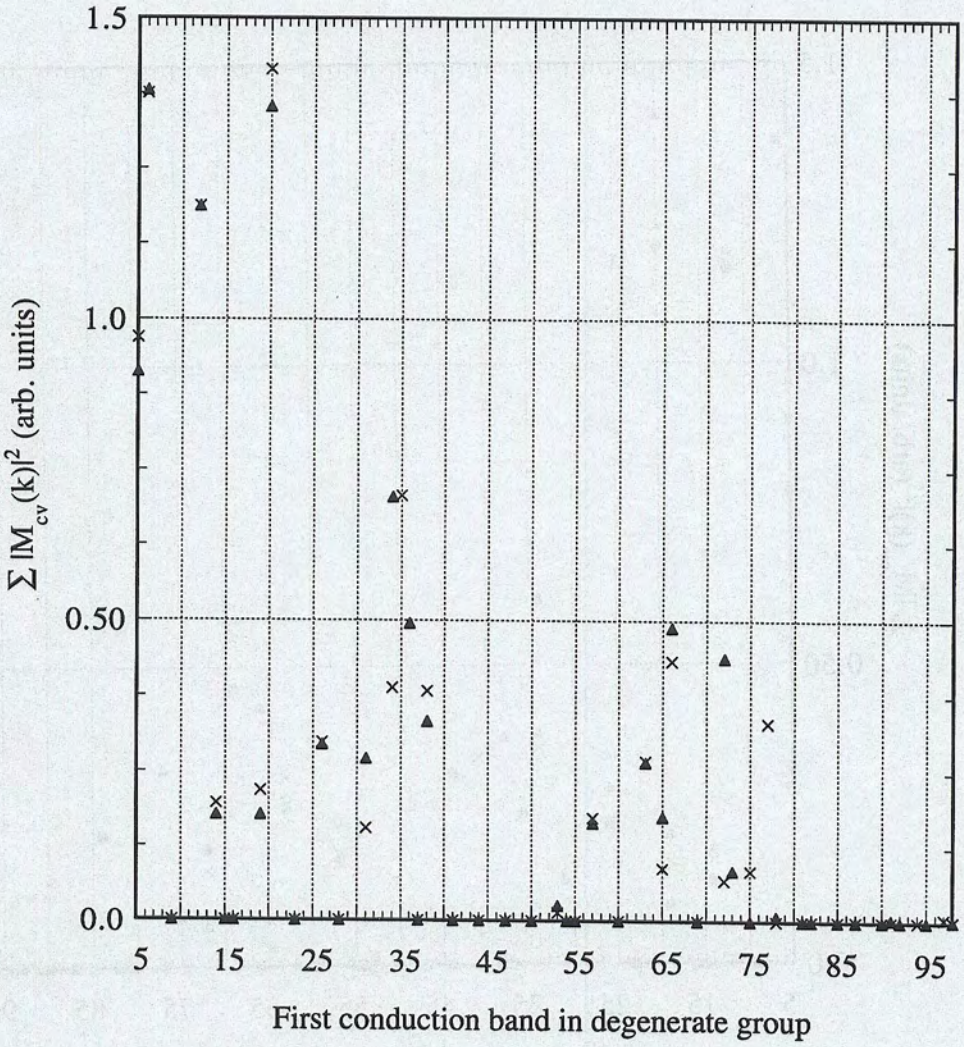


Figure A.7: Comparison of optical matrix elements $|M_{cv}(\Gamma)|^2$ of Li_2O for the first 100 bands using only the Γ point in the BZ sampling for the sMTO configuration at the experimental lattice constant according to Eq. (A.39). ▲: KB, ×: non KB.

