PhD-FSTC-2019-63
The Faculty of Sciences, Technology and Communication

DISSERTATION

Defence held on 30/09/2019 in Luxembourg

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG

EN PHYSIQUE

by

Fulvio PALEARI
Born on 16 May 1990 in Carate Brianza (Italy)

TITLE OF THE DISSERTATION

First-principles approaches to the description of indirect absorption and luminescence spectroscopy: exciton-phonon coupling in hexagonal boron nitride

Dissertation defence committee

Prof Dr Ludger Wirtz, dissertation supervisor
Professor, Université du Luxembourg

Dr Lucia Reining
Permanent researcher, Laboratoire des Solides Irradiés, École Polytechnique, Palaiseau

Prof Dr Susanne Siebentritt, Chairman
Professor, Université du Luxembourg

Prof Dr Matthieu Verstraete
Professor, Nanomat Lab, Université de Liège

A-Prof Dr Jorge Íñiguez
Professor, Luxembourg Institute of Science and Technology and Université du Luxembourg

UNIVERSITY OF LUXEMBOURG

DOCTORAL THESIS

First-principles approaches to the description of indirect absorption and luminescence spectroscopy: exciton-phonon coupling in hexagonal boron nitride

Author: Fulvio Paleari
Supervisor: Prof. Ludger Wirtz

Reviewer: Prof. Jorge Íñiguez
Reviewer: Dr. Lucia Reining
Reviewer: Prof. Susanne Siebentritt
Reviewer: Prof. Matthieu Verstraete

Theoretical Solid State Physics Group
Physics and Materials Science Research Unit

UNIVERSITÉ DU LUXEMBOURG

List of publications

1 T. Galvani, **F. Paleari**, H. Miranda, A. Molina-Sánchez, L. Wirtz, S. Latyl, H. Amara, and F. Ducastelle, *Excitons in boron nitride single layer*, Phys. Rev. B. **94**, 125303 (2016).

2 L. Sponza, H. Amara, C. Attaccalite, S. Latyl, T. Galvani, **F. Paleari**, L. Wirtz, and F. Ducastelle, *Direct and indirect excitons in boron nitride polymorphs: A story of atomic configuration and electronic correlation*, Phys. Rev. B. **98**, 125206 (2018).

3 **F. Paleari**, T. Galvani, H. Amara, F. Ducastelle, A. Molina–Sánchez, and L. Wirtz, *Excitons in few–layer hexagonal boron nitride: Davydov splitting and surface localization*, 2D Materials **5**, 4 (2018).

4 **F. Paleari**, H. Miranda, A. Molina–Sánchez, and L. Wirtz, *Exciton-Phonon Coupling in the Ultraviolet Absorption and Emission Spectra of Bulk Hexagonal Boron Nitride*, Phys. Rev. Lett. **122**, 187401 (2019)

5 D. Sangalli, A. Ferretti, H. Miranda, C. Attaccalite, I. Marri, E. Cannuccia, P. Melo, M. Marsili, **F. Paleari**, A. Marrazzo, G. Prandini, P. Bonfà, M. Atambo, F. Affinito, M. Palummo, A. Molina–Sánchez, C. Hogan, M. Grüning, D. Varsano, and A. Marini, *Many–body perturbation theory calculations using the yambo code*, J. Phys. Condens. Matter **31**, 32 (2019)

Abstract

The development of novel optoelectronic applications crucially depends on the detailed understanding of light–matter interaction in the candidate materials. From a theoretical point of view, this task is especially difficult in the case of quasi–2D semiconductors, since their optical response is dominated by strongly bound excitons and many–body perturbation theory (MBPT) must be employed together with first–principles computer simulations. The case of hexagonal boron nitride (hBN) is interesting because its large band gap and high absorption/emission efficiency make it amenable for the fabrication of UV emitting devices. However, the specific microscopic mechanisms that govern the appearance of complex fine structures in the optical spectra of different types of hBN samples (monolayers, few–layers, bulk samples) are poorly understood, leading to discrepancies between experimental and theoretical results. In this Thesis, we first show how the interlayer interaction in multilayers leads to a Davydov splitting of the excitonic states of single–layer hBN. We study and characterize the absorption spectra in single layer, multilayer and bulk hBN systems, focusing on the exciton symmetry and optical activity. We show that in multilayers, Davydov splitting leads to a surface localization of the lowest–lying optically active excitons.

These additional spectral features still cannot explain, by themselves, the experimentally measured optical fine structure. Therefore, we calculate exciton dispersion curves in order to search for the existence of indirect excitons with lower energy than the lowest direct exciton. We find that in bilayer hBN the nature of the optical gap (direct) changes with respect to the single–particle gap (indirect). In contrast, in bulk hBN (and thicker few–layers of hBN) both optical and quasiparticle gaps are indirect, i.e., they display a pronounced minimum in the excitonic dispersion curve between Γ and K.

If the lowest exciton is indirect, as in bulk hBN, then phonon–assisted transitions become relevant for the description of the optical spectra. The reliable *ab initio* description of exciton–phonon coupling in indirect absorption and emission is the main focus of this Thesis. We have tackled the problem with two approaches. In the static approach, we calculate the coupling of excitons with phonons in a supercell via a finite–displacement method. The supercell is commensurate with the k–point corresponding to the minimum of the exciton dispersion. In this way, we are able to reproduce the rich fine structure in the luminescence spectrum of bulk hBN in good agreement with experiment and to explain it in terms of exciton–phonon coupling.

The finite–displacement approach is supposed to work for indirect optical spectra in any material where the minimum of the exciton dispersion gives the dominant phonon–assisted contribution. Since in the absorption case this is not always true, we have employed a second, perturbative approach for the description of indirect absorption spectra which also includes dynamical effects and a microscopic treatment of the coupling. We have implemented the required many–body quantities (namely, the exciton–phonon coupling matrix elements and self–energy) in the `Yambo` many–body code. We expect this new method to allow us to overcome many of the theoretical limitations of previous approaches to indirect absorption.

Contents

List of Abbreviations

BO	**B**orn–**O**ppenheimer
BSE	**B**ethe–**S**alpeter **e**quation
BZ	**B**rillouin **Z**one
CL	**C**athodo**l**uminescence
DFT	**D**ensity **F**unctional **T**heory
DFPT	**D**ensity **F**unctional **P**erturbation **T**heory
DW	**D**ebye–**W**aller
eh	Electron–**h**ole
em	Electro**m**agnetic
ep	Electron–**p**honon
EPC	Electron–**p**honon **c**oupling
ExcP	Exciton–**p**honon
FM	**F**an–**M**igdal
GF	**G**reen's **f**unction
HPC	**H**igh–**p**erformance **c**omputing
IBZ	**I**rreducible **B**rillouin **z**one
il	Inter**l**ayer
ip	**I**n–**p**lane
IP	**I**ndependent–**p**article
KS	**K**ohn–**S**ham
LDA	**L**ocal **d**ensity **a**pproximation
LT	**L**ongitudinal–**t**ransverse
MBPT	**M**any–**B**ody **P**erturbation **T**heory
NFE	**N**early–**f**ree **e**lectrons
RPA	**R**andom **p**hase **a**pproximation
PL	**P**hoto**l**uminescence
PLE	**P**hoto**l**uminescence **e**xcitation
QE	**Q**UANTUM **E**SPRESSO
QP	**Q**uasi**p**article
sBSE	**S**tatic **B**ethe–**S**alpeter **e**quation
sscGW	**S**emi–**s**elf–**c**onsistent GW
TBEM	**T**ight–**b**inding **e**xcitonic **m**odel
TDA	**T**amm–**D**ancoff **a**pproximation
WL	**W**illiams–**L**ax

Dedicated to my colleagues who, after many years of research work, are now fighting to secure a stable employment.

Introduction

The interaction of light with matter is perhaps one of the most important natural phenomena, its effects extending from the cosmological to the microscopic scale, and from living organisms to inanimate objects. Down to the microscopic scale, the interaction takes place via the atomic constituents, electrons and nuclei. If we imagine two quantised, isolated energy levels with energies $E_i < E_f$, with an electron occupying level i, the *optical absorption* process takes place when the electron interacts with an incoming electro–magnetic field, whose quanta carry the energy $\hbar\omega = E_f - E_i$. The electron has a certain probability to absorb this energy and move to the level f, leaving i empty. Then, the system is in an unstable *excited state*. The excited electron will soon *relax* back to level i by *emitting* a photon with energy $\hbar\omega$. This simple one–electron, two–level picture is considerably altered in a condensed matter system, like a crystal, which contains $\sim 10^{23}$ particles per cm^3: the electrons are moving around the atomic nuclei which are arranged in an organized, symmetric structure called lattice. In this case, the mutual interactions between the particles complicate the response of the system to an incoming electro–magnetic field, with various many–body phenomena possibly shaping it (e.g., metal or insulating character, magnetism, electronic correlations and screening, electron–lattice interaction, ...).

This thesis is concerned with the theoretical calculation of optical spectra in hexagonal boron nitride, a prototypical layered material, taking into account the coupling of optical excitations with lattice vibrations. We introduce the topics below.

Theoretical spectroscopy

Spectroscopy is the study of the response of matter to external radiation, an electro–magnetic (em) field. This radiation brings the system into an *excited state*. In particular, *optical* spectroscopy is concerned with the response of the matter system to incoming em waves of frequencies (or equivalently wavelengths) broadly spanning the range between infrared and ultraviolet light. Normally, the information about the excited states of the system are encoded in a frequency–dependent *response function*, from which an optical spectrum can be obtained.[1]

In general, when light (coming, for example, from a laser beam) encounters the surface of a material, it is partly reflected and partly transmitted within. Inside the system, as light propagates in the presence of electronic and ionic charge densities, it may be scattered or absorbed. In the latter case, and if our system is semiconducting, only em waves with energy $\hbar\omega$ larger than the *energy gap* of the material may be absorbed. The system (at least in the absence of defects) will be transparent to the remaining light. Such light can then be collected in a detector upon exiting the

[1] In our case, we will always consider the response of the system to be *linear*, that is, its interaction with the external field is weak enough to be described by first–order perturbation theory. Similarly, the interation between the electrons and nuclei inside the material will be weak enough so that they may be initially considered as separate systems.

system and compared with the initial incoming field: the missing frequencies are the absorbed ones. This is the simplest way to obtain an *absorption spectrum*.[1] Clearly, the intensity, number, and energy of the spectral features depend on the probability of light absorption. Therefore, by analysing absorption spectra we are able to obtain a wealth of information on the electronic structure of the system. For example, the onset of the absorption spectrum is an approximate measure of the energy gap of the material: depending on its value, the material might become a valuable core component of a transistor, LED, solar cell, or be a transparent substrate to other crystals, *et cetera*. Furthermore, the intensities of the spectral peaks hint at the strength of the coupling between the electronic wave functions and the incoming em wave, as well as at the electronic density of states: this may give some insight into the possible usage of the material in an efficient light absorbing or emitting device. We can then conclude that optical spectra open a window into the inner, microscopic features of the system, revealing the quantum interplay between its particle constituents.

When the energy carried by the light is transferred to a system previously in its ground state (i.e. at equilibrium, undisturbed by external probes), we say that the system enters an excited state, or that an excitation is created. The simplest conceptualisation of such an excitation is the picture in which an electron is promoted from a valence band (an allowed energy–momentum level for the system at rest) to a conduction band (an allowed higher–energy level), that is the creation of an *electron–hole pair*. Conversely, suppose now that a non–monochromatic laser field excites many electrons to various conduction bands, or that a high–frequency monochromatic field sends an electron well above the energy gap. The excited–state relaxation dynamics of these electrons, taking place afterwards, will complicate the description of the system. However we can imagine, in the simplest case, that the excited carriers will relax *fast* to the respective band edges, and only then *recombine* (i.e. restoring the ground state of the system) with a certain probability. When this happens, the system must lose energy, and if the recombination is radiative a photon with the appropriate frequency will be emitted. This process is called *photoluminescence*. The frequency–resolved detection of the emitted photons constitutes a *luminescence* (or more broadly *emission*) spectrum. Luminescence spectra are in general more difficult to interpret than absorption ones, because they depend on the great variety of relaxation processes that may take place when the system is pushed out of equilibrium. However, since luminescence usually takes place only in a localised region of the sample (close to where the laser excites the system), these experiments are in most cases much easier to carry out than absorption ones. Furthermore, some features that are not present or negligible in the absorption spectrum may become dominant in the luminescence one (and viceversa), so that a comparison of the results from both techniques allows for a more complete insight into the electronic structure of the system and the physical mechanisms regulating the excitation process. Materials with high luminescence yield, for example, are well–suited for the fabrication of light–emitting devices such as lasers.

If the electron–electron interaction inside the system is highly screened, the electrons' behavior approaches that of non–interacting particles moving in an electrostatic potential generated by the "frozen" atomic nuclei. If this is not the case, additional effects have to be considered to correctly describe the process of electron–hole creation. For example, when an electron jumps to an excited state, it causes a redistribution of the remaining electrons in response to its change of position and energy. In particular, the electron will drag (or be pushed by) a cloud of nearby particles, affecting its motion, energy, scattering rates *et cetera*. We call the excited electron

"dressed" by its cloud a *quasiparticle*. A quasiparticle may also be formed by the excited electron plus its interaction with the nearby quantised lattice vibrations, i.e. the *phonons*. Indeed, the atomic nuclei are not frozen in their equilibrium positions, but oscillate around them in a type of collective motion determined by their own interaction with the electronic charge density. This picture may be additionally complicated when the electron non–negligibly interacts with its corresponding hole: in fact, the hole acquires an effective positive charge because of the electron's departure from its energy level in the valence band. The presence of an electron–hole interaction means that electron and hole motions are correlated. This can be in general conceptualised by considering the electron and the hole (which in general have different effective masses in a solid) as orbiting around each other. However, this is actually a collective excitation made of many different electron–hole pairs excited at the same time, and we call it *exciton*. If the exciton is highly localised in momentum (i.e. reciprocal) space, then it approaches the single electron–hole pair picture. In this thesis, we are going to investigate all these effects and in particular focus on the excitons and their further interaction with phonons as our main topic: in fact, we will study *indirect* optical excitations.

Since in general the wave vector of light in the optical range is much smaller that the one typically characterising the electronic wave functions in a solid — the latter's order of magnitude being determined by the crystal lattice constant — an electron excited from the valence to the conduction band only changes its energy but not its momentum. However, it is certainly possible that the bottom of the conduction band and the top of the valence one lie at different momenta (like in silicon — in this case the energy gap is *indirect*). Despite this, an electron–hole pair can still be optically created by a second–order process involving phonons. In fact, although the phonon energies are in general rather small compared to the electronic energy gap, their crystal wave vectors are of the same magnitude as the electron ones. The electron can then simultaneously (i) absorb light and (ii) be knocked "sideways" by a phonon, which provides the necessary momentum transfer to bridge the indirect gap. This is a phonon–assisted optical transition. Naturally, the reverse transition is also possible (phonon–assisted recombination). The probability of an indirect transition, being the product of the probabilities of both electron–light and electron–phonon scattering, is in general much lower than that of a direct transition. The indirect spectral signature, however, will be enhanced with respect to its direct counterpart in a luminescence experiment: this is because in this case the indirect transitions lie at lower energy. A scheme of direct and indirect optical transitions is displayed in Fig. 1

The necessity to obtain a theoretical equivalent to the experimental observables for realistic, complex materials suggests that they should be modeled using computer simulations. Indeed, the quantum mechanical equations governing the interactions of a many–body system can be solved numerically at different levels of approximation. In general, these techniques are termed *first–principles* or *ab initio*,[2, 3] in the sense that they do not require external fitting parameters and models, but only the number and type of charged particles present in the system along with their mutual Coulomb interaction. The most commonly employed of such first–principles techniques, useful to calculate the electronic ground state of the system and an approximation of its band structure, is called density functional theory (DFT).[4] Phonons can also be obtained in a DFT–based formalism by density functional perturbation theory (DFPT).[5, 6] Finally, quasiparticle corrections[7] and excitons[8] are obtained by many–body perturbation theory (MBPT), where the results of DFT simulations are used as input data. As more sophisticated first–principles simulations require more and more computational efforts, the availability of large–scale

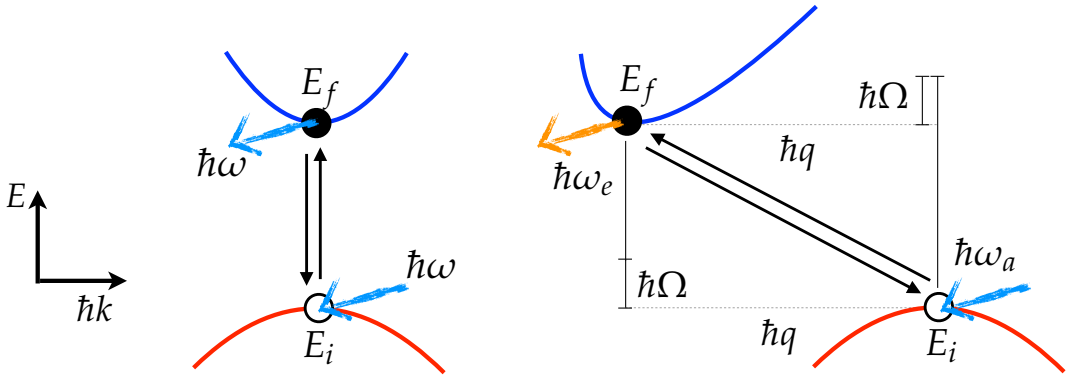


FIGURE 1: Two–band scheme for optical transitions (direct transtions: left, indirect: right). The valence band is marked with a red line, the conduction band with a blue line. An electron (black dot) is excited from valence to conduction, leaving a hole (white circle) behind. The presently formed electron–hole pair can also recombine via the inverse process. In the direct case, the absorption or emission of a photon with energy $\hbar\omega$ (blue arrow) enables the transition. In the indirect case, the photon provides most of the energy to bridge the band gap, but only the concomitant emission of a phonon (dotted arrow, energy $\hbar\Omega$) can account for the needed momentum $\hbar q = \hbar|k_f - k_i|$. In this case the energies of the absorbed and emitted photons are different (blue and orange arrows, respectively), leading to an energy shift between absorption and luminescence spectra.

high–performance computing (HPC) facilities is crucial for the undertaking of what we may also denote as "computational experiments".

In this thesis, we develop the necessary tools to perform computational indirect absorption and luminescence spectroscopy. We test this on a particularly interesting material, described in the following.

Boron nitride: a simple structure leads to complex spectral features

In the wake of the experimental discovery of graphene, the single layer of graphite, by Novoselov, Geim *et al.* in 2004,[2][10] an intense research effort — both experimental and theoretical — was started to investigate the unusual properties enjoyed by layered materials because of electronic quantum confinement along the stacking direction,[11] particularly semiconducting ones.[9] Among these materials, boron nitride holds the largest energy gap, deep in the UV range (see Fig. 2). Boron nitride in its hexagonal form (hBN) is stable at room temperature and ambient pressure.[12] Being made by stacked BN layers arranged in a hexagonal, honeycomb structure dictated by strong $sp2$ covalent bonds, and weakly coupled to each other by van der Waals forces and the overlapping p_z orbitals, hBN acts as the insulating counterpart of graphene/graphite.[13] It can be grown from solutions at high–pressure, using solvents that contain both B and N,[14, 15] or by epitaxial techniques such as chemical vapour deposition and molecular beam epitaxy.[16, 17] Few–layer samples can also be obtained by epitaxy or by exfoliating the bulk material (however their production has proven to be difficult due to easy defect creation[18]).

[2] K. S. Novoselov and A. K. Geim were awarded the 2010 Nobel Prize in Physics "for groundbreaking experiments regarding the two-dimensional material graphene." (https://www.nobelprize.org/prizes/physics/2010/summary/).

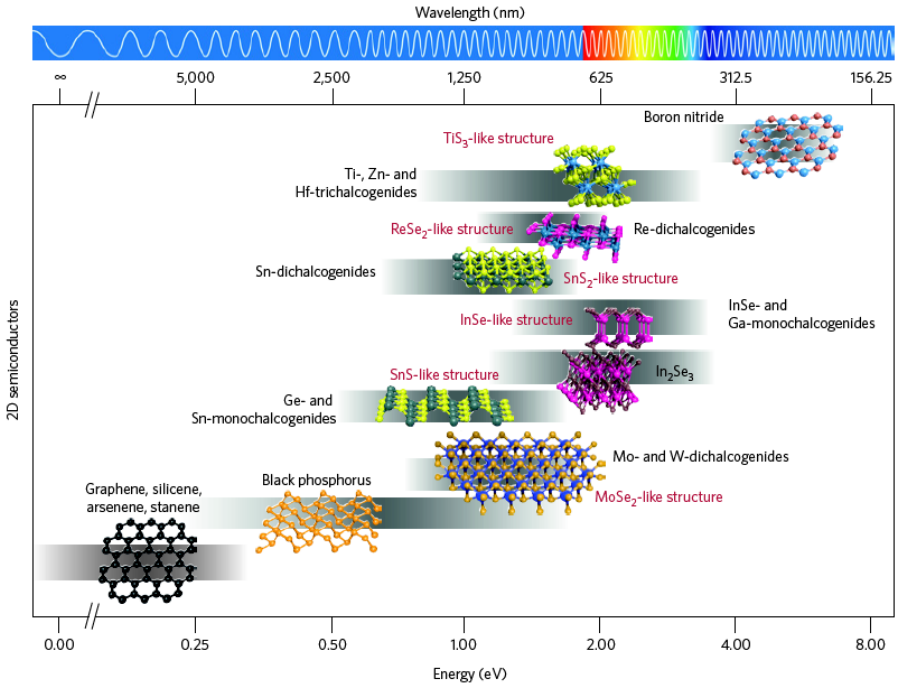


FIGURE 2: The "zoo" of 2D semiconductors. Figure taken from Ref. [9]. The layered materials represented in the graph are ordered from left to right according to increasing band gap energy. They start with graphene (0 eV) and end with hBN (> 6 eV)

Indeed, this material may appear quite simple at first sight: the required ingredients are just a unit cell composed of two light and easy–to–model atoms, a highly symmetric planar lattice structure (as shown in Fig. 2; see also Fig. 2.1), and only one relevant valence and conduction band, both spin–degenerate. Yet, the physical phenomena emerging from this uncomplicated structure are unusual, intriguing, and often difficult to grasp. Indeed, the interesting physical properties of hBN have been shown to be highly valuable for many applications, including graphene–based electronic devices (where hBN is used as an encapsulating substrate, flattening graphene and improving its electron mobility)[19, 20] and electrical insulating applications in vacuum and/or high temperature technology (in nanotube form).[16, 21, 22] Additionally, point defects in hBN have been proven to act as bright single–photon emitters and the pristine material, viewed from the perspective of infrared nanophotonics, can host hyperbolic phonon polaritons.[17] However, in the context of this thesis we are interested in the ultraviolet spectroscopy of hBN, since this material is also promising for optoelectronic applications in the deep UV range. In fact, its large band gap, in combination with the observed high luminescence yield of high–quality samples, makes hBN amenable for the fabrication of ultraviolet (UV) emitters.[14, 16, 21, 23] UV emitters have a variety of applications, arguably the simplest possible one being information storage technology (optical storage, i.e. laser–based data reading and writing). They may be also useful for environmental protection, since they can trigger redox reactions in TiO_2–based nanomaterials (i.e. photocatalysis): this process is able to remove organic pollutants from air– or water–based environments. Another important field of application is the medical one, since intense UV light can be used, for example, to sterilise and purify water. In order to envisage efficient BN–based UV emitters, however, the optical properties of bulk and few–layer hBN must be thoroughly understood: we will offer several contributions towards this goal in

this thesis.

From the experimental point of view, a variety of techniques have been used to investigate the optical properties of hBN, including cathodo– and photoluminescence, deep inelastic x–ray scattering, Raman spectroscopy and two–photon spectroscopy.[13, 14, 18–20, 24, 25] Theoretically, the state–of–the–art MBPT techniques — namely the quasiparticle and excitonic frameworks — have been employed in numerical simulations.[21–23, 26–29] From the theoretical works, it is clear that bulk hBN features an indirect quasiparticle gap, with the difference in energy with the smallest direct transition from the valence to the conduction band being small compared to the gap energy (about 0.5 eV). Excited electrons in hBN, however, are subjected to a highly anisotropic screening of their Coulomb interaction by the other electrons: the screening is high in the plane of the hexagonal layer, but very low in the rest of the volume, where the charge density is much reduced. This condition gives rise to strong electron–hole interactions. Indeed, the optical properties of hBN are found to be dominated by excitonic effects both experimentally and theoretically, and a description in terms of single–particle band structures is not possible. The binding energy of the lowest–lying "direct" exciton in BN systems is of the order of 1 eV, a value that is two to three orders of magnitude larger than in 3D semiconductors, and still much larger than in other layered materials such as transition metal dichalcogenides.[30]

Experimentally, the appearance of a rich fine structure (i.e. the presence of multiple spectral peaks in a small energy window) in the excitonic luminescence spectra of BN systems raised interesting questions about the true nature of intrinsic optical excitations in hBN[13, 18, 19, 24, 25, 31, 32] (see also Fig. 3.1). In fact, this complicates the experimental interpretation of absorption/luminescence spectra, where additional obfuscating effects due to defects, temperature, and low resolution (spectroscopy in the deep UV range is difficult) also play a role. Various explanations have been proposed for this fine structure: an additional relaxation of the lattice geometry taking place after optical excitation (an excitonic Jahn–Teller effect), thus splitting the degeneracy of excitonic states and thereby creating multiple peaks;[24] interaction with defect states, as well as breaking of the symmetry of the hexagonal lattice due to their presence;[13, 27] phonon sidebands arising from exciton recombinations. The latter interpretation has been convincingly advanced by Cassabois, Valvin and Gil in order to interpret their photoluminescence (PL) spectra. They present a case for phonon–assisted recombination of an *indirect* exciton in bulk hBN, albeit using the single–particle picture in their intepretation.[19] This conclusion, however, carries the implication that also the *optical* gap in hBN, that is, the energy gap at the exciton rather than single–particle level, has indirect and not direct character, despite the large binding energy theoretically predicted for the "direct" exciton. Indeed, phonon replicas have been observed in the cathodoluminescence (CL) spectra of few–layer hBN,[18] which would indicate the presence of a strong electron–phonon interaction in this material, something that we will also establish from the theoretical side at the end of Chapter 2. The energy position of the "indirect" peak in Ref. [19], however, is very close to the "direct" one predicted by theoretical simulations. The simulations also disagree with Ref. [19] about the nature of the excitons (either strongly or weakly bound), and finally, the few experimental data available on absorption do not show the expected mirror symmetry with the luminescence process. Certainly, a phonon–assisted recombination could reduce the probability that an emitted photon is immediately self–absorbed, since the two processes would then happen at different photon frequencies: this allows to extract more light from the material and can contribute to the observed high luminescence yield. However, the exciton–phonon

coupling should also be extremely efficient in order to justify the aforementioned high yield. In conclusion, the question of the nature of the optical gap in hBN, as well as that of the competition between indirect and direct optical absorption processes, had remained open.

Scope of this thesis

The goal of this thesis is to understand, calculate and explain the origin of the fine structures in the optical spectra of hBN systems, both in multilayer and bulk form. In particular, the coupling of finite–momentum excitons with phonons, leading to indirect contributions to optical absorption and luminescence, will be the main focus. The only way to definitively interpret the puzzling experimentally observed spectral features of hBN samples is to include the right physical mechanisms in the analysis. First, we must consistently introduce an "excitonic" point of view, and proceed in our spectroscopy analysis from there. This is accomplished numerically by systematically using the Bethe–Salpeter equation (BSE) for interacting electron–hole pairs as the starting point of our investigations.

A thorough understanding of the structure of the excitonic wave functions in hBN, including their symmetries and therefore their optical activity, can already provide some insight into the spatial localisation of the excitations: this is one source of fine structures in systems where the lattice periodicity along the stacking direction is broken (i.e. few–layer samples). In order to carry out this analysis we consider monolayer, bilayer, trilayer and "penta"–layer BN systems. The first–principles study of finite–momentum excitons is a relatively new[33, 34] and until recently very unexplored avenue. Experimentally, excitons are still often pictured as "defect–like" energy levels situated within the band gap. However, also interacting electron–hole pairs possess a dispersion relation, and its features might be crucial to explain absorption and luminescence spectra. For example, the radial ΓK segment of the hexagonal Brillouin zone (BZ), that in BN can discriminate between indirect and direct absorption (because it hosts the indirect minimum of the *single–particle* conduction band), had not been explored before. We show that it is in fact possible for the bilayer to still maintain a "*direct*" optical gap at the exciton level, with a direct–to–indirect crossover happening at higher layer number. Such a crossover may also be engineered by compressing the system in the planar direction. Additionally, we uncover and analyse the excitons corresponding to electronic transitions towards "nearly–free electron" (NFE) states,[35] which are usually ignored in the theoretical spectroscopy of BN but could provide additional minima to the exciton dispersion and potentially greatly reduce the luminescence yield for very thin films.

Once we have understood both the exciton and phonon dispersions and properties, we start delving into the electron-phonon (ep) and exciton–phonon (ExcP) interactions. We first test the strength of the coupling with phonons in the monolayer system and find that a strong ($\sim$ 0.5 eV) renormalisation of both the single–particle band gap and the exciton absorption spectrum is induced. Therefore, the ExcP coupling cannot be discarded in the description of the optical properties of BN systems.

Switching to the bulk system, we then investigate the indirect contribution to the absorption and luminescence spectra. Although some very recent advances have been made in the *ab initio* description of indirect absorption at the independent–particle (IP) level,[36, 37] these are few and far between, and even on the theoretical side the many–body description of indirect absorption is still lacking: in fact, it usually only relies on approximations of a textbook second–order perturbation theory

formula that has severe limitations. We build upon the existing treatments in order to extend their applicability to the exciton case.

The main achievement of the thesis consists in the reproduction of the experimentally observed luminescence spectrum: we obtain it by developing an approach to compute the phonon–assisted sidebands based on finite displacements of the atomic nuclei according to the phonon eigenvectors. We analyse the excitons at the wave vector q of the indirect gap. The evaluation of the ExcP coupling strengths is performed in a static approximation by a second–order expansion of the optical response function with respect to the atomic displacements within a simulation supercell commensurate with q. This expression reduces to the product of the mean square displacement of the atoms with the second derivative of the excitonic oscillator strengths. Our previously established knowledge of the exciton symmetries and dispersion relation allows us to compute the exciton–phonon–light selection rules, which are fulfilled by our computational results. Detailed balance arguments then allow us to obtain the luminescence spectrum from the absorption one if certain conditions are met. In this way, not only we obtain a good agreement with experiment: we also fully elucidate the underlying mechanisms regulating luminescence in this system and explain why the absorption and emission spectra are different: two different exciton states dominate the two different processes.

Finally, we employ a microscopic description of the ExcP coupling which retains dynamical effects, following an approach recently put forward in the context of plasmons.[38] We introduce it by analogy with simple single–particle models that can be solved exactly; in particular, we find that it is possible to obtain the exact solution for the (indirect) absorption spectrum of a two–level system featuring both electron–phonon and electron–hole interaction. The extension of the model to realistic cases is based on a dynamical correction to the BSE: the calculation of the dynamical part of the phonon–mediated electron–hole interaction at first order is expressed in terms of an exciton–phonon self–energy. This approach treats the ExcP coupling in a more sophisticated way and in principle permits the description on the same footing of both direct and indirect components of the absorption spectrum: this is because it allows for a full integration over all exciton momenta in reciprocal space. We have implemented the ExcP self–energy in the many–body code `Yambo` (after discussing its link to the optical response functions of interest, focusing in particular on the mixing of "longitudinal" and "transverse" excitons), and some preliminary results are shown at the end of the manuscript.

Additional citations to the relevant literature are provided within the main body of the manuscript for contextual consistency.

Structure of the thesis

Chapter 1 is an introduction to the theoretical background of first–principles calculations. We describe first DFT in Sec. 1.2 and then the Green's function (GF) formalism used to evaluate the dynamical electronic screening and provide quasiparticle corrections to the DFT band energies (Sec. 1.3). We subsequently describe the connection of the many–body response function thus defined to the optical observables, in particular the absorption coefficient, in Sec. 1.4. This is followed by the explanation of the BSE, the equation that permits to compute the excitons (Sec. 1.5), and finally by the formalism employed for the calculation of phonons and ep coupling (Sec. 1.6).

In Chapter 2 we start the analysis of the excitons in hBN, from monolayer to the bulk system. Particular attention is devoted to the excitonic energy levels in monolayer (Sec. 2.2),[39] and subsequently to the optical absorption of multilayer systems (Sec. 2.3): we study the exciton symmetry and related selection rules and find that the lowest–lying excitons are always localised on the surface layers, therefore providing an initial source of fine structure in the spectral features.[40] The second part of the Chapter (Sec. 2.4) concerns the study of the exciton dispersion curves in the bilayer and trilayer systems by solving the BSE at finite momentum. We study the character of the optical gap (direct in bilayer, indirect in trilayer), its robustness with respect to strain and compression of the lattice, and the effects of the NFE states on the exciton dispersion.

Chapter 3 is contains the main results of the thesis.[41] We start Chapter 3 with an introduction to the problem of luminescence in bulk hBN by discussing the experimental results (Sec. 3.1). We also perform a computational study of the ep coupling on the monolayer, including phonon–mediated renormalisation of the exciton energies and oscillator strengths (Sec. 3.2); in obtaining the latter results, we compare two different computational approaches. Then, we continue with an in–depth description of indirect absorption at the IP level in Sec. (3.3). Finally, we introduce our approach to obtain phonon–assisted sidebands from the static derivatives of the *excitonic* response function (Sec. 3.4). Here we also compute the ExcP coupling selection rules and subsequently we perform the calculations to obtain the final spectra for bulk hBN.

Chapter 4 deals with the microscopic, perturbative treatment of the ExcP coupling. We introduce it by analogy with simple single–particle models that can be solved exactly (Sec. 4.1). The full ExcP self–energy is described in Sec. 4.2.1 and the results of the `Yambo` implementation are shown, along with our final microscopic expression for phonon–assisted absorption including excitons.

The main text is complemented by a large body of Appendices (From A to J), variously discussing simplified models, tips for numerical simulations, code development, convergence parameters, and providing a reference for mathematical expressions and derivations.

Chapter 1

Basic theory and implementations

1.1 Introduction

In this Chapter we will cover the theoretical foundations and approaches that constitute the state of the art in Theoretical Solid–State Physics / Quantum Many–Body Physics / Materials Science[1] that are relevant for this thesis. The intent is to present a treatment that is both concise, organic and reasonably self–contained (nonetheless previous basic knowledge of solid–state physics is required[43]). In particular, we will establish the notation used throughout the thesis and put an emphasis on what is needed for computational research using *ab initio* simulations. By *ab initio* or first-principles simulation we refer to a computer simulation that takes as input only the atomic and electronic positions of a crystal or molecule, as well as the values of the various physical constants. It produces as output a quantity or observable that can be directly compared with experimental results without relying on adjustable parameters. However, the possibility to describe or not a certain physical property depends on the approximations that need to be introduced in order to make the computations feasible.

In the following, we will in general use atomic units (for example identifying frequencies and energies), but exceptions will be made for some expressions. Likewise, we will put the electronic charge e to unity as well, except for the initial expressions. We will mainly follow Refs. [3, 44–53] (additional references will be provided within the body of the Chapter). We start with the general matter Hamiltonian for a crystal:

$$\hat{\mathcal{H}} = \hat{T}_{ion} + \hat{V}_{ion}(\{\mathbf{R}\}) + \hat{T}_e + \hat{V}_{e-e}(\{\mathbf{r}\}) + \hat{V}_{e-ion}(\{\mathbf{r}\},\{\mathbf{R}\}). \tag{1.1}$$

Here, $\hat{T}_{ion}$ and $\hat{T}_e$ are the kinetic energy operators for ions and electrons, while $\hat{V}_{ion}$, $\hat{V}_e$ and $\hat{V}_{e-ion}$ represent the coulombic ion–ion, electron–electron and electron–ion interactions. The set of the ionic positions is labeled by $\mathbf{R}$ and the one of the electronic positions by $\mathbf{r}$. We already stipulate the following conventions and approximations: (i) relativistic corrections are neglected (because we will not be concerned with spin-orbit coupling in this work); (ii) the spin degrees of freedom are always implied (because we will not deal with spin-polarised systems); (iii) by 'ion' we mean the composite of an atomic nucleus and the core electrons localised around it (under the assumption that they can be treated on the same footing), and by 'electron' we mean a valence electron.

Here the ionic and electronic dynamics are coupled. However, we are not interested in the slow–moving ions, and we can partially decouple them from the electrons: in fact, the ions will only react to a time–averaged potential generated by the rapidly changing electronic density, while the electrons will mainly feel the presence

[1] The proper name to use for this field of Physics has changed during the course of the last century and is currently mostly left to the author, as seen in Ref. [42].

of an 'external' static potential determined by the set of ionic positions at equilibrium, $\{\mathbf{R}^0\}$. This is called adiabatic or Born–Oppenheimer (BO) approximation and leads us to the following Hamiltonian:

$$\hat{\mathcal{H}}_{BO} = \hat{\mathcal{H}}^0_{ion}(\{\mathbf{R}\}) + \hat{\mathcal{H}}_e(\{\mathbf{r}\}, (\{\mathbf{R}^0\})) + \hat{\mathcal{H}}_{e-ion}(\{\mathbf{r}\}, \{\mathbf{\Delta R}\}). \tag{1.2}$$

In the following, we will focus on the interacting electronic system with the ions fixed at their equilibrium positions ($\hat{\mathcal{H}}_e$). The term $\hat{\mathcal{H}}_{e-ion}$ describes the effect that lattice vibrations (i.e. displacements $\mathbf{\Delta R}$ of the ions from their equilibrim positions) have on the electronic system: this electron–lattice interaction will be reintroduced later as a (small) perturbation, and we will do the same with the electron–light interaction due to an external electro–magnetic field.

Now we have

$$\begin{aligned}\hat{\mathcal{H}}_e &= \hat{T}_e + \hat{V}_{e-e}(\{\mathbf{r}\}) + \hat{V}_{e-ion}(\{\mathbf{r}\}, \{\mathbf{R}_0\}) \\ &= \sum_i \frac{p_i^2}{2m_i} + \frac{1}{2}\sum_{i\neq j}^{N_e} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i,j}^{N_e,N_n} \frac{Z_j e^2}{|\mathbf{r}_i - \mathbf{R}_j^0|},\end{aligned} \tag{1.3}$$

where $\{\mathbf{p}\}$ is the set of electron momenta while $-e$, m and Z_j are the electron charge, electron mass and the effective atomic number of the j_{th} ion, respectively. N_e and N_n are the total numbers of electron and nuclei.

The first task in the study of a crystal is then to fully characterise its electronic properties by solving the related Schrödinger (eigenvalue) equation of the interacting many-electron system

$$\hat{\mathcal{H}}_e \Psi^{MB}(\mathbf{r}_1, \dots \mathbf{r}_{N_e}) = E\Psi^{MB}(\mathbf{r}_1, \dots \mathbf{r}_{N_e}), \tag{1.4}$$

where the superscript 'MB' means 'many-body'. The complicated eigenfunction Ψ^{MB} (which is a function of $\sim 10^{23}$ variables) contains in principle all the information about the electronic system (both ground state and excitations). However, the computational effort required to diagonalise the full $\hat{\mathcal{H}}_e$ prevents us from accessing this information directly, even for isolated systems. Therefore, we first turn to DFT in order to find a suitable way to simplify the problem: this will allow us to obtain the electronic ground state.

1.2 Density functional theory

DFT was first developed in 1964 by Hohenberg and Kohn[54] and then Kohn and Sham[55], and has since become one of the most popular tools of the computational material scientist, quantum chemist, and solid–state physicist. Its foundation is the possibility of expressing the total energy of a system as a functional $E[\rho_0]$ of the ground–state electron density $\rho_0(\mathbf{r})$.[2] The physical meaning of $\rho(\mathbf{r})$, which is denoted quantum–mechanically as the *diagonal one–body density matrix*, is connected to the probability of finding one particle at point $\mathbf{r}$. For the interacting electron system of Eq. (1.4), it is defined as the expectation value of the density operator $\hat{\delta}(\mathbf{r}) =$

[2] Technically, we may distinguish between the *particle* density $n(\mathbf{r})$ and the *charge* density $\rho(\mathbf{r}) = -en(\mathbf{r})$. We will not do so in the remainder of this manuscript.

$\sum_i^{N_e} \delta(\mathbf{r} - \mathbf{r}_i)$:

$$\begin{aligned} \rho(\mathbf{r}) &= \int \mathrm{d}^3 r_1 \dots \mathrm{d}^3 r_{N_e} \Psi^{*\,MB}(\mathbf{r}_1, \dots \mathbf{r}_{N_e}) \hat{\delta}(\mathbf{r}) \Psi^{MB}(\mathbf{r}_1, \dots \mathbf{r}_{N_e}) \\ &= N_e \int \mathrm{d}^3 r_2 \dots \mathrm{d}^3 r_{N_e} |\Psi^{MB}(\mathbf{r}, \mathbf{r}_2 \dots \mathbf{r}_{N_e})|^2. \end{aligned} \tag{1.5}$$

Here, most of the information contained in Ψ^{MB} has been integrated away, and $\rho(\mathbf{r})$ is just a function of three spatial variables. Nevertheless, this will be enough if we are interested in just the ground–state properties (i.e., expectation values) of the electronic system.

1.2.1 Theory

The first Hohenberg–Kohn theorem states that there is a one–to–one correspondence between ρ_0 and the external potential applied to the system. In the case of Eq. (1.3), the external potential is $\hat{V}_{e-ion}(\{\mathbf{r}\}, \{\mathbf{R}_0\})$, which is periodic according to the symmetries of the crystal lattice. The second Hohenberg–Kohn theorem guarantees that we can write the ground–state energy of the system, for any fixed $\hat{V}_{e-ion}$, as a unique functional of the ground–state electron density. Therefore we have

$$\begin{aligned} E_0 \equiv \langle \Psi_0^{MB} | \hat{\mathcal{H}}_e | \Psi_0^{MB} \rangle &= \langle \Psi_0^{MB} | \hat{T}_e + \hat{V}_{e-e} + \hat{V}_{e-ion} | \Psi_0^{MB} \rangle \\ &= \langle \Psi_0^{MB} | \hat{T}_e + \hat{V}_{e-e} | \Psi_0^{MB} \rangle + \int \mathrm{d}^3 r v_R(\mathbf{r}) \rho_0(\mathbf{r}) \\ &= T[\rho_0] + V_H[\rho_0] + V_{xc}[\rho_0] + \int \mathrm{d}^3 r v_R(\mathbf{r}) \rho_0(\mathbf{r}), \end{aligned} \tag{1.6}$$

where for $\hat{V}_{e-ion}$ we used the fact that a local operator can be written as a linear functional of the one–body density matrix and $v_R(\mathbf{r}) = -\sum_j e^2 Z_j |\mathbf{r} - \mathbf{R}_j^0|^{-1}$ is the external (i.e. ionic) Coulomb potential. The rest can be broken up in a kinetic part T, a classical charge–charge electrostatic interaction V_H (called Hartree functional), and an 'exchange–correlation' part that encodes the quantum properties of the fermion–fermion interaction. The Hartree functional is then

$$V_H[\rho_0] = \frac{1}{2} \int \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 \frac{\rho_0(\mathbf{r}_1) \rho_0(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}, \tag{1.7}$$

but we still need to find a way to evaluate T and V_{xc}. The crucial insight comes from the Kohn–Sham theorem: it is possible to write a single–particle Schrödinger equation with an effective, local potential $v_{eff}(\mathbf{r})$ in such a way that the ground–state electron density of this auxiliary system is the same as the one of the original, interacting system. The non–interacting Schrödinger equation, called Kohn–Sham equation, is

$$\left[-\frac{1}{2} \boldsymbol{\nabla}_{\mathbf{r}}^2 + v_{eff}(\mathbf{r}) \right] \varphi_i(\mathbf{r}) = \epsilon_i \varphi_i(\mathbf{r}), \tag{1.8}$$

and the electron density of the auxiliary system ρ_s, equal to ρ_0, can be expressed in terms of the eigenfunctions φ_i as

$$\rho_0(\mathbf{r}) \equiv \rho_s(\mathbf{r}) = \sum_i^{N_e} \varphi_i^*(\mathbf{r}) \varphi_i(\mathbf{r}). \tag{1.9}$$

We call ϵ_i the Kohn–Sham (KS) energies and φ_i the KS wave functions (the system contains N_e electrons which contribute to the ground–state electron density; note that Eqs. (1.8) may be solved also for $i > N_e$: these will represent unoccupied KS states). We can now explicitly compute the kinetic energy functional of the non–interacting system, T_s, and move the remaining unknown part into the exchange–correlation functional with the following replacement: $V_{xc} \to V_{xc} + T - T_s$. Now the energy of the system will be given by $E_0 = T_s + V_H + \int \mathrm{d}^3 r v(\mathbf{r})\rho_0(\mathbf{r}) + V_{xc}$. As for the effective potential, we have

$$\begin{aligned} v_{eff}(\mathbf{r}) &= v_R(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}), \\ v_H(\mathbf{r}) &= \int \mathrm{d}^3 r_1 \frac{\rho_0(\mathbf{r}_1)}{|\mathbf{r}_1 - \mathbf{r}|}, \\ v_{xc}(\mathbf{r}) &\equiv \frac{\delta V_{xc}[\rho_0]}{\delta \rho_0}. \end{aligned} \tag{1.10}$$

Here, v_H and v_{xc} are the Hartree and exchange–correlation potentials, respectively. So far, DFT is still exact (in the same sense that Eq. (1.4) is exact: an interacting electron system in the BOapproximation for the ions, with no relativistic effects or external electro-magnetic fields). In order to compute $E_0[\rho_0]$, we first solve the KS equations, obtain the φ_i, then build ρ_s. However, as it is clear from Eq. (1.10), the KS equations themselves depend on ρ_s, therefore the solution must be obtained with a self-consistent (sc) procedure: an initial guess $\rho_s^{(0)}$ is given (for example an electron density composed of pure atomic orbitals), Eqs. (1.8) are solved, a new $\rho_s^{(1)}$ is found and plugged back in Eqs. (1.8). The procedure ends when $E_0^{(n)} - E_0^{(n-1)}$ is smaller than the required accuracy.