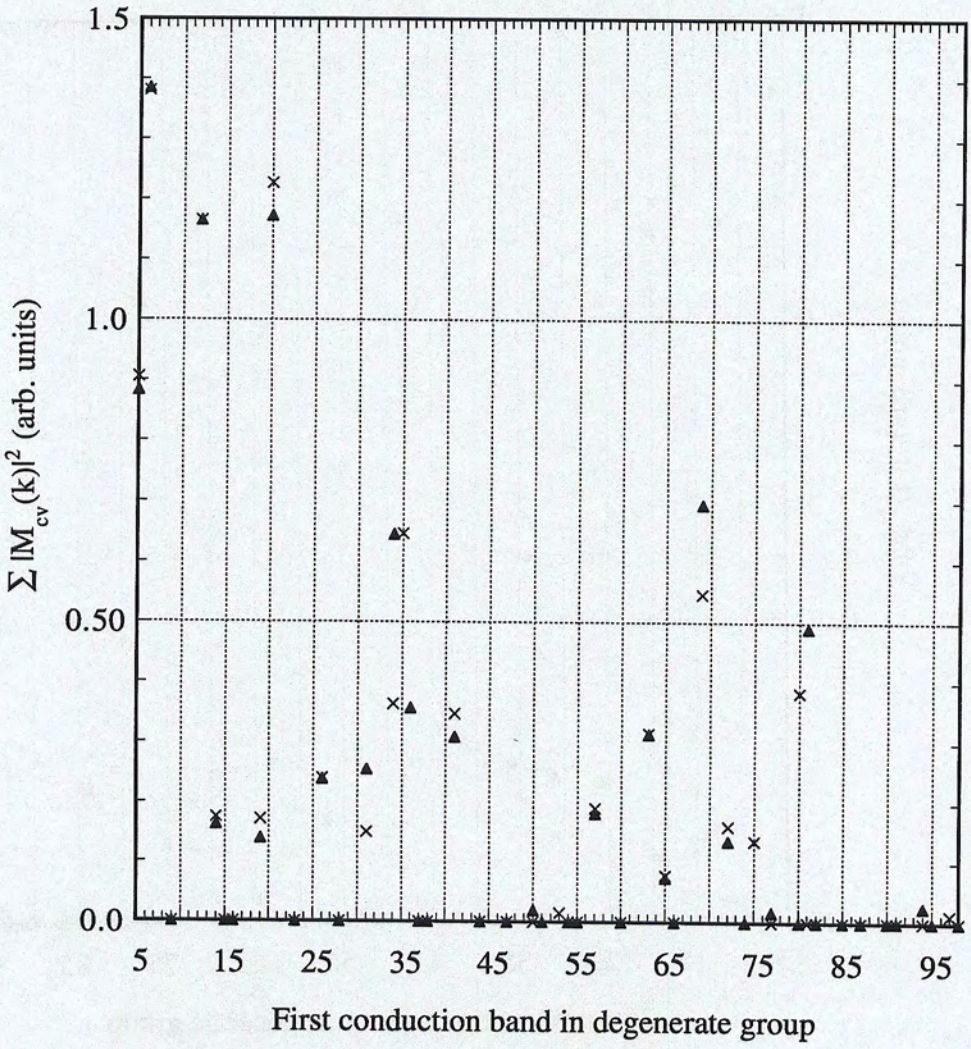


Figure A.8: Like Fig. A.7, but for the hMTO configuration.

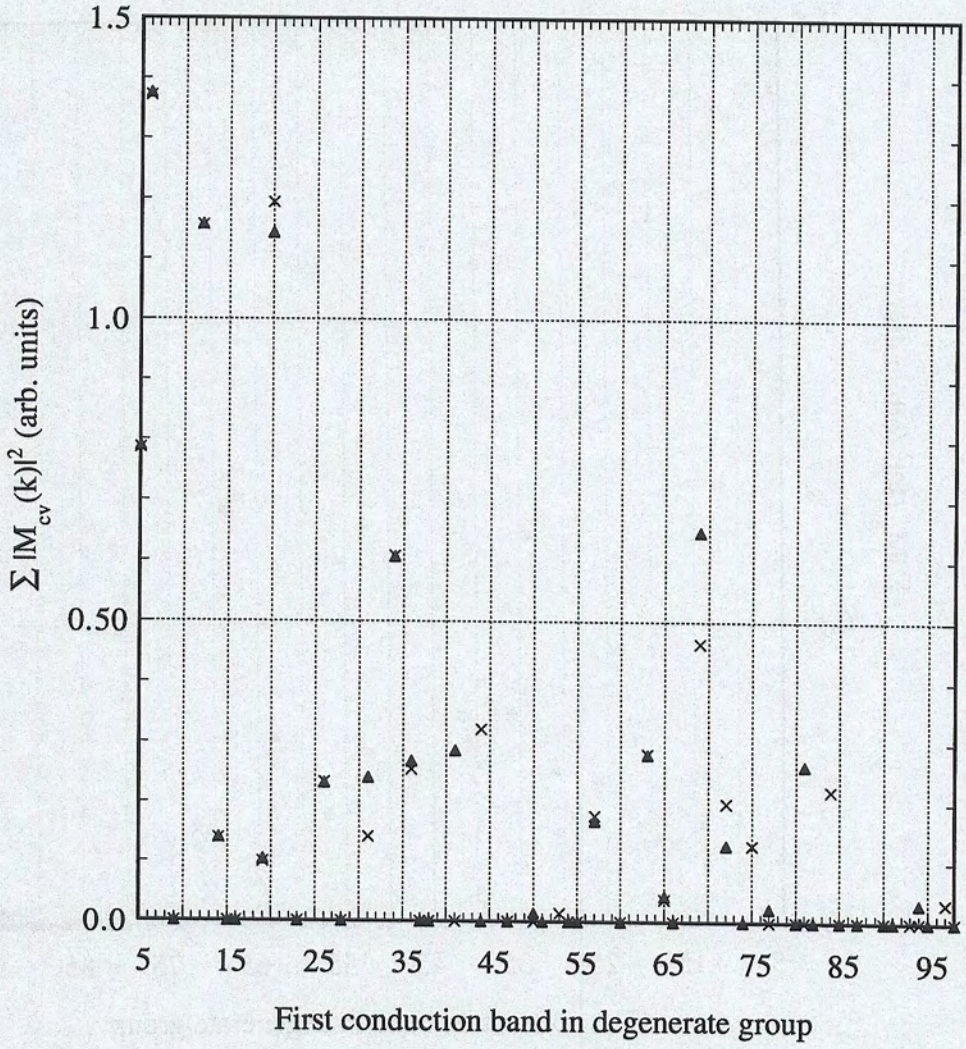


Figure A.9: Like Fig. A.7, but for the BHSO configuration.

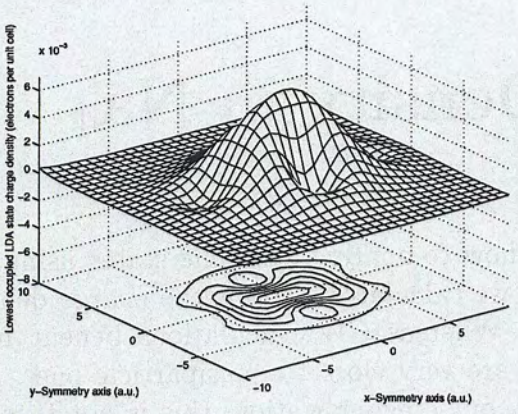
Appendix B

Excited Charge Density in Na_4

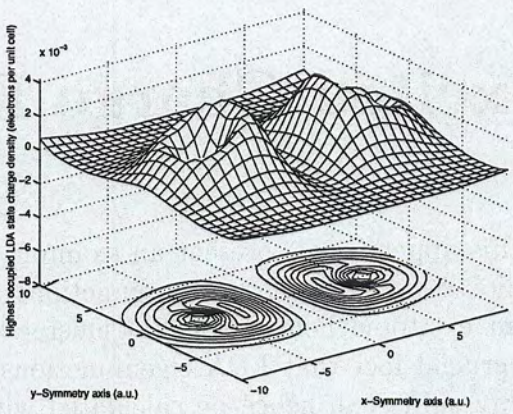
In this appendix we present an example how to analyze excitonic states using the explicit knowledge of the wavefunctions: we calculate and plot the charge density of an excitonic hole in an Na_4 cluster. Whereas *GW* calculations benefit from the crucial fact that LDA eigenfunctions are very close to quasiparticle ones, and energy shifts can hence be calculated without diagonalization, this is not true for the exciton problem: a diagonalization, or, equivalent, matrix inversions or repeated multiplications, cannot be avoided.