In order to proceed with these calculations, however, we have to find suitable approximations for the unknown functional V_{xc}, with the reliability of DFT as a computational tool depending crucially on the type of approximation chosen. In this work we will use the local density approximation (LDA), the simplest and most physically motivated for a crystal, because we are only going to use DFT as a starting point: complicated many–body effects related to electronic correlations and screening will be included later via MBPT.[3] Within LDA, the non–local exchange–correlation functional is approximated as a local functional; this functional is given, at each point $\mathbf{r}$, by the exchange–correlation energy ϵ^{HEG} per particle of the homogeneous electron gas with density $\rho(\mathbf{r})$:

$$V_{xc}[\rho] \equiv \int \mathrm{d}^3 r \rho(\mathbf{r}) \epsilon^{HEG}[\rho(\mathbf{r})]. \tag{1.11}$$

1.2.2 Calculations in practice

In our case, the KS eigenstates φ_i describe non–interacting electrons moving in a periodic, infinite crystal lattice (i.e. subject to a periodic effective potential), therefore they can be identified with the Bloch functions from solid–state theory. They are characterised by their discrete translational invariance due to the periodic crystal lattice, and are therefore best treated in momentum space. For this reason, it is useful to expand the φ_i in the plane–wave basis $\{\mathrm{e}^{\mathrm{i}(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}/\sqrt{\Omega}\}$: here Ω is the volume of the reciprocal unit cell and $\mathbf{K} = \mathbf{k} + \mathbf{G}$ is the momentum of the electron, with $\mathbf{k} \in \mathrm{BZ}$ being the crystal momentum and $\mathbf{G}$ a reciprocal lattice vector. Now, for the crystal

[3] It is important to note that both at the DFT and at the G_0W_0 levels (for the latter see Sec. 1.3.3), we might need to replace LDA with other, more accurate approximations depending on the system and physical properties under investigation.

quantum numbers we have $i \to nk$, with k labeling the crystal momenta and n being the band index (we will use v for valence bands and c for conduction ones), leading to

$$\varphi_{nk}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{nk}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{\mathbf{G}} c_{nk}(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}} \tag{1.12}$$

where $u_{nk}(\mathbf{r}) = u_{nk}(\mathbf{r}+\boldsymbol{\tau})$ ($\boldsymbol{\tau}$ being a lattice translation vector) is a lattice–periodic function and, for the electron density of Eq. (1.9), to

$$\rho_0(\mathbf{r}) = \sum_{vk} |u_{vk}(\mathbf{r})|^2. \tag{1.13}$$

It is clear from Eqs. (1.12) and (1.13) that the solution of the KS equations in a DFT computer simulation yields the $c_{nk}(\mathbf{G})$ coefficients. The infinite sums over k and $\mathbf{G}$ pose a problem of numerical convergence: a cutoff to both sums must be chosen in such a way that the final $E_0[\rho_0]$ does not depend on it. In the case of the sum over k–vectors, this is achieved by performing the sum over a (usually regular) mesh of k–points spanning the entire BZ, whose density is then increased until convergence is achieved.[4] In the following, we will switch back and forth from integral to discrete sum when discussing an integration over the BZ: $\int_{BZ} d^3k/\Omega_{BZ} F(k) \leftrightarrow N_k^{-1} \sum_k^{N_k} F_k$, where N_k is the number of points in the k–mesh. In the case of the sum over G–vectors, a cutoff is usually chosen in the form of an energy value $\hbar^2|\mathbf{k}+\mathbf{G}|^2/2m$. When convergence is achieved, the self–consistent solution of the KS equations will yield the ground–state electron density and total energy of the system.[5] At this point, several additional non–self–consistent calculations may be performed for any arbitrary nk indices, in order to obtain quantities like the band structure along high–symmetry directions, density of states, real–space representations of the Bloch functions *et cetera*. All the DFT calculations presented in this thesis have been performed with the QUANTUM ESPRESSO (QE) simulation package.[56]

1.2.3 Shortcomings of DFT

Historically, DFT has been very successful in predicting a wide range of material properties while remaining computationally cheap. In spite of that, the access to ground–state properties only and the reliance on approximated treatments of many–body effects entail some crucial shortcomings. Below we list the three main flaws of DFT and how we have dealt with them in this thesis.

(i) *Test charge outside the matter system.* The effective potential $v_{eff}(\mathbf{r})$ is responsible for confining the electrons inside the crystal; however, it decays exponentially in the vacuum region outside the crystal and does not represent correctly its interaction with an external charge. This unphysical behavior (there should be a long–range decay) is due to the approximations of V_{xc}. For our purposes this is an advantage because in systems with lowered periodicity (e.g. two–dimensional materials), we will still be able to consider fully periodic simulation cells; this is achieved by stacking repeated copies of the system with enough vacuum space separation between them such that their spurious mutual interaction is rendered negligible. We will see that this is not the case for many–body response functions beyond DFT.

[4] In most cases the symmetry operations under which the system is invariant, i.e., the elements of the point group of the crystal plus time reversal, are employed in order to dramatically reduce the spanned portion of the BZ. The smallest possible volume thus identified in reciprocal space is called irreducible Brillouin zone (IBZ).

[5] The forces acting on the ions due to the electronic potential may also be computed, and it is then possible to optimise the geometry of the crystal lattice via structural relaxation.

(ii) *Dispersion forces (i.e. Van der Waals interactions).* These interactions are not captured by standard DFT methods. Although corrective approximations are available, we will not deal with this problem in this work. Since we will be investigating layered materials where Van der Waals interactions are important for the bonding of one layer to another, we will fix the interlayer distances to the known experimental values instead of relying on DFT structural relaxations.[6]

(iii) *Band gap energy and single–particle states.* When talking about the band structure of a material in the context of DFT, what we mean is the k-dispersion of the KS eigenvalues, obtained solving Eqs. (1.8) for both occupied and unoccupied states (the valence and conduction bands, respectively). However, these eigenvalues have no rigorous physical meaning: in principle there is no guarantee that an interacting condensed matter system could be effectively described within a single–particle framework, and the whole solid–state concept of a band structure (i.e. non–interacting electrons in a periodic potential) could very well be meaningless. In practice, this description works well for all materials where electronic correlation could be treated as a perturbation (e.g., it may compare reasonably well to experimental ARPES measurements.[57]) In this case, a more accurate treatment of the electron–electron interaction beyond DFT leads to the concept of *quasiparticles*: the system is still described in a single–particle framework, but the band energies are shifted with respect to the KS eigenvalues, and furthermore acquire a finite broadening which indicates that the single–particle states now have a finite lifetime (i.e. they are not true eigenstates of the condensed matter system). From these considerations, it follows that DFT is not able to correctly describe the band gaps of insulating and semiconducting systems. In fact, the addition of one electron to the interacting system ($N_e \to N_e + 1$), now occupying the bottom of the conduction band, induces a variation of the electron density with respect to the ground–state one, and consequently there is a variation in the KS eigenvalues themselves. Therefore we have the following relation between the band gap as defined in DFT–KS and the "true" band gap that may be seen in photoemission / inverse photoemission experiments:[1]

$$E_{gap}^{KS} \equiv \epsilon_{H+1}(N_e) - \epsilon_H(N_e) < \epsilon_H(N_e + 1) - \epsilon_H(N_e) \equiv E_{gap}^{TRUE}, \tag{1.14}$$

where the subscript H labels the highest occupied KS state. Section 1.3.3 is concerned with the treatment of the band gap underestimation problem via the quasiparticle description of the electronic structure.

In addition to the previous point, since our goal is to investigate theoretical spectroscopy, we have to consider that an optical absorption experiment creates a neutral excitation in the form of an electron–hole pair (a valence electron is given enough energy to jump to a conduction state, with the overall electron number remaining equal to N_e). Such an experiment would result in yet another gap, the *optical gap*, which might be very different from the DFT–KS or quasiparticle band gaps when the interaction between the excited, negatively charged electron and the positively charged hole left in the valence band is poorly screened. This leads to the concept of *excitons*, bound electron–hole states, that will be the main focus of this thesis. The *ab initio* theory of excitons is laid out in Section 1.5.

[6] The LDA underestimates the interlayer distance. In many cases, the magnitude of this underestimation is numerically very close to the neglected Van der Waals contribution.

1.3 Many–body perturbation theory

All the results presented in this thesis are obtained using a MBPT framework, which we outline in this Section.

1.3.1 Single–particle Green's function

First of all, we switch to a second–quantisation description of the condensed matter system. We introduce the fermionic creation and annihilation operators $\hat{a}_i$ and $\hat{a}_i^\dagger$, with the anticommutation properties $\{\hat{a}_i, \hat{a}_j^\dagger\} = \delta_{ij}$ and $\{\hat{a}_i, \hat{a}_j\} = \{\hat{a}_i^\dagger, \hat{a}_j^\dagger\} = 0$. These operators act on the fermionic ground state by annihilating and creating a fermion, respectively. In a crystal, by fermionic ground state we mean that all the bands lying below the Fermi energy are occupied, all the ones lying above are empty. By expanding these operators into the position basis $|\mathbf{r}\rangle$ we obtain the field operators (here written in the Heisenberg picture):

$$\begin{aligned} \hat{\Psi}(\mathbf{r}, t) &= \sum_i \varphi_i(\mathbf{r}) \mathrm{e}^{-\mathrm{i}\hat{H}_0 t} \hat{a}_i, \\ \hat{\Psi}^\dagger(\mathbf{r}, t) &= \sum_i \varphi_i^*(\mathbf{r}) \mathrm{e}^{\mathrm{i}\hat{H}_0 t} \hat{a}_i^\dagger. \end{aligned} \tag{1.15}$$

The single–particle wave functions $\varphi_i(\mathbf{r}) = \langle \mathbf{r}|i\rangle$ and energies ($\hat{H}_0 |i\rangle = \epsilon_i |i\rangle$) appear in these expressions and we have $[\hat{\Psi}(\mathbf{r}_1, t), \hat{\Psi}^\dagger(\mathbf{r}_2, t)]_- = \delta(\mathbf{r}_1 - \mathbf{r}_2)$. The connection with the previous Section can be made at the level of the one-body density matrix, Eq. (1.5), which can be written in second quantisation as

$$\rho(\mathbf{r}) = \langle \Psi_0^{MB}| \hat{\Psi}^\dagger(\mathbf{r}, t) \hat{\Psi}(\mathbf{r}, t) |\Psi_0^{MB}\rangle . \tag{1.16}$$

($|\Psi_0^{MB}\rangle$ is the ground state of the many–body system). Now this expression can be generalised by considering a version which is non–local in both space and time coordinates:

$$\begin{aligned} \mathrm{i}\mathcal{G}(\mathbf{r}, t; \mathbf{r}', t') &= \langle \Psi_0^{MB}| \hat{\mathcal{T}} \left[\hat{\Psi}(\mathbf{r}, t) \hat{\Psi}^\dagger(\mathbf{r}', t') \right] |\Psi_0^{MB}\rangle \\ &= \langle \Psi_0^{MB}| \hat{\Psi}(\mathbf{r}, t) \hat{\Psi}^\dagger(\mathbf{r}', t') |\Psi_0^{MB}\rangle \, \theta(t - t') \\ &- \langle \Psi_0^{MB}| \hat{\Psi}^\dagger(\mathbf{r}', t') \hat{\Psi}(\mathbf{r}, t) |\Psi_0^{MB}\rangle \, \theta(t' - t). \end{aligned} \tag{1.17}$$

Here $\hat{\mathcal{T}}$ is the time–ordering operator. The first term of the difference is now connected to the probability that a particle being initially created at space–time coordinates $(\mathbf{r}', t')$ (i.e. a conduction electron) will propagate to $(\mathbf{r}, t)$ before being annihilated. The second term instead starts with a particle initially annihilated at $(\mathbf{r}, t)$, therefore a valence hole is created; it is connected to the probability that it will propagate to $(\mathbf{r}', t')$ before being annihilated. Therefore, the physical meaning of $\mathcal{G}(\mathbf{r}, t; \mathbf{r}', t')$ is that of a single–fermion *propagator*.

The mathematical meaning is instead that of the *Green's function* (GF) of the time–dependent Schrödinger equation for the electronic problem, with the following relations being formally valid:

$$\begin{aligned} \left[\mathrm{i}\partial_t - \hat{\mathcal{H}}_e\right] \mathcal{G}(\mathbf{r}, t; \mathbf{r}', t') &= \delta(\mathbf{r} - \mathbf{r}')\delta(t - t'), \\ \mathcal{G}(\mathbf{r}, \mathbf{r}', \omega) &= \left[\omega - \hat{\mathcal{H}}_e\right]^{-1}. \end{aligned} \tag{1.18}$$

The second line is justified by the fact that if the Hamiltonian does not depend on time, then $G(\mathbf{r},t;\mathbf{r}',t') = G(\mathbf{r},\mathbf{r}';t-t')$: this allows the Fourier transform into the frequency domain. Furthermore, if the system is also translationally invariant in space we have $\mathcal{G}(\mathbf{r},\mathbf{r}';t-t') = \mathcal{G}(\mathbf{r}-\mathbf{r}',t-t')$, which allows the Fourier transform into momentum space, so that finally $\mathcal{G}(\mathbf{r}-\mathbf{r}',t-t') \to \mathcal{G}(\mathbf{K},\omega)$. Note that now the electron density from Eqs. (1.5) and (1.16) can be written in terms of the GF as[7]

$$\rho(\mathbf{r}) = -\lim_{t'\to t} \mathrm{i}\mathcal{G}(\mathbf{r},t;\mathbf{r},t') \tag{1.19}$$

Indeed, using the GF we can write an exact expression for the ground–state total energy[8] equivalent to Eq. (1.6),

$$E_0 = -\mathrm{i}\frac{V}{2}\lim_{\eta\to 0^+}\int \frac{\mathrm{d}^3k\mathrm{d}\omega}{(2\pi)^4}\left[\hbar\omega + \frac{\hbar^2k^2}{2m} + v_R(\mathbf{K})\right]\mathrm{e}^{\mathrm{i}\omega\eta}\mathcal{G}(\mathbf{K},\omega), \tag{1.20}$$

which is known as the Migdal–Galitzki formula (V is the volume of the unit cell). The first two terms come from the electron–electron interaction and the kinetic energy, respectively, with the third one accounting for the external potential.

The single–particle GF permits us to go beyond ground–state expectation values and access directly the true single–particle excitation energies (as seen in Eq. (1.14)). We can write $\mathcal{G}(\mathbf{k},\omega)$ in the so–called Lehmann representation to make the excited states appear explicitly:

$$\begin{aligned}\mathcal{G}(\mathbf{k},\omega) = V\sum_n &\frac{\langle\Psi_0^{MB}|\,\hat{\Psi}(\mathbf{0})\,|\Psi_{nk}^{MB};N_e+1\rangle\,\langle\Psi_{nk}^{MB};N_e+1|\,\hat{\Psi}^\dagger(\mathbf{0})\,|\Psi_0^{MB}\rangle}{\omega - [E_n(N_e+1) - E_0(N_e)] + \mathrm{i}\eta} \\ + V\sum_n &\frac{\langle\Psi_0^{MB}|\,\hat{\Psi}(\mathbf{0})\,|\Psi_{n-k}^{MB};N_e-1\rangle\,\langle\Psi_{n-k}^{MB};N_e-1|\,\hat{\Psi}^\dagger(\mathbf{0})\,|\Psi_0^{MB}\rangle}{\omega - [E_0(N_e) - E_n(N_e-1)] - \mathrm{i}\eta}.\end{aligned} \tag{1.21}$$

Here $|\Psi_{n\pm k}^{MB};N_e\pm 1\rangle$ is the n–th many–body excited state for a system of $N_e\pm 1$ particles, where the created (annihilated) particle has momentum $\pm\mathbf{k}$; $E_n(N)$ is the total energy of the n–th excited state of the N–particle system. It is clear that the total energy differences in the denominators describe the change in the system after the creation of one conduction electron (first term) and of one valence hole (second term). We can express these differences in terms of single–particle energies ϵ_n as

$$\begin{aligned}E_n(N_e+1) - E_0(N_e) &= [E_n(N_e+1) - E_0(N_e+1)] - [E_0(N_e+1) - E_0(N_e)] = \epsilon_n(N_e+1) - \mu \\ E_0(N_e) - E_n(N_e-1) &= [E_n(N_e) - E_0(N_e-1)] + [E_0(N_e+1) - E_n(N_e-1)] = \mu - \epsilon_n(N_e-1)\end{aligned} \tag{1.22}$$

(the chemical potentia μ becomes the Fermi energy E_F at zero temperature). Therefore, the single–particle GF displays poles at the excitation energies of the system.

However, in order to have access to these quantities, the many–body problem has to be solved exactly (i.e. the quantum electron–electron interaction has to be treated exactly), which is not possible. Therefore, we will instead consider the GF of a non–interacting electron system, G^0, then introduce the interaction as a perturbation, and ultimately find the new, approximate interacting GF denoted G. Our G^0 will not

[7] The equal–time GF is ill–defined, therefore we introduce an infinitesimal value $\eta \equiv t'-t$ in such a way that $\lim_{t'\to t} \equiv \lim_{\eta\to 0^+}$.

[8] This is also true for the ground–state expectation value of any single–particle operator.

be the free–electron GF, but rather the one relative to the KS equations (i.e. non–interacting electrons moving under the influence of an effective potential, see Eq. (1.8)): in this way we will be able to use the DFT–computed eigenvalues and wave functions, which represent a much better starting point for perturbation theory than the free electrons. Therefore G^0 will be diagonal once expanded in the KS basis of Bloch states $|nk\rangle$:

$$\begin{aligned} G^0(\mathbf{r},\mathbf{r}';t-t') &= \sum_{nk,\,n'k'} \langle \mathbf{r}|nk\rangle \, \langle nk|\, G^0(t-t')\, |n'k'\rangle \, \langle n'k'|\mathbf{r}'\rangle \\ &= \sum_{nk} \langle \mathbf{r}|nk\rangle \, G^0_{nk}(t-t') \, \langle nk|\mathbf{r}'\rangle \\ &= \sum_{nk} \varphi_{nk}(\mathbf{r})\varphi^*_{nk}(\mathbf{r}')G^0_{nk}(t-t'). \end{aligned} \tag{1.23}$$

We refer to the function $G^0_{nk}(t-t')$ as the Green's function for a particle of band n and (crystal) momentum $\mathbf{k}$. Since so far we are dealing with a non–interacting system, $G^0_{nk}(\omega)$ takes the simple form listed below:

$$G^0_{nk}(\omega) = \frac{1-f_{nk}}{\omega-\epsilon_{nk}+\mathrm{i}\eta} + \frac{f_{nk}}{\omega-\epsilon_{nk}-\mathrm{i}\eta}. \tag{1.24}$$

For a gapped material at zero temperature, the occupation factors f_{nk} become $f_{vk}=1$ and $f_{ck}=0$: therefore it might be convenient to write G^0_{nk}, in both time and frequency domains, in the following way:

$$\begin{aligned} G^0_{ck}(t_{21}) &= -\mathrm{i}\theta(t_2-t_1)\,\mathrm{e}^{\mathrm{i}\epsilon_{ck}(t_1-t_2)} = \quad t_1 \xrightarrow{\qquad ck \qquad} t_2 \\ G^0_{ck}(\omega) &= \frac{1}{\omega-\epsilon_{ck}+\mathrm{i}\eta} \\ G^0_{vk}(t_{21}) &= \mathrm{i}\theta(t_1-t_2)\,\mathrm{e}^{\mathrm{i}\epsilon_{vk}(t_1-t_2)} = \quad t_1 \xrightarrow{\qquad vk \qquad} t_2 \\ G^0_{vk}(\omega) &= \frac{1}{\omega-\epsilon_{vk}-\mathrm{i}\eta}. \end{aligned} \tag{1.25}$$

Note that in these expressions we introduced a graphic representation for the propagator of either a conduction electron or a valence hole.

We can rewrite the exact GF in a way that makes G_0 explicit and allows for a direct comparison with the DFT case (here we group the space–time variables together as $x \to (\mathbf{r},t)$):

$$\begin{aligned} \mathcal{G}(x,x') &= G^0(x,x') + \int \mathrm{d}^4x_1\mathrm{d}^4x_2 G^0(x,x_1)\Sigma^{xc}(x_1,x_2)\mathcal{G}(x_2,x) \\ \mathcal{G}(\mathbf{K},\omega) &= G^0(\mathbf{K},\omega) + G^0(\mathbf{K},\omega)\Sigma^{xc}(\mathbf{K},\omega)\mathcal{G}(\mathbf{K},\omega). \end{aligned} \tag{1.26}$$

This is Dyson's equation. We see now that the non–interacting Green's function is dressed by a complicated effective potential $\Sigma^{xc}(x,x')$, called the self–energy. The self–energy, which is in general non–local in both space and time, represents an improvement with respect to the DFT treatment, in which $\Sigma^{xc}_{DFT}(x,x') = v_{xc}[\rho_0](\mathbf{r})\delta(t-t')$. The non–locality in space is lost in DFT because of the need to express the exchange–correlation functional in terms of the local particle density $\rho(\mathbf{r})$. The non–locality in time, as we shall see, refers to the time needed by the system to respond to the presence of an extra charge, which leads to a frequency–dependent (i.e., *dynamical*) screening of the electron–electron interaction. This can be captured

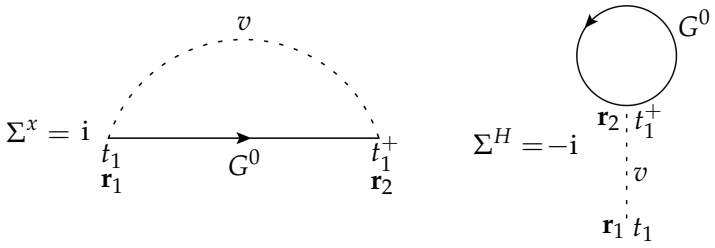


FIGURE 1.1: Schematic representation of the bare exchange (left) and Hartree (right) self–energy contributions, Eqs. (1.27) and (1.28) respectively.

efficiently by working with the self–energy. Note that, since in our definition the KS energies entering G^0 already contain an approximated correction due to v_{xc}, in order to avoid double–counting and maintain consistency it is always implied that this quantity is subtracted from the self–energy operator, i.e. $\Sigma^{xc}(x_1, x_2) \to \Sigma^{xc}(x_1, x_2) - v_{xc}(x_1, x_2)\delta(x_1, x_2)$. In order to keep the notation light we will write down this difference explicitly only when required by the context.

Thanks to MBPT, we can now try to expand $\Sigma^{xc}(x_1, x_2)$ in powers of the interaction (i.e. $v(x_{21}) = \delta(t_1 - t_2)v(\mathbf{r}_{21}) = \delta(t_1 - t_2)|\mathbf{r}_1 - \mathbf{r}_2|^{-1}$) to obtain a suitable approximation of the exact function. To first order, the only contribution is

$$\Sigma^x(x_1, x_2) = \mathrm{i}v(\mathbf{r}_{21}) \lim_{t_2 \to t_1^+} G^0(\mathbf{r}_1, \mathbf{r}_2; t_1 - t_2). \tag{1.27}$$

This is not the full Σ^{xc} yet, because the interaction is still static. This contribution, called Σ^x, is the exchange or Fock term, i.e. the quantum part of the electrostatic Coulomb interaction that arises due to the antisymmetric property of the fermionic many–body wave functions. Since the Coulomb interaction is not time–dependent, it means that at this level of approximation the presence of a perturbation (the addition or removal of an electron to the system) does not really cause a modification of the energy levels (although total energies and therefore band gaps in the DFT sense are affected). Predictably, this approximation is not sufficient in solids because the dynamical electronic screening plays an important role: the electron density of the system collectively reacts, polarises, oscillates, *et cetera* in response to an external perturbation, and the frequencies associated to this response in turn modulate the effects of the perturbation, typically reducing the strength of the interaction. This is why a simple Hartree–Fock approach usually fails for solids, while DFT works better: it adds an approximation to the electronic correlations responsible for screening effects. Note that the classical or Hartree part of the electrostatic interaction (v_H in Eq. (1.10), already contained in G^0) can also be written in terms of the GF using Eq. (1.19):

$$\Sigma^H(x_1, x_2) = -\mathrm{i}v(\mathbf{r}_{21}) \lim_{t_2 \to t_1^+} G^0(\mathbf{r}_2, \mathbf{r}_2; t_1 - t_2) \tag{1.28}$$

A schematic representation of these two terms is depicted in Fig. 1.1.

In order to obtain a more meaningful description of the condensed matter sytem we then have to find a way to properly describe the electronic screening.

1.3.2 Screening functions

Let us consider a time–dependent perturbation to our electronic system in the form of an external potential $\phi_{ext}(\mathbf{r}, t)$ due to external charges added to the system. We

already know that the additional term perturbing the Hamiltonian can be written as a functional of the electron density, i.e.

$$\hat{\mathcal{H}}'(t) = \int \mathrm{d}^3 r \hat{\rho}(r) \phi_{ext}(\mathbf{r}, t). \tag{1.29}$$

This *longitudinal* perturbation (the external electric field due to the presence of additional charge being parallel to the propagation direction) induces a variation $\delta\rho$ in the electron density. In turn, $\delta\rho$ generates an induced potential ϕ_{ind} in such a way that the total potential acting on the system is now $\phi_{tot} = \phi_{ext} + \phi_{ind}$. We wish to determine $\delta\rho$ up to linear order, since we consider ϕ_{ext} to be a weak perturbation.

In linear response theory, we can employ Kubo's formula for the first–order variation of any observable $A(t)$ – i.e., the expectation value of the corresponding operator $\hat{A}(t)$ – under the action of $\hat{\mathcal{H}}'(t)$:

$$\begin{aligned} \delta \left\langle \hat{A}(t) \right\rangle &= \int \mathrm{d}t' C^R_{A\mathcal{H}'}(t, t'), \\ \delta \left\langle \hat{A}(\omega) \right\rangle &= C^R_{A\mathcal{H}'}(\omega). \end{aligned} \tag{1.30}$$

$C^R_{A\mathcal{H}'}(t, t')$ contains the *response function*, which mathematically assumes the form of a retarded ($t > t'$) correlation function between $\hat{A}$ at time t and $\hat{\mathcal{H}}'$ time t':[9]

$$C^R_{A\mathcal{H}'}(t, t') = -\mathrm{i}\theta(t - t') \left\langle \left[\hat{A}(t), \hat{\mathcal{H}}'(t') \right] \right\rangle_0 \tag{1.31}$$

(the Fourier transform being possible if $C^R_{A\mathcal{H}'}(t, t') = C^R_{A\mathcal{H}'}(t - t')$). The subscript 0 indicates that the expectation value is relative to the ground–state of the unperturbed electronic system $\hat{\mathcal{H}}_e$ (i.e., at linear order we are not dealing with out–of–equilibrium quantities).

In the case of the electron density induced by Eq. (1.29) we obtain

$$\delta\rho(x) = \delta \left\langle \hat{\rho}(x) \right\rangle = \int \mathrm{d}t' C^R_{\rho\mathcal{H}'}(x, x') = \int \mathrm{d}^4 x' C^R_{\rho\rho}(x, x') \phi_{ext}(x'), \tag{1.32}$$

with

$$\chi^R(x, x') \equiv C^R_{\rho\rho}(x, x') = -\mathrm{i}\theta(t - t') \left\langle \left[\hat{\rho}'(x), \hat{\rho}'(x') \right] \right\rangle_0, \tag{1.33}$$

being the *retarded*, longitudinal charge–charge response function (with $\hat{\rho}' = \hat{\rho} - \langle \hat{\rho} \rangle_0$ in this case). The perturbation acts at an earlier time t'; we measure its effects on the system at a later time t. It is worth recalling here that MBPT gives us a mathematical tool for the evaluation of *time–ordered* correlation functions, i.e. defined with the same time structure as the single–particle GF in Eq. (1.17). Therefore, we define a time–ordered version of the charge–charge response as

$$\mathrm{i}\chi(x, x') = \left\langle \mathcal{T} \left[\hat{\rho}'(x) \hat{\rho}'(x') \right] \right\rangle_0. \tag{1.34}$$

When the time–ordered (T) response has been computed, we can easily extract its retarded (R) counterpart: for example in the frequency domain the following identity for correlation functions holds:

$$C^R_{AB}(\mathbf{r}_1, \mathbf{r}_2; \omega) = \theta(\omega) C^T_{AB}(\mathbf{r}_1, \mathbf{r}_2; \omega) + \theta(-\omega) C^{T*}_{AB}(\mathbf{r}_1, \mathbf{r}_2; \omega). \tag{1.35}$$

From Eq. (1.16) we know that $\hat{\rho}(x) = \hat{\Psi}^\dagger(x)\hat{\Psi}(x)$, and let us recall the relationship

[9] Here we are implying that the perturbation is switched on at an indefinite earlier time $t_i \to -\infty$ and that the response decays with increasing $t - t'$, as it physically makes sense.

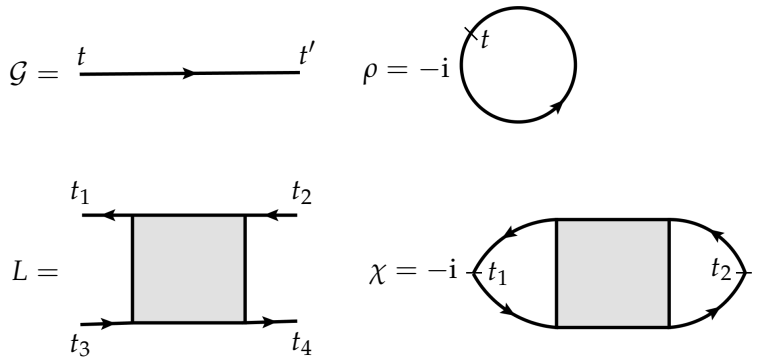


FIGURE 1.2: Graphical schemes of the electronic many–body functions. Top: single–particle GF $\mathcal{G}$ and particle density ρ from Eq. (1.16). Bottom: two–particle correlation function L and susceptibility χ from Eq. (1.37).

that links ρ and $\mathcal{G}$: we can find that an analogous relationship exists between χ and the *two–particle correlation function*. In particular, we introduce the two–particle Green's function G_2 as follows:

$$\begin{aligned} G_2(x_1, x_2, x_3, x_4) &\equiv (-\mathrm{i})^2 \left\langle \mathcal{T}\left[\hat{\psi}(x_1)\hat{\psi}(x_2)\hat{\psi}^\dagger(x_3)\hat{\psi}^\dagger(x_4)\right]\right\rangle_0, \\ L(x_1, x_2, x_3, x_4) &\equiv -G_2(x_1, x_2, x_3, x_4) + G(x_1, x_3)G(x_2, x_4). \end{aligned} \tag{1.36}$$

G_2 describes the joint propagation of two (interacting) particles in the condensed matter system, which are created at x_3, x_4 and annihilated at x_1, x_2 respectively. The two–particle correlation function L is the same except for the removal of the uncorrelated part GG. We additionally constrain the time orderings in L so that it describes the propagation of electron–hole pairs, not just any two particles. Therefore, at both the earlier and later pair of times we will always have one electron and one hole. If we further require that the electron and the hole are created and annihilated together (i.e. the processes of electronic excitation or recombination), we have to set $x_3 = x_1^+$ and $x_2 = x_4^+$, which leads us to the charge–charge response function. In short, we have

$$\mathrm{i}\chi(x, x') = L(x, x'^+, x, x'^+); \tag{1.37}$$

The new induced electron density will give rise to the induced potential

$$\phi_{ind}(x) = \int \mathrm{d}^4x' v(x - x')\delta\rho(x') = \int \mathrm{d}^4x'\mathrm{d}^4x'' v(x - x')\chi^R(x', x'')\phi_{ext}(x''). \tag{1.38}$$

Now the total potential $\phi_{tot} = \phi_{ext} + \phi_{ind}$ can be expressed in terms of the external potential if we just define an additional function:

$$\varepsilon^{-1}(x, x') = \delta(x - x') + \int \mathrm{d}^4x'' v(x - x'')\chi^R(x'', x') = \frac{\delta\phi_{tot}(x)}{\delta\phi_{ext}(x')}. \tag{1.39}$$

This is the inverse microscopic (longitudinal) dielectric function, and its physical meaning is clarified by the last term in the equalities: it expresses the rate of variation of the total potential with respect to the external one, i.e. it is a measure of the screening of the external interaction by the induced potential. Now we can finally write

$$\phi_{tot}(x) = \int \mathrm{d}^4x' \varepsilon^{-1}(x, x')\phi_{ext}(x'). \tag{1.40}$$

This makes clear that in order to describe the electronic screening of an external potential we need the dielectric function, which in turn is determined by the response function χ.

As mentioned above, the way in which an electron density responds to an external field can be perturbatively described in terms of neutral excitations of the system. We are working in a single–particle (KS) framework and the simplest form of such an excitation involves the formation of an electron–hole (eh) pair, whose simultaneous propagation is described by a two–particle correlation function. At the zero order of approximation, this excitation will be described by a non–interacting eh pair, created at time t_1 and destroyed at time t_2:

$$\chi^0(x_1, x_2) = -\mathrm{i}G^0(x_{21})G^0(x_{12}), \tag{1.41}$$

while in Fourier space we can write

$$\chi^0(\mathbf{Q}, \mathbf{Q}', t_{21}) \equiv \chi^0_{GG'}(\mathbf{q}, t_{21}) = \int \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 \mathrm{e}^{-\mathrm{i}(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}_1 + \mathrm{i}(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}_2} \chi^0(\mathbf{r}_1, \mathbf{r}_2, t_{21}). \tag{1.42}$$

Here, we have $\mathbf{Q} = \mathbf{q} + \mathbf{G}$ and the response function is written as a tensor in the reciprocal lattice vectors. It is a function of the energy and crystal momentum $\mathbf{q} = \mathbf{k} - \mathbf{k}' + \mathbf{G}_0 \in BZ$ transferred to the eh pair during the excitation process. This process represents the excitation of an electron from KS energy ϵ_{vk-q} to KS energy ϵ_{ck}. The off–diagonal terms in $\chi^0_{GG'}(\mathbf{q})$ represent the contributions to the short–range response coming from the microscopic spatial inohomogeneities of the system (the so–called local–field effects, see Sec. 1.4). In other words, they are present owing to the discreteness of the lattice structure.

The response function $\chi^0(\mathbf{r}, \mathbf{r}')$ takes the following explicit form when expanded in a basis of eh transitions (which we will call *transition basis* henceforth):

$$\chi^0(\mathbf{r}, \mathbf{r}'; t_{21}) = \sum_{cvk} \varphi_{ck}(\mathbf{r})\varphi^*_{vk-q}(\mathbf{r})\varphi_{vk-q}(\mathbf{r}')\varphi^*_{ck}(\mathbf{r}')\chi^0_{cvkq}(t_{21}). \tag{1.43}$$

In order to simplify the notation, we change the indices of the transition basis as $cvkq \to \mathcal{K}q$, with $\mathcal{K}$ labeling a single vck vertical transition between valence and conduction bands. Now, taking Eq. (1.41) for $\chi^0_{\mathcal{K}q}$ in the new transition basis, in which it is diagonal, gives the simple forms

$$\begin{aligned}
\chi^0_{\mathcal{K}q}(t_{21}) &= -\mathrm{i}G^0_{vk-q}(t_{21})G^0_{ck}(t_{12}) = \\
&= -\mathrm{i}\theta(t_2 - t_1)\mathrm{e}^{\mathrm{i}\Delta\epsilon_{\mathcal{K}q}(t_1 - t_2)} + \mathrm{i}\theta(t_1 - t_2)\mathrm{e}^{\mathrm{i}\Delta\epsilon_{\mathcal{K}q}(t_1 - t_2)}, \\
\chi^0_{\mathcal{K}q}(\omega) &= \frac{1}{2\pi}\int \mathrm{d}\omega' G^0_{vk-q}(\omega')G^0_{ck}(\omega + \omega') \\
&= \frac{1}{\omega - \Delta\epsilon_{\mathcal{K}q} + \mathrm{i}\eta} - \frac{1}{\omega + \Delta\epsilon_{\mathcal{K}q} - \mathrm{i}\eta},
\end{aligned} \tag{1.44}$$

which are formally analogous to their single–particle counterparts, Eqs. (1.25), and allow for a similar physical intepretation, this time in terms of a *polarisation* propagator with a resonant ($\mathrm{i}\eta$) and antiresonant ($-\mathrm{i}\eta$) term. We have further defined $\Delta\epsilon_{\mathcal{K}q} \equiv \epsilon_{ck} - \epsilon_{vk-q}$ as the poles of the polarisation function and from now on we will drop the space–time representation and focus on the momentum–frequency domain. Therefore, we insert Eq. (1.43) into the Fourier transform Eq. (1.42), finally obtaining

$$\chi^0_{GG'}(\mathbf{q}, \omega) = \sum_{\mathcal{K}} \varrho^*_{\mathcal{K}q}(\mathbf{G})\varrho_{\mathcal{K}q}(\mathbf{G}')\chi^0_{\mathcal{K}q}(\omega). \tag{1.45}$$

In the above equation we have defined a crucial integral to be numerically evaluated by *ab initio* codes: the screening matrix element

$$\varrho_{\mathcal{K}q}(\mathbf{G}) = \langle ck| \, \mathrm{e}^{\mathrm{i}(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} \, |vk-q\rangle \,, \tag{1.46}$$

which represents the coupling of the electrons with a longitudinal electric field carryng both energy and momentum.

From Fig. 1.3 we see how χ^0 gives a first–order correction to the bare Coulomb interaction v. It is actually possible to improve on this and consider an infinite sum of non–interacting eh pairs at all orders in v: this is a geometric series and therefore exactly summable, leading to a huge improvement in the description of the screening. This is called Random Phase Approximation (RPA) and it represents the most commonly used screening method when electron–hole interaction is neglected. We first obtain the Dyson's equation

$$\chi^{RPA}_{GG'}(\mathbf{q},\omega) = \chi^0_{GG'}(\mathbf{q},\omega) + \sum_{G_1} \chi^0_{GG_1}(\mathbf{q},\omega) v_{G_1}(\mathbf{q}) \chi^{RPA}_{G_1G'}(\mathbf{q},\omega). \tag{1.47}$$

here we have used $v_G(\mathbf{q})$, i.e. the Fourier transform of the bare Coulomb interaction for a crystal,

$$v(|\mathbf{r}-\mathbf{r}'|) = \sum_{qG} \mathrm{e}^{\mathrm{i}(\mathbf{q}+\mathbf{G})\cdot(\mathbf{r}-\mathbf{r}')} \frac{4\pi\delta_{GG'}}{|\mathbf{q}+\mathbf{G}|^2} = \sum_{qG} \mathrm{e}^{\mathrm{i}(\mathbf{q}+\mathbf{G})\cdot(\mathbf{r}-\mathbf{r}')} v_G(\mathbf{q}). \tag{1.48}$$

Equation (1.47) can now be inverted to give, formally, $\chi^{RPA} = [1 - v\chi^0]^{-1}\chi^0$, and from this the RPA inverse dielectric function $\varepsilon^{-1}_{GG'}(\mathbf{q},\omega)$ can be constructed according to Eq. (1.39), written concisely as $\varepsilon^{-1} = 1 + v\chi^{RPA}$. Finally, the screened Coulomb interaction is given, in accordance with Eqs. (1.40) and (1.47), by $W(\mathbf{r},\mathbf{r}';t-t') = \varepsilon^{-1}(\mathbf{r},\mathbf{r}';t-t')v(|\mathbf{r}-\mathbf{r}'|)$ so that now we get a non–local, frequency–dependent interaction in place of the bare Coulomb one:

$$\begin{aligned} W(\mathbf{r},\mathbf{r}';t-t') &= \sum_q \sum_{GG'} \mathrm{e}^{\mathrm{i}(\mathbf{q}+\mathbf{G})\cdot\mathbf{r} - \mathrm{i}(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'} \mathrm{e}^{-\mathrm{i}\omega(t-t')} \frac{\varepsilon^{-1}_{GG'}(\mathbf{q},\omega)}{|\mathbf{q}+\mathbf{G}||\mathbf{q}+\mathbf{G}'|} \\ W_{GG'}(\mathbf{q},\omega) &\equiv \frac{\varepsilon^{-1}_{GG'}(\mathbf{q},\omega)}{|\mathbf{q}+\mathbf{G}||\mathbf{q}+\mathbf{G}'|}. \end{aligned} \tag{1.49}$$

Equipped with a good description of the screened interaction, we can now go back to the problem of the electronic self–energy.

1.3.3 Quasiparticle corrections I: electron–electron interaction

We can modify the self–energy of Eq. (1.27) by making a first–order expansion of the GF with respect to the screened interaction W instead of the bare one v:

$$\begin{aligned} \Sigma^{xc}(x_1,x_2) &= \mathrm{i}W(x_1,x_2)G^0(x_1,x_2), \\ \Sigma^{xc}(\mathbf{k},\omega) &= \frac{1}{(2\pi)^4} \sum_{GG'} \int \mathrm{d}^3q \mathrm{d}\omega' W(\mathbf{q},\omega') G^0(\mathbf{k}-\mathbf{q},\omega-\omega') \end{aligned} \tag{1.50}$$

(whe) This term (the lowest–order approximation to Σ) can be summed at all orders in the expansion of the GF (in exactly the same way as for Eq. (1.47), see Fig. 1.4(a))

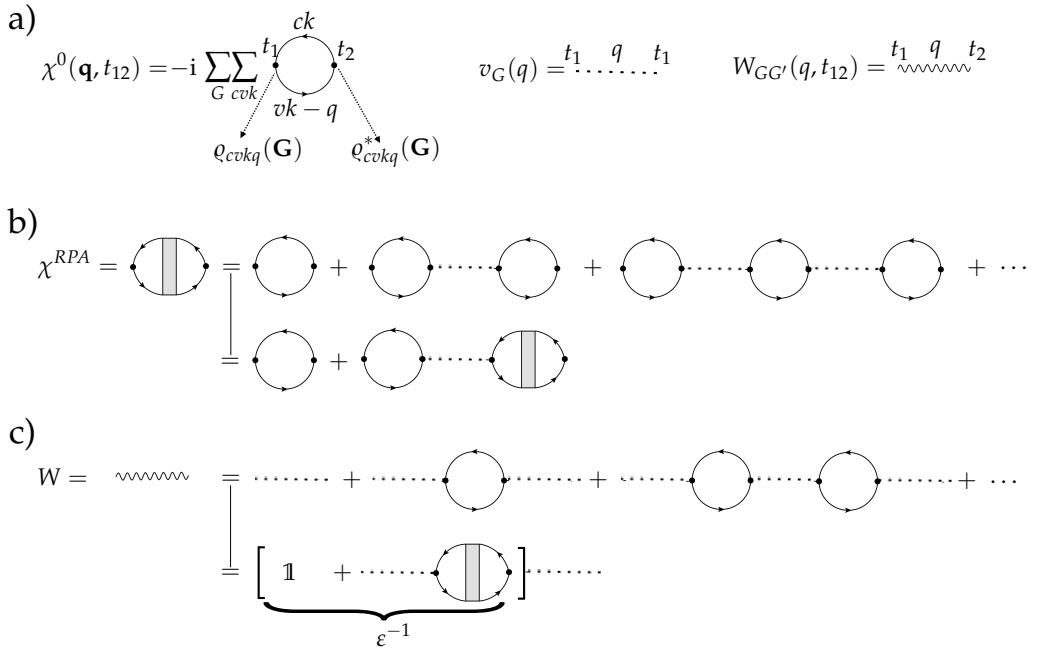


FIGURE 1.3: Graphical representation of the main equations for the electronic screening. a) Eqs. (1.43) / (1.45) (left), Eq. (1.48) (Fourier transform, center) and Eq. (1.49) (Fourier transform, right). All the indices and summations are explicitly given. b) Eq. (1.47). c) Eq. (1.44), $W = \varepsilon^{-1} v$, as the RPA–screened Coulomb interaction (equivalently, Eq. (1.49)). Eq. (1.39), $\varepsilon^{-1} = 1 + v\chi^{RPA}$, is also implied in the last line.

leading to a Dyson's equation for the new, *quasiparticle* (QP) Green's function:

$$G(\mathbf{k}, \omega) = G^0(\mathbf{k}, \omega) + G^0(\mathbf{k}, \omega)\Sigma^{xc}(\mathbf{k}, \omega)G(\mathbf{k}, \omega). \tag{1.51}$$

In fact, G is now describing the propagation of an electron or a hole "dressed" by the electronic interactions contained in Σ^{xc} (a *quasi*electron and a *quasi*hole). The poles of G will give us both the QP energies and the energies of additional satellite structures. Currently, we are interested in the former, which correct the KS eigenvalues and allow for a more realistic description of the band gap, as well as for a more accurate band structure as a whole. This method is called the GW approximation (GWA), since we obtained an approximated self–energy composed of G^0 and W. The Dyson's equation (1.51) represents a summation of self–energy contributions at infinite order in W. Some contributions at each order greater than the first are nonetheless neglected. Such terms at second order are shown in Fig. 1.4(b) as an example. A subset of the neglected terms, the so–called self–consistency diagrams, can be included by (re)computing G self–consistently, but it is not guaranteed that this will lead to better results: in most cases, the GW method consists of a one–shot calculation with the DFT eigenvalues and wave functions as a starting point (in this case we say that we are employing the G_0W_0 method[7]).[10]

If we want to compute the correction to a KS state $|nk\rangle$, we then take the corresponding matrix element of G in the KS basis, and by inverting the Dyson equation

[10]The one–shot G_0W_0 method has been very successful in accurately describing the band gap of a vast class of materials. We will not be concerned with its failures in this thesis, although we will mention that it underestimates the actual band gap of BN systems.

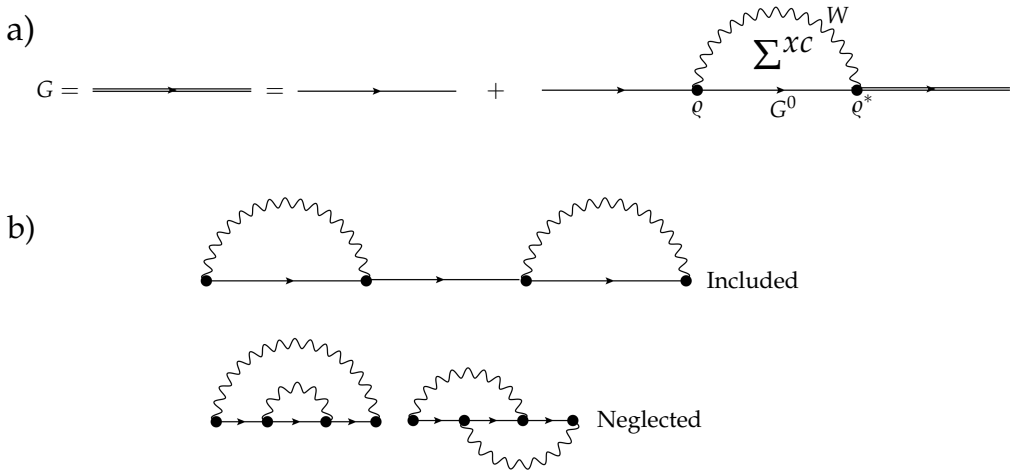


FIGURE 1.4: Graphical representation of the Dyson equation for the GW self–energy. a) Eq. (1.51), $G = G^0 + G^0\Sigma^{xc}G$. b) Terms contributing to the QP Green's function G at second order in the screened interaction W. The top one is included in Eq. (1.51), the bottom two are neglected.

we obtain (recall that $G^{0,-1}_{nk} = \omega - \epsilon_{nk}$)

$$G_{nk}(\omega) = \frac{1}{\omega - \epsilon_{nk} - \langle nk| \Sigma^{xc} |nk\rangle}, \tag{1.52}$$

and the focus of the *ab initio* calculation lies in the computation of the complex quantity $\Sigma^{xc}_{nk}(\omega) \equiv \langle nk| \Sigma^{xc}(\omega) |nk\rangle$.[11] This can be broken up into the sum of two integral expressions, one for the static exchange Σ^x and the other for the dynamical correlations part Σ^c. The exchange part is responsible for the opening of the band gap. The correlation part instead closes the gap, however its contribution to the QP correction is typically smaller than that of Σ^x.

$$\begin{aligned}
\Sigma^x_{nk}(\omega) &= -\sum_v \sum_G \int \mathrm{d}^3q \frac{4\pi}{|\mathbf{q}+\mathbf{G}|^2} |\varrho_{nvkq}(\mathbf{G})|^2, \\
\Sigma^c_{nk}(\omega) &= \mathrm{i}\sum_m \sum_{GG'} \int \mathrm{d}^3q \mathrm{d}\omega' \frac{4\pi\, \varepsilon^{-1}_{GG'}(\mathbf{q},\omega')}{|\mathbf{q}+\mathbf{G}||\mathbf{q}+\mathbf{G}'|} \varrho_{nmkq}(\mathbf{G}) \varrho^*_{nmkq}(\mathbf{G}') G^0_{mk-q}(\omega-\omega').
\end{aligned} \tag{1.53}$$

All the MBPT simulations performed in this thesis were run with the `Yambo` code.[58, 59] In order to obtain converged and reliable results we need to pay attention to Eqs. (1.47) and (1.53). First of all both the k–mesh and the q–mesh need to be dense enough to properly evaluate the corresponding integrals. In general the two meshes coincide and their convergence is handled at the same time: the convergence requirements are in general much stricter than in a DFT calculation and one needs either exceptionally dense grids within the (I)BZ or advanced numerical integration methods (the largest values used in this thesis are $42 \times 42 \times 1$ and $36 \times 36 \times 6$ for the 5–layer and bulk hBN, respectively). This is one of the main reasons why *ab initio* MBPT is in general much more computationally expensive than

[11] Here only the diagonal components of Σ^{xc} in the single–particle basis are considered. They represent the perturbative corrections to the KS energies. Recall that we are implicitly subtracting the DFT exchange–correlation potential from this expression.

DFT, and allows treatment of much smaller systems (of the order of just ~ 10 atoms per unit cell). An additional difficulty is the sum over all empty KS bands appearing in both equations. A cutoff number must be chosen, and in general it is very system–specific. Nonetheless, a large number of empty bands is typically to be computed with high accuracy at the DFT level in order to have reliable results (the required energy range above the bottom of the conduction band can be of the order of ~ 100 eV). A third requirement, which is especially heavy from a memory perspective, is the size of the G–tensors. Here the cutoff values are normally given in energy units and they regulate, among other things, how many *local–field components* of the dielectric function are taken into account. Lastly, in the Σ^c calculation, the frequency integral involving the dielectric function can be performed directly (in which case a discrete grid of ω–values must be carefully converged), or a model dielectric function like the plasmon–pole approximation could be used and the integral perfomed analytically (in which case the reliability of the model should be tested).

Once the self–energy has been obtained for state $|nk\rangle$, we can define and calculate the QP correction. We see from Eq. (1.52) that it is given by the nonlinear equation

$$\omega - \epsilon_{nk} - \Sigma^{xc}_{nk}(\omega) = 0 \tag{1.54}$$

(with $\Sigma^{xc}_{nk} = \Sigma^R_{nk} + \mathrm{i}\Sigma^I_{nk}$), which is generally linearised by expanding its real part around the original, real, KS eigenvalue, and then solved numerically:

$$\begin{aligned} &\omega - \epsilon_{nk} - \Sigma^{xc}_{nk}(\epsilon_{nk}) - \left.\frac{\partial \Sigma^R_{nk}(\omega)}{\partial \omega}\right|_{\omega=\epsilon_{nk}} (\omega - \epsilon_{nk}) = 0, \\ &Z \equiv \left[1 - \left.\frac{\partial \Sigma^R_{nk}(\omega)}{\partial \omega}\right|_{\omega=\epsilon_{nk}}\right]^{-1}, \end{aligned} \tag{1.55}$$

with Z called the renormalisation factor. With this definition, we can write the final form of the QP GF and of the QP band energy:

$$\begin{aligned} G_{nk}(\omega; Z) &= \frac{Z}{\omega - \epsilon_{nk} - Z\Sigma^{xc}_{nk}(\epsilon_{nk})}, \\ E_{nk} &= \epsilon_{nk} + Z\Sigma^{xc}_{nk}(\epsilon_{nk}). \end{aligned} \tag{1.56}$$

Therefore, the correction to the DFT band energies is $Z\Sigma^{xc}_{nk}(\epsilon_{nk})$. The Z–factor is a measure of the single–particle character of the system. If $Z = 1$, the electron addition/removal spectra (given by ImG) shows a single peak at the QP energy. The width of the peak has the physical meaning of the inverse lifetime of the single–particle state – due to electron–electron scattering – and is proportional to Σ^I. Typically, if the system is not strongly correlated, $Z \lesssim 1$: in this case the missing spectral strength is transferred to additional new structures called satellites. However, it should be noted that a GW–type self–energy gives rise to a single satellite at the wrong energy with respect to experimental observations. We will become concerned with satellites (only the ones due the electron–phonon interaction) near the end of this thesis, but so far we ignore them. If $Z \ll 1$, this is a signal that the material in question is not well described from a single–particle perspective, and the theoretical framework has to shift away from standard MBPT. In order to be able to consistently

model the electron–hole interaction between single–particle states, while incorporating the QP corrections, we further define

$$\overline{G}_{nk}(\omega) \equiv \frac{1}{\omega - E_{nk}} = \frac{1}{\omega - \epsilon_{nk} - Z\Sigma^{xc}_{nk}(\epsilon_{nk})}. \tag{1.57}$$

1.4 Optics

We will now start dealing with optical excitations: the response of the electronic system to an external electro–magnetic (em) field, e.g. a laser field. The response is described in terms of neutral excitations of the system that can be created (optical absorption) or destroyed (radiative emission/recombination, luminescence). Let us start with the former case: which is the basic physical mechanism that permits us to understand optical absorption in a gapped material? Is it just the excitation of an electron from the top of the valence band to the bottom of the conduction one, upon absorption of the laser energy? We will see that more complex structures are in general needed to model optical excitations if the electronic screening is weak, and in many cases a description based on the band structure, useful as it is to conceptualise the absorption process, must be formally abandoned.

The coupling of the electronic system to an external weak em field can be described in linear response theory. The treatment is analogous to the one already adopted for the screening (Eq. (1.29)), however now the external perturbation contains both a longitudinal and transverse component and must be described in terms of both a scalar and a vector potential $\phi_{ext}(x)$ and $\mathbf{A}_{ext}(x)$. The external perturbation becomes, at linear order in the incoming fields:

$$\hat{\mathcal{H}}'(t) = \int \mathrm{d}^3 r \hat{\rho}(\mathbf{r})\phi_{ext}(\mathbf{r},t) - \frac{1}{c}\int \mathrm{d}^3 r \mathbf{A}_{ext}(\mathbf{r},t)\cdot\hat{\mathbf{j}}(\mathbf{r}), \tag{1.58}$$

where $\hat{\mathbf{j}} = \frac{\hbar}{2im}[\hat{\Psi}^\dagger(\mathbf{r})\boldsymbol{\nabla}\hat{\Psi}(\mathbf{r}) - (\boldsymbol{\nabla}\hat{\Psi}^\dagger(\mathbf{r}))\hat{\Psi}(\mathbf{r})]$ is the current density operator, since now an induced current will also be generated inside the material alongside the induced charge. Actually, we can immediately choose the *Coulomb gauge* for the em field, i.e. $\boldsymbol{\nabla}\cdot\mathbf{A}_{ext}(\mathbf{r},t) = 0$. This choice of gauge has the effect of separating the *optical character* of the em interaction: now the scalar potential will describe an instantaneous, purely longitudinal field (as we assumed in the screening section), while the vector potential will be equivalent to a purely transverse field. As the scalar potential gives rise to a longitudinal induced charge, which leads to Eq. (1.34) for the charge–charge response function χ, in turn, the vector potential induces a transverse current via the (transverse) current–current response function $\overleftrightarrow{\chi}$:[12]

$$\mathrm{i}\overleftrightarrow{\chi}(x_1,x_2) = \langle \mathcal{T}\left[\hat{\mathbf{j}}'(x_1)\hat{\mathbf{j}}'(x_2)\right]\rangle_0 \tag{1.59}$$

(with $\hat{\mathbf{j}}' = \hat{\mathbf{j}} - \mathbf{j}$). The connection of this function to the two–particle correlation function is more cumbersome than in the case of χ, since we have $\overleftrightarrow{\chi}(x_1,x_2) \propto (\boldsymbol{\nabla}_{\mathbf{r}_2} - \boldsymbol{\nabla}_{\mathbf{r}_4})L(x_1,x_2,x_3,x_4)$. Therefore, in general, $\overleftrightarrow{\chi}(x_1,x_2)$ is not calculated explicitly, as we will see later.

[12] Actually, there is also a mixed longitudinal–transverse response that leads to a transverse induced charge and a longitudinal induced current. We will not introduce the corresponding response functions here for simplicity (for further details, see Ref. [46]). This omission does not impact the discussion in this section.

Let us now start to develop the charge–charge response χ for the case of optical absorption. In this case, the external field is *macroscopic*: it typically has wavelengths that extend from the UV (180 nm / 7 eV) to the IR (1200 nm / 1 eV) range, so that the characteristic length over which it varies is much larger than that of the electronic system. For example, if q is the momentum carried by the laser field and a is the lattice parameter of the crystal (< 1 nm), we may write the following inequality: $qa \ll 1$. This is the *optical limit*, which in the q–dependent formulas will appear as $\lim_{q\to 0}$. What about the induced field ϕ_{ind}: is it macroscopic as well? The answer is no, since the crystal is discrete and so the electronic density is not homogeneous: then the induced fields experience relevant variations at the microscopic level. These so–called *local–field effects* add up in ϕ_{ind} and consequently in $\phi_{tot} = \phi_{ext} + \phi_{ind}$. Mathematically, they are described by all the components of the $\chi_{GG'}$ tensor with $\mathbf{G} \neq \mathbf{0}$ (i.e. short–range contributions in real space). In turn, the experimental observables are macroscopic quantities: as the incoming *macroscopic* field impinges on the system, the outcoming *macroscopic* (i.e. $\mathbf{G} = \mathbf{0}$) field is measured.

In order to proceed, we need to introduce yet another response function, the *proper* one $\overline{\chi}$. Considering any invertible Dyson equation like $\chi = \chi^0 + \chi^0 K \chi$ with $K = v + \dots$ (for example $K = v$ in the RPA case of Eq. (1.47)), we can always write the useful relation

$$\chi^{-1} = \left(\chi^0\right)^{-1} - K. \tag{1.60}$$

At this point we break the Coulomb interaction into two parts, $v = v_0 + \overline{v}$, where v_0 is the long–range component with $\mathbf{G} = \mathbf{0}$, while $\overline{v}$ contains all the $\mathbf{G} \neq \mathbf{0}$ components. We can now define the proper response function as

$$\overline{\chi}^{-1} \equiv \left(\chi^0\right)^{-1} - \overline{v}. \tag{1.61}$$

In the RPA case, $\overline{\chi}^{RPA}$ satisfies the same Dyson equation (1.47) as χ^{RPA}, but now the internal sum over G–vectors must have $G_1 \neq 0$. Furthermore, $\overline{\chi}$ and χ are connected by the following invertible relations:

$$\begin{aligned} \chi &= \overline{\chi} + \overline{\chi} v_0 \chi, \\ \overline{\chi} &= \chi - \chi v_0 \overline{\chi}. \end{aligned} \tag{1.62}$$

In order to obtain the macroscopic quantities that describe optical absorption we need to set $G = 0$, pass to the retarded functions and take the optical limit, therefore we have (this time writing explicitly the dependencies)

$$\begin{aligned} &\overline{\chi}^R_{00}(\mathbf{q},\omega) = \frac{\chi^R_{00}(\mathbf{q},\omega)}{1 + v_0(\mathbf{q})\chi^R_{00}(\mathbf{q},\omega)} = \frac{\chi^R_{00}(\mathbf{q},\omega)}{[\varepsilon^{-1}(\mathbf{q},\omega)]_{00}} \implies \\ &\left[\varepsilon^{-1}(\mathbf{q},\omega)\right]_{00} = \frac{1}{1 - v_0(\mathbf{q})\overline{\chi}^R_{00}(\mathbf{q},\omega)}, \end{aligned} \tag{1.63}$$

where in the second equality on the first line we have used the matrix version of Eq. (1.39) to introduce the microscopic dielectric function. The notation using square brackets is meant to emphasize that we are taking the first element of the *inverse* function, and not the inverse of the of the first element. Now we introduce the central definition of the *macroscopic* dielectric function, the quantity that will be connected to the experimentally observable absorption spectrum:

$$\varepsilon_M(\omega) = \lim_{q\to 0} \frac{1}{[\varepsilon^{-1}(\mathbf{q},\omega)]_{00}}, \tag{1.64}$$

which permits us to finally write

$$\varepsilon_M(\omega) = 1 - \lim_{q\to 0} \frac{4\pi}{q^2} \overline{\chi}^R_{00}(\mathbf{q},\omega). \tag{1.65}$$

The above equation is well–behaved in the small–q limit because

$$\overline{\chi}^R_{00}(\mathbf{q},\omega) \underset{q\to 0}{\sim} q^2 \overline{\chi}^R_{00}(\omega) : \tag{1.66}$$

this can be understood by considering the small–q expansion of the screening matrix elements of Eq. (1.46). The macroscopic dielectric function can be equivalently evaluated by simply inverting Eq. (1.39), $\varepsilon^{-1} = \mathbb{1} + \lim_{q\to 0} v\chi^R$, however this matrix inversion depends on all the GG' components of the tensor.

It is now time to make a somewhat obvious remark. The laser field impinging on the crystal during an absorption experiment is a purely transverse one, carried by $\mathbf{A}_{ext}$ which is orthogonal to the direction of propagation of the laser wave. No external charge distribution is generally inserted into the system ($\phi_{ext} = 0$), something that would justify the use of the longitudinal charge–charge response. Instead, a purely transverse incoming field can in general generate excitations with both longitudinal and transverse components. In a system with cubic symmetry, only a purely transverse current density will be generated.[60] Therefore, we would have to consider a dielectric *tensor*, $\overleftrightarrow{\varepsilon}_M[\omega; \overleftrightarrow{\chi}(\mathbf{q},\omega)]$, expressed in terms of the current–current response function. This quantity has in principle nine components and is considerably more difficult to evaluate than Eq. (1.65). Therefore, the following assumption may be made at this point: (i) The system has cubic symmetry, (ii) we are always at the optical limit ($q \to 0$). Condition (i) ensures that the transverse components of $\overleftrightarrow{\chi}(\mathbf{q},\omega)$ are purely on the diagonal, while the longitudinal ones are off–diagonal; together with condition (ii), it ensures that the dielectric tensor is a scalar function; condition (ii) also ensures that this scalar function is equal to $\varepsilon_M(\omega)$. In short,

$$\lim_{q\to 0} \overleftrightarrow{\varepsilon}_M(\mathbf{q},\omega) = \overleftrightarrow{\mathbb{1}} \varepsilon_M(\omega), \tag{1.67}$$

therefore the transverse response of the system is known when $\overline{\chi}$ is known. Paraphrasing Ref. [60] and [46], this means that, given conditions (i) and (ii): *the transverse response of a (cubic) gapped system is given by the longitudinal response of the same system but with the Coulomb interaction v replaced by $\overline{v}$.* The computational literature strongly relies on the calculation of $\varepsilon_M(\omega)$ to access the optical properties of systems of any symmetry, giving mostly accurate results. This is because the longitudinal response can describe the effects of external transverse fields even in non–cubic crystals (provided that condition (ii) still holds), if the momentum $\mathbf{q}$ of the incoming light is orthogonal to one of the *principal axes* of the dielectric tensor.[61] In other words, if we consider a frame of reference in which the dielectric tensor is diagonal, and the incoming electric field (orthogonal to the light momentum) is parallel to one of the corresponding principal axes, then $\lim_{\mathbf{q}\to\mathbf{0}} [\varepsilon_M(\mathbf{q},\omega)]_{\alpha\alpha}$ (where α labels the principal axis of choice) is still given by Eq. (1.65). This is condition (iii). In the course of this thesis we will assume condition (iii), or rather its consequence, to always be approximately valid, given that we are interested in hexagonal layered crystals with incoming light polarised in the layer plane. However we will briefly go beyond condition (ii) at the end, when discussing the effects of finite–q transitions mediated by phonons on the optical absorption spectrum.

Having discussed the response of the system to an em field, we can now look for solutions to the macroscopic Maxwell's equations for a monochromatic wave penetrating the material (**n** being the polarisation versor)

$$\mathbf{A}_{ext}(\mathbf{r},t) = \mathbf{n}A_0(t)\left(\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}} + \mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\mathbf{r}}\right), \tag{1.68}$$

leading to the well–known exponential attenuation of its intensity I along the direction of propagation z: $I(\omega)/I(0) = \mathrm{e}^{-\alpha(\omega)z}$. The frequency–dependent absorption coefficient $\alpha(\omega)$ is the quantity to be determined by MBPT and computed *ab initio*. Let us note down some useful relations: first, the macroscopic dielectric function has a real and imaginary part: $\varepsilon_M(\omega) = \varepsilon_1(\omega) + \mathrm{i}\varepsilon_2(\omega)$. Note that $\varepsilon_1(0)$ is the static dielectric constant. Then, the complex refractive index $n(\omega) = n_1(\omega) + \mathrm{i}n_2(\omega)$ can be obtained from $\varepsilon_M(\omega)$ because the two are connected by the relation $\varepsilon_M(\omega) = n^2(\omega)$. Finally, the absorption coefficient is given by

$$\alpha(\omega) = \frac{\omega}{c}\frac{\varepsilon_2(\omega)}{n_1(\omega)}. \tag{1.69}$$

Therefore, $\alpha(\omega)$ is completely determined by $\varepsilon_2(\omega)$, which yields the absorption spectrum and is the final quantity we aim to compute in this thesis.

We can clarify the last two paragraphs by computing the optical absorption in the independent–particle case, using the zero–order polarisation $\chi^0 = -\mathrm{i}G^0G^0$ defined in Eq. (1.41). This example will further emphasise the need to introduce an addional layer of theory (the electron–hole interaction) to obtain an accurate description of this physical process. In this case, the analytical form of $\overset{\leftrightarrow 0}{\chi}$ is identical to χ^0 as given in Eqs. (1.33) and (1.44) (provided that we are in a cubic sytem or condition (iii) holds). However, the matrix elements from Eq. (1.46) are different: since in the Coulomb gauge $[\mathbf{p}, \mathbf{A}_{ext}] = 0$ (with $\mathbf{p} = -\mathrm{i}\boldsymbol{\nabla}$), then the transverse field in the perturbing Hamiltonian (Eq. (1.58)) leads to the *optical* matrix elements

$$D_{\mathcal{K}q}(\mathbf{G}) = \frac{\mathbf{n}\cdot\langle ck|\,\mathrm{e}^{\mathrm{i}(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}\mathbf{p}\,|vk-q\rangle}{m(\epsilon_{ck}-\epsilon_{vk-q})} \tag{1.70}$$

(this is to be compared with the screening matrix elements in Eq. (1.46)). We can compare the microscopic dielectric functions obtained from χ^0 and $\overset{\leftrightarrow 0}{\chi}$, (called $\varepsilon^{\text{long}}$ and $\varepsilon^{\text{trans}}$ respectively), considering for simplicity the resonant $\omega > 0$ part (the only one relevant for absorption):[13]

$$\begin{aligned}
\varepsilon^{\text{long}}(\mathbf{q},\omega) &= 1 + \frac{8\pi}{q^2}\frac{e^2}{V}\sum_{GG'}\sum_{q}\sum_{\mathcal{K}}\frac{\varrho^*_{\mathcal{K}q}(\mathbf{G})\varrho_{\mathcal{K}q}(\mathbf{G}')}{w-\Delta\epsilon_{\mathcal{K}q}+\mathrm{i}\eta},\\
\varepsilon^{\text{trans}}(\mathbf{q},\omega) &= 1 + \frac{8\pi e^2}{V}\sum_{GG'}\sum_{q}\sum_{\mathcal{K}}\frac{D^*_{\mathcal{K}q}(\mathbf{G})D_{\mathcal{K}q}(\mathbf{G}')}{w-\Delta\epsilon_{\mathcal{K}q}+\mathrm{i}\eta}.
\end{aligned} \tag{1.71}$$

We approach the optical limit: we first take the long–wavelength Fourier components $\mathbf{G} = \mathbf{G}' = \mathbf{0}$ and obtain the modulus squared of the long–range optical and screening matrix elements, $|D_{\mathcal{K}q}|^2$ and $|\varrho_{\mathcal{K}q}|^2$. We then expand the matrix elements at first order in $\mathbf{q}$: in the screening case the first non–zero term is $\propto |\mathrm{i}\mathbf{q}\cdot\langle c|\,\mathbf{r}-\mathbf{r}'\,|v\rangle\,|^2 =$

[13] The factor of 2 coming from the spin summation is explicitly taken into account in these expressions.

$q^2|\mathbf{n}\cdot\langle c|\,\mathbf{r}\,|v\rangle\,|^2$. In the optical case, the first order is provided already by the momentum operator $\mathbf{p}$ and we can simpy approximate the phase factor to one: this is the so–called dipole approximation. We call the resulting matrix element *dipole matrix element* (DME), $d_{\mathcal{K}}$, in the velocity gauge; furthermore we introduce the following useful relation to convert it to the length gauge:[14]

$$d_{\mathcal{K}} \equiv \frac{\mathrm{i}\hbar}{m}\mathbf{n}\cdot\frac{\langle ck|\,\mathbf{p}\,|vk\rangle}{\Delta\epsilon_{\mathcal{K}}} = \mathbf{n}\cdot\langle ck|\,\mathbf{r}\,|vk\rangle\,. \tag{1.72}$$

All the above considerations show that

$$\lim_{q\to 0}\varepsilon^{\mathrm{trans}}(\mathbf{q},\omega) = \lim_{q\to 0}\varepsilon^{\mathrm{long}}(\mathbf{q},\omega) = \varepsilon_M(\omega), \tag{1.73}$$

and we can write the absorption spectrum in the independent–particle case as

$$\varepsilon_2(\omega) = \frac{8\pi^2 e^2}{V}\sum_{\mathcal{K}}|d_{\mathcal{K}}|^2\delta(\omega-\Delta\epsilon_{\mathcal{K}}). \tag{1.74}$$

We have obtained the textbook result of Fermi's golden rule from time–dependent perturbation theory.

If this were enough to obtain spectra matching with experiments, there would be little need for the complicated theoretical framework that we laid out until now. This formula casts the absorption process in terms of single–particle, "vertical" transitions of electrons that absorb the incoming energy of the laser field. When a transition is allowed by symmetry (i.e. the integral $d_{\mathcal{K}}$ is not identically zero), it appears as a peak in the absorption spectrum, whose intensity is proportional to the probability of the transition taking place. The absorption edge of the system is at the (direct) band gap energy $E_g = \min_{\mathcal{K}}\{\Delta\epsilon_{\mathcal{K}}\}$. Clearly, in most cases the experimental absorption edge lies higher in energy: this is because we did not account for the QP correction and we are using the DFT band gap. By replacing G^0 with the $\overline{G}$ obtained from a G_0W_0 calculation (recall Eq. (1.57)) we can obviate this problem, as now $\chi^0 = -\mathrm{i}\overline{G}\overline{G}$. In this way, we obtain Eq. (1.74) with $\Delta E_{\mathcal{K}}$ instead of $\Delta\epsilon_{\mathcal{K}}$.[15] This might work for materials where the electronic interaction is highly screened, but in many cases we now find ourselves facing the opposite problem: the theoretical absorption edge now lies too high in energy! Replacing χ^0 with $\overline{\chi}^{RPA}$ will not help us lower the band gap of the system, therefore the missing part of the theory is not merely a more advanced single–particle description of the response function. We need to include electron–hole interaction in our theory. Conduction electron and valence hole propagations are correlated, and if their interaction is strong enough they can form a bound state whose binding energy lies below the single–particle band gap. Such a structure is called *exciton*: it is not merely a below–band gap "defect–like" state, but rather it requires a complete reframing of our description of the system. We are dealing with a collective excitation that incorporates *all* electronic transitions of the system, spanning in principle the whole energy–momentum phase space. To elucidate this point, we will anticipate the form taken by ε_2 when the electron–hole

[14] This relation is strictly valid only if the unperturbed Hamiltonian is local, but it can be extended to the non–local case by adding an additional term. The `Yambo` code contains the correction for the non–local case.

[15] However, note that the spectral intensities will change if the energies are shifted, since $\varepsilon_2(\omega)$ must respect the f–sum rule $\int_0^{+\infty}\mathrm{d}\omega\omega\varepsilon_2(\omega) =$ constant. In general, we will use arbitrary units for the spectral intensities.

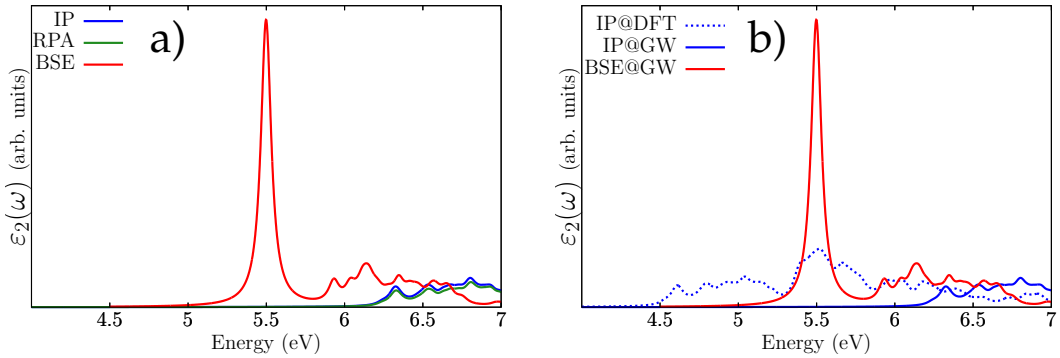


FIGURE 1.5: Comparison of $\varepsilon_2(\omega)$ calculated according to different approximations and levels of theory for bulk hexagonal boron nitride. a) The independent–particle (IP) result, Eq. (1.74), is shown for χ^0 in blue and for $\overline{\chi}^{RPA}$ in green. In both cases the peaks are at the single–particle transition energies (the two functions share the same pole structure) and the absorption onset is at the QP band gap. The BSE result, Eq. (1.75), is instead shown in red: it shows a substantial oscillator strength transfer to the first excitonic state at 5.5 eV and the absorption spectrum is completely altered. b) Comparison between the IP result at the DFT level (dashed blue), at the GW level (blue) and at the GW+BSE level (red). The DFT onset is around 4.5 eV, but the GWA corrects the band gap by about 1.6 eV. The binding energy of the lowest energy exciton is 0.7 eV and sits around the midpoint between the DFT and GW absorption onsets.

interaction is included via MBPT (see the next Section):

$$\varepsilon_2(\omega) = \frac{8\pi^2 e^2}{V} \sum_{\lambda} \left| \sum_{\mathcal{K}} \overline{A}_{\lambda}^{\mathcal{K}} d_{\mathcal{K}} \right|^2 \delta(\omega - E_{\lambda}) \tag{1.75}$$

Here, the possible excitations we are summing over are the excitons λ, and each exciton is made of a linear combination of all single–particle transitions $\mathcal{K}$. The weights $\overline{A}_{\lambda}^{\mathcal{K}}$ allow for the change from transition to exciton basis. They are computed, alongside the exciton energies E_{λ}, by solving the Bethe–Salpeter equation (BSE). Some numerical examples pertaining to the present discussion are presented in Fig. 1.5.

1.5 Excitons

In order to take into account excitonic effects, we need to consider a perturbation expansion of the response function χ in terms of electron–hole interactions. We want to consider these interactions at first order and in a consistent way with the single–particle corrections (i.e., the GW self–energy). We can start from the inverted Eq. (1.32) (in the time–ordered case):[16]

$$\chi(12) = \frac{\delta\rho(1)}{\delta\phi_{ext}(2)}. \tag{1.76}$$

Here for simplicity we made the replacement $1 \to x_1$, $2 \to x_2$, ... for the space–time coordinates. It can be shown[3] that this equation leads to the Bethe–Salpeter

[16] In order to find more information about this functional approach we direct the reader to Ref. [3].

equation (BSE) for χ, written as

$$\begin{aligned}\chi(12) &= \chi^0(12) + \int \mathrm{d}(3456)\chi^0(134)K(3456)\chi(562),\\ K(3456) &= \frac{\delta\Sigma(35)}{\delta\mathcal{G}(46)},\end{aligned} \tag{1.77}$$

whose graphical representation is shown in Fig. 1.6(a). We have a Dyson–like expression with the non–interacting contribution χ^0 and then the interacting part: an electron–hole pair is created at 1, then the electron experiences the interaction at 3, 4 and the hole at 5, 6 before recombining at 2. Both in χ^0 and in the evaluation of the kernel K we consider the QP Green's function $\overline{G}$ obtained from the GW approximation (Eq. (1.57)). Therefore, $\mathcal{G} \to \overline{G}$ and $\Sigma = \Sigma^H + \mathrm{i}\overline{G}W$ with W obtained in the RPA (Eqs. (1.28) and (1.49)). The kernel is then given by

$$\begin{aligned}K(3456) &= \frac{\delta\Sigma^H(35)}{\delta\overline{G}(46)} + \frac{\delta\Sigma^{xc}(35)}{\delta\overline{G}(46)} = -\mathrm{i}\delta(35)\delta(46)v(36) + \mathrm{i}\frac{\delta\overline{G}(35)W(35)}{\delta\overline{G}(46)}\\ &\simeq -\mathrm{i}\delta(35)\delta(46)2v(|\mathbf{r}_3-\mathbf{r}_6|)\delta_t(t_1-t_6) + \mathrm{i}\delta(36)\delta(45)W(35),\end{aligned} \tag{1.78}$$

where in the second line we neglected $\frac{\delta W}{\delta\overline{G}} = W\frac{\delta\chi^{RPA}}{\delta\overline{G}}W$ because it is of higher order in W. One last simplification remains: recall that $\overline{G}$ is by definition the GW Green's function without its satellite structures, i.e., the effect of the dynamical part of the interaction is only included at the level of the QP energies. For consistency reasons, we then take the static approximation for the screened interaction within K, $W(35) = W(\mathbf{r}_3,\mathbf{r}_5)\delta_t(t_3-t_5)$. This means that upon external perturbation, the electron density of the system polarises instantaneously. The static approximation allows us to invert Eq. (1.77), because now the time integrations will be reduced to just one time, and in the frequency domain the equation will only depend on the incoming frequency, without an integration over the internal frequency of the screening. In the final part of this thesis we will try to go beyond the static approximation when including phonons into the kernel. Note that in order to compute $\overline{\chi}$ instead of χ it is enough to consider a kernel $\overline{K}$ where the bare Coulomb interaction v constituting the first of the two terms is replaced with $\overline{v}$ (see Eqs. (1.60) and (1.61)).

Before inverting the BSE, let us write down explicitly the two components of the kernel in the transition basis:

$$\begin{aligned}&-\mathrm{i}K^q_{\mathcal{K}\mathcal{K}'} = -V^q_{\mathcal{K}\mathcal{K}'} + W^q_{\mathcal{K}\mathcal{K}'},\\ &W^q_{\mathcal{K}\mathcal{K}'} = \langle\mathcal{K}|W|\mathcal{K}'\rangle = \int \mathrm{d}^3r\mathrm{d}^3r'\varphi_{v_1k_1-q}(\mathbf{r})\varphi^*_{c_1k_1}(\mathbf{r}')W(\mathbf{r},\mathbf{r}')\varphi^*_{v_2k_2-q}(\mathbf{r})\varphi_{c_2k_2}(\mathbf{r}'),\\ &V^q_{\mathcal{K}\mathcal{K}'} = \langle\mathcal{K}|V|\mathcal{K}'\rangle = 2\int \mathrm{d}^3r\mathrm{d}^3r'\varphi_{v_1k_1-q}(\mathbf{r})\varphi^*_{c_1k_1}(\mathbf{r})v(|\mathbf{r}-\mathbf{r}'|)\varphi^*_{v_2k_2-q}(\mathbf{r}')\varphi_{c_2k_2}(\mathbf{r}').\end{aligned} \tag{1.79}$$

The first term represents an attractive, screened interaction between electron and hole. This term is responsible for the formation of bound states lying at lower energy than the single–particle band gap and is often called "direct" eh interaction; in order to emphasize its origin from the derivative of Σ^{xc}, it may also be called "screened exchange" term. The second term exchanges particle and hole indices and is repulsive, short–range and mediated by the bare Coulomb interaction. It has the effect of reducing the exciton binding energy and is generally called "exchange" interaction, although this can be quite confusing since it comes from the Hartree part of

the self–energy, not the exchange part; therefore it may also be called "Hartee" term to emphasize its origin. In order to make the connection with the previously computed *ab initio* quantities, we take the Fourier transform of Eq. (1.79) and write the final form of the "direct" and "exchange" matrix elements in terms of the screening matrix elements,

$$\begin{aligned} W^q_{\mathcal{K}\mathcal{K}'} &= \frac{1}{N_q\Omega}\sum_{q_W}\sum_{GG'}\frac{\varepsilon^{-1}(\mathbf{q}_W,0)}{|\mathbf{q}_W+\mathbf{G}||\mathbf{q}_W+\mathbf{G}'|}\varrho^*_{v_1v_2k_1-qq_W}(\mathbf{G})\varrho_{c_1c_2k_1q_W}(\mathbf{G}'), \\ V^q_{\mathcal{K}\mathcal{K}'} &= \frac{2}{N_q\Omega}\sum_{G}\frac{4\pi}{|\mathbf{q}+\mathbf{G}|^2}\varrho^*_{c_1v_1k_1q}(\mathbf{G})\varrho_{c_2v_2k_2q}(\mathbf{G}). \end{aligned} \tag{1.80}$$

Note that the screened interaction depends on its own internal momentum $\mathbf{q}_W = \mathbf{k}_2 - \mathbf{k}_1$ because of the presence of the microscopic inverse dielectric function, and that the screening matrix elements couple different transitions in W and V. The BSE with kernel $K = -V + W$ is schematically shown in Fig. 1.6(b). Up to this point, we implicitly considered only *resonant* electron–hole transitions, i.e., those with positive $E_c - E_v$ energy differences. However, when summing over electron–hole pairs $\mathcal{K}$, in principle we have to include also anti–resonant transitions, i.e. when $E_v - E_c < 0$. In order to make this more explicit, we use the following definition: given a transition labeled by $\mathcal{K}$, then $\overline{\mathcal{K}}$ will label the corresponding transition with the opposite resonant character (i.e. if $\mathcal{K} = v_1c_1k_1$, then $\overline{\mathcal{K}} = c_1v_1k_1$). Now we can introduce a useful symmetry relation for the kernel:

$$\left[K^q_{\mathcal{K}\mathcal{K}'}\right]^* = K^{-q}_{\overline{\mathcal{K}}\overline{\mathcal{K}}'}. \tag{1.81}$$

We now need to find $\chi^q_{\mathcal{K}\mathcal{K}'}(\omega)$. In order to proceed, we rewrite the inverted Eq. (1.77), $\chi(\omega) = [\mathbb{1} - \chi^0(\omega)K]^{-1}\chi^0(\omega)$, using its *resolvent*:[17]

$$\chi^q_{\mathcal{K}\mathcal{K}'}(\omega) = [H - \mathbb{1}z]^{-1}_{\mathcal{K}\mathcal{K}'}\,\Delta f^q_{\mathcal{K}}, \tag{1.82}$$

where $z = \omega + \mathrm{i}\eta$, $\Delta f^q_{\mathcal{K}}$ is a difference between occupation factors (if $\mathcal{K}$ is resonant then $\Delta f^q_{\mathcal{K}} = f_{c_1k_1} - f_{v_1k_1-q} = -1$ and $\Delta f^q_{\overline{\mathcal{K}}} = 1$), and H is a matrix that always has real eigenvalues, although it may not be hermitian if anti–resonant transitions are considered. We call H *excitonic* Hamiltonian, since it is given by

$$H^q_{\mathcal{K}\mathcal{K}'} = \Delta E_{\mathcal{K}q}\delta_{\mathcal{K}\mathcal{K}'} - \mathrm{i}\Delta f^q_{\mathcal{K}'}K^q_{\mathcal{K}\mathcal{K}'}, \tag{1.83}$$

(if $\mathcal{K}'$ is resonant then $\Delta f^q_{\mathcal{K}'} = -1$ and $\Delta f^q_{\overline{\mathcal{K}}'} = 1$).

In order to solve the excitonic problem we have to diagonalise the Hamiltonian H. Its diagonal part consists of single–particle QP transition energies (here $\delta_{\mathcal{K}\mathcal{K}'} = \delta_{c_1c_2}\delta_{v_1v_2}\delta_{k_1k_2}$): if electron–hole interaction is neglected, this will lead us back to Eq. (1.74). It is the non–diagonal part, which contains the electron–hole interaction kernel, that forces us to reconsider the excitation problem: the eigenvalues E^q_λ of H then represent the energy of the excitonic state λ, which is the real two–particle excitation energy. Analogously, the eigenvectors of H, expressed as $A^{\mathcal{K}}_{\lambda q} \equiv \langle \mathcal{K}q|\lambda\rangle$,

[17] Here and henceforth, recall that we are explicitly considering electron–hole transitions only – i.e. $\mathcal{K} = (v_1c_1k_1)$, $\mathcal{K}' = (v_2c_2k_2)$ – for a gapped material at zero temperature. The following discussion can be generalised by considering any two–particle transitions and adding explicitly the temperature–dependent band occupation factors. In this case, the two–particle correlation function L, which is more general than χ, should be used.

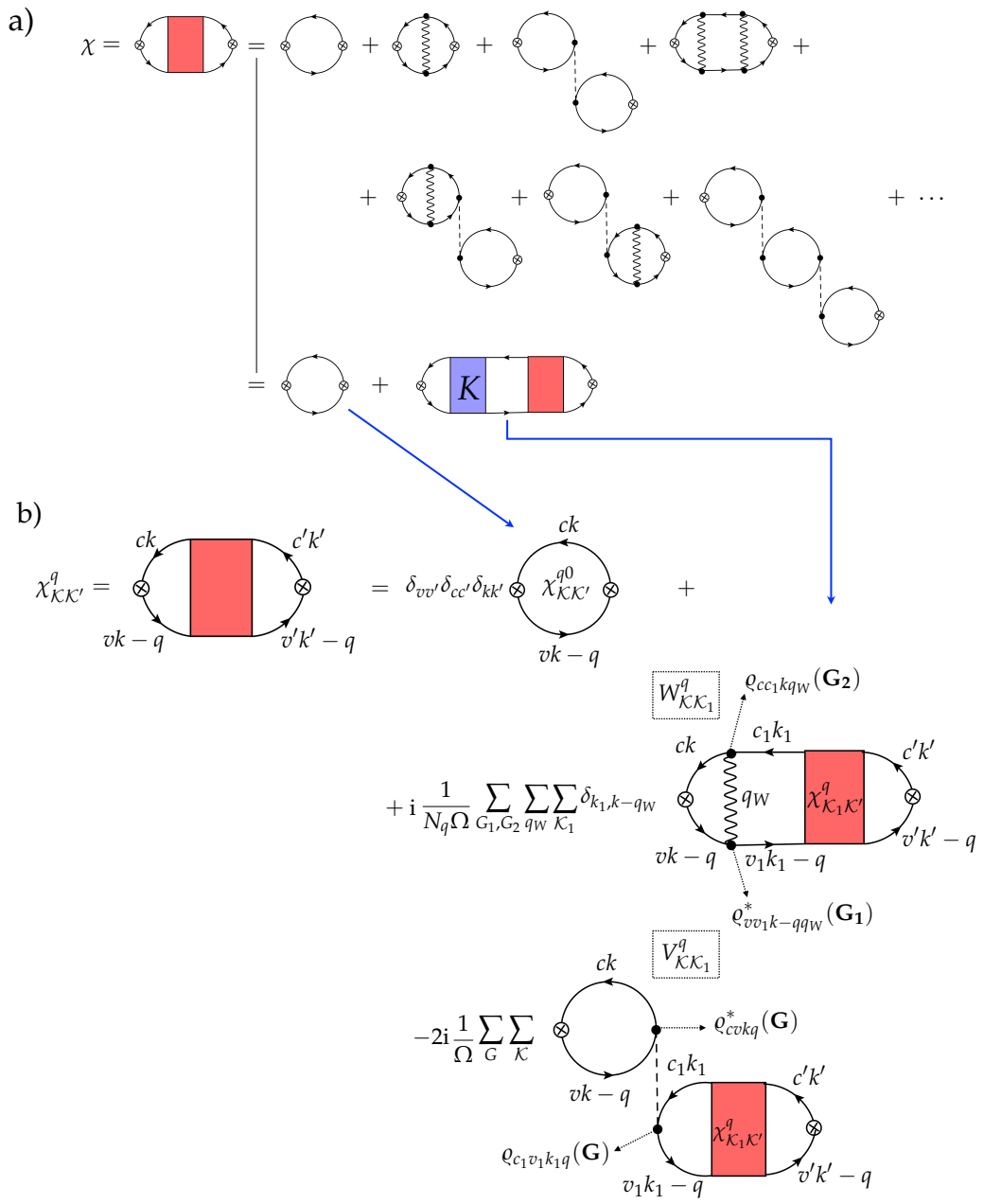


FIGURE 1.6: Graphical representation of the static Bethe–Salpeter equation. a) The contributions to the electron–hole interaction included in the kernel K, leading to Eq. (1.77) $\chi = \chi^0 + \chi^0 K \chi$. The wiggly and dotted lines are the statically screened and bare Coulomb interactions, respectively. b) Eq. (1.77) before its inversion (which leads to Eq. (1.82)). All the indices and summations, including those for the kernel $K = W - V$, based on Eq. (1.80), are explicitly written. The black dots represent the screening matrix elements, while the white crossed dots become the dipole matrix elements at the optical limit.

are connected to the exciton wave functions, and we have

$$\sum_{\mathcal{K}'} H^q_{\mathcal{K}\mathcal{K}'} \langle \mathcal{K}' q | \lambda \rangle = E^q_\lambda \langle \mathcal{K} q | \lambda \rangle \,. \tag{1.84}$$

Note that the eigenvectors relative to $\overline{\chi}$ will be denoted $\overline{A}$ instead of A. If the eigenvectors are not orthogonal we can define an overlap matrix $N_{\lambda\lambda'} = \sum_{\mathcal{K}} A^{\mathcal{K},*}_{\lambda} A^{\mathcal{K}}_{\lambda'}$ (which reduces to $\delta_{\lambda\lambda'}$ in the orthogonal case) and their completeness relation becomes

$$\sum_{\lambda\lambda'} A^{\mathcal{K}}_{\lambda} N^{-1}_{\lambda\lambda'} A^{\mathcal{K}'*}_{\lambda'} = \delta_{\mathcal{K}\mathcal{K}'}. \tag{1.85}$$

Then, the spectral representation of χ can be written in terms of the spectrum of H, giving in particular for its retarded part (i.e. positive exciton energies)

$$\begin{aligned} \chi^{q,R}_{\mathcal{K}\mathcal{K}'}(\omega) &= \sum_{\lambda\lambda'} A^{\mathcal{K}}_{\lambda q} N^{-1}_{\lambda\lambda'} A^{\mathcal{K}'*}_{\lambda' q} \chi^{q,R}_{\lambda\lambda'}(\omega), \\ \chi^{q,R}_{\lambda\lambda'} &= \frac{\delta_{\lambda\lambda'}}{\omega - E^{q}_{\lambda} + \mathrm{i}\eta}. \end{aligned} \tag{1.86}$$

Here $\chi^{q,R}_{\lambda\lambda'}(\omega) = \chi^{R}_{\lambda}(\mathbf{q},\omega)$ is the propagator written in the excitonic basis, in which it is diagonal. Now we just need to express χ^R as a tensor in the G–vectors (Eq. (1.43)),

$$\chi^{R}_{GG'}(\mathbf{q},\omega) = \sum_{\mathcal{K}\mathcal{K}'} \sum_{\lambda} A^{\mathcal{K}}_{\lambda q} \varrho^{*}_{\mathcal{K}q}(\mathbf{G}) A^{\mathcal{K}'*}_{\lambda q} \varrho_{\mathcal{K}'q}(\mathbf{G}') \chi^{R}_{\lambda}(\mathbf{q},\omega) \tag{1.87}$$

(here we assumed $N_{\lambda\lambda'} = \delta_{\lambda\lambda'}$ for simplicity). Finally, if we switch from $\chi^R_{GG'}$ to $\overline{\chi}^R_{GG'}$ and take the optical limit ($G = G' = 0$, $q \to 0$), we can use Eq. (1.65) to access the macroscopic dielectric function $\varepsilon_M(\omega)$, and then we obtain Eq. (1.75) as its imaginary part: this gives us the absorption spectrum including excitonic effects. The exciton wave functions can also be represented in real space as six–dimensional objects depending on both the electron and hole coordinates, given by

$$\psi^{q}_{\lambda}(\mathbf{r}_e,\mathbf{r}_h) = \sum_{vck} A^{vck}_{\lambda q} \varphi_{ck}(\mathbf{r}_e) \varphi^{*}_{vk-q}(\mathbf{r}_h), \tag{1.88}$$

with $|\psi^q_\lambda(\mathbf{r}_e,\mathbf{r}_h)|^2$ giving the electron–hole charge distribution. Similarly, in order to analyse which transitions are the most important for a certain exciton λ, the k–resolved excitonic weights $\sum_{cvk} |A^{cvk}_{\lambda q}|^2$ in the BZ can be considered.