The sodium tetramer consists of four Na atoms in a plane [67]. DFT-LDA, *GW* and exciton calculation have been performed in the past by Onida *et al.* [48]. They have shown that the inclusion of excitonic effects is crucial in order to obtain an absorption spectrum in agreement with experiment.

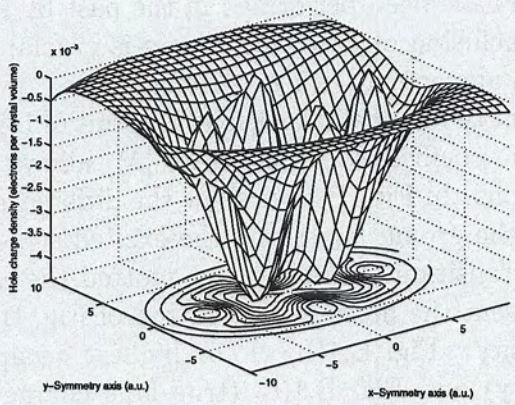
The decomposition of the optical transitions in terms of LDA electron–hole pairs, Eq. (3.72), given by the coefficients of the state $|\lambda\rangle$, allows to identify the strong mixing of LDA transitions. In Fig. B.1 we show the difference in the charge density of the cluster between the ground state and the excited state corresponding to the first optically allowed transition. Here we have plotted the negative contributions, Eq. (3.73), which represent the hole. A comparison of Fig. B.1(a) (lowest occupied LDA state charge density), Fig. B.1(b) (LDA highest occupied molecular orbital (HOMO) charge density) and Fig. B.1(c) (true hole charge of the excited state) shows that Figs. B.1(a) and B.1(b) both strongly contribute to the hole, illustrating that the excited state would be very badly approximated by a single LDA electron–hole pair.



(a) Charge density of the lowest occupied LDA state.



(b) Charge density of the highest occupied LDA state.



(c) Charge density of the true hole.

Figure B.1: Hole charge density of the first optically excited state of a sodium tetramer. The four Na atoms lie in the plane $z = 0$ at $(\pm 2.86, 0)$ a.u. and $(0, \pm 5.83)$ a.u.

Bibliography

[1] J. C. Phillips, in *Solid State Physics*, edited by F. Seitz, D. Turnbull, and H. Ehrenreich (Academic, New York, 1966), Vol. 18, p. 56.

[2] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[3] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[4] L. Hedin, Phys. Rev. B **139**, A796 (1965).

[5] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).

[6] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. B **37**, 10159 (1988).

[7] C. S. Wang and B. M. Klein, Phys. Rev. B **24**, 3417 (1981).

[8] G. E. Engel and B. Farid, Phys. Rev. B **46**, 1812 (1992).

[9] V. I. Gavrilenko and F. Bechstedt, Phys. Rev. B **55**, 13416 (1996).

[10] W. Hanke and L. J. Sham, Phys. Rev. B **12**, 4501 (1975).

[11] W. Hanke and L. J. Sham, Phys. Rev. Lett. **43**, 387 (1979).

[12] W. Hanke and L. J. Sham, Phys. Rev. B **21**, 4656 (1980).

[13] L. J. Sham and T. M. Rice, Phys. Rev. **144**, 807 (1966).

[14] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York, 1971).

[15] L. H. Thomas, Camb. Philos. Soc. **23**, 542 (1927).

[16] E. Fermi, Z. Phys. **48**, 73 (1928).

[17] R. M. Dreizler and E. K. U. Gross, *Density Functional Theory – An Approach to the Quantum Many-Body Problem* (Springer-Verlag, Berlin, 1990).

[18] S. Albrecht, Diplomarbeit, Technische Universität München, 1996.

[19] F. Finocchi, Ph.D. thesis, Università degli Studi di Roma "La Sapienza", 1992.

[20] W. Kohn and P. Vashishta, in *Theory of the Inhomogeneous Electron Gas*, edited by S. Lundqvist and N. H. March (Plenum, New York, 1983).

[21] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[22] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[23] S. Albrecht, G. Onida, and L. Reining, Phys. Rev. B **55**, 10 278 (1997).