Before moving on, let us take a look at the explicit matrix structure of $H(\mathbf{q})$. In particular, under which conditions are we allowed to restrict H to the resonant part only? This is a relevant question because the diagonalisation of the excitonic Hamiltonian, as we shall see, is a computationally demanding task. We have:

$$H(\mathbf{q}) = \begin{array}{c} \\ \mathcal{K} \\ \overline{\mathcal{K}} \end{array} \begin{array}{c} \begin{array}{cc} \mathcal{K}' & \overline{\mathcal{K}}' \end{array} \\ \begin{pmatrix} \Delta E_{\mathcal{K}q} + \mathrm{i}K^{q}_{\mathcal{K}\mathcal{K}'} & -\mathrm{i}K^{q}_{\mathcal{K}\overline{\mathcal{K}}'} \\ \mathrm{i}K^{-q}_{\mathcal{K}\overline{\mathcal{K}}'} & -\left[\Delta E_{\mathcal{K}q} + \mathrm{i}K^{-q}_{\mathcal{K}\mathcal{K}'}\right] \end{pmatrix} \end{array} \tag{1.89}$$

where we have rewritten the second row using Eq. (1.81). We can see that resonant (top left) and anti–resonant (bottom right) transitions are coupled by the off–diagonal elements. However, their contribution in the absorption case is usually very small and only non–negligible in the case of metals: therefore they are neglected. This is called Tamm–Dancoff approximation (TDA). It reduces the size of the kernel because now the resonant and anti–resonant blocks may be diagonalised separately, and the latter may be computed in terms of the former, albeit with opposite momentum. The two blocks then coincide at $\mathbf{q} = 0$: since in general the *ab initio*

solution of the BSE is done in the context of optical absorption, this is the only kernel usually computed and diagonalised by the numerical codes.[8] If one is instead interested in extracting information about the exciton dispersion, then the build–up and subsequent diagonalisation of multiple $H(\mathbf{q})$ matrices is in principle needed on a certain mesh of q–points. Finally, in order to substantiate the discussion in this Section with an example, the analytical solution for H in the case of a two–level system is presented in App. A.

1.5.1 Calculations in practice with the BSE and `Yambo`

The usual flow of a many–body calculation starts from the ground–state properties, then moves on to QP corrections, and finally goes on to the excitonic states. In short: DFT $\rightarrow$ GW $\rightarrow$ BSE. The RPA screening can be obtained in the context of a GW calculation or, if a model dielectric function is used for the GW step, its static part can be computed before a BSE run. In the case of the calculations presented in this thesis, the DFT energies and wave functions computed with QE are read by the `Yambo` code. `Yambo` produces several databases containing the data required either for subsequent calculations or for analysis (QP energies and lifetimes, spectral functions, components of χ^{RPA}, exciton eigenvectors and energies, *et cetera*).

For a BSE calculation, new convergence requirements must be satisfied in addition to those pertaining to the GW and screening parts (see Secs. 1.3.2 nad 1.3.3). In particular, an independent convergence test on the k–mesh must be performed in order to ensure the accuracy of the exciton energies and spectra. Also with regards to Eq. (1.83), the number of c and v states to include in the excitonic Hamiltonian is crucial; usually this number is not very high, as the contributions to the various exciton states tend to be localised around few bands. Fulfilling these requirements means rapidly increasing the size of H by adding more and more transitions $\mathcal{K}$, something that can very easily lead to memory problems during the calculations. Additionally, the necessary numbers of G-vectors in the sums of the V and W components of the kernel must also be carefully checked. Regarding the diagonalisation of H, different methods are available depending on its size. In the best–case scenario of a tractable problem, a full diagonalisation may be performed, giving access to all exciton eigenstates and energies. This is presently the only possibility for finite–q calculations. If the kernel size is too big, an iterative solution called Haydock's method (based on the Lanczos algorithm)[62] is available: this gives access to the absorption spectra but not to the exciton wave functions. Recently, with the addition in `Yambo` of the SLEPC library,[63] it is possible to perform an iterative solution valid also for the eigenstates: in this case the number of excitonic states λ to be converged (starting from the lowest–bound one) is set in the input.

Parallelisation of the calculations is possible and the k, q, c and v loops can be efficiently parallelised in all the main sections of the code. In general, the parallelisation over k–points increases the speed of the calculations, while the one over bands allows for better memory distribution.

The implementation of the solution of the BSE at finite momentum has been recently completed.

1.5.2 Additional caveats for two–dimensional materials

The *ab initio* simulation of a 2D material is especially delicate. First of all, due to the lack of periodicity along the direction perpendicular to the layer plane (here and

henceforth called the stacking direction), it is necessary to construct simulation supercells where the repeated copies of the system along the stacking direction are separated by enough vacuum space so as to not to interact with each other. This is easy in the case of DFT, but when computing the many–body χ, the presence of the bare Coulomb interaction $v(\mathbf{r})$ means that the vacuum intersystem separation becomes another parameter to be converged. Unfortunately, the other convergence parameters (k–grid and, sums over G vectors, number of unoccupied states) depend on it.[64] Additionally, due to the long–range character of v, true convergence is often impossible before the calculation becomes unfeasibly heavy.[26] In order to alleviate this issue and speed convergence up, a truncation of the Coulomb potential in real space (Coulomb cutoff) is performed: $v(\mathbf{r})$ is put to zero outside a user–chosen geometric region (e.g. sphere, cylinder, box).[65] This allows to obtain converged results with relatively small intersystem separations ($\sim$ 20 Å in the case of the systems studied in this thesis).

Another issue is the numerical instability of the q–integrals over a quasi–2D BZ involving the Coulomb interaction $v(\mathbf{q}+\mathbf{G}) \propto |\mathbf{q}+\mathbf{G}|^{-2}$. To avoid divergences, the random integration method (RIM) should be applied to these Coulomb integrals: for each q of the discretised integration q–grid, the Coulomb interaction is replaced with a Monte Carlo integral over a large (user–defined) number of random Q–points, as: $v^{RIM}(\mathbf{q}+\mathbf{G}) = \int_R \mathrm{d}^3Q\, v(\mathbf{q}+\mathbf{G}+\mathbf{Q})$ (where R is a small user–defined volume in reciprocal space). This allows for a smoother function and a well–behaved q integration.

1.6 Lattice vibrations

This Section is concerned with the interaction of the electronic system with the ions. We will describe the lattice vibrations, i.e. the phonons, in a framework that is useful for Density Functional Perturbation Theory (DFPT) calculations, and then switch to a many–body description. The treatment is mainly adapted from Refs.[5, 49, 66, 67].

1.6.1 Dynamical matrix and phonon dispersion

We first consider the ionic system in the BO approximation. Here the ions will interact with each other and be subject to the potential due to the electronic system evaluated at the static ionic positions. According to Eq. (1.2), we can write for the ionic system:

$$\langle \hat{\mathcal{H}}_{ion}(\{\mathbf{R}\})\rangle = \langle \hat{\mathcal{H}}^0_{ion}(\{\mathbf{R}\})\rangle + E[\{\mathbf{R}^0\}] + \langle \hat{\mathcal{H}}_{e-ion}(\{\mathbf{r}\},\{\mathbf{\Delta R}\})\rangle \tag{1.90}$$

where $E[\{\mathbf{R}^0\}]$ is the DFT total energy. The interacting potential $\hat{\mathcal{H}}_{e-ion}$ will give rise to collective excitations of the ionic system in the form of lattice waves or phonons. Actually, $\hat{\mathcal{H}}_{e-ion}$ can be identified with the variation of $E[\{\mathbf{R}\}]$ when the atoms are displaced from their equilibrium positions. Since in our case this interaction will not be as strong as to break the BO approximation, we can perform a Taylor expansion in the atomic positions up to the second order: this is the *harmonic approximation*.

Therefore we obtain

$$\begin{aligned} E(\{\mathbf{R}\}) =& E[\{\mathbf{R}^0\}] + \langle \hat{\mathcal{H}}_{e-ion}(\{\mathbf{r}\},\{\mathbf{\Delta R}\})\rangle = \\ & E[\{\mathbf{R}^0\}] + \sum_{Li\alpha} \left.\frac{\partial E(\{\mathbf{R}\})}{\partial R^0_{Li\alpha}}\right|_0 (R_{Li\alpha} - R^0_{Li\alpha}) + \\ & \frac{1}{2}\sum_{\substack{Li\alpha \\ Mj\beta}} \left.\frac{\partial^2 E(\{\mathbf{R}\})}{\partial R^0_{Li\alpha}\partial R^0_{Mj\beta}}\right|_0 (R_{Li\alpha} - R^0_{Li\alpha})(R_{Mj\beta} - R^0_{Mj\beta}) + \dots \end{aligned} \tag{1.91}$$

Several remarks are now in order. The atomic positions depend on three indices: the Greek letter identifies the Cartesian direction; the lowercase Latin letter identifies the ion within the unit cell of the crystal (therefore it runs over the crystal basis); the uppercase latin letter counts the repeated unit cells (therefore we may call it "supercell index"). In other words, the position of an ion in the crystal is expressed as $\mathbf{R}_{Li} = \mathbf{r}_i + \boldsymbol{\tau}_L$, where $\mathbf{r}_i$ is restricted to one unit cell and $\boldsymbol{\tau}_L$ is a lattice vector. The derivatives appearing in Eq. (1.91) are evaluated at the equilibrium ionic positions: therefore the first derivative vanishes and we are left with the second–order term only. In general,

$$F_{Li\alpha} \equiv -\frac{\partial E(\{\mathbf{R}\})}{\partial R^0_{Li\alpha}} \tag{1.92}$$

is the restoring force that the electron system exerts upon the (Li)–th ion when it is displaced along the α direction. Analogously,

$$C^{Li\alpha}_{Mj\beta} \equiv \frac{\partial^2 E(\{\mathbf{R}\})}{\partial R^0_{Li\alpha}\partial R^0_{Mj\beta}} = -\frac{\partial F_{Li\alpha}}{\partial R^0_{Mj\beta}}. \tag{1.93}$$

If ion (Li) is displaced along α, and ion (Mj) is displaced along β, then $C^{Li\alpha}_{Mj\beta}$ is the variation in the restoring force acting upon (Li) induced by the displacement of (Mj). Let us now define the atomic *displacements* $\mathbf{u}_{Li} \equiv \mathbf{R}_{Li} - \mathbf{R}^0_{Li}$. Now by considering the ionic momenta defined as $\mathbf{P}_{Li} \equiv M_i \dot{\mathbf{u}}_{Li}$ and a constant energy shift E_{shift} coming from the interaction energy of the ionic system at $\{\mathbf{R}^0\}$, we can write the ionic Hamiltonian as

$$\langle \hat{\mathcal{H}}_{ion}(\{\mathbf{R}\})\rangle \simeq E_{shift} + \frac{1}{2}\sum_{Li\alpha} M_i \dot{u}^2_{Li\alpha} + \frac{1}{2}\sum_{Li\alpha,Mj\beta} C^{Li\alpha}_{Mj\beta} u_{Li\alpha} u_{Mj\beta}. \tag{1.94}$$

The size of the force constant matrix C is then $3NN_{at}$, with N being the number of unit cells within a periodically repeated supercell and N_{at} the basis of the crystal. However, because of the unit–cell periodicity of the crystal, the force constants are invariant under a translation of a lattice vector $\boldsymbol{\tau}$, i.e. they only depend on the difference $\boldsymbol{\tau}_I = \boldsymbol{\tau}_L - \boldsymbol{\tau}_M$, which allows for the Fourier transform into reciprocal space. In this way we can solve the problem just in a single unit cell of the crystal, replacing the sums over the supercell index with q–sums in the BZ:

$$D^{i\alpha}_{j\beta}(\mathbf{q}) = \frac{1}{\sqrt{M_i M_j}} \sum_I C^{Ii\alpha}_{j\beta} e^{i\mathbf{q}\cdot\boldsymbol{\tau}_I}. \tag{1.95}$$

This is the dynamical matrix. The lattice vibrations will be the solutions to the equation of motion for Eq. (1.94), which is

$$M_i \ddot{u}_{Li\alpha} = -\frac{\partial \langle \hat{\mathcal{H}}_{ion}(\{\mathbf{R}\})\rangle}{\partial u_{Li\alpha}} = -\sum_{Mj\beta} C^{Li\alpha}_{Mj\beta} u_{Mj\beta}. \tag{1.96}$$

These can be found by first diagonalising D for each $\mathbf{q}$, so that we obtain:

$$\sum_{i\alpha} D^{i\alpha}_{j\beta}(\mathbf{q}) \xi^{\lambda q}_{i\alpha} = \omega^2_{q\lambda} \xi^{\lambda q}_{j\beta}. \tag{1.97}$$

Here, the eigenvectors $\boldsymbol{\xi}^{\lambda q} = \left(\xi^{\lambda q}_{1x}\xi^{\lambda q}_{1y}\xi^{\lambda q}_{1z}|\xi^{\lambda q}_{2x}\ldots\xi^{\lambda q}_{N_{at}z}\right)^T$ are the normal modes of the oscillating system which are orthonormal,

$$\sum_{i\alpha} \xi^{\lambda q *}_{i\alpha} \xi^{\mu q}_{i\alpha} = \delta_{\lambda\mu}, \qquad \sum_{\lambda} \xi^{\lambda q *}_{i\alpha} \xi^{\lambda q}_{j\beta} = \delta_{ij}\delta_{\alpha\beta}. \tag{1.98}$$

The normal modes describe the phonons in terms of $3N_{at}$ collective, periodic oscillations of the crystal with momentum q that are independent of each other, rather than in terms of single–ion displacements. The smaller the momentum, the larger the phonon periodicity in real space. Now $\omega_{q\lambda} \geq 0$ is the frequency associated with the phonon branch λ at momentum q (compare this with the Section about the excitonic Hamiltonian, Eq (1.83)). The original atomic displacements u (the solutions of the equation of motion) are obtained from the normal modes. They have the form of standing waves and can be obtained in any real–space unit cell by a basis change with the proper supercell phase factors:

$$\begin{aligned} u_{Li\alpha}(t) &= \frac{1}{\sqrt{NM_i}} \sum_{\lambda q} e^{i\mathbf{q}\cdot\boldsymbol{\tau}_L} \xi^{\lambda q}_{i\alpha} \bar{z}^{\lambda q}(t), \\ \bar{z}^{\lambda q}(t) &= \frac{1}{2}\left[A_{q\lambda}(T)\exp^{-i\omega_{q\lambda}t} + A^*_{q\lambda}(T)\exp^{i\omega_{q\lambda}t}\right]. \end{aligned} \tag{1.99}$$

Here we have used the following properties: $\omega_{\lambda -q} = \omega_{\lambda q}$, $\xi^{\lambda q *}_{i\alpha} = \xi^{\lambda -q}_{i\alpha}$ and $\bar{z}^{\lambda q *}_{i\alpha} = \bar{z}^{\lambda -q}_{i\alpha}$. Furthermore, the temperature–dependent amplitudes of the oscillations are given by[68] $A_{q\lambda}(T) = \sqrt{2k_B T}/\omega_{q\lambda}$, k_B being the Boltzmann constant.

In summary, the relevant information about crystal lattice vibrations can be accessed by first computing the interatomic force constants (IFC) and then diagonalising the dynamical matrix.

1.6.2 Density functional perturbation theory

On the matter of computing the IFCs, again we can use DFT as a starting point, turning to DFPT.[5, 6] We give here only a brief account of the logic behind DFPT and of the main quantities that result from an *ab initio* phonon calculation. In order to compute the second derivatives of the total energy, we start with the first derivatives and apply the Hellmann–Feynman theorem to transform the derivative of the

expectation value into the derivative of the operators (see Eq. (1.6)),

$$\begin{aligned}\frac{\partial E(\{\mathbf{R}\})}{\partial R^0_{Ii\alpha}} &= \langle \Psi_0^{MB}| \frac{\partial \hat{\mathcal{H}}_e}{\partial R^0_{Ii\alpha}} |\Psi_0^{MB}\rangle = \langle \Psi_0^{MB}| \frac{\partial \hat{V}_{e-ion}}{\partial R^0_{Ii\alpha}} |\Psi_0^{MB}\rangle \\ &= \int \mathrm{d}^3 r \frac{\partial v_R(\mathbf{r})}{\partial R^0_{Ii\alpha}} \rho_0(\mathbf{r}),\end{aligned} \tag{1.100}$$

where the second step stems from the fact that only the external potential depends explicitly on the ionic positions. This leads to

$$C^{Ii\alpha}_{j\beta} = \int \mathrm{d}^3 r \left[\frac{\partial^2 v_R(\mathbf{r})}{\partial R^0_{Ii\alpha} \partial R^0_{0j\beta}} \rho_0(\mathbf{r}) + \frac{\partial v_R(\mathbf{r})}{\partial R^0_{Ii\alpha}} \frac{\partial \rho_0(\mathbf{r})}{\partial R^0_{0j\beta}} \right]. \tag{1.101}$$

At this point, the task is reduced to the computation of the first derivatives of the ground–state electron density. These are computed in a self–consistent scheme similar to the standard DFT one, but this time yielding the derivatives of the KS wave functions via the ones for the effective potential $\partial_R v_{eff}$. These latter quantities will also yield the electron–phonon coupling matrix elements. This method, although quite cheap computationally–wise compared to those based on atomic displacements,[49] is still much more computationally expensive than a standard DFT run, as all the $(Ii\alpha)$ variations of ρ and v_{eff}, which in general lower the symmetry of the crystal, have to be computed.

1.6.3 $\mathbf{q} \to \mathbf{0}$ limit

The system will always have three acoustic phonon branches, and if $N_{at} \geq 2$ it will also feature $3N_{at} - 3$ high–frequency optical branches. Additionally, in a layered material we can distinguish between out–of–plane modes, where the ionic oscillations are orthogonal to the layer plane, and planar modes. Among the planar modes we have the longitudinal and transverse modes, in which the direction of propagation of the phonon is either parellel or orthogonal, respectively, to their momentum $\mathbf{q}$. When the momentum goes to zero, the phonon periodicity tends to infinity: this has consequences for both acoustic and optical phonons. The acoustic phonons have a linear dispersion for small momenta, and respect the acoustic sum rule (which is usually enforced by hand in numerical calculations to avoid spurious negative frequencies):

$$\sum_j C^{Ii\alpha}_{j\beta}(|\boldsymbol{\tau}_I| \to \infty) = 0. \tag{1.102}$$

This is a consequence of the translational invariance of the crystal, that must allow for zero–frequency, rigid translations.

Let us consider a non–2D material (i.e. any crystal that is not just a 2D atomically thin sheet) which is also polar. The latter condition is realised when there is a permanent dipole moment within its unit cell: this can happen for example when there is a difference in electronegativity between the atoms in the crystal basis, and the effects of the resulting spatially inhomogeneous charge distribution are not cancelled by symmetries. In this case, when $\mathbf{q} \to \mathbf{0}$, a macroscopic electric field arises due to dipole–dipole interactions. The optical modes, whose frequency at $\mathbf{q} = \mathbf{0}$ is finite, will be affected: $\omega_{0\lambda}$ will depend on the direction along which the momentum goes to zero. This induces a splitting between longitudinal and transverse modes at small $\mathbf{q}$, called LO–TO splitting, with the LO frequency becoming larger. This splitting can only be captured by adding a non–analytic part C^{NA} to the matrix C, so that

$C \to C(\mathbf{q}=\mathbf{0}) + C^{NA}(\mathbf{q}\to\mathbf{0})$.[6] We list below the corresponding expression for completeness:

$$C^{NA\,i\alpha}_{\quad j\beta}(\mathbf{q}\to\mathbf{0}) = \frac{4\pi}{V}\frac{(\mathbf{q}\cdot\mathbf{Z}^*_{i\alpha})\,(\mathbf{q}\cdot\mathbf{Z}^*_{i\alpha})}{\sum_{\alpha\beta} q_\alpha \varepsilon^\infty_{\alpha\beta} q_\beta}. \tag{1.103}$$

Here $\varepsilon^\infty_{\alpha\beta}$ is the electronic contribution to the dielectric permittivity tensor and $\mathbf{Z}^*_{i\alpha} = \partial F_{i\alpha}/\partial\mathbf{E}$ the Born effective charge (also computed via DFPT); $\mathbf{E}$ is the macroscopic electric field. There is an analogy between the LT phonon splitting and the appearance of a LT *exciton* splitting at $\mathbf{q}=\mathbf{0}$. Recall Section 1.5. Some of the (purely transverse) degenerate excitonic states resulting from the excitonic Hamiltonian for $\overline{\chi}$ are split if the one for χ is used instead. This is because χ contains the longitudinal, macroscopic component of the Coulomb interaction. However "longitudinal" excitons cannot be seen in absorption experiments because the incoming field is purely transverse. When a longitudinal perturbation is applied, instead, some states will split depending on the direction along which $\mathbf{q}$ approaches $\mathbf{0}$.[18]

1.6.4 Quantisation of the phonons

For most applications where electron–phonon coupling is needed, the quantisation of lattice vibration must be considered (for example, in order to account for zero–point motion effects). We do this by transforming the normal coordinates $\overline{z}$ in Eq. (1.99) into operators, according to the following relation:

$$\hat{\overline{z}}^{\lambda q}(t) = \sqrt{\frac{\hbar}{2\omega_{q\lambda}}}\left[\mathrm{e}^{-\mathrm{i}\omega_{q\lambda}t}\hat{b}_{q\lambda} + \mathrm{e}^{\mathrm{i}\omega_{q\lambda}t}\hat{b}^\dagger_{-q\lambda}\right], \tag{1.104}$$

where $\hat{b}_{q\lambda}$ ($\hat{b}^\dagger_{q\lambda}$) is the annihilation (creation) operator of a phonon with momentum q and branch index λ. These operators follow the commutation relations $[\hat{b}_{q\lambda}, \hat{b}^\dagger_{q'\lambda'}] = \delta_{qq'}\delta_{\lambda\lambda'}$ and $[\hat{b}_{q\lambda}, \hat{b}_{q'\lambda'}] = [\hat{b}^\dagger_{q\lambda}, \hat{b}^\dagger_{q'\lambda'}] = 0$, and allow to rewrite the ionic Hamiltonian (minus the constant energy shift) as a collection of quantum harmonic oscillators:

$$\hat{\mathcal{H}}_{ion} = \sum_{q\lambda}\hbar\omega_{q\lambda}(\hat{b}^\dagger_{q\lambda}\hat{b}_{q\lambda} + \frac{1}{2}). \tag{1.105}$$

We can thus see that in the BO and harmonic approximations, lattice vibrations are quantised as perfect bosons.

The definition for the phonon propagator is then that of a bosonic Green's function. In the present treatment, the field operators we are concerned with are then the atomic displacement operators $\hat{u}_{Li\alpha}(t)$ obtained combining Eqs. (1.99) and (1.104). We then write the propagator as (compare with Eqs. (1.17) and (1.23))

$$\begin{aligned}
&\mathrm{i}\mathcal{D}(Li\alpha,t;Mj\beta,t') = \langle\Omega_T|\,\hat{\mathcal{T}}\left[\hat{u}_{Li\alpha}(t)\hat{u}^\dagger_{Mj\beta}(t')\right]|\Omega_T\rangle \\
&= \frac{1}{N}\sum_{\substack{\lambda q\\ \lambda' q'}} \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\boldsymbol{\tau}_L}\mathrm{e}^{-\mathrm{i}\mathbf{q}'\cdot\boldsymbol{\tau}_M}\overline{\xi}^{\lambda q}_{i\alpha}\overline{\xi}^{\lambda' -q'}_{j\beta}\,\langle\Omega_T|\,\hat{\mathcal{T}}\left[\hat{\overline{z}}^{\lambda q}(t)\hat{\overline{z}}^{\lambda q\dagger}(t')\right]|\Omega_T\rangle \\
&\equiv \frac{1}{N}\sum_{\substack{\lambda q\\ \lambda' q'}} \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\boldsymbol{\tau}_L}\mathrm{e}^{-\mathrm{i}\mathbf{q}'\cdot\boldsymbol{\tau}_M}\overline{\xi}^{\lambda q}_{i\alpha}\overline{\xi}^{\lambda' -q'}_{j\beta}\,\mathrm{i}D_{\lambda q}(t,t').
\end{aligned} \tag{1.106}$$

[18] furthermore, new collective excitations (plasmons, i.e., electron density oscillations) will become visible and can be captured using χ instead of $\overline{\chi}$.

Note that we have defined $\overline{\xi}_{i\alpha}^{\lambda q} \equiv \sqrt{\hbar/2M_i\omega_{q\lambda}}\xi_{i\alpha}^{\lambda q}$ for brevity. Here, we have considered explicitly a temperature–dependent phonon ground state, Ω_T, as occupation factors determine the phonon populations and therefore affect the probability of electron–phonon scattering processes. We will deal with the non–interacting bosonic propagator written in the basis of normal modes (where it is diagonal), which we called $D_{\lambda q}$. It assumes the following form

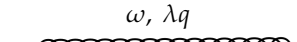
$$\begin{aligned} \mathrm{i}D_{\lambda q}(t_{21}) = (1+n_{q\lambda}(T)) & \left[\theta(t_1-t_2)\mathrm{e}^{-\mathrm{i}\omega_{q\lambda}(t_1-t_2)} + \theta(t_2-t_1)\mathrm{e}^{\mathrm{i}\omega_{q\lambda}(t_1-t_2)}\right] \\ n_{q\lambda}(T) & \left[\theta(t_1-t_2)\mathrm{e}^{\mathrm{i}\omega_{q\lambda}(t_1-t_2)} + \theta(t_2-t_1)\mathrm{e}^{-\mathrm{i}\omega_{q\lambda}(t_1-t_2)}\right], \end{aligned} \tag{1.107}$$

and its Fourier transform is

$$D_{\lambda q}(\omega) = [1+n_{q\lambda}(T)]\frac{2\omega_{q\lambda}}{\omega^2-\omega_{q\lambda}^2+\mathrm{i}\eta} - n_{q\lambda}(T)\frac{2\omega_{q\lambda}}{\omega^2-\omega_{q\lambda}^2-\mathrm{i}\eta}. \tag{1.108}$$

These equations contain the Bose–Einstein distribution function for phonons, $n_{q\lambda} = (\mathrm{e}^{\omega_{q\lambda}/k_BT}-1)^{-1}$. It is worth to point out that the first term in the above equation will be involved in electron–phonon scattering processes where a phonon is emitted by the electron system, while the second term will describe phonon absorption. We can see that even at zero temperature – when no phonons are present in the system – the quantum zero–point motion always allows for the probability of a *spontaneous phonon emission*. In general these zero–point contributions can be very important for the correct description of the electronic system (e.g. they may cause a reduction in the band gap energy with respect to the DFT or GW values). Furthermore, once the effects of quantum lattice vibrations have been incorporated into the electronic system at $T = 0$ (via the perturbation theory of electron–phonon interaction), the extension to finite T is often trivial.At $T = 0$, the phonon propagator becomes (compare with Eqs. (1.25) and (1.44))

$$D_{q\lambda}(\omega) = \frac{1}{\omega-\omega_{q\lambda}+\mathrm{i}\eta} - \frac{1}{\omega+\omega_{q\lambda}-\mathrm{i}\eta} = \quad . \tag{1.109}$$



This propagator will act on the electronic system as an effective dynamical interaction, in a similar way to the screened electron–electron interaction, and will produce its own kind of quasiparticle corrections. Because it permits a finite momentum transfer between electron states, it will also lead to the appearance of "diagonal" electronic transitions (for example finite–q eh pairs). Finally, a static limit to $D_{q\lambda}(\omega)$ can be defined in this way:

$$D_{q\lambda}^{st}(t_{21}) \equiv D_{q\lambda}(\omega=0)\delta(t_{21}) = -\frac{2}{\omega_{q\lambda}}\delta(t_2-t_1). \tag{1.110}$$

1.6.5 Electron–phonon coupling

We are interested in describing the coupling between a Kohn–Sham electron and a phonon mode. Therefore, we take the Kohn–Sham Hamiltonian, $H^{KS}(\{\mathbf{R}^0\}) = -\boldsymbol{\nabla}_{\mathbf{r}}^2/2 + v_{eff}(\mathbf{r},\{\mathbf{R}^0\})$ and we expand the effective potential to first order in the lattice displacements. Now, the electron–phonon coupling Hamiltonian will be given by

$$H_{ep}^{(1)} = \sum_{Li\alpha}\frac{\partial v_{eff}}{\partial R_{Li\alpha}}u_{Li\alpha}. \tag{1.111}$$

This expression can be written in terms of the normal coordinates in second quantisation as

$$H_{ep}^{(1)} = \sum_{nm\lambda kq} g_{nmk}^{\lambda q} \hat{a}_{nk}^{\dagger} \hat{a}_{mk-q} (\hat{b}_{\lambda q} + \hat{b}_{\lambda -q}^{\dagger}) \tag{1.112}$$

(note, for completeness, that the KS Hamiltonian in second quantisation is $H^{KS} = \sum_{nk} \epsilon_{nk} \hat{a}_{nk}^{\dagger} \hat{a}_{nk}$). Here, the electron–phonon coupling (EPC) matrix element g must be defined: it controls the strength of the electron–phonon interaction and plays the same role that the screening matrix element (Eq. (1.46)) plays in the case of the electron–electron interaction. It is given by, using the notation by Giustino [49],

$$\begin{aligned} g_{nmk}^{\lambda q} &\equiv g_{nm}^{\lambda}(\mathbf{k}, \mathbf{k}-\mathbf{q}) = \sum_i \frac{\boldsymbol{\xi}_i^{\lambda q} \cdot \boldsymbol{\gamma}_{mni}(\mathbf{k}, \mathbf{k}-\mathbf{q})}{\sqrt{2M_i \omega_{q\lambda}}}, \\ \boldsymbol{\gamma}_{mni}(\mathbf{k}, \mathbf{k}-\mathbf{q}) &= \sum_L \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot(\mathbf{r}-\boldsymbol{\tau}_L)} \langle \overline{nk} | \frac{\partial v_{eff}}{\partial \mathbf{r}_i}\Big|_{\mathbf{r}-\boldsymbol{\tau}_L} | \overline{mk-q} \rangle . \end{aligned} \tag{1.113}$$

Here the overlines in the bra and ket states indicate that the integral is evaluated in the unit cell of the crystal using the lattice–periodic parts $u_{nk}(\mathbf{r})$ of the KS wave functions (Eq. (1.12)). Note that this complicated expression can be simply recast as

$$g_{nmk}^{\lambda q} = \langle \overline{nk} | \, \Delta_{q\lambda} v_{eff} \, | \overline{mk-q} \rangle , \tag{1.114}$$

which emphasises its physical meaning, by defining an effective derivative $\Delta_{q\lambda}$. By evaluating higher–order derivatives of v_{eff} we can in principle compute higher–order couplings, however these are usually neglected apart from the second–order term

$$\Lambda_{mn}^{q\lambda, q'\lambda'} = \frac{1}{2} \langle \overline{nk} | \, \Delta_{q\lambda} \Delta_{q'\lambda'} v_{eff} \, | \overline{mk-q-q'} \rangle \propto \sum_i \langle \overline{nk} | \frac{\partial^2 v_{eff}}{\partial \mathbf{r}_i \partial \mathbf{r}_i'} | \overline{mk-q-q'} \rangle . \tag{1.115}$$

This term plays a role in the correction of the band energies, appearing in Eq. (1.117) below. In particular, we will be interested only in the diagonal second–order coupling ($m \to n$, $\mathbf{q}' = -\mathbf{q}$, $\lambda' \to \lambda$), which can be obtained by the first–order coupling at $\mathbf{q} = \mathbf{0}$.[69]

A final remark is in order about which kind of physics is captured by the calculation of the EPCs from a DFT standpoint. Clearly, this is a vast improvement with respect to the "bare" coupling given simply by $\partial_R v_R$ (the derivatives with respect to the bare electron–ion interaction), because it includes corrections due to the interacting nature of the electron system (i.e. screening and "vertex" corrections), albeit at the DFT level. Yet, Eq. (1.114) does not coincide with the *true* result for a first–order EPC coming from a rigorous many–body treatment, because the electronic screening is only at the static RPA level and the non–local, dynamical vertex corrections become local and static in the approximations for v_{xc}. A detailed description of these subtleties goes beyond the scope of this thesis and can be found in Ref. [49]: in this thesis, we will only deal with DFT–based EPCs.

1.6.6 Quasiparticle corrections II: electron–phonon interaction

The quasiparticle correction of the electronic band structure due to electron–phonon interaction is treated along the same lines as that for the electron–electron interaction (*GW* approximation) seen in Sec. 1.3.3. This time, however, instead of the electronic screening we have the phonon propagator, which is the displacement–displacement

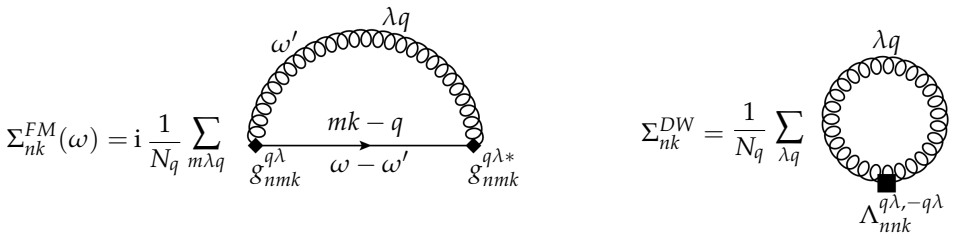


FIGURE 1.7: Schematic representation of the Fan–Migdal (left) and Debye–Waller (right) contributions to the electron–phonon self–energy (Eqs. (1.116) and (1.117) respectively. The black diamond and square represent the electron–phonon coupling matrix elements.

correlation function and permits the description of the polarisation of the electronic system due to lattice vibrations. We obtain a Dyson equation for the quasiparticle Green's function G in terms of the KS Green's function G^0 and the electron–phonon self–energy. Since we are using the bare phonon propagator of Eq. 1.108, instead of the screened Coulomb interaction, now the frequency integral in Eq. 1.50 can be explicitly evaluated in the electron–phonon case, leading in the KS basis to

$$\Sigma^{FM}_{nk}(\omega;T) = \frac{1}{N_q}\sum_{m\lambda q}|g^{q\lambda}_{nmk}|^2\left[\frac{n_{q\lambda}(T)+1-f_{mk-q}}{\omega-\epsilon_{mk-q}-\omega_{q\lambda}-\mathrm{i}\eta}+\frac{n_{q\lambda}(T)+f_{mk-q}}{\omega-\epsilon_{mk-q}+\omega_{q\lambda}-\mathrm{i}\eta}\right]. \tag{1.116}$$

This is the Fan–Migdal (FM) self–energy,[70, 71] expressed as the first–order dynamical correction to the electronic state (nk) due to electron–phonon interaction. It features a sum over the electronic states m as well as over the phonon branches and momenta. In most cases, an additional self–energy term is found to add a non–negligible shift to the bands: this is a second–order, static correction called Debye–Waller (DW) self–energy[72] and given by

$$\Sigma^{DW}_{nk}(T) = \frac{1}{N_q}\sum_{\lambda q}\Lambda^{q\lambda,-q\lambda}_{nnk}\left[2n_{q\lambda}(T)+1\right]. \tag{1.117}$$

The DW term is evaluated with the second–order, diagonal EPC matrix element. The electron–phonon self–energy is schematically depicted in Fig. 1.7. These expressions for the electron–phonon self–energy constitute the so–called Allen–Heine–Cardona (AHC) theory.[73–75] We proceed in the same way as in the GW case and we obtain for the QP Green's function G, spectral function A, and QP energies E the following expressions:

$$\begin{aligned} G_{nk}(\omega,T) &= \left[\omega-\epsilon_{nk}-\Sigma^{FM}_{nk}(\omega,T)-\Sigma^{DW}_{nk}(T)\right]^{-1}, \\ A_{nk}(\omega,T) &= \frac{1}{\pi}\mathrm{Im}G_{nk}(\omega,T), \\ E_{nk}(T) &= \epsilon_{nk}+Z_{nk}(T)\left[\Sigma^{FM}_{nk}(\epsilon_{nk},T)+\Sigma^{DW}_{nk}(T)\right] \end{aligned} \tag{1.118}$$

where the renormalisation factor Z is defined according to Eq. (1.55).

By computing the electron–phonon self–energy, we are able to access the temperature dependence of the band structure and most importantly of the band gap, as well the electron–phonon scattering lifetimes ($\propto [\mathrm{Im}\Sigma^{FM}]^{-1}$) for the electronic states, which will depend on the strength of the coupling $|g|^2$ to the various phonon

modes and on their occupation number ($n_{q\lambda}$). It is important to remark that a significant effect may also be seen at $T = 0$, due to quantum zero–point motion. From the computational point of view, generally the EPC matrix elements $g^{q\lambda}_{nmk}$ and phonon frequencies $\omega_{q\lambda}$ are evaluated thanks to DFPT,[19] while the self–energy is computed with YAMBO. The convergence of the summation over electronic states in Eq. (1.116) is not too strict: as a rule of thumb, the states m contained in an energy window of a few LO phonon frequencies above and below the (nk) state to be corrected are sufficient. The convergence of the q–summation is instead more problematic, with the self–energy expressions easily becoming numerically unstable at small q, where the acoustic phonon frequencies are going to zero (note that at $q = 0$, the acoustic modes are removed from the summation. For more details, see Ref. [76]). In general, a partial cancellation between the DW and FM terms at small q tends to alleviate this problem, and in addition techniques to correct this issue are used: for example the interpolation of the EPC matrix elements to a very fine q–mesh,[77] or using a large number of randomized q–points instead of a regular grid.

[19] If a large number of q–points is involved, the calculations are better managed with appropriate scripts that allow for control over each step; one example would be the automatic resubmission in case of a failure at a specific q, while the rest of the calculation proceeds without exiting.

Chapter 2

Boron nitride and excitons

This chapter is partly based on our publications [39] and [40]. Some of the text and figures contained in this Chapter are adapted from these references.

In this chapter, we are going to look at the excitonic properties of BN systems, from monolayer to bulk. After characterising the excitonic series in these systems in terms of symmetry and optical activity, we will study in detail how layer stacking impacts properties like the energy gap and the absorption spectrum. Then, we will broaden our analysis by solving the BSE at finite momentum for bilayer, trilayer and bulk and investigating the resulting exciton dispersion curves.

2.1 Hexagonal Boron nitride

The experimental lattice parameters in the bulk material are $a = 2.496$ Å (in–plane) and $c = 3.305$ Å (interlayer distance).[78] The values corresponding to freestanding multilayer and monolayer systems are currently not known exactly. Previous first–principle calculations have shown that the quasi–particle band gap of bulk hBN is around 6.5 eV[21], so that this material is often considered a "large gap semiconductor".[14, 19] This is due to the difference in electronegativity between nitrogen and boron atoms: the highest–energy valence electron tends to be localized around the nitrogen sites, and likewise the Bloch function corresponding to the bottom of the conduction band is localised on the boron atoms, as shown in Fig. 2.1.

We start our analysis with the band structure. Figure 2.1 (right side) shows that, unlike graphitic systems, the band–edge KS / Bloch wave functions in BN have a strong overlap ($> 99\%$) with the atomic p_z orbitals they originate from in a simple tight–binding picture. The relevant energy bands at the BZ edge are indeed of π (valence) and π^* (conduction) character, and may be conceptualised as localised p_z orbitals along the full KM symmetry direction and up to the crossing with the σ and σ^* bands around the midpoint of the ΓM and ΓK lines (see the band structures in the left column of Fig. 2.2). This permits us to think of an optical excitation from valence to conduction band as a "hopping" in real space from N to B atoms.[39] Although the name "Van der Waals materials" is often employed as a synonym to "layered materials", it is important to emphasise that even though the Van der Waals interaction is crucial for the bonding of the stacked layers, there is also a non–negligible amount of hybridization between vertically aligned p_z orbitals. Its importance depends on the stacking order[79] and in this thesis we will be concerned with the most stable one, the so–called AA′ stacking, where nitrogen and boron atoms sit exactly on top of each other in alternating order. In this case, the p_z hybridization lowers the bottom of the conduction band at the M point, making the quasiparticle band gap indirect (approximately from K to M) for all systems except the monolayer, where it is direct at K (this band gap transition is a common feature in many semiconducting

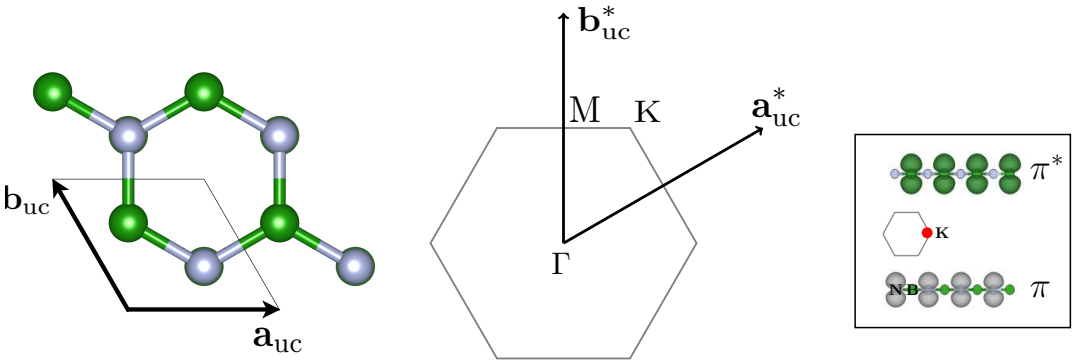


FIGURE 2.1: Hexagonal boron nitride. Boron atoms are green, nitrogen ones are white. Left: hexagonal unit cell (top view). Center: hexagonal BZ (Wigner–Seitz cell) with high–symmetry points (top view). Right: the Kohn–Sham Bloch wave intensities for the monolayer, at the valence band maximum (bottom) and conduction band minimum (top), both at the K point, appear as B/N–localised p_z atomic orbitals, respectively.

layered materials). Additionally, increased screening has the effect of lowering the quasiparticle gap with the addition of each new layer (this will be the case also for exciton binding energies) . The energy difference between the direct and indirect gap is largest in the bulk material, where it is around 0.5 eV (Fig. 2.3).

From Fig. 2.2 (left column) it may also be seen that a large number of parabolic conduction bands are present around the Γ point. Upon the GW correction, the bottom of the conduction band actually seems to lie at the Γ point, in correspondence with the σ^* bands, for all systems except bulk, meaning that the QP correction for the σ–type bands is lower than the one for the π–type ones. This feature is obtained also if hybrid functionals are used instead of the GW approximation.[80, 81] These states are a combination of: (i) states with mixed σ^* and π^* character (this happens after the σ^*–π^* band crossing), which show a "nearly–free–electron" (NFE) localisation type and can retain up to 30% π^* character; (ii) vacuum states that slide down in energy due to the high amount of vacuum space included in the computational supercell (if the intersystem separation along the stacking direction where infinite, they would form a free–electron continuum). As the density of electronic states around Γ increases in the case of multilayers, many (avoided) band crossings start to appear, leading to band mixing. In these cases, our G_0W_0 calculation leads to an unnatural steepness of some bands (see for example Fig. 2.2(d)). We believe that in order to accurately reproduce the bands in this region of the BZ a fully self-consistent GW calculation (where the electronic wave functions are also updated, instead of being kept fixed at the DFT–LDA level) should be performed. However, for now we shall not be concerned with these states since they do not participate in the optical absorption process (i.e. direct, $q \to 0$ transitions). This is because they are either forbidden by selection rules ($\pi \to \sigma^*$, $\sigma \to \pi^*$ transitions) or the actual excitonic states are well above the QP band gap energy ($\sigma \to \sigma^*$), or the weight of the transitions is negligible (transitions to NFE, vacuum states). The region in BZ that contributes the most to optical absorption is the one along KM ($\pi \to \pi^*$), where the conduction bands are almost flat and consequently the density of states is large, although the high-energy region up to the π^*-σ^* crossing is also relevant (more details about the relevant transition energy window are given in App. C). The NFE might nonetheless become relevant if non–vertical (i.e. indirect, $q \neq 0$) transitions are considered (e.g. from

K to Γ): therefore we will discuss such states in more detail (including "NFE–type" excitons) in Sec. 2.4. We will now turn to the study of excitons in few–layer systems: the BSE absorption spectra from monolayer (1L) to five–layer (5L) hBN are shown in Fig. 2.2, right column).

Considering the poor screening of the electron–hole interaction, we expect the optical properties of hBN to be dominated by excitonic effects, as it is indeed found both experimentally and theoretically. In particular, existing theoretical calculations predict strongly bound direct excitons: in bulk hBN the exciton binding energy is around 700 meV.[21, 23, 27] Quasi–2D, few–layer systems display much stronger excitonic effects with respect to their bulk counterpart due to the electronic screening being reduced even more. Additionally, there is a complex interplay between optically active (bright) and inactive (dark) excitons in hBN–based systems[21, 23, 82] that we will elucidate in the next Section (the BSE absorption spectra from monolayer (1L) to five–layer (5L) hBN are shown in Fig. 2.2, right column). Experimentally, there has been much debate around the direct or indirect nature of the *optical* gap in bulk hBN.[14, 19, 24, 83, 84] The most recent works on the subject present a case for phonon–assisted creation of an indirect exciton: we will deal with this in the next Chapter. For now, it is sufficient to summarise our findings in the following way: an *excitonic fine structure* (i.e. the presence of multiple spectral peaks in a small energy window) characterizes the lowest–bound hBN excitations, appearing because of two independent reasons: (i) Davydov splitting of the excitonic levels (Sec. 2.3) and (ii) phonon–assisted sidepeaks (Chapter 3).

2.2 Optical excitations: monolayer

Let us first take a look at the excitonic series in the monolayer,[39] as seen in Fig. 2.2(e). This systematic analysis will establish the framework that we will use to look at excitations in more complicated systems, as well as showcase some useful general properties of excitons in BN sheets.

The DFT–LDA direct band gap is 4.5 eV but the G_0W_0 correction brings it to 7.3 eV. Analogously, a giant binding energy of 1.9 eV is observed for the lowest–bound exciton (main peak in the absorption spectrum) upon solution of the BSE. This is a doubly degenerate exciton, whose wave function intensity (Eq. (1.88)) is shown in Fig. 2.4(1) in real space (left), while the exciton weights in reciprocal space are displayed on the right. In real space, we fix the hole position $\boldsymbol{r}_h$ just above a nitrogen atom: this is physically the most relevant position as can be seen from Fig. 2.1 (right side). The plot then represents the probability to find the electron at position $\boldsymbol{r}_e$ if the hole is located at $\boldsymbol{r}_h$. As expected, the electron density is centered on the boron atoms, with a high probability — about 30% — on the first nearest neighbours. This means that the exciton appears well localised in real space (so that we might expect it to be delocalised in reciprocal space), however longer–range low–probability tails are also present. Therefore, when we look at the relevant transitions in reciprocal space, we see that most of the excitonic weights come from the area around the K point (i.e. the direct band gap), while the BZ edge segments also give non–negligible contributions. Note that the presented shape of the ground–state exciton remains the same when we pass to few–layer and bulk systems, where electron and hole are found to remain mostly confined to the same layer (*in–plane* ip exciton).

The second peak in the absorption spectrum corresponds to a doubly degenerate exciton with a 1 eV binding energy, then we have two non–degenerate *dark* states, and finally the third peak is again doubly degenerate. All these excitons (number

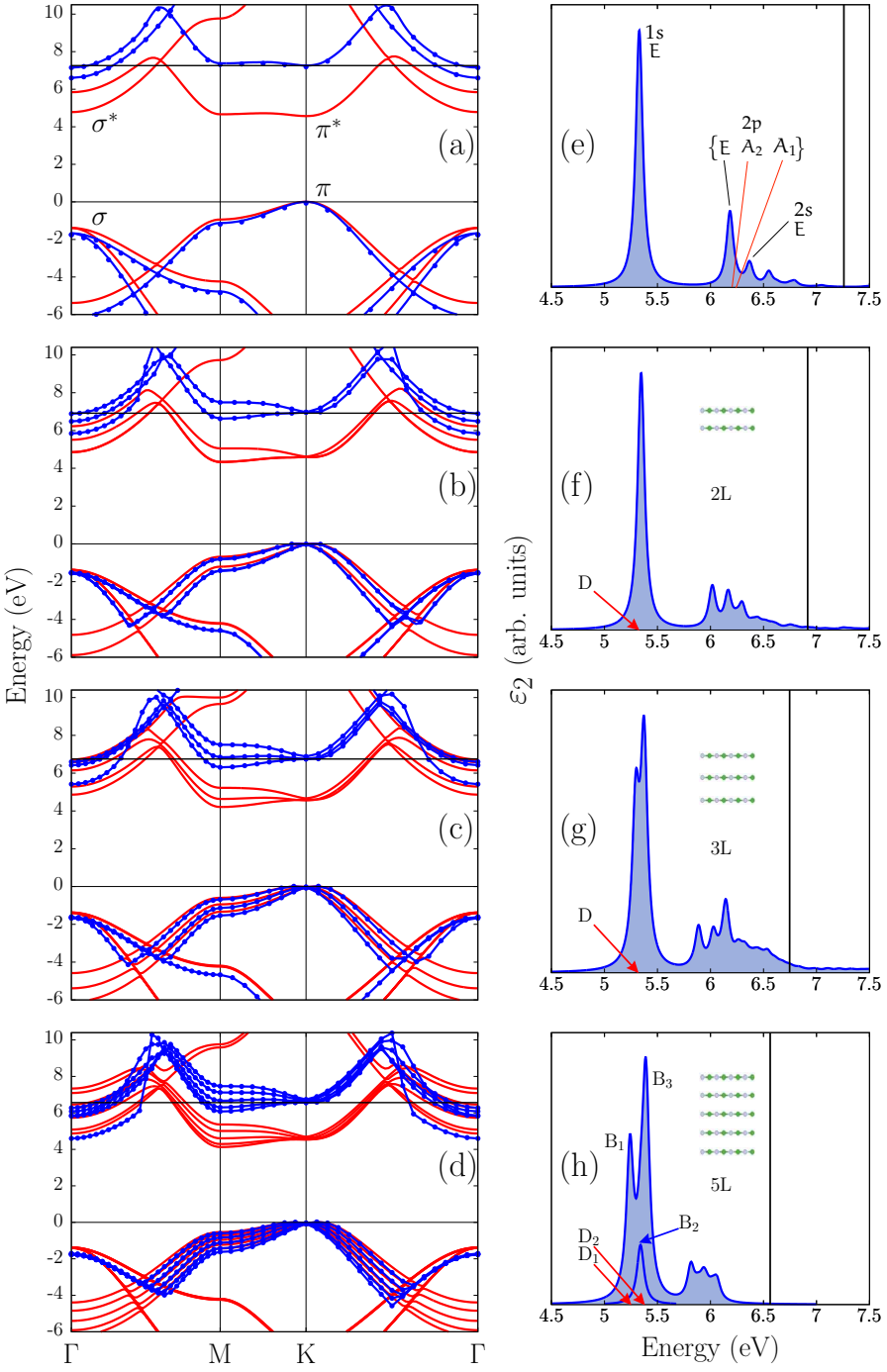


FIGURE 2.2: Many–body results for monolayer [(a),(e)], bilayer [(b),(f)], trilayer [(c),(g)] and pentalayer [(d),(h)] hBN. Left panels: DFT–LDA (red) and GW (blue) band structures. In (a) the π and σ bands are labeled. Right panel: imaginary parts of the dielectric functions. The vertical lines represent the GW direct band gaps. The red arrows in (f), (g), (h) indicate the positions of low–energy dark (D) excitonic states. In (e), the first five excitons are labeled according to their symmetry representations (see text and Ref. [39]). In (h), the bright peaks (B) are labeled for later comparison with Fig. 2.7, and an additional bright exciton (B_2), which is hidden in the main two–peak structure, is highlighted. Note that in the case of the pentalayer, the BSE was solved for states only up to the energy of ~ 6 eV.

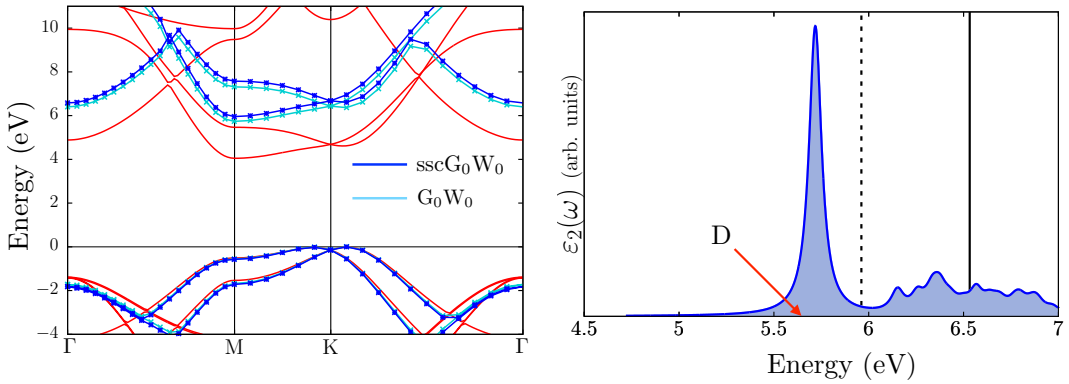


FIGURE 2.3: Many–body results for bulk hBN. Left: DFT–LDA (red), G_0W_0 (light–blue) and "semi–self–consistent GW" ($sscG_0W_0$, blue, see text on page 59) band structures. Right: imaginary part of the dielectric function. The indirect and direct GW band gaps are labeled with a dashed and solid black vertical line, respectively, and the red arrow labels the dark (D) partner to the main bright exciton.

1 to 5) are represented in real and reciprocal space on the left and right columns of Fig. 2.4, respectively. Since these are all excitons originating from $\pi \rightarrow \pi^*$ band transitions, the electron density is always mostly localised on boron atoms. The first observation that we can make is that broadly speaking the excitons become more delocalised in real space / localised in reciprocal space with decreasing binding energies. The ground–state exciton is more localised (a few unit cells); the other excitons, lying 0.9 eV above and therefore much more weakly bound, extend across multiple unit cells. In order to model the excitonic series in BN systems, our collaborators[1] developed a tight–binding excitonic model (TBEM) using localised atomic orbitals and a purely 2D electrostatic potential, called Keldysh potential.[85, 86] The model requires fitting from *ab initio* data only for a few parameters.[2] The TBEM results are in remarkable agreement with the first–principle calculations and, among other things, permit us to shed more light on the symmetries of the excitonic wave functions. More information can be found in Refs. [39, 40].

We can indeed see from Fig. 2.4 that the charge distribution of the BN excitations is determined by the underlying symmetry of the crystal lattice. Starting now, simple group–theoretical arguments for the description of excitonic states will be made throughout this thesis. This is not commonly done in the case of excitons, at least when first–principle calculations are involved: we will nonetheless see that information about the optical activity of the excitations, and even about their possible coupling with phonon modes, can be obtained before running any computer simulation and provide a clear reference to assess the validity of *ab initio* results. For the 2D BN sheet, the point group symmetry is C_{3v} which allows for three irreducible representations: the identity representation A_1, another monodimensional representation, A_2, which differs from A_1 because it is odd with respect to σ_v reflections, and the two–dimensional representation E. Since in our plots the hole is fixed in a symmetrically invariant position, we can use them to assess the symmetry of the full exciton wave functions. We can immediately assign excitons number 1, 2 and 5, which are doubly degenerate, to the E representation. Among the non–degenerate states, the identity

[1] Thomas Galvani, Hakim Amara, Sylvain Latil, and François Ducastelle.

[2] The fitting parameters are the excitonic hopping integral and the screening length of the Keldysh potential. In the case of multilayers, two additional parameters are needed, being the interlayer hopping integral and interlayer screening length.

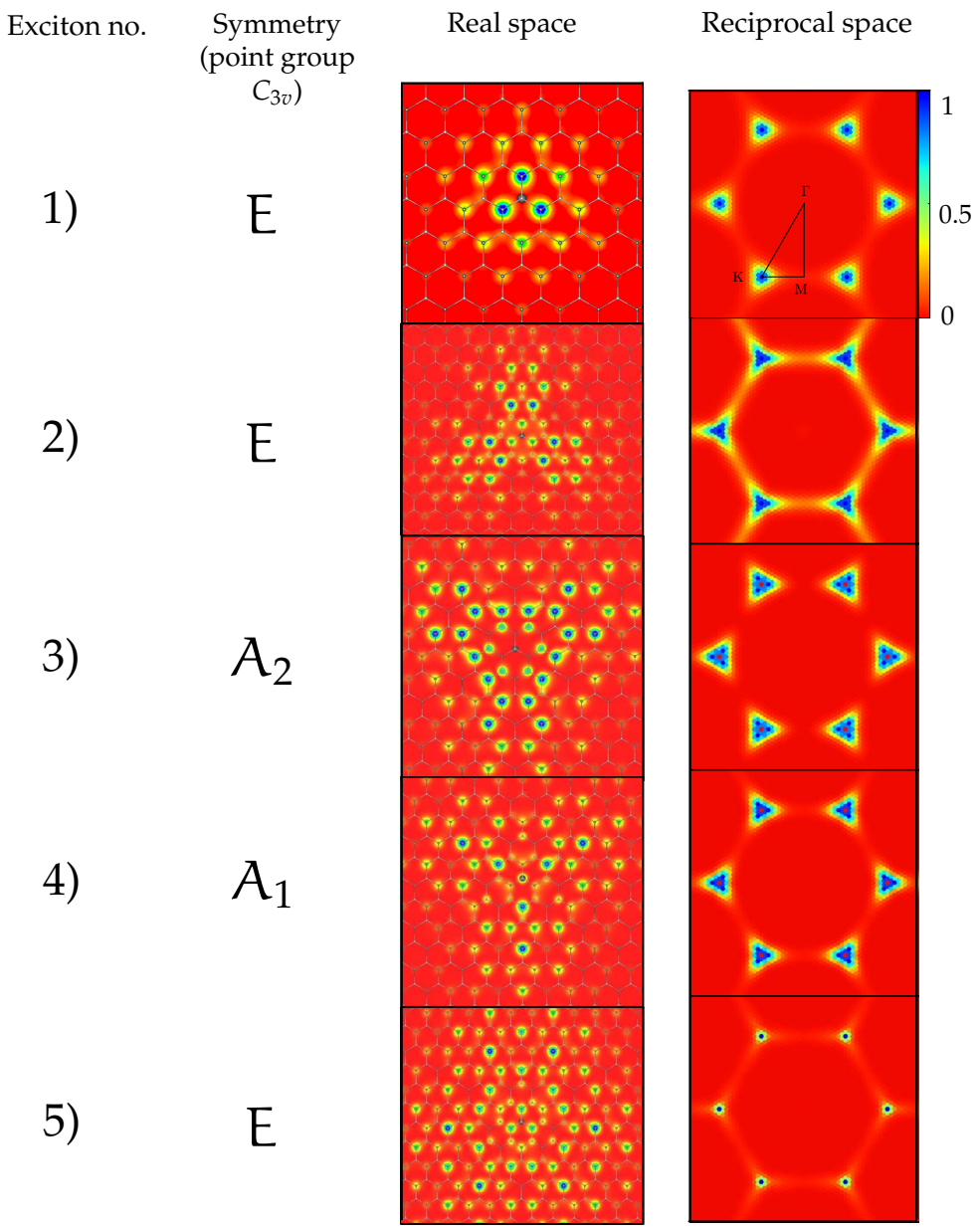


FIGURE 2.4: First five excitons in the monolayer hBN series. The numerical ordering goes from lowest– to highest–bound. Left column: wave function intensities in real space $|\psi_\lambda(\mathbf{r}_e)|^2$ (the hole is fixed above a N atom). Right column: exciton weights $\sum_{cvk} |\overline{A}_\lambda^{vck}|^2$ in reciprocal space. The symmetry representation of each state according to the lattice point group C_{3v} is also listed.

A_1 is the only one that can allow some charge density on the hole site (because it is even with respect to σ_v mirrors), therefore exciton number 4 transforms like A_1. Finally, exciton number 3 will transform as A_2 (these symmetries are confirmed by the TBEM). It is instructive to note that since the repulsive exchange contribution to the BSE kernel acts at short–range, exciton A_1, where electron and hole may overlap, is the one most affected by it. Indeed, by rerunning the BSE calculations with the exchange term set to zero, we find that the binding energy of excitons 1, 2, 3 and 5 is negligibly affected (therefore it may even be safely neglected if average accuracy is required). In the case of exciton 4, however, the exchange contribution accounts for

a 28% reduction in the binding energy.

A final remark concerns the optical activity of the excitations. In our calculations, since we are considering ultra–thin materials, the incoming electro–magnetic field is always polarised in the plane of the layer, which means that the incoming wave vector is parallel to the stacking direction, so that the light impinges on the layer plane directly from above. Hexagonally symmetric systems are isotropic in the plane, therefore the x and y light polarisation directions are equivalent. This means that the dipole matrix element transforms like the E representation, and acting on the exciton "vacuum" state,[3] which has the full A_1 symmetry, can only excite E states. Indeed, all the visible peaks in the monolayer absorption spectrum are doubly degenerate E excitons.

2.3 Optical excitations: multilayers

We will now investigate multilayer systems, focusing on the lowest–bound E excitons.

Let us turn our attention to the right column of Fig. 2.2. We can check that when more layers are added, the main, doubly–degenerate "monolayer" excitons localised on each new layer have different energies. The absorption spectrum in the bilayer case (f) is similar to the one of the monolayer, but now a dark exciton (shown by the red arrow) appears before the main peak. This state is optically invisible, but is nonetheless present as a solution of the BSE. The bright and dark states form a "Davydov pair", which becomes a triplet in the trilayer case (g) with two bright excitons and a dark one in the middle. The pentalayer (h) shows two bright peaks as well, but a low–intensity third one (shown with a superimposed Lorentzian) is hidden between them. The "Davydov multiplet" is completed by the presence of two dark excitons (red arrows). Therefore, in the case of 3L and 5L, the dominant excitonic feature in the spectrum displays a visible fine structure.

The concept of Davydov splitting, originally developed to describe the energy levels in clusters of identical molecules,[87] can be applied to molecular crystals,[88] but then naturally also to layered materials consisting of identical layers stacked on top of one another. For example, Davydov splitting of phonon frequencies is observed in transition metal dicalchogenides few–layer systems.[89–92] Considering an ideal monolayer, we may take into account an excitonic state S with degeneracy m. If we start adding more layers to the system, but we keep them far enough from each other as to not interact, S becomes a state with degeneracy Nm where N is the number of layers. Now, if the N layers are brought closer together and start interacting, the degeneracies may be lifted by the interlayer coupling and we might have N m-fold degenerate states forming a "Davydov multiplet". In bulk hBN, for example (Fig: 2.3), we have a Davydov pair (as the number of layers per unit cell are equivalent to the case $N = 2$) with an energy separation of 0.06 eV and both with a large binding energy of 0.7 eV. However, only one state is optically allowed and contributes to the strong exciton peak in the absorption spectrum, while in the multilayer case dark and bright states are found to be alternating within the multiplets. We will discuss here the structure of the Davydov multiplets for increasing layer number and their optical activity; we will show that the lowest–lying excitations turn out to be localised on the surfaces.

[3] That is, when no excitons are present: all the valence bands are filled and all the conduction ones are empty.

Exciton	1 (×2)	2 (×2)	3	4	5 (×2)	6	7 (×2)	8
Bind. En.	−1.644	−1.614	−1.17	−1.16	−1.02	−1.00	−0.94	−0.90
Bright	no	yes	no	no	no	no	yes	no
Symm.	E_g	E_u	A_{1g}	A_{1u}	E_g	A_{2g}	E_u	A_{2u}
Local.	ip	ip	il	il	ip	il	ip	il

TABLE 2.1: Bilayer excitons. The binding energies are in eV. The optical activities and symmetries of the states are also listed, as well as their localisation in terms of being in–plane (ip) or interlayer (il).

2.3.1 Bilayer

The bilayer excitons can be characterised in the same way as for the monolayer. Their symmetry can be checked with the TBEM, but also directly from the *ab initio* result by exctracting the *phase* of the exciton wave functions, not just the intensity, in order to gather the full information about the parity with which they transform under certain symmetry operations. The results are presented in Table 2.1 up to exciton 8 of the series, and an explicit example of this procedure, both for doubly degenerate and non–degenerate states, will be given below.

We have the following two main differences with the monolayer case. (i) A new class of excitations is now possible, namely the *interlayer* il (or charge–transfer) excitons, where the interacting electron and hole are mostly confined on different layers. The wave function analysis shows that excitons 3, 4, 6 and 8 have il character, while excitons 1, 2, 5 and 7 have ip character. (ii) At variance with the monolayer case, the bilayer possesses the inversion symmetry $\mathcal{I}$ (its point group is called D_{3d}): all the original monolayer excitons have undergone Davydov splitting into pairs of even (g) and odd (u) states with respect to $\mathcal{I}$. The pairs need not be adjacent in energy and in fact are given by $(1,2)$, $(3,8)$, $(4,6)$ and $(5,7)$. Pairs $(1,2)$ and $(5,7)$ correspond to the splitting of the first two excitons of the monolayer. As for the optical activity, the dipole matrix element transforms as a vector, therefore it is odd under inversion (E_u). Consequently, only doubly degenerate E_u excitons with odd parity with respect to inversion symmetry can be bright. The first exciton (transforming as E_g) is thus dark, and the main peak of the absorption spectrum comes from the second state. Note that the same qualitative considerations about the first two excitons can be done for the bulk system as well.

If we look at the wave function intensities of the lowest–bound Davydov pair, shown in the top frames of Fig. 2.5(a) and (b), the two states appear indistinguishable (they have the shape of the lowest-bound E monolayer exciton). However, state $S = 1$ is optically dark, whereas state $S = 2$, which lies 0.06 eV above, is bright. This suggests that $S = 1$ should be *even* under inversion symmetry and $S = 2$ should be *odd*. Thus, the complete symmetry analysis requires to visualize the phase of the excitonic wave function.

Since these states are doubly degenerate, we start the analysis by presenting a simpler case, the non–degenerate (dark) state $S = 3$. Its intensity is shown in the top frame of Fig. 2.5(c). This is an il exciton: if the hole is fixed on one layer, the electron density is distributed on the other (which is the only layer shown in the Figure, labeled Layer 1). In the middle frame of Fig. 2.5(c) we present a phase–intensity plot of the same exciton.[4] The phase is remarkably constant on each atom and, as

[4]The values of the phase are shown in the areas with intensity greater than 5% of its maximum value.

expected, any two adjacent boron sites are separated by a node of the wavefunction (the phase difference is π). In the bottom frame of Fig. 2.5(c) we show the same plot, but now the hole is fixed at a position $r'_h = \mathcal{I}(r_h)$. The resulting electron density is now localized on the opposite layer (Layer 2) with respect to the previous case.

We can immediately see that the phase distribution does not change in the two cases: state $S = 3$ is even under inversion symmetry, and we can assign it to the A_{1g} representation of point group D_{3d} of bilayer hBN. In order to find its Davydov partner, we look for an il state with the same symmetry, but odd under inversion (i.e. belonging to representation A_{2u}). We find that it is state $S = 8$, represented in Fig. 2.5(d) and listed in Table 2.1, with a considerable Davydov splitting of 0.27 eV.

We are now ready to go back to the doubly–degenerate states $S = 1$ and $S = 2$. In order to fully represent the phase information, it is necessary to rotate the two complex wavefunctions in the degenerate subspace until they are (almost) fully real or fully imaginary.[5]

In Fig. 2.5(a) and (b) we select one such wave function for each state (panel (a) for $S = 1$ and (b) for $S = 2$), and we plot a linecut of the intensity along the three boron atoms that are nearest neighbours to the nitrogen above which the hole is fixed. These are the sites where most of the intensity is found. The value of the phase (which rotates along the linecut) is shown in a color scale. In analogy with panels (c) and (d), the corresponding wavefunctions under inversion symmetry are plotted in the bottom frames of Fig. 2.5(a) and (b). We also show sections of the phase–intensity plots for the leading peak in the insets. We can clearly see how $S = 1$ is even (E_g, optically forbidden) and $S = 2$ is odd (E_u, optically active) under inversion symmetry. This symmetry analysis is confirmed by the TBEM results.

2.3.2 Multilayers

The variation of the (π) band gaps in hBN as a function of layer number is displayed in the top two frames of Fig. 2.6(a). The indirect gap (orange) and minimimum direct gap (teal) are shown both in the DFT–LDA case (upper frame) and after the GW correction (lower frame). The energy of the bottom of the conduction band at M is lowered every time the number of hybridized layers is increased, reducing the indirect gap. On the other hand, the value of the direct band gap is only negligibly affected by layer stacking at the DFT level.

For both gaps the GW correction to the DFT values is huge ($\gtrsim$ 2 eV). As the screening environment evolves from quasi–2D to 3D with layer stacking, the GW gaps decrease, converging to the bulk value. In particular, in the case of the minimum direct gap (the leading contributor to optical absorption), the DFT calculation is completely unable to capture the increase in screening along the stacking direction with every added layer, giving a constant value of 4.56/4.53 eV from monolayer to bulk. After the GW correction, the gap in bulk (at 6.24 eV) is lower than the gaps in monolayer and pentalayer by 1 and 0.3 eV, respectively.

The two bottom frames of Fig. 2.6(a) are concerned with excitonic states. In the upper one, the binding energies of the lowest–bound Davydov multiplet are plotted in green (dark excitons are in gray). In monolayer and pentalayer, the binding energies are 1.93 and 1.32 eV, respectively, as opposed to 0.7 eV in the bulk. By looking at the absorption spectra, we can see that the effects due to the reduction in binding energy and to the shrinking GW band gap tend to cancel: in fact, the absolute peak

[5] In particular, if ψ_a and ψ_b are the two degenerate wave functions, with the transformation $(\psi_a \pm \psi_b)/\sqrt{2}$ both of the new functions behave in the same way with respect to inversion symmetry as the full exciton.

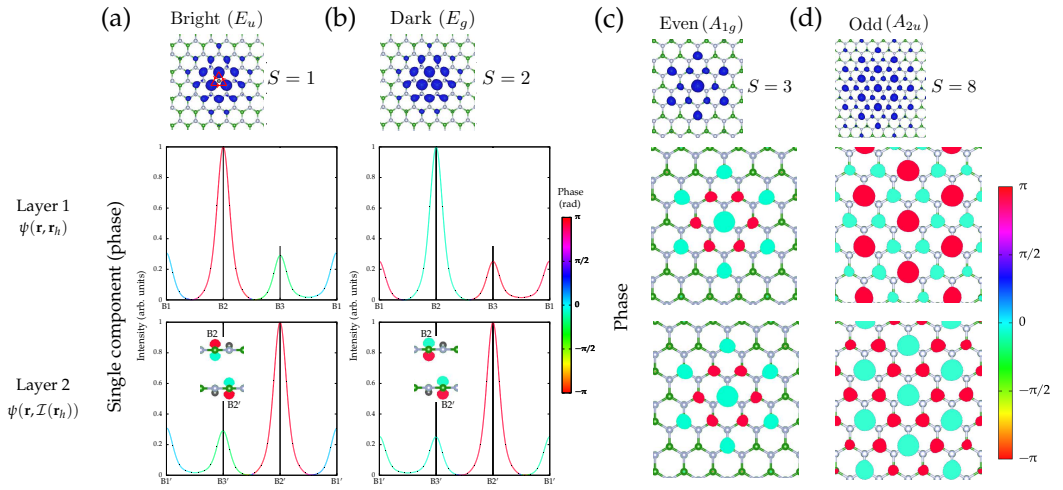


FIGURE 2.5: Excitons in bilayer hBN under inversion symmetry. Panel (a) / (b): excitonic state $S = 1$ (even, dark) / $S = 2$ (odd, bright). Panel (c) / (d): excitonic state $S = 3$ (A_{1g}) / $S = 8$ (A_{2u}). In the top frames of (a) and (b) the intensities of the full doubly degenerate states on the BN lattice (hole fixed on the central nitrogen) are shown. Below, the choice of one appropriate component wavefunction in the degenerate subspace (see text) permits the representation of the phase of this excitonic state. In the middle and bottom frames the phase is plotted for the electron distribution when the holes are fixed in two positions related by inversion symmetry $\mathcal{I}$ (layer 1 and layer 2). The intensity is plotted along the triangle formed by the borons which are nearest–neighbours to the hole nitrogen (B1 and B1$' = \mathcal{I}$(B1), B2 and B2$' = \mathcal{I}$(B2), and B3 and B3$' = \mathcal{I}$(B3): red triangle in the top frame). The phase, which rotates around the path, is shown in color scale as in the other plots. The insets show a section of the phase–intensity plots on both layers, relative to the boron with the largest intensity. In the top frames of (c) and (d) the intensities of the wave functions of the non–degenerate states are portrayed (since these are il excitons, the hole layer is not shown). The phase–intensity plots are shown in the middle and bottom frames. For each exciton, two wavefunctions connected by inversion symmetry ($\mathbf{r}'_{\text{hole}} = \mathcal{I}(\mathbf{r}_{\text{hole}})$) are depicted, showing their respective parity [(c): even, (d): odd].

positions, shown as red (bright) and gray (dark) circles in the lower frame, are almost constant, averaging around 5.3 eV. The position of the bulk excitons is around 5.5 eV.

Figure 2.6(b) provides a scheme of the Davydov splitting from bilayer to bulk for the lowest–bound exciton E. We make the following observations: (i) dark and bright states alternate, and (ii) in tri– and pentalayer we have a bright–dark couple at lower energy, while the rest of the multiplet lies above. These latter states correspond to inner or "bulk–like" excitons (see the following), therefore they should be compared with the bulk excitons. The bright–dark couple is made of surface excitons that have no counterpart in the bulk crystal and their relative intensity decreases to negligible values for increasing number of layers.

The bottom frame of Fig. 2.6(a) shows that the energy of the bright inner peaks increases with layer number, which leads to the bulk values. However, this increasing trend might be related to the numerics of the G_0W_0 approach, which has been known to underestimate large band gaps.[93] Indeed, experimentally the first exciton in multilayer and bulk systems is usually found around 6 eV,[19, 32] while in *ab initio* calculations a value around 5.5 eV is obtained. In order to elucidate this point, we performed simulations on monolayer and bulk using a semi self–consistent GW

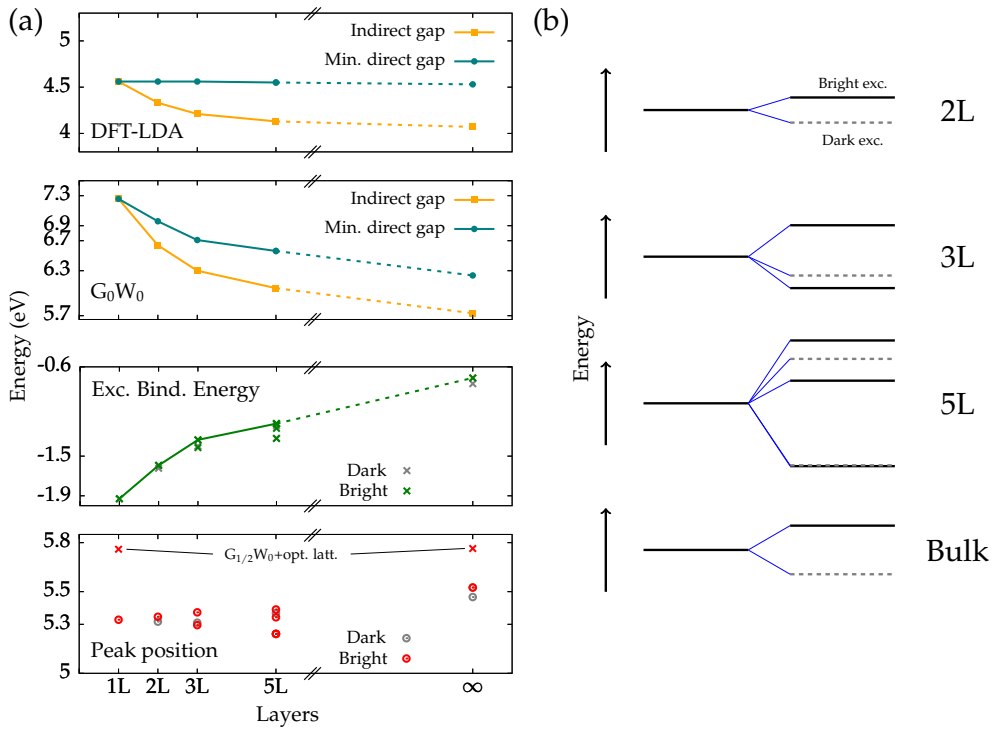


FIGURE 2.6: Panel (a): DFT band gaps (top), G_0W_0 band gaps (second to top), binding energies of the lowest–bound excitons (second to bottom) and absolute excitonic peak positions (bottom) are shown as a function of the layer number. The orange squares label the band gap, the teal circles label the minimum direct gap, the green (gray) crosses / red (gray) empty circles label the binding energies / absolute peak positions of the bright (dark) excitons. The red crosses in the bottom frames represent calculations made with a semi self–consistent GW approach and with the optimized lattice constant for the monolayer. Panel (b): scheme of the Davydov splitting in energy of the lowest–bound excitons for N–layer systems ($N = 2, 3, 5, \infty$) starting from the energy of N isolated monolayer excitons. Solid black (dashed gray) lines represent bright (dark) states. The energy separation within the multiplets is in scale for all systems.

scheme (labeled sscGW or $G_{1/2}W_0$), updating the band energies in G during subsequent G_0W_0 runs until convergence. We obtain an additional correction to the band gap and peak positions of monolayer and bulk by 0.34 and 0.22 eV, respectively. We also used the LDA–optimized lattice constant for the monolayer (2.479 Å) instead of the bulk one (2.496 Å), which accounts for another 0.1 eV increase in the peak energy. The final band gap for the monolayer is 7.69 eV, and its main excitonic peak is now almost at the same energy as the bulk one (5.76 eV, red crosses in Fig. 2.6(a)).

In conclusion, additional refinements in the calculations (e.g., fully self–consistent GW[94] and using the "true" experimental few–layer lattice constants) may lead to an inversion of the trend and show peak energies that are both higher and *decreasing* towards the bulk value.

The various layers stop being equivalent when $N > 2$, because now we have two *surface* layers and $N - 2$ *inner* ones. If we restrict our attention to the lowest–bound Davydov multiplets (the only ones relevant for optical transitions), we know that they have ip character and that their spatial localisation *within the layer plane* is the same. However, now eh pairs on the surface layers will be considerably less screened than they would be on the inner ones. Additionally, less nitrogen to boron interlayer hopping sites are present on the surface, which according to the TBEM

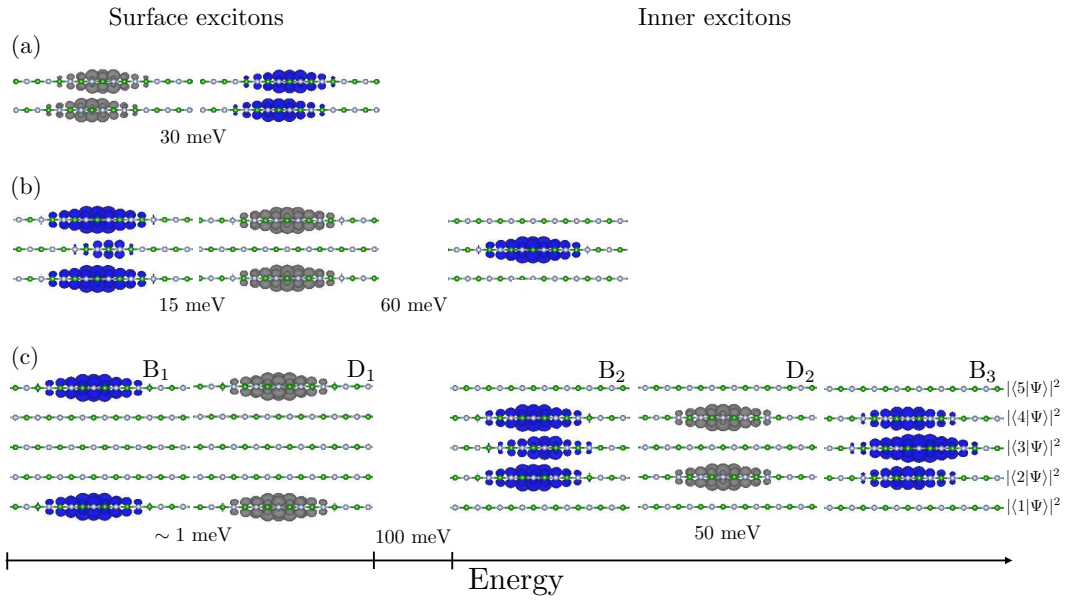


FIGURE 2.7: *Ab initio* simulations: lowest–bound Davydov multiplets in (a) bilayer, (b) trilayer and (c) pentalayer hBN. Each wavefunction is plotted for N fixed positions of the hole (N is the layer number) as explained in the text. The intensities shown in blue (gray) belong to bright (dark) excitons. The energies of the splittings between and within surface and inner exciton subsets are also given in meV. Bright (B) and dark (D) states are labeled in (c) for comparison with Figs. 2.2.

lowers the overall exciton kinetic energy. As a consequence, we might expect a larger binding energy for excitations formed on the surface. If that is the case, the asymmetric energy separation within the Davydov multiplets (see Fig. 2.6(b)) can be explained in terms of the energy gap between surface and "bulk–like" excitons. The multilayer TBEM can be recast in this case into a finite linear chain model where each chain site represents one layer[40] and gives the same predictions, additionally providing the selection rules for the ip Davydov multiplets as follows: if N is even, the system has inversion symmetry. Then, *odd* states are *bright* and *even* states are *dark*. If N is odd, the crucial symmetry is the mirror symmetry with respect to the central layer. In this case, *even* states are *bright* and *odd* states are *dark*.

Let us now check if the full *ab initio* results confirm these arguments. In order to investigate the spatial localisation of the excitations let us consider the excitonic wave function $\psi_\lambda(\mathbf{r}, \mathbf{r}_h)$ for a state λ belonging to the lowest–bound Davydov multiplet. We identify the hole subscript with the layer index: now $h \in [1, N]$ denotes the fixed position of the hole in a specific layer. We can then plot the quantity $\sum_i^N |\psi_\lambda(\mathbf{r}, \mathbf{r}_i)|^2$ (i.e., an intensity plot for N different hole positions, one on each layer) for bilayer, trilayer and pentalayer hBN. The results are shown in Fig. 2.7: the bright excitons are portrayed in blue, the dark ones in gray. The Figure allows us to easily verify that the selection rules stated above are rigorously respected: for example, no intensity is allowed on the central layer of tri– and pentalayer for the dark states, since they are odd.

Most importantly, there is indeed a separation within the multiplets between low–energy surface states and high–energy inner states. Note that the leading peaks in the imaginary part of the dielectric function – see Fig. 2.2(g) and (h) – come from

the excitons that are mostly localized on the central layers; however, during the relaxation process that takes place after the excitation is formed by optical absorption, the electron–hole pairs will diffuse towards the surface.

Let us now define the *oscillator strength* of an exciton state λ as $f_\lambda = |\sum_{\mathcal{K}} \overline{A}^{\mathcal{K}}_\lambda|^2$ (i.e., if the exciton lifetime is not considered, the intensity of the peak in the absorption spectrum). An interesting observation is that this quantity depends on the layer number: more precisely, the ratio between the oscillator strengths of inner and surface excitons is determined by N. The TBEM permits to derive the following implicit relation between the two quantities:[40]

$$\frac{1}{N-1}\cot^2\left(\frac{\pi}{2(N-1)}\right) = \frac{f_{m^*}}{f_O} \tag{2.1}$$

where the subscript O refers to the surface state and m^* to the brightest inner state. This relation can be approximated to an explicit one for large N.

2.3.3 Conclusion

The emergence of low–lying surface excitations in multilayer systems acts as a source of fine structure in their optical spectra, as can be seen in Fig. 2.2(g) and (h). This may help in the interpretation of experimental results,[32] while providing for a general understanding of the effects of layer stacking on the optical properties (e.g., Davydov splitting, dark vs bright excitons, *et cetera*), which we may now take for granted. Yet, these results also call for additional investigations. For example, we want to ascertain if excitations with large wave vectors may also be relevant to the optical spectra: we will do this in the remainder of the Chapter.

2.4 Excitons at finite momentum

So far we have solved the BSE at $q = 0$, obtaining excitons that lie at the Γ point of the first BZ and correspond to zero momentum transfer ($\mathbf{q} = \mathbf{k}' - \mathbf{k}$). Their binding energies, for BN systems, are large enough that the lowest–bound exciton always lies lower than the bottom of the conduction band even when the quasiparticle band gap is indirect. Therefore it may be argued that the *optical* gap of these systems is "direct" in the sense that the absorption spectrum would be dominated by the Γ excitons regardless of the underlying band structure.[6] However excitons have a momentum dispersion as well, and even though finite–q excitons cannot be created with optical light, they may become important in emission processes like luminescence: an excitation is created at $q = 0$, it quickly relaxes to the bottom of its dispersion curve $\overline{\mathbf{q}}$ and later it may not annihilate (electron–hole recombination) without the absorption or emission of phonons with wave vector $\overline{\mathbf{q}}$. It then becomes relevant to study the finite–q solutions of the BSE. Experimentally, excitonic states are often conceptualised as energy levels lying within the band gap of a system, downshifted with respect to the conduction band energies. However a rigid downshift of the band structure implies that the the q–dependence of BSE kernel $K^q_{\mathcal{K}\mathcal{K}'}$, Eq. (1.80), is very weak so that $K^q_{\mathcal{K}\mathcal{K}'} = K^0_{\mathcal{K}\mathcal{K}'} + O(q) \simeq K^0_{\mathcal{K}\mathcal{K}'}$, something that is in general not true.

Since finite–q BSE calculations have become available in many–body codes,[34] several studies involving bulk hBN and monolayer have been performed. These

[6] In the following, we will say that the optical gap is "direct" if the Γ point is the minimum of the exciton dispersion, "indirect" if it lies elsewhere.

works focused on x–ray scattering[29] (where finite–momentum excitations can be created without the help of a phonon) or probed the low–q behavior of the dispersion curve.[28, 95] The results we will present in this Section are preliminary,[7] nonetheless they can shed more light on the nature of finite–momentum excitations in BN systems and their possible consequences for the interpretation of luminescence spectra in multilayers[32] (we will treat the bulk luminescence separately in Chapter 3). Such studies were previously conducted only at the independent–particle level.[96]

Our calculations were performed using the proper response function $\overline{\chi}$ (i.e. describing "transverse" excitons) instead of χ (see Secs. 1.4 and last paragraph of 1.6.3). A comparison between the two can be found in Ref. [28] and we will come back to it in Chapter 4. For now, suffice it to say that based on Eq. (1.60), $\chi^{-1} = \overline{\chi}^{-1} - v_0$, we know that $\overline{\chi}$ and χ tend to coincide away from the low–q regime: at low q, instead, some energy levels that are degenerate in the spectrum of $\overline{\chi}$ may split by acquiring a longitudinal component in the spectrum of χ (they remain invisible to optical absorption at Γ since the external probe is a transverse field). Likewise, the exciton coefficients $A^{\mathcal{K}}_{\lambda}$ and $\overline{A}^{\mathcal{K}}_{\lambda}$ may be different at low q (yet, we know from Sec. 1.4 that at the optical limit the same macroscopic dielectric function may be obtained, by different relations, from both χ and $\overline{\chi}$).

2.4.1 Bulk

The exciton dispersion in the bulk system is displayed, starting with the first two (doubly degenerate) exciton states at Γ, in Fig. 2.8(a) along every high–symmetry direction in the 3D BZ.[8] We can see that their degeneracy is lifted, away from Γ, in the regions where the finite–q point groups do not allow two–dimensional irreducible representations. The most important feature of the dispersion is shown in Fig. 2.8(b): its minimum lies along the ΓK direction, not at Γ. The fine structure with two minima and a local maximum between them (the latter being at $|\Gamma K|/2$) is directly inherited from the single–particle valence bands which feature a pair of maxima close to the K point (Fig. 2.3). Here, the lowest–lying exciton at Γ (the dark E_{2u} state according to the D_{6h} point group of bulk hBN) splits into the non–degenerate $i1$ and $i2$ branches. The electron–hole interaction "flattens" the dispersion with respect to the IP case, so that now the difference between the "direct" and "indirect" optical gaps is between 0.1 and 0.15 eV (the former value if we take into account the dark Γ state E_{2u} as the direct minimum; the latter if we consider its bright Davydov partner E_{1g}). It is worh noting that such a difference is of the same order as the optical phonon frequencies in bulk hBN ($\sim$ 0.1 / 0.18 eV). In the IP case the direct–to–indirect band gap energy difference was instead about 0.5 eV.

These results shows that indirect, phonon–mediated transitions must be involved in the creation and annihilation of low–energy excitations in bulk hBN: in particular, the fine structure reported in luminescence spectra should be interpreted as a signature of phonon–assisted exciton annihilation. In Chapter 3 we will compare first–principle calculations of such spectra to the experimental results.

Another interesting question now arises. Suppose that we have a system where the direct–to–indirect IP energy difference is smaller than in the bulk case: will it be possible in this case for the optical gap to remain "direct"? If so, the IP and BSE treatments will give radically different predictions about how a luminescence spectrum should look like. In particular, our hypothesis is that there might be a crossover

[7] A publication is in preparation.

[8] The computational cost of this calculation is roughly N_q times the cost of a standard $q = 0$ BSE calculation for optical absorption.