[24] R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).

[25] M. S. Hybertsen and S. G. Louie, Phys. Rev. Lett. **55**, 1418 (1985).

[26] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **32**, R7005 (1985).

[27] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. Lett. **56**, 2415 (1986).

[28] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. B **36**, 3497 (1987).

[29] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. B **35**, R4170 (1987).

[30] L. J. Sham and M. Schlüter, Phys. Rev. Lett. **51**, 1888 (1983).

[31] J. P. Perdew and M. Levy, Phys. Rev. Lett. **51**, 1884 (1983).

[32] M. Levy, J. P. Perdew, and V. Sahni, Phys. Rev. A **30**, 2745 (1984).

[33] C.-O. Almbladh and U. von Barth, Phys. Rev. B **31**, 3231 (1985).

[34] L. J. Sham and M. Schlüter, Phys. Rev. B **32**, 3883 (1985).

[35] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

[36] M. Rohlfing, P. Krüger, and J. Pollmann, Phys. Rev. B **48**, 17791 (1993).

[37] M. Rohlfing, P. Krüger, and J. Pollmann, Phys. Rev. B **56**, R7065 (1997).

[38] D. J. Chadi and M. L. Cohen, Phys. Rev. B **8**, 5747 (1973).

[39] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).

[40] S. L. Altmann, *Band Theory of Solids: An Introduction from the Point of View of Symmetry* (Clarendon, Oxford, 1995).

[41] R. Car and M. Parrinello, in *Simple Molecular Systems at Very High Density*, edited by A. Polian, P. Loubeyre, and N. Boccara (Plenum, New York, 1989), p. 455.

[42] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **35**, 5585 (1987).

[43] P. Singhal and J. Callaway, Phys. Rev. B **14**, 2347 (1976).

[44] A. R. Williams and U. von Barth, in *Theory of the Inhomogeneous Electron Gas*, edited by S. Lundqvist and N. H. March (Plenum, New York, 1983), p. 189.

[45] S. L. Adler, Phys. Rev. **126**, 413 (1962).

[46] N. Wiser, Phys. Rev. **129**, 62 (1963).

[47] R. W. Godby and R. J. Needs, Phys. Rev. Lett. **62**, 1169 (1989).

[48] G. Onida, L. Reining, R. W. Godby, R. Del Sole, and W. Andreoni, Phys. Rev. Lett. **95**, 818 (1995).

[49] G. B. Bachelet and N. E. Christensen, Phys. Rev. B **31**, 879 (1985).

[50] M. S. Hybertsen and S. G. Louie, Solid. State Commun. **51**, 451 (1984).

[51] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **30**, 5777 (1984).

[52] A. Zunger and A. J. Freeman, Phys. Rev. B **15**, 5049 (1977).

[53] E. K. U. Gross, E. Runge, and O. Heinonen, *Many-Particle Theory* (Adam Hilger, Bristol, 1991).

[54] L. Hedin and S. Lundqvist, in *Solid State Physics*, edited by F. Seitz, D. Turnbull, and H. Ehrenreich (Academic, New York, 1969), Vol. 23, p. 1.

[55] G. C. Wick, Phys. Rev. **80**, 268 (1950).

[56] V. M. Galitskii and A. B. Migdal, Sov. Phys. JETP **7**, 96 (1958).

[57] R. Del Sole and R. Girlanda, Phys. Rev. B **48**, 11789 (1993).

[58] W. Jones and N. H. March, *Theoretical Solid State Physics* (John Wiley, London, 1973).

[59] F. Bassani and G. Pastori Parravicini, in *Electronic States and Optical Transisitions in Solids*, edited by R. A. Ballinger (Pergamon, London, 1975).

[60] A. F. Starace, Phys. Rev. A **3**, 1242 (1971).

[61] B. Adolph, V. I. Gavrilenko, K. Tenelsen, and F. Bechstedt, Phys. Rev. B **53**, 9797 (1996).

[62] G. Strinati, Phys. Rev. B **29**, 5718 (1984).

[63] G. Strinati, Phys. Rev. Lett. **49**, 1519 (1982).