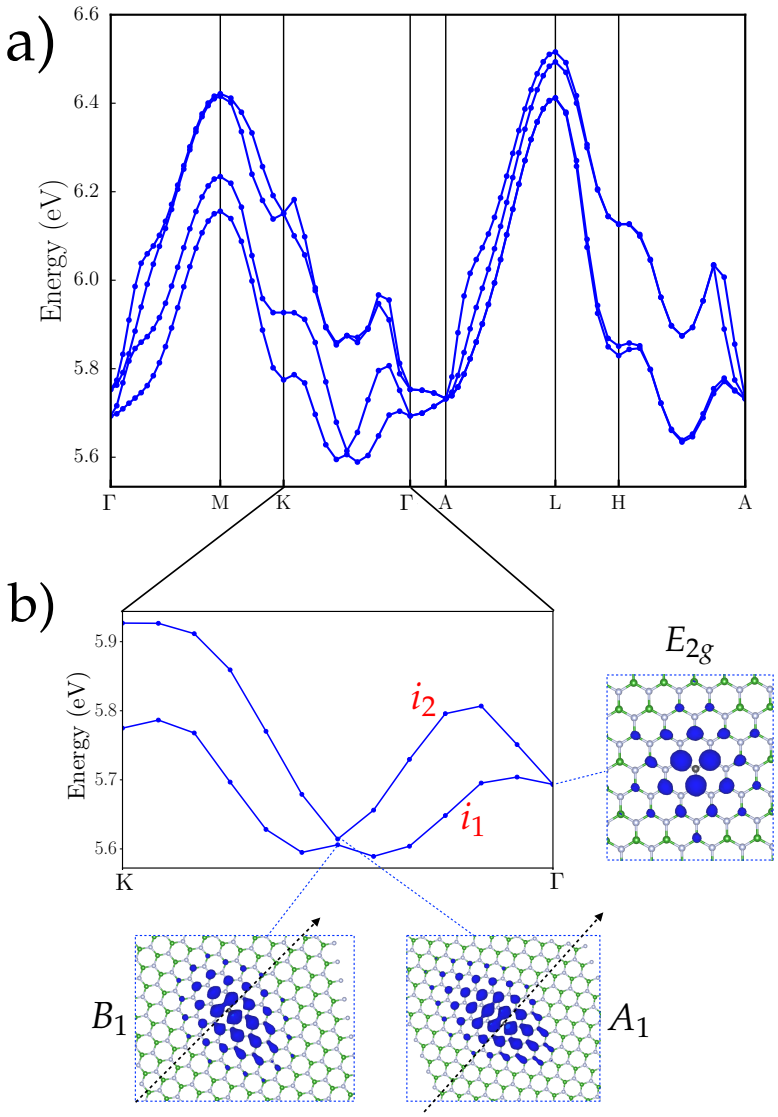


FIGURE 2.8: Exciton dispersion for bulk hBN (DFT–LDA + sscGW + BSE). (a) The dispersion of the lowest–bound Davydov pair in the full BZ. (b) Zoom of the ΓK segment of the BZ for the lowest–bound exciton. The wave function intensities are shown for $q = 0$ and $q = |\Gamma K|/2$; in the latter case the black arrows denote the **q** direction (symmetry labels are present, to be discussed in Chapter 3).

between a direct and phonon–assisted luminescence regime as more and more BN layers are stacked in a few–layer system (we know already that in monolayer, where no p_z orbital hybridisation occurs, the minimum of the exciton dispersion remains at the Γ point[97]). We will test this hypotesis in bilayer and trilayer hBN.

2.4.2 Bilayer and trilayer

There are two types of exciton dispersion branches relevant to few–layer systems. The first type is composed of excitons formed by $\pi \rightarrow \pi^*$ electronic transitions; this is the only type relevant for the bulk case because it is periodic along the stacking

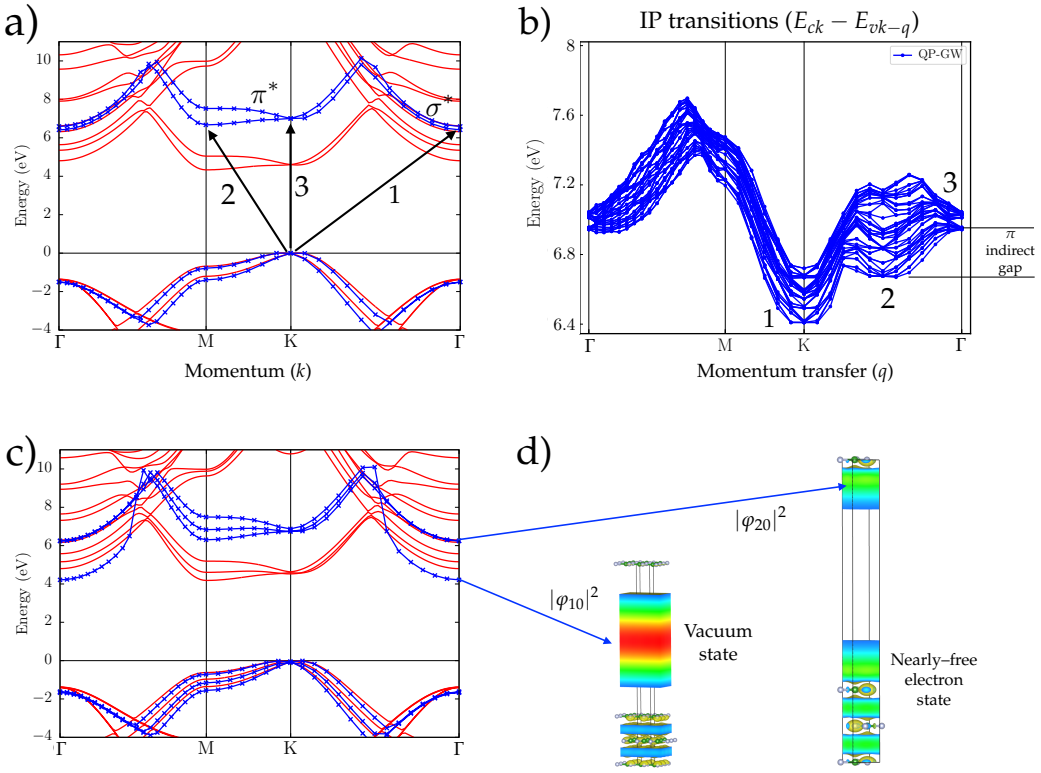


FIGURE 2.9: Single–particle transitions in bilayer [(a), (b)] and trilayer hBN [(c), (d)]. [(a), (c)]: band structures (red: DFT–LDA, blue: G_0W_0). The most important transitions are shown with arrows in the bilayer case. (b) IP transition energies $\Delta E_{\mathcal{K}q}$ for the bilayer. The direct and indirect band gaps due to $\pi \to \pi^*$ transitions are emphasised. (d) Lowest–lying conduction Kohn–Sham states at Γ (trilayer). The KS wave function intensities $|\varphi_n k(\mathbf{r})|^2$ are shown in the simulation supercell for a vacuum and a NFE state (left and right, respectively — see text for the characterisation of the states).

direction.[9] The second type is composed of excitons formed by $\pi \to \sigma^*$ electronic transitions; these were not relevant for absorption at $q = 0$, but are important at finite q. Consider for example the transition energies in the bilayer when no electron–hole attraction is present, as shown in Fig. 2.9(b) (this is in other words a plot of the quantity $\Delta E_{\mathcal{K}q}$). We see the $\pi \to \pi^*$ transitions being the lowest ones at Γ (direct gap, K → K transitions, labeled 1) and $|\Gamma K|/2$ (indirect π gap, K → M transitions, labeled 2). However the indirect, K → Γ transitions (labeled 3) provide an even lower minimum in the dispersion at the K point.

Are these transitions physically relevant? The corresponding conduction bands at Γ have a parabolic, free–electron–like behavior (which also explains why they experience a smaller GW correction than the π bands). Indeed, some of these states are spurious bands relative to free electrons propagating in the vacuum between repeated copies of the system in the simulation supercell (first wave function in Fig. 2.9(d)). These are vacuum states which are not related to the BN system. There are, however, also states where a delocalised electron lies both in the interstitial regions between the layers and $1 - 2$ Å above the surfaces (second wave function in Fig. 2.9(d)). These are the so–called nearly–free electrons (NFE), which are well known

[9] Provided that the polarisation of the incoming light is parallel to the layer plane. In any case the σ–type excitons in bulk are far higher in energy than the π–type ones.

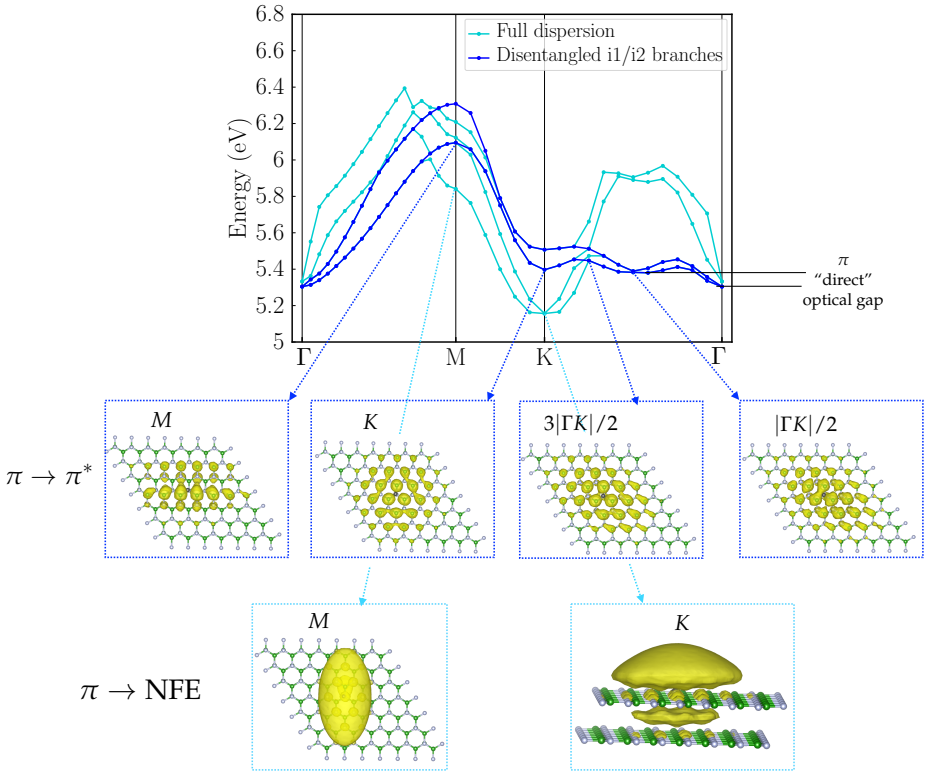


FIGURE 2.10: Exciton dispersion in bilayer hBN. The blue curves represent the lowest–energy *disentangled* branches relative to excitons formed by $\pi \to \pi^*$ electronic transitions. Some of the corresponding wave function intensities along the high–symmetry directions in the BZ are plotted (connected by blue arrows). The light–blue curves represent all the other branches, including the $\pi \to$ NFE transitions. Some of the wave function intensities belonging to excitons of the latter type are plotted (connected by light–blue arrows). The "direct" optical gap of the π–type excitons is emphasised.

to be a physically relevant feature of these systems (for example BN monolayer and nanotubes as well as graphitic systems),[35, 98–102] therefore they should be properly analysed, especially since their role in the exciton picture is not known. They should be almost dark for optical spectroscopy, because their $\pi \to \sigma^*$ transition matrix element is small: it is only non–zero because the σ^* states are in reality partially hybridised with the π states.[10]

Therefore, we organise the presentation of the results as follows: we focus first on the analysis of the $\pi \to \pi^*$ exciton branches, and in a second step we discuss the $\pi \to$ NFE ones.

π–type exciton branches

We can disentangle the two branch types by just following the evolution of the excitonic wave function along the BZ: we know that the π excitons are in–plane and that their shape is constrained by the point group of the wave vector q.[97] We do so in Fig. 2.10 starting from the lowest–lying Γ exciton E_u. Its π–type dispersion branch is shown in blue, and its wave function intentities in different areas of the BZ are shown below. We see immediately that in the bilayer case the same structure around $|\Gamma K|/2$ that in the bulk case represented the minimum of the dispersion lies instead

[10] This can be verified by projecting the σ^* KS wave functions at Γ onto the atomic p_z orbitals.

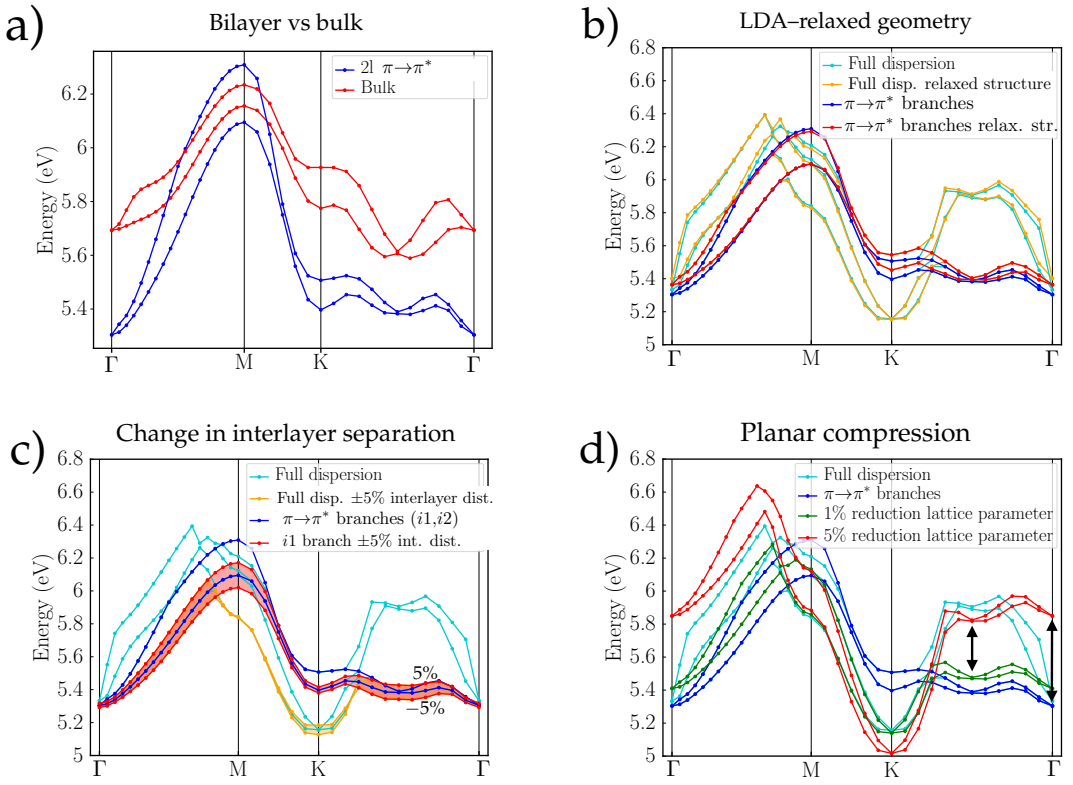


FIGURE 2.11: Exciton dispersion in bilayer hBN for different values of the lattice parameters. The blue and light–blue curves have the same meaning as in Fig. 2.10. (a) Comparison with the corresponding dispersion branch in the bulk system (red). (b) Comparison of the same dispersion branches (red for π–type, orange for the rest) with the LDA–relaxed lattice geometry. (c) Comparison of the same dispersion branches – same color scheme as in (b) – with $\pm 5\%$ interlayer separation distance. (d) Comparison of the same dispersion branches with 1% and 5% planar compression (green and red, respectively).

above the Γ excitons. This means that for π excitons, the relevant ones for optical absorption, the optical gap is in fact "direct", in contrast to the QP gap. This effect is due to the electron–hole interaction, which acts here more strongly for $q = 0$. The bottom branches at K and M (light–blue color in Fig. 2.10) are composed of NFE–type excitons and provide the true minimum of the dispersion at K. The electron distribution of the resulting excitons appear, when the hole is fixed above a nitrogen atom, to be localised in the same regions as the original NFE states. Yet, most of their intensity (up to 5% of the maximum value) is localised within few unit cells from the hole. This demonstrates the bound character of these excitons.

A comparison of the lowest–lying branches between bulk and bilayer can be seen in Fig. 2.11(a).[11] These two systems have the same unit cell and differ only the periodicity along the stacking direction (there is no periodicity in the bilayer). The comparison makes it clear that in the bilayer the interlayer interaction is too weak to induce an "indirect" optical gap (yet, it is strong enough to provide and indirect QP gap if the eh interaction is neglected).

So far we have fixed the lattice parameters to the experimental bulk values. In

[11] Here we are more interested in comparing the dispersion curves than the absolute exciton energies, since we know that the G_0W_0 band gap underestimation error might affect absolute exciton energies differently in bilayer and bulk.

this case, the energy difference between "direct" and "indirect" minima of the π–type dispersion is just 0.08 eV (and 0.055 eV from the bright Γ state E_g). In Fig. 2.11(b), we perform the same calculation using the LDA–relaxed lattice parameters ($\sim$ 1% reduction in the planar lattice constant a, $\sim$ 4% reduction in the interlayer distance d). The bottom π–type branches for the new calculations are shown in red, the rest in orange. We see that in this case the "indirect" structure at $|\Gamma K|/2$ is almost unaffected, while the energy of the Γ excitons is noticeably increased. Now, the local minimum at $|\Gamma K|/2$ lies between the dark E_u and the bright E_g states at Γ. If we keep a fixed to the bulk value, but we let d vary, we obtain the plot in Fig. 2.11(c) where the red curves refer to a $\pm 5\%$ variation of d with respect to the bulk value (blue curve). We see that in this case the Γ excitons do not change while the structure at $|\Gamma K|/2$ varies slightly. We can thus take the dispersion curves at $\pm 5\%$ d as the "errorbars" of our calculation with respect to the "true" van der Waals–dictated interlayer separation. We also notice, based on Figs. 2.11(b) and (c) that the combined effects of compression along a and along d tend to cancel away from Γ. Finally, we fix d to the bulk value and we reduce a by 1 and 5%, obtaining the plot (green and red curves) in Fig. 2.11(d). We can see that here the effect of the compression, which strengthens the covalent $sp2$ planar bonds, is to dramatically increase the exciton energies at Γ, reversing the ordering between "direct" and "indirect" optical gap (further investigation will be needed to better elucidate this result).

In order to analyse the trilayer dispersion, we must get rid of the low–lying "vacuum" band featured in Fig. 2.9(c): we do so by setting its energies back to the DFT values (thus neglecting the pathological negative GW correction) before moving on to the BSE calculations. In this way, we finally obtain the dispersion relation depicted in Fig. 2.12. The red lines represent the disentangled π–type excitons: we can see that starting from the Davydov triplet at Γ, where we have three doubly–degenerate states (the first two of which are the surface states from Sec. 2.3.2), the exciton levels split into six non–degenerate states. Here, since the two branches originating from the second (dark) surface exciton split farther apart in energy than the others, we can identify two "triplets" of roughly parallel branches, the triplets being about 0.4 eV apart at $|\Gamma K|/2$. In the trilayer, the bright surface exciton at Γ and the minimum of the π–type dispersion at $|\Gamma K|/2$ are roughly at the same energy, with the optical gap being slightly "indirect" for bulk lattice parameters. The "true" minimum of the dispersion is at the K point, due, as in the bilayer case, to NFE–type excitons (orange lines in Fig. 2.12).

Effect of NFE–type branches at K

Let us now consider the effects of the low–lying NFE–type excitons at the K point. As we remarked, $\pi \rightarrow$ NFE electronic transitions are not entirely forbidden because of the hybridisation of the σ^* and π^* bands. We report in Fig. 2.13 a plot of the microscopic dielectric function $\varepsilon_2(\mathbf{q}, \omega)$ for $q = \mathrm{K}$. The bilayer is shown on the left, the trilayer on the right. Note that since we are at large q, these spectra are essentially equivalent to electron–energy loss (EELS) spectra.[12] As expected, the dominant peaks are related to the π–type excitons, i.e. to the branches originating from the lowest–bound Davydov multiplets at Γ, whose onset is marked by the vertical lines on the plot. We nonetheless see a faint signal at the energy of the NFE–type excitons, preceding the largest peaks.

[12] The EELS spectrum, useful to identify collective longitudinal excitations (plasmons), is given by $\mathrm{Im}\varepsilon^{-1}(\mathbf{q}, \omega) = v(q)\mathrm{Im}\chi(\mathbf{q}, \omega)$.

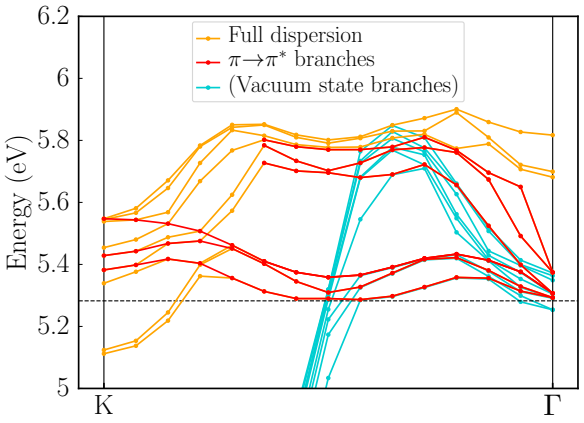


FIGURE 2.12: Exciton dispersion in trilayer hBN. The red curves represent the disentangled branches relative to excitons formed by $\pi \rightarrow \pi^*$ electronic transitions. The orange curves represent all the other branches, including the $\pi \rightarrow$ NFE transitions. The horizontal dashed line marks the minimum of the π–type dispersion. The light–blue curves represent the (spurious) branches obtained when the vacuum state is not removed.

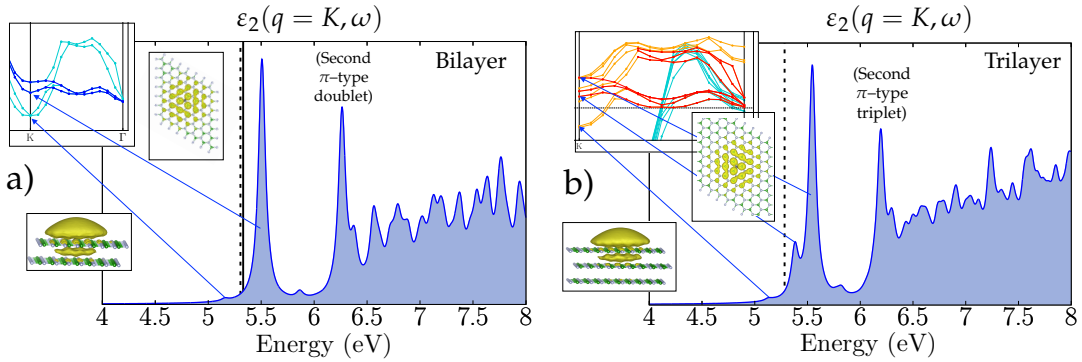


FIGURE 2.13: Microscopic dielectric function at $q = \mathrm{K}$ for (a) bilayer, (b) trilayer hBN. The most relevant features are connected to their origin in the exciton dispersion relations (insets) by blue arrows. The other insets show the wave function intensities of the excitons responsible for such features. The vertical black lines in the main plot mark the energies of the lowest–bound excitons at the Γ point.

2.4.3 Conclusions

We can make the following conclusions. (i) It is possible to have a "direct" optical gap in bilayer hBN, despite having an indirect QP gap (ii) Most likely, there will be a competition between the direct and phonon–assisted $\pi \rightarrow \pi^*$ transitions in absorption, with the leading mechanism depending on the environmental conditions of the bilayer sample (strain, substrate, *et cetera*); while the phonon–assisted transitions will always account for an additional fine structure, having a properly "direct" optical gap will affect the absorption/emission efficiency of UV light. (iii) It is possible, in theory, to change the nature of the band gap by planar strain/compression. (iv) For bulk parameters, the crossover between π–type direct–to–indirect gap already

happens at layer number $N = 3$. (v) In the luminescence case, the role of the low–lying NFE–type excitons should be further investigated. For example, experimentally no luminescence signal could be observed for multilayers below $N = 6$.[32][13] Our results show that this may be due to the fact that the excitations are "trapped" at the minimum at K with very long recombination lifetimes due to the low value of the screening matrix element (to be further multiplied to the electron/exciton–phonon coupling matrix element, reducing the probability of a recombination even more). As the number of layers increases, and the material becomes more bulk–like, the NFE–type branches are lifted up in energy and eventually disappear, allowing recombination from the π–type excitons.

As a final remark, we point out that even with a deceptively simple system like hBN, a complicated description of the optical response arises nonetheless, making it difficult to capture the essential physics by simple models alone (for example, the TBEM only works for π–type states).

[13] Private communication with F. Ducastelle and J. Barjon.

Chapter 3

Indirect optical spectra: many–body approaches

This chapter is partly based on our publication [41]. Some of the text and figures contained in this Chapter are adapted from this reference.

3.1 Introduction: luminescence in bulk hBN

Recent experimental results on bulk hBN luminescence are displayed in Fig. 3.1. The intrinsic luminescence spectrum in bulk hBN is shown from Refs. [24], [83] and [84] (left, center and right, respectively). The characteristic shape composed of two lower–intensity peaks between 5.85 and 5.9 eV, followed by a higher–intensity structure at lower energies (5.75 to 5.80 eV) is present. Note that the intensity of the two main peaks in Ref. [83] is reversed with respect to the others, because in this case the spectral response of the detection system (not negligible in the UV range) was not subtracted. Refs. [24] and [83] are photoluminescence (PL) spectra, while Ref. [84] reports the result of a cathodoluminescence (CL) experiment. In CL, an electron beam is used instead of an optical laser to probe the system, leading to the generation and eventual recombination of electron–hole pairs. The two techniques achieve the same resulting spectra because the experimental conditions are similar, namely: (i) low laser / electron beam power density, (ii) near–band gap excitation, (iii) continuous pumping leading to a steady state between absorption and emission processes. In the bottom panel (b) of Fig. 3.1, the CL/PL spectra (red) are compared to the photoluminescence excitation spectra (PLE, blue). Ref. [19] is on the right, Ref. [84] is on the left. Under the right conditions, PLE spectra are proportional to absorption ones,[1] and albeit at a lower resolution than their PL/CL counterpart, clearly display the main exciton peak at 6.05 eV. The energy difference between emission and absorption spectra is called Stokes shift. In the case of indirect transitions, the two spectra are expected to be roughly symmetric, mirrored around the energy of the indirect transition that is involved in the phonon–assisted excitation/recombination process. However, in the case of bulk hBN the two spectra are not symmetric at all. We will show in the course of this Chapter that this asymmetry is due to the fact that the absorption process is dominated by the *direct* exciton at Γ, while the emission process is dictated by the *indirect* excitonic minimum in the middle of the ΓK segment. Therefore, two different excitonic states determine the two different physical processes of absorption and emission.

[1] In PLE, the luminescence at a specific energy is detected as a function of the excitation energy. For example, the energy of a certain exciton level may be selected. Varying the excitation energy, the PLE spectrum then generally presents the same features of the absorption spectrum, despite some small deviations.

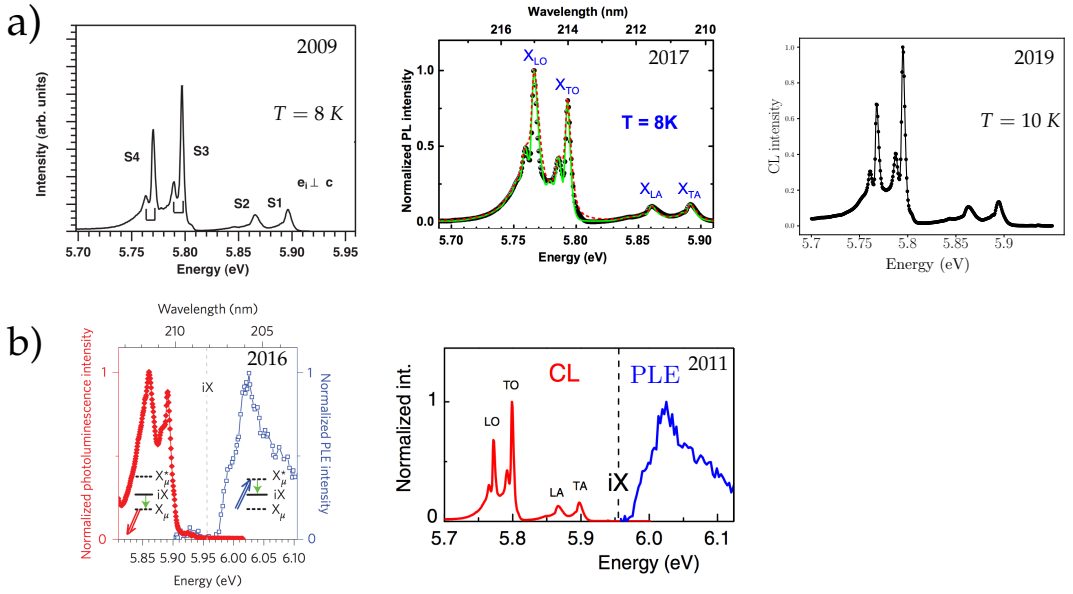


FIGURE 3.1: Luminescence in bulk hBN. (a) Intrinsic luminescence spectrum. Left: PL from Ref. [24]. Center: PL from Ref. [83]. Right: CL from Ref. [84]. (b) Comparison between PL/CL (red) and PLE (blue, proportional to absorption) spectra from Ref. [19] (left) and Ref. [84] (right).

Our theoretical investigation on finite–momentum excitons in hBN, conducted in the previous Chapter, provides strong support for the recent experimental interpretation in terms of phonon–assisted optical excitations, since the minimum of the exciton dispersion is found to lie away from the Γ point. Moreover, from the above discussion and the previous Chapter it is clear that modifications of the electronic/excitonic spectral structure due to exciton–phonon interactions cannot be ruled out. Indeed, the *ab initio* theoretical description of phonon–assisted optical absorption is an important issue in condensed matter physics, that has yet to be tackled fully. In order to correctly evaluate the relative intensity of the indirect/direct optical absorption processes in bulk hBN, for example, exciton–phonon coupling needs to be included. In this way, the efficiency of hBN as an UV emitter — or the way to tailor it towards the desired effect — could be established. Refs. [103] and [36] provide an example of the theoretical calculation of a phonon–assisted optical absorption spectrum (of GaN and silicon, respectively), although only in the independent–particle picture and confined to the energy window where only indirect transitions contribute. We will discuss Ref. [36] in more detail in the next Section. The Wialliams–Lax method discussed in Sec. 3.2 (Refs. [37, 104, 105]) provides a unified approach to describe on the same footing lattice–dependent band features and phonon–assisted transitions. However, it is computationally expensive if the BSE is involved while also restricted to static approximations, therefore imprecise close to the absorption onset. Moreover, various models of the exciton–phonon coupling including dynamical effects are available in the literature[82, 106–111] (with Ref. [112] specifically tackling luminescence), although their direct test on real materials from first principles is, again, computationally demanding. Therefore, first–principles calculations of indirect absorption including excitons are still missing, and we bridge this gap by providing the first such calculations in this Chapter. It is worth noting that layered materials in which phonons may play an important role for optical spectroscopy, like hBN or transition metal dichalcogenides like MoS_2[91]

(which, with the exception of the monolayer form, displays an indirect band gap), are the perfect candidates for this kind of analysis.

In general, starting from the form given in Eq. (1.75) for the excitonic macroscopic dielectric function, we expect it to be modified by the exciton–phonon interaction in the following way:

$$\varepsilon_2(\omega) = \frac{8\pi}{V}\sum_{\lambda}(1-R_\lambda)\left|\sum_{\mathcal{K}}\mathcal{A}^{\lambda}_{\mathcal{K}}d_{\mathcal{K}}\right|^2 \mathrm{Im}\left([\omega-(E_\lambda+\Delta E_\lambda)+i\eta]^{-1}\right)+\mathcal{F}(\{E_{\lambda q},\omega_{\mu q}\}). \tag{3.1}$$

Here, $\mathcal{A}^{\lambda}_{\mathcal{K}}$ can be equal to either $A^{\lambda}_{\mathcal{K}}$ or $\overline{A}^{\lambda}_{\mathcal{K}}$. The quantity R_λ represents a renormalisation factor determining how much oscillator strength is transferred from the direct peaks to the phonon–assisted parts of the spectra, given here by the unknown functions $\mathcal{F}$ depending in principle on all the exciton and phonon energies. The complex energy correction ΔE_λ is the dressing of the exciton energy due to exciton–phonon interaction. We will investigate this quantity in Sec. 3.2 for the monolayer case. In this way we can show that phonon–assisted processes in BN systems are not only *possible*, by virtue of the indirect optical gaps demonstrated in the previous Chapter, but also *likely*, since our results in the aforementioned Section point to a strong electron–phonon coupling in BN. Afterwards, when switching to the bulk system, we will the ignore ΔE_λ, and assume that the exciton energies remain always at their "bare" values. This will not affect our results because we are more interested in studying the shape of the luminescence spectrum, while knowing already that its absolute energy position will be incorrect anyway because the GW approximation underestimates the band gap (as mentioned in Secs. 2.3.2 and 3.2.1).

We will thus focus on a way to reproduce the $\mathcal{F}$ functions, the phonon–assisted contribution to the optical spectrum, using state–of–the–art computational many–body techniques: in particular we will compute the derivatives of ε_2 with respect to static atomic displacements (Sec. 3.4). This will be enough to understand and reproduce the results shown in Fig. 3.1, leading to the main result of this thesis.

We also anticipate that in Chapter 4, after identifying the phonon–assisted part of the absorption spectrum as the *satellites of the excitonic spectral function*, we will propose a method, based on very recent theoretical developments,[38, 113] to include an "exciton–phonon self–energy" in the *ab initio* simulations: in this way we can overcome the previously mentioned static approximation and obtain the full absorption spectrum including the R_λ factor.

3.2 Monolayer: testing electron– and exciton–phonon coupling

As far as optical spectroscopy is concerned, the alteration of the electronic system due to (perturbatively weak) electron–phonon interactions can be separated in two main effects: the temperature dependence of band structures (seen in Sec. 1.6.6) and the possibility of indirect optical transitions (e.g. indirect absorption).[114] Although the two effects are intimately linked (both depend on the electron–phonon self–energy, and in particular on the strength of the electron–phonon coupling matrix elements), they can usually be treated separately as the band structure "renormalisation" causes an overall shift in the absorption spectra, while the indirect transitions modify its shape. We will be mainly concerned with the latter effect in the remainder of this Chapter and in Chapter 4. Nonetheless, in this Section we will present the results of electron–phonon (ep) calculations performed for monolayer

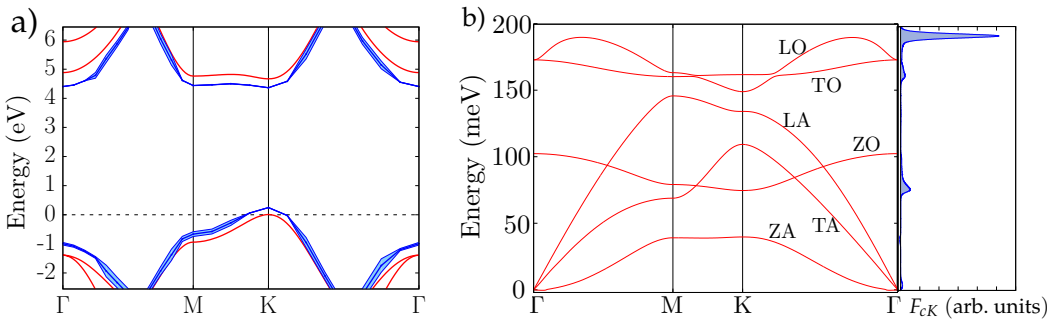


FIGURE 3.2: Electron–phonon coupling in monolayer hBN. (a) Red: DFT–LDA band structure. Blue: Corrected quasiparticle band structure including the ep self–energy Σ^{FM}_{nk}. The correction is shown for the top valence and bottom conduction bands. The blue shaded areas, of width $2\text{Im}\Sigma^{FM}_{nk}$, are inversely proportional to the electron–phonon lifetimes. (b) Phonon dispersion. To the right, the Eliashberg function $F_{cK}(\omega)$ related to the bottom of the electronic conduction band.

hBN according to Sec. 1.6.6. This allows us to gauge the strength of the ep coupling in this system, and to gain new indications about the magnitude of the error that the GWA makes, in its G_0W_0 forms, in correcting the band gaps of BN–type systems with respect to experimental observations. Furthermore, we will try to obtain a phonon–mediated correction to the *exciton* energies as well ("exciton–phonon" coupling). These latter results are not to be considered definitive, for reasons that will be clarified in the following, but they are useful to get an idea of what to expect in BN–type systems, where exceptionally strongly bound excitons plus very high optical phonon frequencies (because of the low atomic masses of B and N) may lead to quite noticeable effects. Indeed we find that this is the case.

3.2.1 Band gap renormalisation

The phonon dispersion in monolayer hBN is shown in Fig. 3.2(b). Since we have only two atoms per unit cell, three optical branches are obtained in addition to the three acoustic ones. The maximum frequency of the of the longitudinal optical phonon is close to 190 meV (46 THz, 1530 cm^{-1}), and the strong overbending of this mode[115] leading to the LO–TO degeneracy at the Γ point (no LO–TO splitting at $q = 0$ is possible in purely 2D systems[116]) is reproduced correctly using the method presented in Ref. [117]. We also compute ep coupling matrix elements (Eqs. (1.114) and (1.115)) with DFPT and obtain the ep QP correction from the self–energy (Eqs. (1.116), (1.117) and (1.118)) calculated with the `Yambo` code. At low frequency (< 50 meV), the FM and DW terms tend to cancel, and the q–integrals can be made stable by using a random integration method (Sec. 1.5.2) close to the Γ point. The effect on the band structure is displayed in Fig. 3.2(a) at zero temperature: in red, the DFT–LDA bands are shown; the phonon–corrected bands are in blue, while the width of the shaded areas is given by $2\text{Im}\Sigma^{FM}$ (and therefore it is inversely proportional to the quasiparticle lifetime). We notice immediately a huge band gap renormalisation, by 0.522 eV, comparable with that of diamond.[118] This is apparent already at zero temperature, therefore entirely due to quantum zero–point motion. This result puts even more into question the accuracy of the electronic G_0W_0 correction to the band gap: since the main exciton peak is expected experimentally around 6 eV with a binding energy between 0.7 (bulk) and 1.9 (monolayer) eV, this puts the single–particle band gaps to values ranging from 6.7 to 7.9 eV. So

far, by resorting to the sscGW technique, the band gap underestimation would be of about 0.2 eV as seen in Sec. 2.3.2. To this we should now add a redshift of ~ 0.5 eV due to electron–phonon effects, which must be explicitly added to the theory but is always present in the real system. Because of this consideration, we can roughly estimate the band gap underestimation of the GW method to be between 0.5 and 1 eV for BN systems.

As for the lifetimes, let us simply note that they become longer as less and less electron–phonon scattering channels become available for the electrons, until, at the band gap, they become infinite (of course the electron–hole recombination probability is not considered here). This is because at zero temperature, an electron can only scatter from a band state to another by spontaneously emitting a phonon, which means (in the case of a conduction electron) lowering its energy. However, at the band gap, there are no lower energy states available, and therefore no scattering is possible. We see that this condition applies through all the KM segment of the conduction band due to its flatness.

Next, we look at which phonons play the most important role in the band gap renormalisation, i.e. the ones that couple most strongly with the electronic states. On the right side of Fig. 3.2(b) we show the Eliashberg function $F_{cK}(\omega)$ related to the bottom of the conduction band (π/p_z electron at the K point). The Eliashberg function can be understood as a phonon density of states weighted with the ep coupling strength to a particular electronic state. In our case it is given by:

$$F_{nk}(\omega) = \frac{1}{N_q} \sum_{q\lambda} \left(\sum_m |g^{q\lambda}_{nmk}|^2 + \Lambda^{q\lambda,-q\lambda}_{nnk} \right) \left[2n_{q\lambda}(T) + 1 \right] \delta(\omega - \omega_{q\lambda}). \tag{3.2}$$

We see that the leading contribution to the coupling comes from the LO mode at its maximum frequency, which corresponds to a wave vector roughly around one fourth of the distance between Γ and the zone edge, and therefore is related to phonons with the periodicity of about 8 to 12 times the unit cell. This means that these phonons are probably the most important for the correct description of the absorption edge of the system, also at the exciton level. Their dispersion is dictated, since we are dealing with a polar material, by the long–range electron–phonon interaction that includes the non–analytic part due to the macroscopic longitudinal electric field (dipole–dipole interactions, Sec. 1.6.3).

3.2.2 Redshift of the absorption spectrum

We will now attempt to compute the correction to the *exciton* energies due to their interaction with phonons. We will compare two approaches that modify the excitonic Hamiltonian, Eq. (1.83), in order to account for the effects of lattice vibration. The first one, which we will call *quasiparticle* (QP) method in this Section, is based on Ref. [58]. An analysis of bulk hBN in this framework has been performed in Ref. [58], and a monolayer study appears in Ref. [119] as well. The second approach, which is based on static atomic displacements with calculations in large supercells (see App. B), is called *Williams–Lax* (WL) method and is detailed in Refs. [37, 104, 105]. This is the first time that the WL method is used in conjunction with the BSE.

As we shall see, both approaches have flaws that prevent a complete solution of the problem, but they provide hints as to where such a solution might be found. They also allow us to draw order–of–magnitude conclusions about the importance of lattice vibrations in the description of BN exciton energies. We will be mainly concerned with the description of indirect absorption in the rest of the thesis, but we

consider it useful to present these results here as a demonstration that strong ep and ExcP coupling are present in BN, leading to non–negligible effects. Therefore, this should be kept in mind when attempting an accurate description of single–particle and exciton energies, as well as absorption and emisssion spectra.

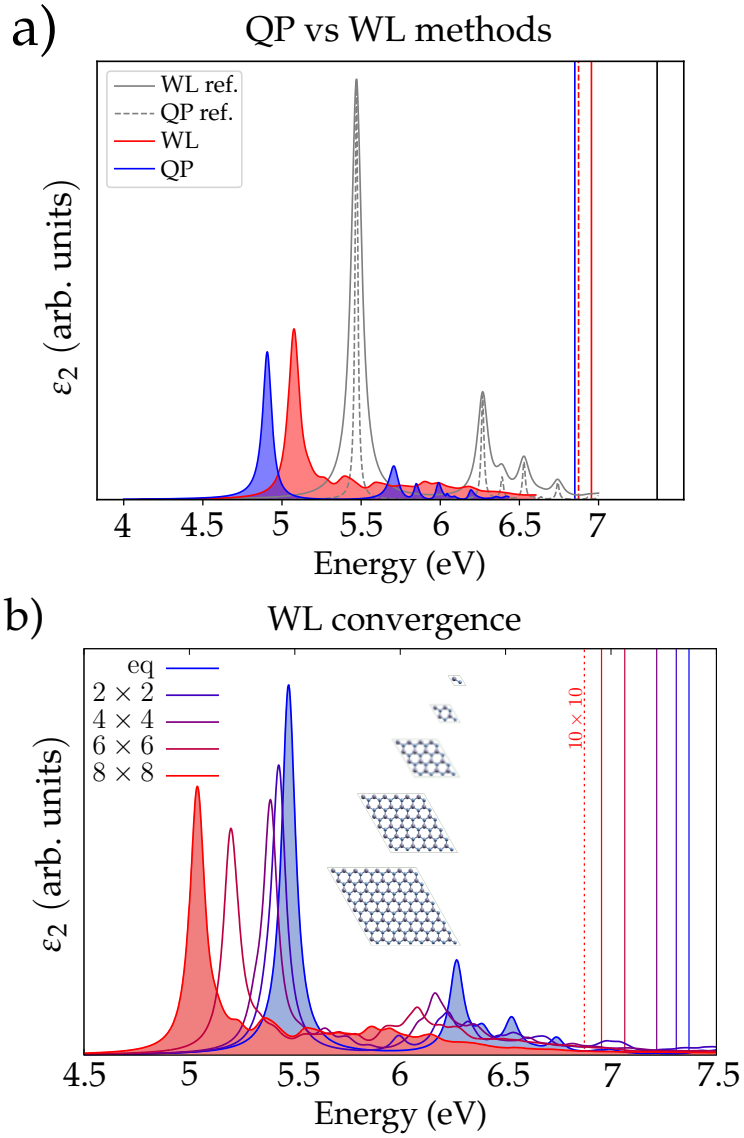


FIGURE 3.3: Zero–temperature phonon dressing of the optical absorption spectrum (ε_2) in monolayer hBN. (a) Comparison between quasiparticle (QP) and Williams–Lax (WL) approaches (see text). Full gray and dashed gray lines: reference calculations in the unit cell for WL and QP methods, respectively. Red line: QP method. Blue line: WL method (8×8 supercell). The vertical lines represent the energy of the GW band gaps. Black: unit cell reference; blue: QP method; full red and dashed red: WL method with 8×8 and 10×10 supercells, respectively. (b) Supercell convergence of the WL method. From blue (1×1) to red (8×8), the spectra for displaced supercells of increasing size (see insets) are shown. The vertical lines represent the corresponding GW band gaps.

Quasiparticle method

In the QP method, we ignore the dependence of the screened interaction and exchange matrix elements on the atomic displacements, and we correct only the

independent–particle (IP), diagonal block of Eq. (1.83). In particular, we include the previously computed ep self–energy corrections to the KS energies in its first term $\Delta E_{\mathcal{K}}\delta_{\mathcal{K}\mathcal{K}'}$. Since these are complex quantities, the excitonic Hamiltonian will now be non–hermitian and yield complex exciton energies $\mathcal{E}_\lambda$. Consequently, the energy–conserving delta function in the absorption spectrum will become a Lorentzian with an *ab initio* full width at half maximum:

$$\pi\delta(\omega - E_\lambda) \to \frac{\mathrm{Im}\mathcal{E}_\lambda}{(\omega - \mathrm{Re}\mathcal{E}_\lambda)^2 + (\mathrm{Im}\mathcal{E}_\lambda)^2}. \tag{3.3}$$

The modified BSE result for optical absorption is displayed by the blue spectrum in Fig. 3.3(a). The frozen–atoms spectrum (obtained with a GW+BSE calculation) is given by the dashed gray line. We find the main exciton energy to be redshifted by 0.56 eV, following the QP band gap redshift (blue vertical line).

What is the validity of these results? Here, we are computing the BSE kernel in the frozen–atoms configuration, which means neglecting any phonon–mediated renormalisation of the electron–hole interaction. The significance of the error introduced in this way is difficult to evaluate: it may be that a correct treatment of the BSE kernel would partially cancel the IP correction to the exciton energies (in fact this happens in the case of the spectral weights of QP spectral functions[120]). Another problem is the finite lifetime of the lowest–lying exciton. According to recent treatments of the exciton–phonon couplings,[107] the lowest–lying excitation, having no state to decay into, should have infinite lifetime. However, as we have already seen, this exciton is composed of a linear combination of electronic transitions around the K point in the BZ. Only the band gap transition at K will have an infinite lifetime (i.e., for this transition the imaginary part of the self–energy is exactly zero). All the other transitions will contribute complex electron and hole energy values whose imaginary parts will not cancel exactly, resulting in a complex exciton energy also for the lowest–bound state. Again, it is possible that including on the same footing the electron–phonon interaction in the BSE kernel will lead to the needed cancellations.[2] Finally, the QP method can only *correct the exciton peaks that are already present in the frozen–atom configuration*. In other words, additional components of the absorption spectrum due to phonon–mediated *indirect* excitations cannot be captured in this way. Another way to see this is considering that the reduction in the overall lattice symmetry due to the atomic displacements along the various phonon normal modes may lead excitons that are dark in the frozen–atoms configuration to acquire some oscillator strength.