[64] F. Bechstedt, K. Tenelsen, and B. Adolph, Phys. Rev. Lett. **78**, 1528 (1997).

[65] S. Albrecht, G. Onida, L. Reining, and R. Del Sole, Comp. Mat. Sci. **10**, 356 (1998).

[66] S. Albrecht, L. Reining, R. Del Sole, and G. Onida, Phys. Rev. Lett. **80**, 4510 (1998).

[67] S. Albrecht, L. Reining, R. Del Sole, and G. Onida, phys. stat. sol. (a) **170**, 189 (1998).

[68] L. X. Benedict, E. L. Shirley, and R. B. Bohn, Phys. Rev. B **57**, R9385 (1998).

[69] L. X. Benedict, E. L. Shirley, and R. B. Bohn, Phys. Rev. Lett. **20**, 4514 (1998).

[70] D. J. Stukel, R. N. Euwema, T. C. Collins, F. Herman, and R. L. Kortum, Phys. Rev. **179**, 740 (1969).

[71] J. P. Walter, M. L. Cohen, Y. Petroff, and M. Balkanski, Phys. Rev. B **1**, 2661 (1970).

[72] H. R. Philipp and H. Ehrenreich, Phys. Rev. **129**, 1550 (1963).

[73] D. E. Aspnes and A. A. Studna, Phys. Rev. B **27**, 985 (1983).

[74] P. Lautenschlager, M. Garriga, L. Viña, and M. Cardona, Phys. Rev. B **36**, 4821 (1987).

[75] S. G. Louie, J. R. Chelikowsky, and M. L. Cohen, Phys. Rev. Lett. **34**, 155 (1975).

[76] J. C. Phillips, Phys. Rev. **136**, A1705 (1964).

[77] B. Velicky and J. Sak, phys. stat. sol. **16**, 147 (1966).

[78] J. E. Rowe, F. H. Pollack, and M. Cardona, Phys. Rev. Lett. **22**, 933 (1969).

[79] R. R. L. Zucca and Y. R. Shen, Phys. Rev. B **1**, 2668 (1970).

[80] G. B. Bachelet, D. R. Hamann, and M. Schlüter, Phys. Rev. B **26**, 4199 (1982).

[81] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).

[82] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).

[83] *Zahlenwerte und Funktionen aus Naturwissenschaften und Technik*, edited by K.-H. Hellwege and O. Madelung (Springer, New York, 1982), Vol. III of Landolt-Börnstein, pt. 17a.

[84] F. D. Murnaghan, Proc. Nat. Acad. Sci. U.S.A. **3**, 244 (1944).

[85] J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B **14**, 556 (1976).

[86] M. Welkowsky and R. Braunstein, Phys. Rev. B **5**, 497 (1972).

[87] G. Lehmann and M. Taut, phys. stat. sol. (b) **54**, 469 (1972).

[88] G. Gilat and N. R. Bharatiya, Phys. Rev. B **12**, 3479 (1975).

[89] N. E. Maddocks, R. W. Godby, and R. J. Needs, Phys. Rev. B **49**, R8502 (1994).

[90] E. L. Shirley, Phys. Rev. B **54**, 16464 (1996).

[91] S. Albrecht, L. Reining, G. Onida, and R. Del Sole, Il Nuovo Cimento **20 D**, 949 (1998).

[92] R. Del Sole and R. Girlanda, Phys. Rev. B **54**, 14376 (1996).

[93] R. R. L. Zucca, J. P. Walter, Y. R. Shen, and M. L. Cohen, Solid State Commun. **8**, 627 (1970).

[94] G. Strinati, H. J. Mattausch, and W. Hanke, Phys. Rev. B **25**, 2867 (1982).

[95] Z. H. Levine and D. C. Allan, Phys. Rev. Lett. **63**, 1719 (1989).

[96] F. Wooten, *Optical Properties of Solids* (Academic, New York, 1972).

[97] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **35**, 5602 (1987).

[98] J. A. Van Vechten and R. M. Martin, Phys. Rev. Lett. **28**, 446 (1972).

[99] J. Donohue, *The Structures of the Elements* (Krieger, Malabar, FL, 1982).

[100] H. J. McSkimin and P. Andreatch, Jr, J. Appl. Phys. **43**, 985 (1972).

[101] A. D. Papadopoulos and E. Anastassakis, Phys. Rev. B **43**, 5090 (1990).

[102] H. R. Phillip and E. A. Taft, Phys. Rev. **136**, A1445 (1964).

[103] S. Albrecht, L. Reining, R. Del Sole, and G. Onida (unpublished).

[104] R. Haydock, Comput. Phys. Commun. **20**, 11 (1980).

[105] M. Rohlfing and S. G. Louie, Phys. Rev. Lett. **81**, 2312 (1998).

[106] E. Fermi, Il Nuovo Cimento **11**, 157 (1934).

[107] M. L. Cohen and V. Heine, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1970), Vol. 24, p. 38.

[108] M. Schlüter, J. R. Chelikowsky, S. G. Louie, and M. L. Cohen, Phys. Rev. Lett. **34**, 1385 (1975).

[109] D. R. Hamann, M. Schlüter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).

[110] H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, 1980).

[111] J. C. Phillips and L. Kleinman, Phys. Rev. **116**, 287 (1983).

[112] G. P. Kerker, J. Phys. C **13**, L189 (1980).

[113] R. W. Shaw and W. A. Harrison, Phys. Rev. **163**, 604 (1967).

[114] J. Friedel, Phil. Mag. **43**, 153 (1952).

[115] A. Messiah, *Quantum Mechanics* (North-Holland, Amsterdam, 1974), Vol. 1, p. 98.

[116] X. Gonze, R. Käckell, and M. Scheffler, Phys. Rev. B **41**, 12264 (1990).

[117] V. Olevano, private communication.

[118] C. E. Johnson, K. R. Kummerer, and E. Roth, J. Nucl. Mater. **155-157**, 188 (1988).

[119] K. Noda, Y. Ishii, H. Ohno, H. Watanabe, and H. Matsui, Adv. Ceram. Mater. **25**, 155 (1989).

[120] N. M. Masaki, K. Noda, H. Watanabe, R. G. Clemmer, and G. W. Hollenberg, J. Nucl. Mater. **212-215**, 908 (1994), and references therein.

[121] A. De Vita, M. J. Gillan, J. S. Lin, M. C. Payne, I. Štich, and L. J. Clarke, Phys. Rev. Lett. **68**, 3319 (1992), and references therein.

[122] P. Vajda and F. Beuneu, Phys. Rev. B **53**, 5335 (1996).

[123] K. Uchida, K. Noda, T. Tanifuji, S. Nasu, T. Kirihara, and A. Kikuchi, phys. stat. sol. (a) **58**, 557 (1980).

[124] W. Rauch, Z. Phys. **116**, 652 (1940).

[125] R. Dovesi, C. Roetti, C. Freyria-Fava, and M. Prencipe, Chem. Phys. **156**, 11 (1991).

[126] L. Liu, V. E. Henrich, W. P. Ellis, and I. Shindo, Phys. Rev. B **54**, 2236 (1996).

[127] K. M. Glassford and J. R. Chelikowsky, Phys. Rev. B **46**, 1284 (1992).

[128] M. Palummo, L. Reining, M. Meyer, and C. M. Bertoni, in *Prodeedings of the 22nd International Conference on the Physics of Semiconductors, Vancouver, Canada, 1994*, edited by D. J. Lockwood (World Scientific, Singapore, 1995), p. 161.

[129] S. Hull, T. W. D. Farley, W. Hayes, and M. T. Hutchings, J. Nucl. Mater. **160**, 125 (1988).

[130] M. P. Surh, H. Chacham, and S. G. Louie, Phys. Rev. B **51**, 7464 (1995).

[131] T. Osaka and I. Shindo, Solid State Commun. **51**, 421 (1984).

[132] X. Gonze, P. Ghosez, and R. W. Godby, Phys. Rev. Lett. **74**, 4035 (1995).

ÉDITÉ PAR
LA DIRECTION DE L'INFORMATION
SCIENTIFIQUE ET TECHNIQUE
CEA/SACLAY 91191 GIF-SUR-YVETTE CEDEX FRANCE