Williams–Lax method

Here we consider the thermal average of the imaginary part of the electronic dielectric function, that can be expressed as[37, 104]

$$\varepsilon_2(\omega; T) = \frac{1}{\mathcal{Z}} \sum_n \mathrm{e}^{-E_n/k_B T} \langle \varepsilon_2(\omega; \{\mathbf{R}\}) \rangle_n. \tag{3.4}$$

Here, the energy n labels a nuclear quantum state evaluated in the Born–Oppenheimer approximation (therefore it is an expectation value of the ionic Hamiltonian involving nuclear wave functions). The partition function is $\mathcal{Z} = \sum_n \exp(-E_n/k_B T)$ and the expectations values $\langle \varepsilon_2 \rangle_n$ of the dielectric function, parametrically depending on

[2] This possibility was discussed during the Workshop on Electron–Phonon Interaction (http://elphon.etsf.eu/) and privately with A. Marini.

the set of ionic coordinates **R**, are taken with respect to each nuclear state n (they are therefore complicated integrals involving all the ionic coordinates). This expression implies a semiclassical approximation of the ionic motion. If we now consider the harmonic approximation, we can write the atomic displacementes in normal coordinates x_λ, where the index λ combines both phonon branches and wave vectors (it can be consider as a real–space index where wave vectors are replaced by additional *folded* phonon branches with the corresponding periodicities, see App. B). Furthermore, in this case Eq. (3.5) can be recast as a product of Gaussian integrals

$$\varepsilon_2(\omega;T)\prod_\lambda \int \mathrm{d}x_\lambda \frac{\mathrm{e}^{x_\lambda^2/2\sigma_{\lambda T}^2}}{\sqrt{2\pi}\sigma_{\lambda T}}\varepsilon_2(\omega;\{x\}), \tag{3.5}$$

where $\sigma_{\lambda T}$ is the thermal average of the squared displacement x_λ^2 and will be given explicitly in Eq. (3.21). This integral has been evaluated stochastically in Ref. [37] by generating a large number of ionic configurations for $\varepsilon_2(\omega;\{x\})$. A final simplification is achieved by noting that in fact only a single, *optimal* ionic configuration is needed to obtain the same result as the full integral,[105] but only in the limit of an infinitely large supercell:

$$\varepsilon_2(\omega;T) = \lim_{N\to\infty} \varepsilon_2(\omega;\{x\}^{\mathrm{opt}}), \tag{3.6}$$

with N being the supercell size[3] and the atomic displacements making up $\{x\}^{\mathrm{opt}}$ given by

$$\Delta R_{i\alpha} = \sum_\lambda (-1)^{\lambda-1}\xi_{i\alpha}^\lambda \frac{\hbar}{2M_i\omega_\lambda}\sqrt{2n_\lambda(T)+1}. \tag{3.7}$$

Equation (3.6) was successfully applied to the IP optical spectra of silicon, diamond and gallium arsenide.

We will now use it at zero temperature (zero–point motion only) on top of BSE calculations for monolayer hBN supercells; the optimal atomic configurations for each N were provided by M. Zacharias. In the limit of infinite supercell, the entire BZ is mapped onto the Γ point, forming a continuum of states. If no atomic displacements are present, i.e. we are in the frozen–atoms case, then the optical absorption spectrum will be exactly the same as the one computed in the unit cell, as no new physics has been added to the system. The additional states folded onto the Γ point can only be dark. However, the reduced symmetry due to the static atomic displacements will allow these finite–q excitations mapped to Γ to acquire some oscillator strength depending on the wave vector (i.e., the long–range periodicity of the displacements). All possible phonon modes are included in the sum in Eq. (3.7), and each supercell size N corresponds to an equivalent wave vector sampling in reciprocal space.

For these calculations we applied a scissor operator to the KS energies in order to reproduce exactly the frozen–atoms GW correction for all the supercells, without the need to run an expensive GW calculation for each N. We are therefore neglecting the variation in the GW *correction* to the energy eigenvalues due to the atomic displacements. More information about the scissor operator can be found in Sec. D.5. The next issue is the calculation of the static RPA screening χ^{RPA}, which becomes unfeasible for $N > 6$ with our computational resources. For this reason, we compute χ^{RPA} in the frozen–atoms unit cell BZ and then *fold* it onto the supercell BZ developing a Python script interfaced with `Yambo`. This is explained in Sec. B.1. In this way, we are

[3] In the case of monolayer hBN, a $N \times N \times 1$ supercelll is associated to each N.

also neglecting the variation in the *screening function* due to atomic displacements. However, we could compare this approximation to the exact value for all N up to 6 and found no appreciable difference. The last step is the BSE one. Here, the WL theory provides a *static* correction to the excitonic Hamiltonian, Eq. (1.83), which affects both single–particle energies and the electron–hole interaction kernel.

The convergence of the WL spectrum with respect to the supercell size N is shown in Fig. 3.3(b). The equilibrium, unit–cell result is in blue and the final correction due to lattice vibrations ($8 \times 8 \times 1$ supercell) is in red. The QP band gaps corresponding to the various supercells are marked by vertical lines (the dotted red line is for a $10 \times 10 \times 1$ supercell). Unfortunately, the BSE results could not be fully converged, as calculations for $N > 8$ turned out to be unfeasible. We notice immediately a strong redshift of both the QP band gap and exciton peak energies (the latter being 0.43 eV), as expected. In the WL method, the broadening of the peak does not come from the imaginary part of the exciton energy, but rather from the presence of the folded copies of the exciton around its main peak position. These folded states are slightly displaced in energy with respect to the central one, and acquire a smaller oscillator strength: their sum "dresses" the main peak (the broadening is mostly visible at the bottom of the peak). At the same time, because of the static displacements, many other states that would be optically forbidden at equilibrium now become visible across the full energy range, generating a richer overall structure with respect to what is possible to obtain with the QP method. We note that while the 2×2 and 4×4 peaks remain very close to the equilibrium one, there is a larger energy step starting with the 6×6 supercell. This is consistent with the observation, made in the previous Section, that phonons with 6–fold periodicity and beyond are the ones that couple significantly with the single–particle transitions that are responsible for this excitonic shift.

Conclusion

Figure 3.3(a) shows a comparison between the phonon–dressed optical spectra obtained with the QP (red) and WL (blue) approaches. The vertical lines represent the QP band gaps (the dashed line being the 10×10 WL calculation, and the black line being the equilibrium reference). In order to compare the spectra, a partial f–sum rule was enforced between reference (gray, $\varepsilon_2^{\mathrm{Ref}}$) and final results separately for both methods, rescaling ε_2 in order to satisfy

$$\int_{\omega_i-\delta}^{\omega_f-\delta} \omega\varepsilon_2(\omega) = \int_{\omega_i}^{\omega_f} \omega\varepsilon_2^{\mathrm{Ref}}(\omega). \tag{3.8}$$

The energy window limits are $\omega_i = 4.5$ eV and $\omega_f = 7$ eV (almost all of the oscillator strength in the unit cell case is contained in this interval), while δ is the value of the peak shift.

It is clear that the dominant contribution to the phonon dressing of the excitons comes from the IP block of the Bethe–Salpeter Hamiltonian, because in the WL case the supercell convergence trend of the BSE follows that of the QP gap: therefore the quasiparticles are dressed *first*, and then the electron–hole interaction binds them into exciton states. However, the single–particle energy variation at each step is bigger than the corresponding exciton peak variation, therefore it seems likely that the contribution from the dressing of the BSE kernel (an electron–hole interaction modified by lattice vibrations) opposes the IP contribution, reducing the redshift.

Certainly, a more rigourous and formal study of the connection of the WL method to a MBPT approach is required to definitively assess the accuracy of this method. Similarly, the QP method needs to be expanded in order to include the BSE kernel (we will give a contribution in this direction in Chapter 4). In this context, a theoretical investigation on the coupling of *excitons* and long–range optical phonons in polar materials (sometimes called Fröhlich coupling[121]) would also be beneficial. Despite all these shortcomings, we can still conclude, based on the results for the monolayer, that one should expect a very strong zero–point redshift of the optical absorption spectrum in BN systems of approximately 0.5 eV.

We will now shift our focus to systems with an indirect energy gap: here, the presence of a strong electron–phonon coupling can considerably change the relevant spectral features, particularly by allowing for the creation of additional excitations, which may dominate the spectral onsets at low energies.

3.3 Basics of indirect absorption

Indirect absorption in an independent–particle (IP) picture is usually conceptualised as a two–step process in which a valence electron interacts first with a photon and then with a phonon — or viceversa. The excited electron, having picked up both energy and momentum, will end up in a conduction band. Therefore, this process can be modeled with second–order time–dependent perturbation theory starting from an IP electronic Hamiltonian interacting with both an external em field and an ionic potential. We will do this now because some of the developments will be relevant to both Sec. 3.4, involving excitons, and Sec. 4.1.2, where a "diagrammatic" approach is used to obtain the same results. Additionally, this is the approach taken in Ref. [36] in order to reproduce, from *ab initio* calculations, the phonon–assisted shoulder in the absorption spectrum of silicon. The expression for the indirect ε_2 found in the above–mentioned reference is reported in App. E as Eq. (E.1). Actually, a careful derivation of this expression (following for example the prescriptions in Chapter XII of the book of Grosso and Pastori Parravicini[43]) reveals that additional terms are required to obtain the "full" formula of the indirect ε_2, and its dependence on the phononic and electronic occupation factors is slightly more complicated. Nonetheless, the reduced formula presented in Eq. (E.1) is enough to describe indirect absorption in silicon. However it may be relevant to discuss the full expression since there might be cases in which it would be needed, for example when the occupation factors of the initial and final electronic states might be comparable (i.e. metals and interband transitions as opposed to the absorption case where $f_{ck} \ll f_{vk-q}$). This is also relevant in order to assess properly under which approximations several simplifications in said expression might be possible. In order to be consistent with Eq. (E.1), we reintroduce here $\hbar$. The phonon frequencies will be denoted with the capital Omega as $\Omega_{q\lambda}$.

3.3.1 Any phonon–assisted transitions

Below we list the starting non–interacting electronic Hamiltonian H_0, the coupling Hamiltonian H_{eL} with a classical em field (up to linear order in the field $\mathbf{A}$, in the Coulomb gauge and in the dipole approximation) and the coupling Hamiltonian

H_{e-ion} with the crystal lattice vibrations via an electron–ion potential:

$$\begin{aligned} H_0(\mathbf{r}) &= \frac{p^2}{2m} + V(\mathbf{r}), \\ H_{eL}(\mathbf{r},t) &= \frac{eA_0}{mc}\mathbf{e}\cdot\mathbf{p}\ \left(\mathrm{e}^{-\mathrm{i}\omega t} + \mathrm{e}^{\mathrm{i}\omega t}\right), \\ H_{e-ion}(\mathbf{r},t) &= \frac{1}{N_q}\sum_{q\lambda}\left(V^{\lambda}_{e-ion}(\mathbf{q},\mathbf{r})\,\mathrm{e}^{-\mathrm{i}\Omega_{q\lambda}t}\hat{b}_{q\lambda} + V^{\lambda,*}_{e-ion}(\mathbf{q},\mathbf{r})\,\mathrm{e}^{\mathrm{i}\Omega_{q\lambda}t}\hat{b}^{\dagger}_{q\lambda}\right). \end{aligned} \tag{3.9}$$

An arbitrary state of the system will be specified by an electronic band state and a phonon number: $|mk\rangle \otimes |n_{q\lambda}\rangle$. It is understood that the phononic occupation functions are produced by the action of H_{e-ion} on the initial state $|n_{q\lambda}\rangle$:

$$\begin{aligned} \hat{b}_{q\lambda}\,|n_{q\lambda}\rangle &= \sqrt{n_{q\lambda}}\,|n_{q\lambda}-1\rangle\,, \\ \hat{b}^{\dagger}_{q\lambda}\,|n_{q\lambda}\rangle &= \sqrt{n_{q\lambda}+1}\,|n_{q\lambda}+1\rangle\,. \end{aligned} \tag{3.10}$$

Now, the perturbation theory is carried out up to second order, but only the terms at first–order simultaneously in both the em field and the phonon number will be retained (these represent phonon–assisted transitions; the rest describe two–photon and two–phonon processes). We also give the following definitions for the matrix elements that describe the couplings:

$$\begin{aligned} \langle jk|\mathbf{e}\cdot\mathbf{p}|ik\rangle &\equiv \mathbf{e}\cdot\mathbf{v}_{ijk} \equiv d_{ijk}, \\ \langle jk|V^{\lambda}_{e-ion}(\mathbf{q},\mathbf{r})|ik-q\rangle &\equiv g^{q\lambda}_{ijk}. \end{aligned} \tag{3.11}$$

A transition $i \to j$ described in this way will undergo a two–step process: first the absorption/emission of a photon brings the electron from i to an intermediate "virtual" state m; then the absorption/emission of a phonon bridges the remainining energy–momentum gap to the final state j. Alternatively, the phonon will be absorbed/emitted first, leading to the intermediate state; the absorption/emission of a photon will then complete the process. The probability of the total $i \to j$ transition will be given by the modulus square of the sum of the two processes, thereby including quantum interference effects.

The complete expression for the imaginary part of the dielectric function is then

$$\begin{aligned} \varepsilon_2(\omega) = C(\omega)\sum_{ijkq\lambda} f_{ik-q}(1-f_{jk})\{ \quad & n_{q\lambda}|T_A|^2\delta(\epsilon_{ik-q}-\epsilon_{jk}+\hbar\omega+\hbar\Omega_{q\lambda}) + \\ & (n_{q\lambda}+1)|T_B|^2\delta(\epsilon_{ik-q}-\epsilon_{jk}+\hbar\omega-\hbar\Omega_{q\lambda}) + \\ & n_{q\lambda}|T_C|^2\delta(\epsilon_{ik-q}-\epsilon_{jk}-\hbar\omega+\hbar\Omega_{q\lambda}) + \\ & (n_{q\lambda}+1)|T_D|^2\delta(\epsilon_{ik-q}-\epsilon_{jk}-\hbar\omega-\hbar\Omega_{q\lambda})\} \end{aligned} \tag{3.12}$$

(the coefficient $C(\omega)$ is given in App. E). Here, the first two lines refer to "resonant" transitions (e.g., Eq. (E.1)) and the last two to "anti–resonant" ones. Ignoring the latter is equivalent to a "Tamm–Dancoff" approximation. The T–factors are the transition amplitudes: each of the four possible processes (as seen in Fig. 3.4(a)) is split into two contributions (for a total of 8 terms) reflecting that the system can reach

the intermediate state both via photon or phonon. Such terms are given by

$$\begin{aligned} T_A &= S_1^{\omega} + S_1^{\Omega}, \\ T_B &= S_2^{\omega} + S_2^{\Omega}, \\ T_C &= S_3^{\omega} + S_3^{\Omega}, \\ T_D &= S_4^{\omega} + S_4^{\Omega}. \end{aligned} \tag{3.13}$$

The S–terms describe all the possible combinations of absorption or emission of one photon and one phonon. Their explicit expression is listed below:

		Photon	Phonon
$S_1^{\omega}(k,q)$	$= \sum_m \frac{d_{imk-q}\, g^{q\lambda}_{mjk}}{\epsilon_{ik-q} - \epsilon_{mk-q} + \hbar\omega + i\eta}$	*Absorbed*	*Absorbed*
$S_1^{\Omega}(k,q)$	$= \sum_m \frac{g^{q\lambda}_{imk}\, d_{mjk}}{\epsilon_{ik-q} - \epsilon_{mk} + \hbar\Omega_{\lambda q} + i\eta}$	*Absorbed*	*Absorbed*
$S_2^{\omega}(k,q)$	$= \sum_m \frac{d_{imk-q}\, \overline{g}^{q\lambda}_{mjk}}{\epsilon_{ik-q} - \epsilon_{mk-q} + \hbar\omega + i\eta}$	*Absorbed*	*Emitted*
$S_2^{\Omega}(k,q)$	$= \sum_m \frac{\overline{g}^{q\lambda}_{imk}\, d_{mjk}}{\epsilon_{ik-q} - \epsilon_{mk} - \hbar\Omega_{\lambda q} + i\eta}$	*Absorbed*	*Emitted*
$S_3^{\omega}(k,q)$	$= \sum_m \frac{\overline{d}_{imk-q}\, g^{q\lambda}_{mjk}}{\epsilon_{ik-q} - \epsilon_{mk-q} - \hbar\omega + i\eta}$	*Emitted*	*Absorbed*
$S_3^{\Omega}(k,q)$	$= \sum_m \frac{g^{q\lambda}_{imk}\, \overline{d}_{mjk}}{\epsilon_{ik-q} - \epsilon_{mk} + \hbar\Omega_{\lambda q} + i\eta}$	*Emitted*	*Absorbed*
$S_4^{\omega}(k,q)$	$= \sum_m \frac{\overline{d}_{imk-q}\, \overline{g}^{q\lambda}_{mjk}}{\epsilon_{ik-q} - \epsilon_{mk-q} - \hbar\omega + i\eta}$	*Emitted*	*Emitted*
$S_4^{\Omega}(k,q)$	$= \sum_m \frac{\overline{g}^{q\lambda}_{imk}\, \overline{d}_{mjk}}{\epsilon_{ik-q} - \epsilon_{mk} - \hbar\Omega_{\lambda q} + i\eta}.$	*Emitted*	*Emitted*

(3.14)

Note that the barred matrix elements, e.g. $\overline{g}^{q\lambda}_{mjk}$, indicate the mj-th element of the conjugate and not the conjugate of $g^{q\lambda}_{mjk}$. Here $\eta \to 0^+$ and if we consider only the first four terms we obtain Eq. (E.1) (except for the structure of the electronic occupation factors).

3.3.2 Net transition rate

The *net* transition rate R is the difference between the $i \to j$ and $j \to i$ transition rates. If i is a valence band and j a conduction one, then the net rate will be the difference between absorption and emission rates (and it will remain an absorption rate if $f_v > f_c$). If we consider the band energies to be independent from their occupations, the $j \to i$ case is found by just swapping the indices $ik - q$ and jk in Eq.

(3.12). In the "Tamm-Dancoff" case (TD) the net absorption rate reads

$$R^{TD}(\omega) = C(\omega) \sum_{ijkq\lambda} \left[|T_A|^2 n_{q\lambda} f_{ik-q} \overline{f}_{jk} - |T_D|^2_{i\leftarrow j} (n_{q\lambda}+1) f_{jk} \overline{f}_{ik-q} \right] \delta(\epsilon_{ik-q} - \epsilon_{jk} + \hbar\omega + \hbar\Omega_{q\lambda}) + \left[|T_B|^2 (n_{q\lambda}+1) f_{ik-q} \overline{f}_{jk} - |T_C|^2_{i\leftarrow j} n_{q\lambda} f_{jk} \overline{f}_{ik-q} \right] \delta(\epsilon_{ik-q} - \epsilon_{jk} + \hbar\omega - \hbar\Omega_{q\lambda}) \quad (3.15)$$

where $i \leftarrow j$ means that the indices $ik-q$ and jk are exchanged in Eq. (3.14) and $\overline{f}_{jk} \equiv 1 - f_{jk}$. In principle, the different denominators prevent exact factorisation of the terms on each of the two lines. However in the optical absorption/emission case the transition amplitudes are often considered slowly varying if only states close to the band minima contribute.[43]

In the simpler case of direct transitions, we instead obtain Eq. (1.74),

$$\varepsilon_2^{\rm dir}(\omega) \propto \sum_{ijk} |d_{ijk}|^2 f_{ik}(1-f_{jk}) \delta(\epsilon_{ik} - \epsilon_{jk} + \hbar\omega), \quad (3.16)$$

with the net rate given simply by

$$R^{\rm dir}(\omega) = \varepsilon_2^{\rm dir} - \varepsilon_2^{\rm dir}|_{i\leftarrow j} \propto \sum_{ijk} |d_{ijk}|^2 (f_{ik} - f_{jk}) \delta(\epsilon_{ik} - \epsilon_{jk} + \hbar\omega). \quad (3.17)$$

3.3.3 Optical absorption and emission

If we are considering optical transitions across a band gap (from valence to conduction), then the resonant lines of Eq. 3.12 describe absorption, the anti–resonant lines describe emission:

$$\begin{aligned} \varepsilon_2^{\rm abs}(\omega) = C(\omega) \sum_{cvkq\lambda} f_{vk-q} \overline{f}_{ck} \{ & n_{q\lambda} |T_A|^2 \delta(\epsilon_{vk-q} - \epsilon_{ck} + \hbar\omega + \hbar\Omega_{q\lambda}) + \\ & (n_{q\lambda}+1) |T_B|^2 \delta(\epsilon_{vk-q} - \epsilon_{ck} + \hbar\omega - \hbar\Omega_{q\lambda}) \}, \\ \varepsilon_2^{\rm em}(\omega) = C(\omega) \sum_{cvkq\lambda} \overline{f}_{vk-q} f_{ck} \{ & n_{q\lambda} |T_C|^2 \delta(\epsilon_{vk-q} - \epsilon_{ck} - \hbar\omega + \hbar\Omega_{q\lambda}) + \\ & (n_{q\lambda}+1) |T_D|^2 \delta(\epsilon_{vk-q} - \epsilon_{ck} - \hbar\omega - \hbar\Omega_{q\lambda}) \}. \end{aligned} \quad (3.18)$$

Figure 3.4(a) schematises the indirect absorption and emission processes between a valence and a conduction state, the processes labeled A, B, C or D in accordance with their respective transition amplitudes T. Note that only T_B and T_D (phonon emission processes) are possible at zero temperature. The resulting optical absorption (blue) and emission (red) spectra for the two levels are depicted in Fig. 3.4(b) (note that the two photon frequencies involved are ω_1 and $\omega_2 = \omega_1 + 2\Omega$).

3.3.4 Static approximation

Let us now restrict ourselves, for simplicity, to the photon absorption case only. We use the static approximation and neglect the phonon frequencies in the transition amplitudes, i.e. $\hbar\Omega_{q\lambda} = 0$, leading to $S_1^\omega + S_1^\Omega = S_2^\omega + S_2^\Omega$. Now, we also take the "resonant" approximation for the denominators where the photon frequency appears explicitly, i.e. $\epsilon_{vk-q} - \epsilon_{mk-q} + \hbar\omega + i\eta \simeq \epsilon_{ck} - \epsilon_{mk-q} + i\eta$ (this is reasonable if we took the static approximation for the phonon frequencies before). Now, for the

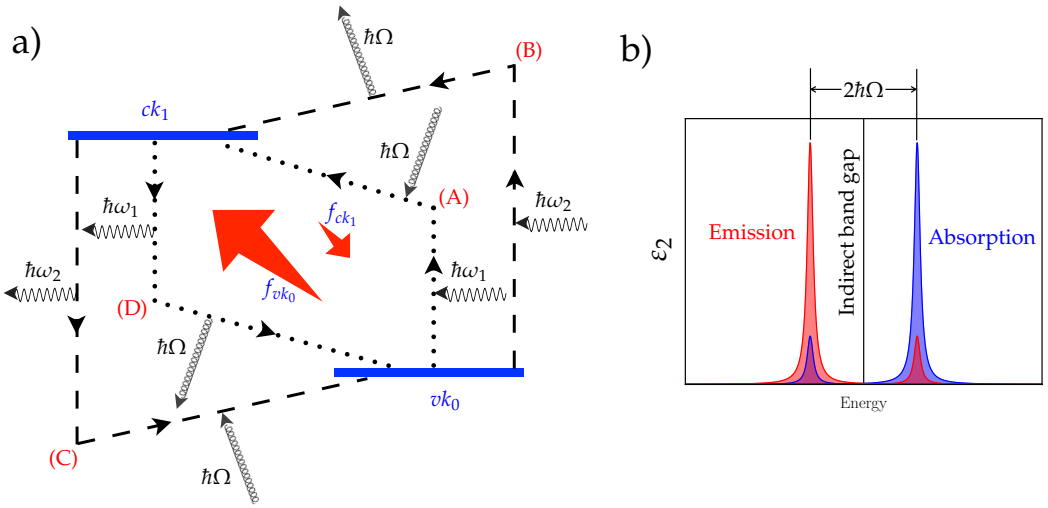


FIGURE 3.4: Phonon–assisted transitions between a valence vk_0 and a conduction ck_1 state. (a) Scheme of the possible transitions, from right to left: photon absorption with phonon emission (B), photon absorption with phonon absorption (A), photon emission with phonon emission (D) and photon emission with phonon absorption (C). The red arrows indicate a net absorption process. (b) Resulting absorption (blue) and emission (red) spectra. The leading peaks are produced by phonon emission.

absorption case, we arrive at the Hall–Bardeen–Blatt theory[114]:

$$\varepsilon_2^{HBB}(\omega) = \sum_{cvkq\lambda} \left| \sum_m \frac{d_{vmk-q} g^{\lambda q}_{mck}}{\epsilon_{ck} - \epsilon_{mk-q} + i\eta} + \sum_m \frac{g^{\lambda q}_{vmk} d_{mck}}{\epsilon_{vk-q} - \epsilon_{mk} + i\eta} \right|^2 (2n_{q\lambda} + 1) f_{vk-q} \overline{f}_{ck} \delta(\epsilon_{vk-q} - \epsilon_{ck} + \hbar\omega + \hbar\Omega_{q\lambda}). \tag{3.19}$$

The HBB expression (albeit with $\hbar\Omega_{\lambda q} = 0$ also in the delta function) can actually be obtained in a different way: that is by taking static derivatives of the direct absorption expression with respect to atomic displacements:[104, 105]

$$\varepsilon_2^{HBB}(\omega; T; \{\Omega\} = 0) = \sum_{q\lambda} \frac{\partial^2 \varepsilon_2^{dir}(\omega; \{R_{\lambda q}\})}{\partial R^2_{\lambda q}} \sigma^2_{\lambda q}(T). \tag{3.20}$$

Here $R_{\lambda q}$ is a set of atomic displacements along phonon mode (λq) and $\sigma^2_{\lambda q}(T)$ is the average squared displacement of an harmonic oscillator (recall that we are considering harmonic phonons),

$$\sigma^2_{\lambda q}(T) = \langle n_{q\lambda} | \hat{x}^2_{q\lambda} | n_{q\lambda} \rangle = \langle n_{q\lambda} | l^2_{\lambda q} (b^\dagger_{q\lambda} + b_{q\lambda})^2 | n_{q\lambda} \rangle = l^2_{\lambda q} (2n_{q\lambda}(T) + 1). \tag{3.21}$$

Here the zero–temperature squared displacement is $l^2_{\lambda q} = \hbar / (2 M_{q\lambda} \Omega_{q\lambda})$. Depending on the normalisation choice for the the phonon eigenvectors, $M_{q\lambda} \equiv M$ is taken as a simple reference mass or, for example, as a weighted "unit–cell" mass $M_{\lambda q} = \sum_i^{N_{\text{ions}}} M_i |\boldsymbol{\xi}_i^{\lambda q}|^2$ (we will use the latter in our numerical simulations).

In the static / HBB case we also have $|T_A|^2 = |T_B|^2 \simeq |T_C|^2 = |T_D|^2$. The only difference between transition rates for optical absorption and emission, once their modulus square is taken, is that in the emission case the positions of the intermediate state energies ϵ_{mk} and ϵ_{mk-q} are swapped between the denominators of S^ω and S^Ω,

something that may in general be safely neglected. In our study about phonon–assisted transitions involving excitons, Sec. 3.4, we will choose an analogous static approach as well.

3.3.5 Luminescence spectra via the van Roosbroeck–Shockley relation

Our goal in the next section will be to model luminescence spectra. However, it is easier for us to obtain *absorption* spectra via the BSE, since in general luminescence can be a quite complicated, essentially out–of–equilibrium process whose full many–body description is presently available only for direct transitions, while non–equilibrium *ab initio* luminescence calculations are still lacking.[122–124] Fortunately, the experimental conditions in which hBN luminescence is measured suggest that various simplifications of the full problem are possible, and eventually allow us to obtain the *spontaneous emission rate* (the physical quantity relevant for experimental luminescence spectra) from the absorption coefficient.[125] In practice this means that it will just be a byproduct of a modified BSE calculation.

The simplifications introduced are the following. We neglect all processes out of equilibrium and assume all excited carriers relaxed to the bottom of the excitonic dispersion curves. We take the probabilities of exciton formation and annihilation to be the same (i.e. we have a *detailed balance* of phonon–assisted creation and annihilation of excitations), a reasonable assumption in our scheme as we will be computing the transition rates in a static approximation. This enables us to employ the *Van Roosbroeck–Shockley (RS) relation*[126] to compute the spontaneous emission rate $R^{\mathrm{sp}}(\omega)$. This entails two commonly held experimental assumptions: (i) the (excited) system is in a steady state, with the contribution from stimulated emission being very small with respect to the spontaneous one; (ii) in the steady state, the occupation functions of the excitonic states involved in recombination processes can be approximated with those at thermal equilibrium. We summarise the RS relation below for both the IP and exciton cases. Note that here we are mainly streamlining the contents of Refs. [125] and [127], where a more exhaustive treatment can be found.

Direct transitions in the independent–particle case

Let us take the net transition rate (per unit energy, per unit volume) between the optical absorption and stimulated emission processes:

$$R'^{,\mathrm{abs}}(\omega) = \mathcal{K}(\omega)\frac{2\pi}{\hbar}\frac{\mathcal{N}(\omega)}{N_k}\sum_{cvk}\mathcal{T}_{cvk}(f_{vk}\overline{f}_{ck} - f_{ck}\overline{f}_{vk})\delta(\epsilon_{ck} - \epsilon_{vk} - \hbar\omega), \tag{3.22}$$

where $\mathcal{T}_{cvk}$ is the squared transition amplitude. We assume a time–independent Fermi–Dirac distribution for electronic occupations in the steady state, $f_{nk} = [1 + \mathrm{e}^{(\epsilon_{nk}-\mu_{e/h})/k_BT}]^{-1}$, with μ_e and μ_h being the chemical potentials for electrons and holes, respectively. In the IP case the transition amplitude can be taken as the optical matrix element in the dipole approximation: $\mathcal{T}_{cvk} = |d_{cvk}|^2$. We can similarly write the expression for the spontaneous emission rate,

$$R^{sp}(\omega) = \mathcal{K}(\omega)\frac{2\pi}{\hbar}\frac{\mathcal{G}(\omega)}{N_k}\sum_{cvk}\mathcal{T}_{cvk}f_{ck}\overline{f}_{vk}\delta(\epsilon_{ck} - \epsilon_{vk} - \hbar\omega). \tag{3.23}$$

The spontaneous emission is only proportional to the photon density of states, $\mathcal{G}(\omega)$, while both absorption and stimulated emission are proportional to the the

photon density per unit energy, $\mathcal{N}(\omega)$. If we define an average photon number, $\overline{\mathcal{N}}$, these two quantities are related by the total photon density $\int \mathcal{N}(\omega)\mathrm{d}\omega = \int \overline{\mathcal{N}}\mathcal{G}(\omega)\mathrm{d}\omega$. The dimensional term $\mathcal{K}(\omega)$ is made of quantities mainly related to the the em field. We now list the relevant expressions for the optical quantities involved.

$$\begin{aligned}
&\text{Dimensional factor} \quad \mathcal{K}(\omega) = \frac{2\pi e^2\hbar^2}{m^2 V}\frac{1}{n_1(\omega)^2\hbar\omega} \\
&\text{Photon density of states} \quad \mathcal{G}(\omega) = \frac{1}{\pi^2c^2\hbar^3}\frac{n_1(\omega)^2(\hbar\omega)^2}{V_g(\omega)} \\
&\text{Group velocity} \quad V_g(\omega) = \frac{c}{n_1(\omega)+\omega\frac{\partial n_1(\omega)}{\partial\omega}} \approx \frac{c}{n_1(\omega)} \\
&\text{Incoming photon flux} \quad \mathcal{F}(\omega) = \mathcal{N}(\omega)\frac{c}{n_1(\omega)}
\end{aligned} \tag{3.24}$$

The absorption coefficient $\alpha(\omega)$ can be written in terms of the absorption rate as $R'^{,\mathrm{abs}}(\hbar\omega) = \mathcal{F}(\omega)\alpha(\omega)$. Finally, we observe that independently from the specific (cvk) transition considered, the following relation always holds:

$$\frac{f_{ck}\overline{f}_{vk}}{f_{vk}\overline{f}_{ck} - f_{ck}\overline{f}_{vk}} = \frac{1}{\mathrm{e}^{(\hbar\omega-(\mu_e-\mu_h))/k_BT}-1} \approx \mathrm{e}^{-(\hbar\omega-\Delta\mu)/k_BT} \tag{3.25}$$

(where in the last step we replaced the resulting Bose–Einstein distribution with a Boltzmann one; this simplification is not necessary). Then, by putting everything together and comparing Eqs. (3.22) and (3.23), we find $R^{\mathrm{sp}}(\omega)$ to be equal to $\mathcal{G}(\omega)V_g(\omega)\alpha(\omega)\mathrm{e}^{-(\hbar\omega-\Delta\mu)/k_BT}$.

This leads us to the RS relation:

$$\begin{aligned}
R^{\mathrm{sp}}(\omega) &= \frac{n_1(\omega)^2(\hbar\omega)^2}{\pi^2c^2\hbar^3}\alpha(\omega)\mathrm{e}^{-(\hbar\omega-\Delta\mu)/k_BT} \\
&= \frac{n_1(\omega)(\hbar\omega)^3}{\pi^2c^3\hbar^4}\varepsilon_2(\omega)\mathrm{e}^{-(\hbar\omega-\Delta\mu)/k_BT},
\end{aligned} \tag{3.26}$$

where for the last equality we have used Eq. (1.69), $\alpha(\omega) = \hbar\omega\varepsilon_2(\omega)/(n_1(\omega)\hbar c)$.

Direct transitions in the exciton case

In this case the absorption coefficient is computed including excitonic effects (just for this Section we label it $\alpha^{\mathrm{exc}}(\omega)$, and equivalently $\varepsilon_2^{\mathrm{exc}}(\omega)$, to avoid confusion); we can always obtain the full refractive index as

$$n_1^{\mathrm{exc}}(\omega) = \sqrt{\frac{1}{2}\sqrt{\varepsilon_1^{\mathrm{exc}}(\omega)^2+\varepsilon_2^{\mathrm{exc}}(\omega)^2}+\varepsilon_1^{\mathrm{exc}}(\omega)}. \tag{3.27}$$

Now the transition amplitude $\mathcal{T}_{cvk}$ in Eqs. (3.22) and (3.23) will be replaced by the excitonic one, $\mathcal{T}_\lambda = |\sum_{cvk}\overline{A}^\lambda_{cvk}d_{cvk}|^2$, with the external sum now running over the exciton index λ. Analogously, the energies of single–particle transitions will be replaced by the exciton energies E_λ. Concerning the occupation functions, if the absorption/emission features are dominated by the creation/annihilation of electron–hole bound pairs, it is sufficient to replace the Bose–Einstein/Boltzmann factor in Eq. (3.26) with a more appropriate term to describe the occupation of excitonic states. We use the Boltzmann distribution $n_B(\hbar\omega) = \mathrm{e}^{-(\hbar\omega-\mu^*)/k_BT}$, with μ^* fixed

to the energy of the lowest–bound exciton. Since below the quasiparticle gap we are dealing with a discrete excitonic spectrum, n_B will be a discrete function taking values for $\hbar\omega = E_\lambda$. Note that the use of a Boltzmann (or equivalently Bose–Einstein) occupation function implies that we are treating excitons as pure bosons, i.e. the excitations of a non–interacting bosonic Hamiltonian with commuting excitonic creation and annihilation operators. This is compatible with our previous assumption of low laser power density, which leads to a low density of excitations inside the material: in this regime excitons may be treated as non–interacting bosons.[107, 128]

Indirect transitions

In this case we have to take into account that the energy of a photon absorbed ($\hbar\omega_a$) and that of a photon emitted ($\hbar\omega_e$) in a process mediated by the same phonon are not the same, and they are both different from the energy of the indirect electronic transition ($\hbar\omega$). In particular, with the help of Figs. 3.4(a) and 3.5 (the latter adapted to the exciton case), we can write the following relations:

$$\begin{aligned} \hbar\omega_e &= \hbar\omega_a \pm 2\hbar\Omega_{q\mu} \\ \hbar\omega_e &= \hbar\omega \pm \hbar\Omega_{q\mu} \\ \hbar\omega_a &= \hbar\omega \mp \hbar\Omega_{q\mu}, \end{aligned} \tag{3.28}$$

(the upper and lower signs referring to the cases of phonon absorption and emission, respectively). We need to write a generalised form of the RS relation that takes these energy differences into account. Focusing on the case of an indirect transition mediated by the *emission* of a single phonon of branch μ and momentum q, the second–order absorption and emission rates (per unit time, energy and volume) can be expressed as (back in the IP case):

$$\begin{aligned} R'^{,\mathrm{abs}}_{\mu q}(\omega_a) = \mathcal{K}(\omega_a)\frac{2\pi}{\hbar}\frac{\mathcal{N}(\omega_a)}{N_k}\sum_{cvk}\mathcal{T}^{(2)}_{cvkq\mu}(n_{q\mu}+1)[f_{vk-q}\overline{f}_{ck} - f_{ck}\overline{f}_{vk-q}] \\ \delta(\epsilon_{ck} - \epsilon_{vk-q} + \hbar\Omega_{q\mu} - \hbar\omega_a), \\ R^{sp}_{\mu q}(\omega_e) = \mathcal{K}(\omega_e)\frac{2\pi}{\hbar}\frac{\mathcal{G}(\omega_e)}{N_k}\sum_{cvk}\mathcal{T}^{(2)}_{cvkq\mu}(n_{q\mu}+1) f_{ck}\overline{f}_{vk-q}\delta(\epsilon_{ck} - \epsilon_{vk-q} - \hbar\Omega_{q\mu} - \hbar\omega_e), \end{aligned} \tag{3.29}$$

where $\mathcal{T}^{(2)} = |T_B|^2 \simeq |T_D|^2$ and the rate dependence on the phonon mode and wave vector has been explicitly indicated.

Then, considering $\alpha(\hbar\omega_a) = R'_{\mathrm{abs}}(\hbar\omega_a)/\mathcal{F}(\hbar\omega_a)$ and writing the frequency–dependent functions explicitly, we can write the final results:

$$\begin{aligned} R^{sp}_{\mu q}(\omega_e) &= \frac{n_1(\omega_e)n_1(\omega_a)(\hbar\omega_e)(\hbar\omega_a)}{\pi^2c^2\hbar^3}\alpha(\omega_a)\mathrm{e}^{-(\hbar\omega-\Delta\mu)/k_BT} \\ &= \frac{n_1(\omega_e)(\hbar\omega_e)(\hbar\omega_a)^2}{\pi^2c^3\hbar^4}\varepsilon_2(\omega_a)\mathrm{e}^{-(\hbar\omega-\Delta\mu)/k_BT}. \end{aligned} \tag{3.30}$$

Or, in the excitonic case (dropping the labels on the frequencies):

$$R^{sp,\mathrm{exc}}_{\mu q}(\omega) \propto n_1^{\mathrm{exc}}(\omega)\omega(\omega - 2\Omega_{q\mu})^2\varepsilon_2^{\mathrm{exc}}(\omega - 2\Omega_{q\lambda})n_B(\omega). \tag{3.31}$$

The emission spectra shown in the next Section are obtained by using the latter equation and summing over all phonon modes with a specific momentum $q = \overline{q}$, the one

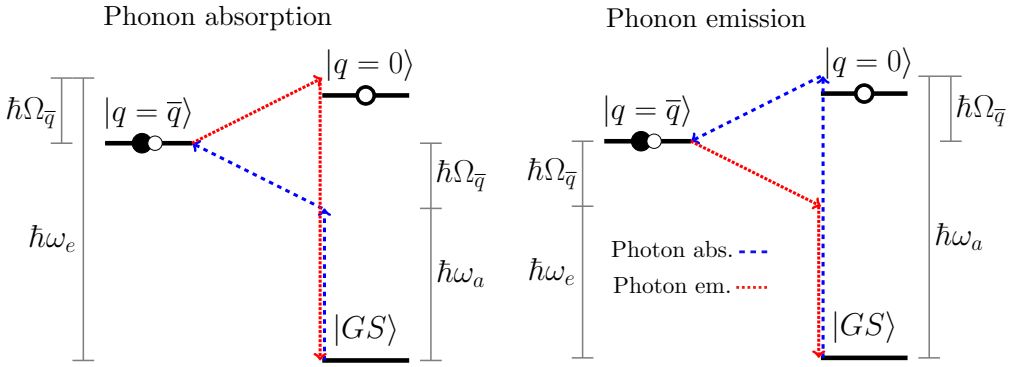


FIGURE 3.5: Simplified transition scheme for a finite–q exciton level at $q = \bar{q}$, displaying the processes of photon absorption (dashed blue arrows) and emission (dotted red arrows) mediated by phonon absorption (left) and phonon emission (right). The "direct" exciton at $q = 0$, lying directly above the excitonic ground state, is also shown with an energy corresponding to the case of bulk hBN.

corresponding to the indirect gap. In the exciton case, the exciton–phonon squared transition amplitude $\mathcal{T}^{(2)}_{\lambda\mu\bar{q}}$ will be computed with a finite–difference method.

3.4 Phonon–assisted optical spectra I: finite–difference corrections to the dielectric function

We will start this Section with an in–depth analysis of the electronic bands, excitons and phonons relevant for indirect transitions in bulk hBN. Then, we will use group theory to infer the selection rules for the exciton–phonon coupling. Finally, we will calculate the optical spectra and present a comparison with experimental results.[41]

3.4.1 Electronic structure and non–diagonal supercell

The relevant part of the band structure (after the sscGW correction) is shown again in Fig. 3.6(a). The direct QP gap is at 6.46 eV, and the indirect one is at 5.96 eV. The direct gap is traditionally identified at the so-called T_1 point[21]: From the Figure, we see that this point lies close to K, along the ΓK symmetry line (to be precise: $|\mathbf{T_1} - \mathbf{K}| \simeq 1/6|\mathbf{K} - \mathbf{\Gamma}|$). However, there are other three points that give a comparable band gap: M, T_2 (along the MK line: $|\mathbf{T_2} - \mathbf{K}| \simeq 1/3|\mathbf{K} - \mathbf{M}| = 1/6|\mathbf{K} - \mathbf{\Gamma}|$), and H, the high–symmetry point directly above K along the out–of–plane direction. In our GW calculations, the band gaps at these points lie in a small 0.1 eV energy interval (with $E_g^H < E_g^{T_2} < E_g^M < E_g^{T_1}$), which corresponds to the accuracy of the GW method. Since we know that GW underestimates the true quasiparticle corrections in hBN, we have to assume that the relative energy differences between these band gaps may change with more refined approximations and/or more accurate calculations. Then, we choose to "average" the true position of the top of the valence band from "around" K to exactly K, by taking $\bar{q} = |\mathbf{K} - \mathbf{M}| = 0.5|\mathbf{K} - \mathbf{\Gamma}|$ as the momentum transfer corresponding to the indirect gap.

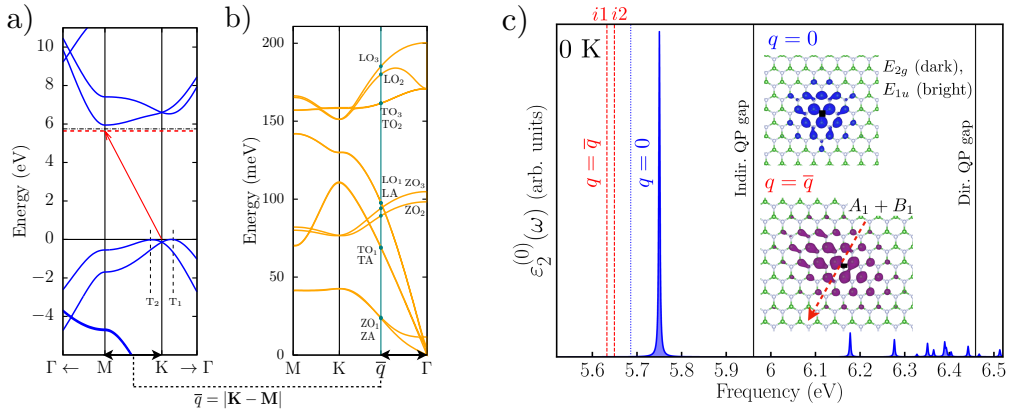


FIGURE 3.6: Electronic, phononic and excitonic properties in bulk hBN (see text). (a) sscGW QP band structure of bulk hBN, displayed in the relevant part of the BZ. The dashed horizontal black line represents the position of the lowest-bound direct exciton, while the dashed red lines correspond to the indirect excitons of momentum $\overline{q}$. The exact valence band maxima are labeled T_1 and T_2. (b) The phonon dispersion in bulk hBN. The teal vertical line highlights the frequencies of the phonon modes with momentum $\overline{q}$. The 12 phonon modes are labeled with the usual notation. (c) The imaginary part of the dielectric function including only transitions at $q = 0$ ($\varepsilon_2^{(0)}(\omega)$ in the text) is shown in blue. The peak broadenings are set to 1.5 meV. The vertical black lines denote the quasiparticle indirect and direct gaps, while the dotted blue line indicates the position of the dark E_{2g} exciton with $q = 0$. The dashed red lines indicate the positions of the B_1 and A_1 excitons at $\overline{q}$ (labeled $i1$ and $i2$, respectively). The excitonic wave functions intensities are plotted in the insets: blue for the $q = 0$ pair, and purple for $q = \overline{q}$ ones (the dashed red arrow labeling the $\overline{q}$ direction).

3.4.2 Excitons

Solving the BSE at $q = 0$, we obtain the two doubly degenerate excitons which form the lowest–lying Davydov pair, as seen in the previous Chapter and depicted in Fig. 3.6(c): the first exciton is at 5.70 eV and is *dark* (dotted blue line). The second exciton at 5.75 eV is instead *bright* (main blue peak). In order to study finite momentum $\overline{q}$ excitons, we perform the same calculations on a non–diagonal hBN supercell, containing 12 atoms per BN layer and chosen such that $\overline{q}$ will be folded onto Γ in its new Brillouin zone, as explained in App. B and Ref. [129]. The supercell we generated for our calculations is represented in Fig. 3.7 and compared with the unit cell. The comparison between the respective reciprocal–space Brillouin zones (BZ) is also shown, emphasizing the folding of the q–point $\overline{q}$ (given in fractional coordinates as $(1/3, -1/6, 0)^T$) and the shape of the supercell BZ (note that both $\overline{q}$ and $-\overline{q}$ are folded onto Γ in this way). Now, two additional excitonic states (dashed red lines in Fig. 3.6(c)) appear at 5.63 and 5.65 eV, respectively, below the lowest–bound direct exciton. These states originate from the splitting, at finite momentum, of the doubly degenerate dark exciton at Γ, as evidenced by the full dispersion calculation displayed in the inset of Fig. 2.8(b). These results are also in agreement with recent Refs. [84, 97]. We denote the two branches, and therefore the two exciton states at $\overline{q}$, as $i1$ and $i2$. These finite–q states are *dark* by themselves but candidates for phonon–assisted absorption and emission.[4] Their wave function intensities are very similar.

[4] This makes clear why we are resorting to a supercell calculation instead of using the full dispersion obtained by solving the finite–q BSE: we need to evaluate the strength of the exciton–phonon coupling and we will do so applying atomic displacements commensurate with the phonon wave vector. The method to perform a full unit–cell calculation of exciton–phonon couplings will be presented in Chapter 4.

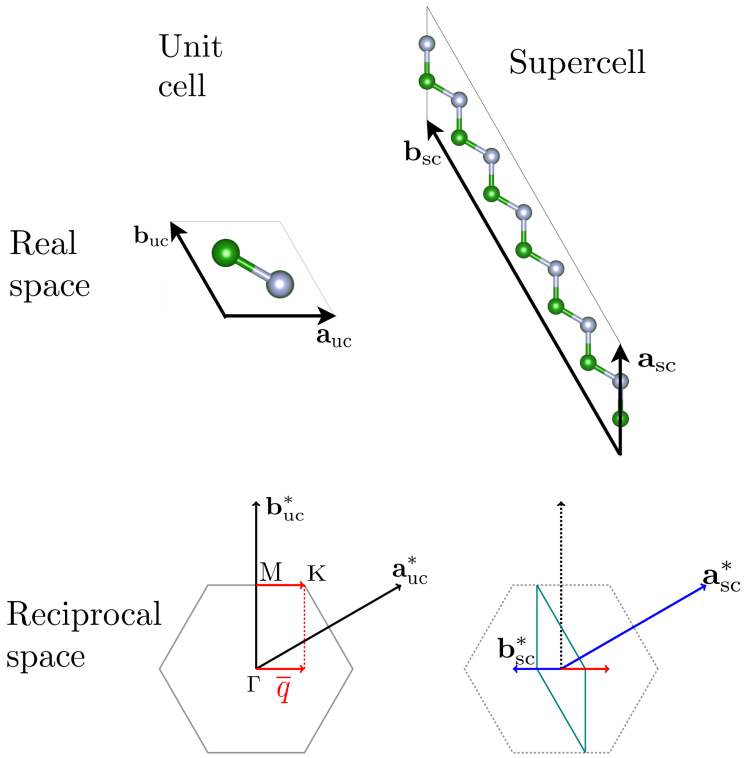


FIGURE 3.7: Bulk hBN Unit cell (uc) versus non–diagonal supercell (sc). Top: representations of the uc and sc used in our calculations. Bottom: schemes of the reciprocal–space BZ in the two cases (the BZ borders are in gray and teal, respectively), showing that in the first Brillouin zone of the supercell, the momentum $\bar{q}$ is folded back onto the Γ point.

One of them is displayed in purple in the second inset of Fig. 3.6(c), showing again a mostly planar distribution as is typical of π–type excitons (some examples obtained by finite–q BSE in the unit cell are displayed in Sec. 2.4). Note that although the wave function in the fixed–hole representation looks approximately distorted along the armchair lattice direction, the full wave function is actually symmetric with respect to the zigzag direction (parallel to the $\bar{q}$–vector) upon rotation around the principal axis of the C_{2v} symmetry group which is oriented in–plane along the $\bar{q}$ direction (see next Section for clarifications on the group theory).

3.4.3 Symmetry of excitons and phonons

In bulk hBN, the point group (including the non–symmorphic point symmetry operations of the space group) is D_{6h}. It contains 24 symmetry operations grouped in 12 classes. It is the group that is also used for the characterizations of perturbations (such as phonons and excitons) of hBN with zero wave–vector (corresponding to the high–symmetry point Γ). In Table 3.1 we report a subsection of the character table focusing on the operations and representations of interest to us. In Fig. 3.8 (left panel) our choice for the Cartesian axes and for the lattice vectors is reported and the three rotation axes belonging to D_{6h} are drawn. Recall that bulk hBN displays AA′ stacking (two inequivalent layers per unit cell): therefore many of the symmetry operations are non–symmorphic. The point group for the symmetry analysis of a perturbation with finite wave vector $\bar{q}$ is a subset of the one at Γ. In the case of $\bar{q}$ lying

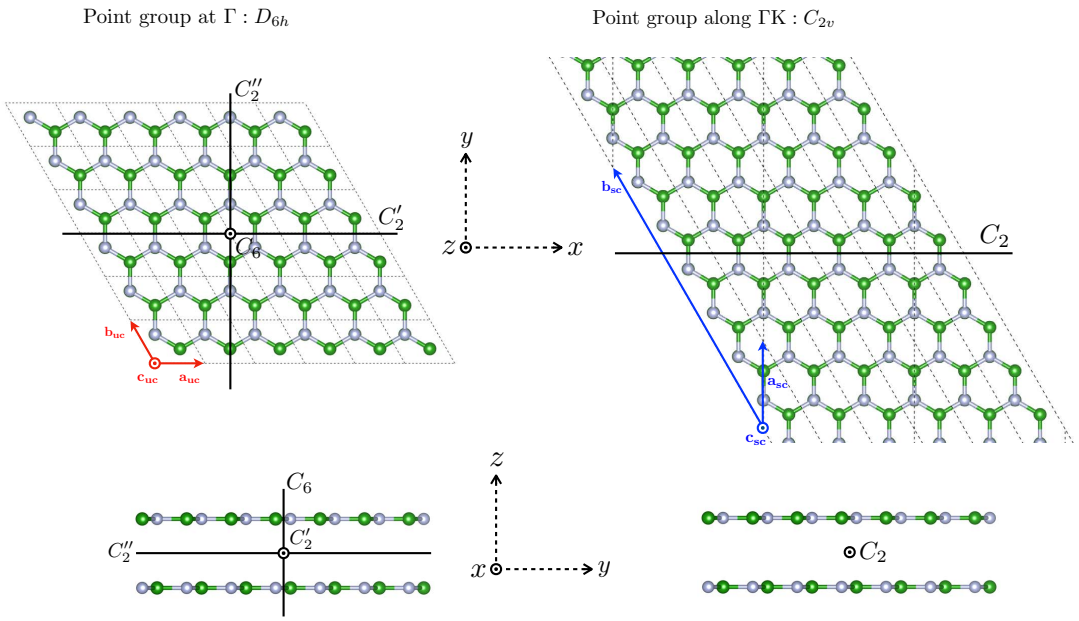


FIGURE 3.8: Symmetries at Γ and $\bar{q}$. Top and side views of the bulk hBN lattice. The borders of the repeated unit cells are shown with dashed black lines. Left: original unit cell for the hexagonal lattice (D_{6h} point group) and lattice vectors. Right: non–diagonal supercell used in the calculations and lattice vectors. The axes of rotation corresponding to the symmetries of the q points of the systems are shown with solid black lines.

on the ΓK line, the point symmetry group is C_{2v} with $C_{2v} \subset D_{6h}$, whose character table of C_{2v} is provided in Table 3.2. The only rotation axis of C_{2v} is drawn in the right panel of Fig. 3.8, showing the crystal lattice as repetitions of the non–diagonal supercell used in our calculations. This axis runs along the zigzag direction, and it is identified by checking the rotational symmetry of the phonon modes, as shown in Fig. 3.9(a).[5] From this we see that the C_2 rotation in C_{2v} coincides with the C_2' rotation in D_{6h}, and we use this to make a connection between the elements of C_{2v} and D_{6h}, shown in Table 3.3. The dipole operator transforms as the $[x, y, z]$ vector and belongs to representations $E_{1u}[x, y] + A_{2u}[z]$ for D_{6h} and $A_1[x] + B_1[y] + B_2[z]$ for C_{2v}. The in–plane component of the dipole transforms accordingly as E_{1u} and $A_1 + B_1$, respectively.

Let us first analyse the two excitons (one dark, one bright) of our system at Γ (Fig. 3.6(c)). The incoming optical light (E_{1u}) interacts with the ground state $|G\rangle$ of the system (which is fully symmetric, A_{1g}) creating an excited state of symmetry $A_{1g} \otimes E_{1u} = E_{1u}$. Therefore, the bright exciton corresponds to the E_{1u} representation (odd under inversion). The dark state is its Davydov partner, thus it must have opposite parity with respect to inversion symmetry (see Sec. 2.3.2 and Ref. [40]) and it corresponds to the E_{2g} representation.

Any irreducible representation of D_{6h} will be a reducible representation of C_{2v}, and therefore can be expressed in terms of a linear combination of irreducible representations of C_{2v}. These are the so–called compatibility relations that we will need to analyse indirect processes from Γ to $\bar{q}$ (in particular, we want to describe the splitting of the E_{1u} and E_{2g} excitons). This applies also to the characters of the representations: if we define $\chi_{\mathcal{G}}(C_k)$ as the character of a (reducible) representation of group

[5] We used the following online tool to visualise the phonon dispersion and lattice displacements corresponding to different q–vectors: http://henriquemiranda.github.io/phononwebsite/phonon.html

D_{6h}	E	$C_2'(x)$	$\sigma_h(xy)$	$\sigma_d(xz)$
A_{1g}	$+1$	$+1$	$+1$	$+1$
A_{2u}	$+1$	-1	-1	$+1$
E_{1g}	$+2$	0	-2	0
E_{2g}	$+2$	0	$+2$	0
E_{1u}	$+2$	0	$+2$	0
E_{2u}	$+2$	0	-2	0

TABLE 3.1: Partial character table for point group D_{6h}.

C_{2v}	E	$C_2(x)$	$\sigma_v(xy)$	$\sigma_v(xz)$
A_1	$+1$	$+1$	$+1$	$+1$
A_2	$+1$	$+1$	-1	-1
B_1	$+1$	-1	$+1$	-1
B_2	$+1$	-1	-1	$+1$

TABLE 3.2: Character table for point group C_{2v}.

$\mathcal{G}$ with respect to symmetry operation C_k, the "wonderful orthogonality theorem" (according to Ref. [130]) for characters establishes the coefficients a_{Γ_i} of the linear combination associated with irreducible representation Γ_i. In our case, the formulas reduce to

$$\begin{aligned} \chi_{D_{6h}}(C_k) &= \sum_{\Gamma_i} a_{\Gamma_i} \chi^{(\Gamma_i)}_{C_{2v}}(C_k) \\ a_{\Gamma_i} &= \frac{1}{4} \sum_k \chi^{(\Gamma_i)}_{C_{2v}}(C_k) \chi_{D_{6h}}(C_k), \end{aligned} \tag{3.32}$$

and we can compute the a_{Γ_i} coefficients using the Tables 3.1, 3.2 and 3.3. We find $A_{2u} \to B_2$ and both E_{1u} and E_{2g} splitting as $A_1 + B_1$, confirming our previous identification of the dipole representations. Therefore, the two indirect excitons labeled as $i1$ and $i2$ with momentum $\overline{q}$ can only have either A_1 or B_1 symmetry.

In fact, as we show below, the same symmetry considerations can be drawn for the phonon dispersion, displayed in Fig. 3.6(b): at Γ, the in–plane phonon modes form Davydov pairs and transform according to the E_{1u} and E_{2g} representation, respectively. Along ΓK these representations reduce to A_1 and B_1. We list in Table 3.4 the results of our DFPT calculations for the symmetries of the phonon modes at $\overline{q}$. These are the phonon modes we are interested in, because they provide the necessary momentum difference to assist the recombination of excitons $i1$ and $i2$ (we will not be concerned with multi–phonon processes). The formation of quasi–degenerate parallel phonon branches can be understood by zooming in on the region close to the

D_{6h}		C_{2v}
C_2'	$\to$	C_2
$\sigma_h(xy)$	$\to$	$\sigma_v(xy)$
$\sigma_d(xz)$	$\to$	$\sigma_v(xz)$

TABLE 3.3: Connection between the elements of C_{2v} and D_{6h}.

Mode	Symm.	Freq. (meV)
LO_3	A_1	183.00
LO_2	B_1	177.63
TO_3	A_1	159.58
TO_2	B_1	159.48
LA	A_1	93.33
LO_1	B_1	93.22
ZO_3	A_2	92.47
ZO_2	B_2	87.53
TO_1	A_1	65.10
TA	B_1	64.72
ZO_1	A_2	22.25
ZA	B_2	21.54

TABLE 3.4: Symmetry of the phonon modes at $\bar{q}$. The modes are listed as Davydov pairs in order of increasing frequency with the lowest–frequency mode at the bottom (compare with Fig. 3.6(b)).

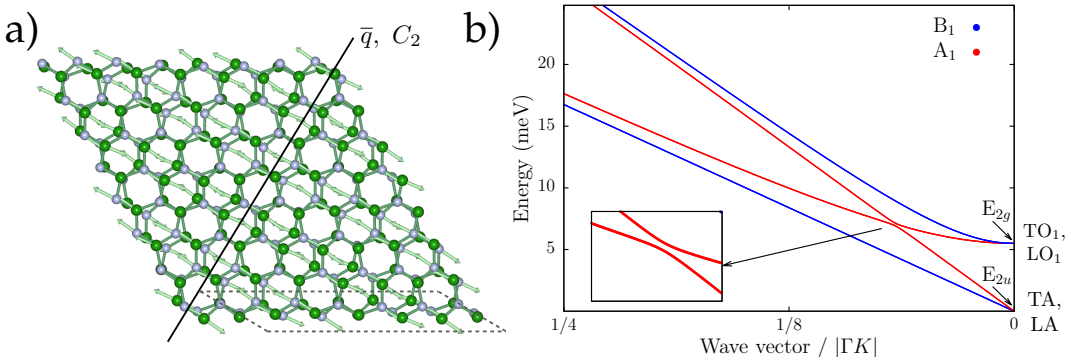


FIGURE 3.9: Phonon information for bulk hBN. (a) Top view of the atomic displacements and forces for the TO_1 phonon mode at one of the six equivalent $\bar{q}$–vectors/C_2 rotation axes (solid black line). The corresponding non–diagonal supercell is represented in dashed gray lines. (b) Interplay between Davydov and symmetry splitting at very low wave vector for the transverse and longidutinal phonon modes (TA, TO_1, LA, LO_1). The dispersion of modes with symmetry A_1 (TO_1, LA) and B_1 (TA, LO_1) is shown in red and blue, respectively.

Γ point as in Fig. 3.9(b). At Γ we see the two degenerate TA and LA modes (E_{2u} symmetry) at zero frequency, as well as their Davydov partner (TO_1 and LO_1 modes, E_{2g} symmetry) 6 meV above. The large value of the splitting is due to the constructive interference of the Fourier components of the inter–layer interaction at zero wave vector. When $q \neq 0$, the degenerate modes further split into two non–degenerate ones of symmetry A_1 and B_1. The two A_1 modes mix via an avoided crossing and then approach their respective B_1 mode. In this way two distinct Davydov pairs are formed, each one with a tiny energy splitting. The low value of the splitting at finite q is due to the destructive interference of the Fourier components of the inter–layer interaction. The relationship between inter–layer interaction and the magnitude of the Davydov splitting along the BZ is shown in App. F with a simple tight–binding model for two interacting linear chains.

3.4.4 Selection rules

In order to analyse the indirect process, we first consider the time–dependent perturbation theory for a model, unspecified "excitonic" Hamiltonian H_0 with exciton–radiation and exciton–lattice interactions as the perturbations: the coupling operators will be d and $g_{\bar{q}}$, respectively (a sum over all phonon modes is implied). We want to qualitatively visualise the phonon–assisted processes leading to the formation or annihilation of a finite–q excitonic state $|\psi_f\rangle$. We only consider phonon emission contributions, and by analogy with Sec. 3.3 we obtain the second–order Fermi golden rule expression for the transition probability per unit time as[6]

$$P_f^{II} = \frac{2\pi}{\hbar}\left|\sum_\alpha \frac{\langle\psi_f|\, g_{\bar{q}}^\dagger\,|\psi_\alpha\rangle\,\langle\psi_\alpha|\, d\,|G\rangle}{E_\alpha - \hbar\omega} + \sum_{\alpha'} \frac{\langle\psi_f|\, d\,|\psi_{\alpha'}\rangle\,\langle\psi_{\alpha'}|\, g_{\bar{q}}^\dagger\,|G\rangle}{E_{\alpha'} - \hbar\Omega_{\bar{q}}}\right|^2 \delta(E_f - \hbar\omega + \hbar\Omega_{\bar{q}}) + \\ \frac{2\pi}{\hbar}\left|\sum_\alpha \frac{\langle\psi_f|\, g_{\bar{q}}^\dagger\,|\psi_\alpha\rangle\,\langle\psi_\alpha|\, d^\dagger\,|G\rangle}{E_\alpha + \hbar\omega} + \sum_{\alpha'} \frac{\langle\psi_f|\, d^\dagger\,|\psi_{\alpha'}\rangle\,\langle\psi_{\alpha'}|\, g_{\bar{q}}^\dagger\,|G\rangle}{E_{\alpha'} - \hbar\Omega_{\bar{q}}}\right|^2 \delta(E_f + \hbar\omega + \hbar\Omega_{\bar{q}}). \tag{3.33}$$

Here $|\psi_\alpha\rangle$ is an intermediate excitonic state with energy E_α and $|G\rangle$ is the ground state of the system. The first term corresponds to the process of photon absorption with phonon emission creating the final excitonic state $|\psi_f\rangle$, while the second term describes the combined photon and phonon emission. The first term in the squared sum represents the creation of a direct virtual exciton $|\psi_\alpha\rangle$ by light as the intermediate step, followed by a scattering to the finite–q state $|\psi_f\rangle$ via phonon emission. The second term adds the contribution of the inverse process, when the virtual state $|\psi_{\alpha'}\rangle$ is created at finite–q by a phonon, and then arrives at the energy of $|\psi_f\rangle$ by absorbing a photon.

For the purpose of finding the selection rules, the two contributions are equivalent and thus we focus on the first one, $\langle\psi_f|\, g_{\bar{q}}^\dagger\,|\psi_\alpha\rangle\,\langle\psi_\alpha|\, d\,|G\rangle$. The final allowed excitons in the energy window that we consider must have A_1 and B_1 symmetry and the first matrix element in the process, $\langle\psi_\alpha|\, d\,|G\rangle$, imposes E_{1u} as the only possible representation for the direct intermediate state $|\psi_\alpha\rangle$. Since $E_{1u} \to A_1 + B_1$ and $g_{\bar{q}}^\dagger$ transforms with the symmetry of the various phonon modes involved, for $g_{\bar{q}}^\dagger\,|\psi_\alpha\rangle$ we have the tensor product $(A_1 + B_1 + A_2 + B_2) \otimes (A_1 + B_1)$. However, $A_1 \otimes (A_2 + B_2) = A_2 + B_2$ and $B_1 \otimes (A_2 + B_2) = B_2 + A_2$, therefore the phonon modes transforming as A_2 or B_2 cannot give the allowed final states and their coupling is forbidden. We see from Table 3.4 that these representations correspond to the out–of–plane Z phonon modes, while the in–plane ones T and L all transform as A_1 or B_1 and therefore are all allowed. If we consider instead incoming light polarised out–of–plane ($A_{2u} \to B_2$), the picture changes and now $(A_1 + B_1 + A_2 + B_2) \otimes B_2 = B_2 + A_2 + B_1 + A_1$, meaning that if the polarization is exclusively out–of–plane only the Z phonon modes can couple to excitons $i1$ and $i2$. This seems to be confirmed by recent experimental results.[131]

3.4.5 Calculation of phonon–assisted optical spectra

Having understood the symmetry constraints for the coupling between finite–q excitons and phonons, we now derive a general expression to calculate these couplings and, thus, phonon–assisted optical spectra. We address this task by (i) using a static

[6] Note that this formula is not meant to give a rigorous, quantitative description of indirect absorption but that it serves only for the symmetry analysis of the involved phonons and excitons.

approximation for the exciton formation probability and (ii) restricting the coupling to harmonic phonons with momentum $\overline{q}$. In our calculations, we focus on the energy region close to the lowest–lying indirect excitons $i1$ and $i2$.

We start by obtaining the complex dielectric function $\varepsilon^{(0)}(\omega)$, given by Eq. (1.75), by a standard BSE calculation in the non–diagonal supercell, with the atoms clamped at their equilibrium positions. We then proceed by considering the Taylor expansion of $\varepsilon(\omega)$ up to second–order in the atomic displacements.[105] This gives a static correction to the equilibrium response $\varepsilon(\omega) \simeq \varepsilon^{(0)}(\omega) + \varepsilon^{\mathrm{st.},(2)}_{\overline{q}}(\omega)$ as detailed below. Let us consider the part of the response function due to excitonic state λ with $T^{\lambda} = \sum_{\mathcal{K}} \overline{A}^{\mathcal{K}}_{\lambda}$:

$$\chi^{\lambda}_{R}(\omega) = \frac{|T^{\lambda}_{R}|^2}{E^{\lambda}_{R} - \hbar\omega + \mathrm{i}\eta}, \tag{3.34}$$

where the subscript R denotes a parametric dependence on lattice displacements, i.e. χ^{λ}_{0} is the frozen–atom response function (we take for simplicity η to be independent of R, an assumption that does not affect the validity of the results). We want to take the Taylor expansion of Eq. (3.34) up to second order in the lattice displacements, therefore as an initial step we need to compute its first derivative and evaluate it at the equilibrium atomic positions:

$$\begin{aligned}\frac{\partial \chi^{\lambda}_{R}(\omega)}{\partial R}\Big|_{R=0} =& \frac{\partial |T^{\lambda}_{R}|}{\partial R}\Big|_{R=0} 2|T^{\lambda}_{R=0}|[E^{\lambda}_{R=0} - \hbar\omega + \mathrm{i}\eta]^{-1} \\ &+ \frac{\partial [E^{\lambda}_{R} - \hbar\omega + \mathrm{i}\eta]^{-1}}{\partial R}\Big|_{R=0} |T^{\lambda}_{R=0}|^2.\end{aligned} \tag{3.35}$$

At this point we note that the oscillator strengths of any finite–q excitons for optical absorption must be zero because of momentum conservation. In the case of a supercell, it means that the excitons being folded onto Γ from a different point in the unit cell are dark if the atoms are clamped at their equilibrium positions. If we label excitons belonging to such subset with λ', this means that $T^{\lambda'}_{R=0} = 0$ and therefore $\partial \chi^{\lambda'}_{R}(\omega)/\partial R\Big|_{R=0} = 0$. This argument is analogous to the one often used in the case of optical absorption in the independent–particle model for the vanishing of the dipole optical matrix elements below the direct band gap.[105, 114]

The same argument applied to the second derivative of χ^{λ}_{R} leads to the vanishing of the terms containing derivatives of the exciton binding energy. The only term that remains is the one containing the second derivative of $T^{\lambda'}_{R}$:

$$\frac{\partial^2 \chi^{\lambda'}_{R}(\omega)}{\partial R^2}\Big|_{R=0} = \frac{\partial^2 |T^{\lambda'}_{R}|^2}{\partial R^2}\Big|_{R=0} [E^{\lambda'}_{R=0} - \hbar\omega + \mathrm{i}\eta]^{-1}, \tag{3.36}$$

This derivative is evaluated numerically using the finite–difference expression

$$\frac{\partial^2 \chi_{R}}{\partial R^2}(\omega) \approx \frac{[\chi(\Delta\mathbf{R};\omega) - 2\chi_0(\omega) + \chi(-\Delta\mathbf{R};\omega)]}{\Delta\mathbf{R}^2}. \tag{3.37}$$

The results, displayed in Fig. D.2 for each phonon mode, confirm the equivalence of the two sides of Eq. (3.36). This means that, numerically, we can obtain the exciton–phonon coupling for the calculation of phonon–assisted absorption/emission both through a finite–difference calculation of the whole response function or through a finite–difference calculation of just the excitonic oscillator strength. In our case, we have $R \to R_{\mu\overline{q}}$, where $R_{\mu\overline{q}}$ refers to a set of atomic displacements according to phonon mode μ with momentum $\overline{q}$ (indicated with a teal vertical line in the phonon

dispersion plot of Fig. 3.6(b)). The normalised phonon displacements (rescaled by the square root of the atomic masses) were used to compute the finite–difference derivatives in order to give to the B and N atom a mass–dependent displacement. A global scaling factor was multiplied by the displacements and converged to 0.0025 Å, which is just above the threshold of numerical noise. The harmonic behaviour of $\varepsilon(\omega)$ with respect to the scaling factor was numerically verified, bringing explicit numerical confirmation of the above equations. We developed a Python tool that handles automatically the generation of the non–diagonal supercell as well as the displacements of its ions starting from phonon eigendisplacements. This is briefly described in App. B.

Now, the second–order correction to the full dielectric function, which adds the contribution of transitions assisted by a single phonon of momentum $\bar{q}$, reads

$$\varepsilon^{\mathrm{st.},(2)}_{\bar{q}}(\omega) = \frac{1}{2}\sum_{\mu}\left[\sum_{i}^{N_{\bar{q}}}\frac{1}{2}\sum_{j}^{2}\frac{\partial^2\varepsilon^{(0)}_{j}(\omega)}{\partial R^2_{\mu\bar{q}_i}}\Big|_{\mathrm{eq}}\right]\sigma^2_{\mu\bar{q}}. \tag{3.38}$$

In this expression j represents the polarisation direction of the incoming light, over which we average, and i labels each of the $N_{\bar{q}} = 6$ equivalent $\bar{q}$-vectors in the BZ, over which we sum. The factor $\sigma^2_{\mu\bar{q}}$ is the thermal average of the squared displacement of a quantum harmonic oscillator, given by Eq. (3.21). As we have seen, the imaginary part of Eq. (3.38) is given by:

$$\mathrm{Im}\frac{\partial^2\varepsilon^{(0)}(\omega)}{\partial R^2_{\mu\bar{q}}}\Big|_{\mathrm{eq}} = \frac{8\pi}{N_k V}\sum_{\lambda'}\frac{\partial^2|T^{\lambda'}|^2}{\partial R^2_{\mu\bar{q}}}\Big|_{\mathrm{eq}}\mathrm{Im}\left\{\frac{1}{\hbar\omega - E^{\lambda'} + \mathrm{i}\eta}\right\}. \tag{3.39}$$

This result allows us to reintroduce the phonon frequency dependence of $\varepsilon^{(2)}_{\bar{q}}(\omega)$ by imposing the correct energy conservation from perturbation theory and distinguishing between phonon emission ($\propto n_{\mu\bar{q}} + 1$) and phonon absorption ($\propto n_{\mu\bar{q}}$): $[2n_{\mu\bar{q}} + 1]/(\hbar\omega - E^{\lambda'} + \mathrm{i}\eta) \to [n_{\mu\bar{q}} + 1]/(\hbar\omega - E^{\lambda'} - \hbar\Omega_{\mu\bar{q}} + \mathrm{i}\eta) + n_{\mu\bar{q}}/(\hbar\omega - E^{\lambda'} + \hbar\Omega_{\mu\bar{q}} + \mathrm{i}\eta)$. Renaming the numerator between square brackets in Eq. (3.38) (including the $1/2$ factors) as $|t^{\mathrm{static}}_{\mu\bar{q}\lambda'}|^2$, since it represents the static formation probability of exciton λ' mediated by a phonon mode μ with momentum $\bar{q}$ and frequency $\Omega_{\mu\bar{q}}$, we obtain the final expression:

$$\varepsilon^{(2)}_{\bar{q}2}(\omega) = \sum_{\lambda\lambda'}|t^{\mathrm{static}}_{\lambda\bar{q}\lambda'}|^2 l^2_{\mu\bar{q}}[n_{\mu\bar{q}} + 1/2 \mp 1/2]\delta(\hbar\omega - E^{\lambda'} \pm \hbar\Omega_{\mu\bar{q}}). \tag{3.40}$$

Here, the upper (lower) sign refers to the process of phonon absorption (emission).

Applying Eq. (3.40) to the description of the process of exciton formation via photon absorption together with phonon emission at $T = 0$ K, and computing the derivatives of $\varepsilon(\omega)$ with finite–difference DFT–BSE calculations in the supercell, we obtain the spectrum in Fig. 3.10(a). It is possible to identify a multi–peak structure associated to the coupling of both the $i1$ and $i2$ excitons to all the in–plane phonon modes, with the higher–energy state $i2$ accounting for most of the oscillator strength. Since the phonon frequencies are close to the energy difference between direct and indirect excitons, the phonon–assisted peaks are distributed around the brightest direct exciton peak with most of the oscillator strength remaining in this narrow energy range. This result suggests that phonon–assisted absorption in hBN is at the origin of the fine structure observed around the brightest exciton peak in absorption experiments.[14, 84] Note, however, that the experimental indirect contribution to

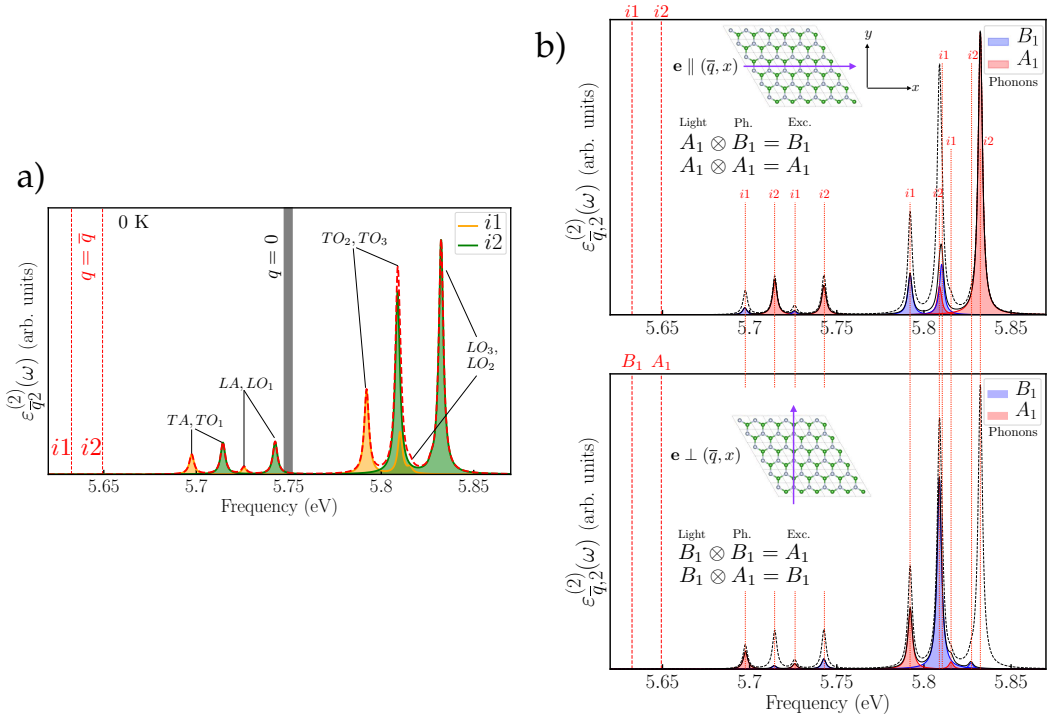


FIGURE 3.10: Phonon–assisted absorption in bulk hBN. (a) Imaginary part of the dielectric function with contributions of transitions at $q = \bar{q}$ mediated by a single phonon (dashed red line, $\varepsilon^{(2)}_{\bar{q}2}(\omega)$ in the text). The orange (green) peaks originate from the phonon couplings to $i1$ ($i2$) and the phonon modes responsible for them are labeled. The peak broadenings are set to 1.5 meV for the temperature of 0 K. The dashed red vertical lines indicate the positions of the two B_1, A_1 excitons at $\bar{q}$ (labeled $i1$ and $i2$). The thick gray vertical line is at the position of the main optically active $q = 0$ exciton. (b) Same as (a), but resolved by direction of light polarization and symmetry of contributing phonon modes. The contribution to absorption is plotted (solid black line) for light polarization $\mathbf{e}$ along the x–direction (top frame) and the y–direction (bottom frame). The dashed black line represents the full result obtained by averaging over the two contributions. The red vertical lines serve as a guide for the eyes to understand which phonon–assisted peaks come from the coupling to excitons $i1$ or $i2$. The color of each phonon–assisted peak indicates the symmetry of the phonon mode responsible for it (B_1: blue, A_1: red).

the optical absorption spectrum will not appear as a series of discrete peaks (as in our approximation), but rather be a continuous function resulting from the integration over all q vectors in the BZ (since they can be connected to the $q = 0$ states by phonons of different momentum). Furthermore, we verify that our first–principles calculations respect the selection rules we have previously determined, as shown in Fig. 3.10(b). This means that the coupling of $i1$ and $i2$ with specific phonon modes depends on the light polarisation direction. In particular, *if light is polarized exclusively along the x/zigzag direction*, i.e., it transforms as A_1 (top frame), then only the TA, LO_1, TO_2 and LO_2 phonon modes (all transforming as B_1, portrayed in blue) can couple to $i1$ forming phonon–assisted peaks, and only the TO_1, LA, TO_3 and LO_3 modes (all transforming as A_1, portrayed in red) can couple to $i2$. Conversely, *if light is polarized exclusively along the y/armchair direction*, i.e., it transforms as B_1 (bottom frame), then only the A_1 phonon modes couple with $i1$, and only the B_1 modes couple with $i2$. This unmistakably shows that exciton $i1$ (the lowest–energy one and therefore the most responsible for the luminescence spectrum) has B_1 symmetry,

while exciton $i2$ transforms as A_1. Additionally, we note that the leading peak for absorption is due solely to the strong coupling between the $i2$ exciton and the LO_3 phonon mode: therefore, this peak completely disappears when light is polarised along the y direction (that is, orthogonal to the $\bar{q}$ vector). Finally, we confirm that the total spectrum is $\pi/3$–periodic by varying the light polarization direction.

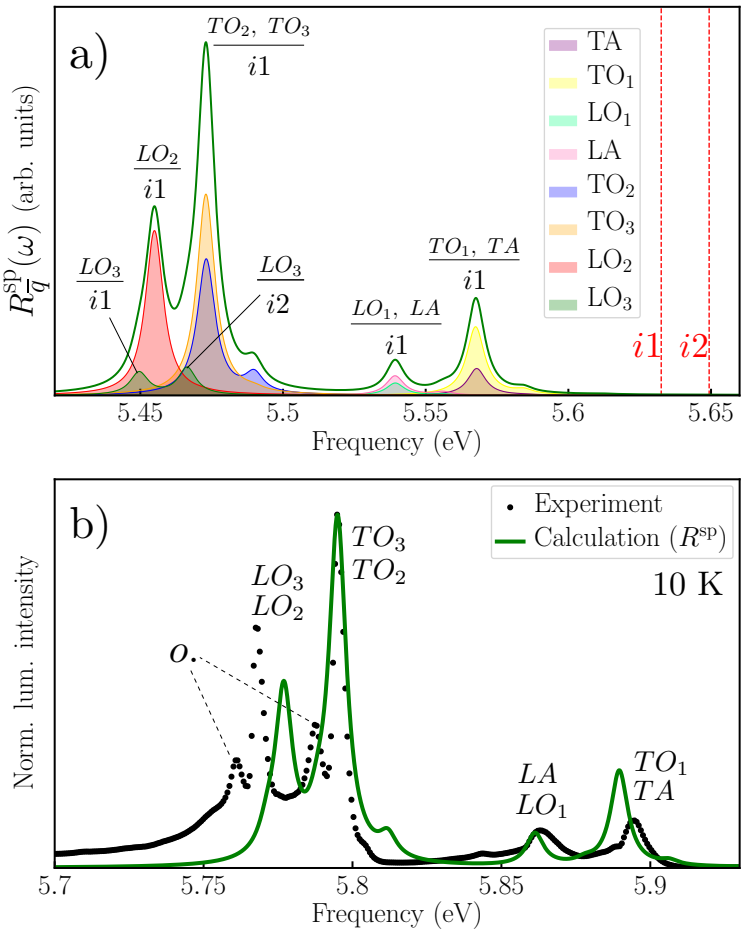


FIGURE 3.11: Phonon–assisted emission in bulk hexagonal boron nitride. (a) Spectral function $R^{sp}_{\bar{q}} = \sum_\lambda R^{sp}_{\lambda\bar{q}}$ (green solid line) at 10 K. The μ–components of the spectrum, belonging to the different phonon modes, are also plotted in various colours. The exciton–phonon couplings are labeled. (b) Comparison with the normalized experimental spectrum (black dots) at 10 K from Ref. [84]. The multi–phonon overtones are denoted as "o.". Note that here our spectrum is blueshifted by 0.322 eV to match the experimental one. The temperature–dependent peak widths are described according to a linear model[109] (parameters taken from the experimental fit, see App. G). The experimental excitonic temperature $T_{exc} = 55$ K (see text) is used.

We now turn to the luminescence spectrum by means of the RS relation, Eq. (3.31), applied on top of our finite–difference dielectric function Eq. (3.40). We know that the energy of an emitted photon differs from the one of an absorbed photon (indirect absorption with phonon emission) by twice the frequency of the phonon involved. We thus define $\varepsilon^{em}_{\lambda\bar{q}}(\omega) \equiv \varepsilon^{(2)}_{\lambda\bar{q}}(\omega - 2\Omega_{\lambda\bar{q}})$. Because of this, each phonon–assisted peak is mirrored with respect to the energy of the excitonic state involved.

The expression for the spontaneous emission rate is:

$$R^{\rm sp}_{\bar{q}}(\omega) = \sum_{\mu} \frac{\omega(\omega + 2\Omega_{\mu\bar{q}})^2}{\pi^2\hbar c^3} n_1(\omega) {\rm Im}\{\varepsilon^{\rm em}_{\mu\bar{q}}(\omega)\} n_B. \tag{3.41}$$

It is sufficient to compute the refractive index from $\varepsilon^{(0)}(\omega)$, which completely determines its slow decay at low frequencies. The exciton occupation function is $n_B(\Delta E_{i1,\lambda}) = {\rm e}^{-(E^\lambda - E^{i1})/k_B T_{\rm exc}}$ (E^{i1} being the energy minimum of the exciton dispersion curve and E^λ the energy of any exciton λ). Because of the large energy difference between E^{i1} and the main direct peak, the latter will always be suppressed by the Boltzmann factor n_B up to room temperature, and therefore it should not be seen in a luminescence experiment. This also means that only the phonon sidebands related to $i1$ and $i2$ will now dominate the spectrum. The results for $R^{\rm sp}_{\bar{q}}(\omega)$ are plotted in Fig. 3.11(a). The energy differences between excitonic levels ($\Delta E_{i1,i2} = 17$ meV) and the phonon frequencies (from 50 to 200 meV) are large, giving rise to a well–spaced peak structure that can be easily resolved. We notice two separated groups of features which are clearly seen in experiment (Figs. 3.6(a) and 3.11(b), black dots): one at higher energy generated by low–frequency phonon emission, and the other at lower energy due to high–frequency optical phonons. The emission spectrum is almost completely dominated by the lowest–bound exciton $i1$ since the occupation factor quenches most of the peaks related to $i2$. However, experiments have shown (Figure 1(c) of Ref. [19]) that in bulk hBN the excitonic temperature T_{exc} that goes into the Boltzmann factor n_B is greater than the lattice temperature T_L. If we set $T_{exc} = T_L = 10$ K in $R^{\rm sp}_{\bar{q}}$, we obtain only the peaks coming from $i1$, while setting T_{exc} to the experimental value of 55 K (as in Fig. 3.11) leads to the appearance of the quenched peaks from $i2$. On the experimental side, the peaks due to high–frequency modes are clearly visible in Fig. 3.11(b) at 10 K: the separation between these peaks thus corresponds to the separation between the LO and TO modes at point $\bar{q}$ in the theoretical calculations, while in experiment it corresponds to the splitting at point $\bar{q} + \Delta q$, where Δq is the error committed by approximating the "true" q–point $|T_1 - M|$ with $\bar{q} = |K - M|$. We attribute to this discrepancy the difference of $\sim$ 8 meV between the theoretical and experimental peak separations. Our error in q–space is approximately 12 % of ΓK, with $\Delta q = 0.01$ Å^{-1}. The discrepancy in the position of the LO peak is consistent with the true minimum of the excitonic dispersion lying on the K side of $\bar{q}$. This is confirmed by both our computational results, Sec. 2.4, and those of Ref. [84]. In our results, the quasi–degenerate transverse modes TO_2 and TO_3 couple with similar strength to $i1$, whereas the longitudinal mode LO_3 has a higher frequency than LO_2 and a much weaker coupling to $i1$ (recall that in the absorption case, the coupling of mode LO_3 with $i2$ is the strongest one). As expected, we cannot capture the various satellite peaks in Fig. 3.11(b), which experimentally are assigned to multi–phonon processes involving zone–center shear phonon modes.[132]

Therefore, the structure of the emission spectrum can be understood in terms of a Davydov pair of finite–q excitons coupling with different strengths to the various in–plane phonon modes. We can further add that the asymmetry observed experimentally in the Stokes shift, Fig. 3.1(b), is due to the fact that the states mainly involved in the absorption and emission processes are not the same. In absorption, the E_{2u} state at Γ remains the most visible,[7] while the emission signal depends on the lowest–lying exciton level at finite–q, $i1/B_1$, lying 0.12 eV below the direct exciton.

[7] Experimental position of the Γ exciton: 6.05 eV; theoretical LDA+sscGW+BSE position: 5.75 eV.

3.4.6 Conclusion

The elucidation of the luminescence spectrum of bulk hBN represents the main result of this thesis. We solve a long–debated issue in the spectroscopy community by uncovering how electron–hole bound pairs and lattice vibrations concur in the response to an electromagnetic field. We show that such a description is essential for the correct interpretation of the optical spectra and can be obtained by modelling the exciton–phonon coupling in a computationally feasible way.

The method developed here for hBN can be applied, in principle, to any material with indirect gap. It is particularly suited for materials with strong excitonic effects as they often occur in layered materials. For example, it could be applied to investigate bilayers of MoS_2 as well as heterobilayers made of different transition metal dichalcogenides, or even to the recently reported case of monolayer WSe_2.[133] The method will work particularly well when the excitonic density of states, multiplied by the Boltzmann factor, is strongly peaked at the energy of the indirect $\bar{q}$ exciton. The full q–integration can then be replaced by just summing over excitons of wave vector $\bar{q}$, since all other contributions to the emission spectrum will be suppressed.

Despite the success of our static method, several interesting generalisations are possible, namely: (i) performing a full q–integration in the BZ, so that a complete indirect absorption spectrum may be reconstructed; (ii) obtaining a microscopic description of the ExcP coupling in terms of the constituent electronic and phononic building blocks, as opposed to computing the derivative of the dielectric function; (iii) overcoming the static approximation (this entails the inclusion of dynamical effects in the BSE). We will answer these questions in the next Chapter.

Chapter 4

Phonon–assisted transitions as satellites of the spectral function

This part of the thesis is the result of collaborations with P. Cudazzo and A. Marini. Results are still preliminary.

Our goal for the last part of this thesis is to find a many–body, dynamical description of the exciton–phonon coupling. In this way, we will be able to perform more refined calculations of the phonon–mediated exciton formation probabilities, to compare directly the strength of the direct and indirect absorption processes and to work in the unit cell, allowing us to obtain the fully q–integrated phonon–assisted absorption spectrum. In other words, we aim to obtain a complete explicit form for the excitonic, macroscopic dielectric function $\varepsilon_2(\omega)$ including phonon contributions, as sketched in Eq. (3.1).

We know from Chap. 1 that, in the case of a single–particle Green's function G corrected with a self–energy via the Dyson equation, its spectral function ImG shows a QP peak at the corrected single–particle energy, followed by a satellite structure. The satellites appear due to the frequency dependence of the self–energy. Here, the key insight is that the structures due to indirect processes in the absorption spectra can also be seen as the *satellites of the exciton spectral function* ImL (or Imχ) in the same way as if we imagine correcting it with a "self–energy" encoding the dynamical interactions between excitons and phonons. Then, the q–integrated satellite structures of Imχ can be identified with the indirect component of the absorption spectrum. Doing this requires considering a frequency–dependent kernel for the BSE, that is usually evaluated only in the static approximation. This is a difficult problem, since a dynamical BSE (see Eq. (1.77)) cannot easily be inverted, and its solution has only rarely been attempted.[134–137] There is yet another complication: as we remarked in Sec. 1.3.3, a GW–type approximation followed by the inversion of the Dyson equation typically yields the wrong shape and position for the satellites, owing to the neglected contributions at higher orders in the Dyson expansion. This problem would be present also for L if we attempted an equivalent treatment.

In order to make these problems clear, we will first analyse the single–particle case with a simple model (that of a localised electron interacting with a phonon) and show how the correct description of the satellites is recovered in this case. By treating the electronic Green's function G in perturbation theory, it can be emphasised how the dynamical part of the interaction (represented by the phonon propagator D) is responsible for the formation of the satellites of the spectral function ImG. The connection with the exact solution of the model is made via the so–called *cumulant ansatz*. In this framework, we will discuss again a simple model, this time a two–level system interacting with a phonon, showing that it leads to the correct description of phonon–assisted absorption (as in Sec. 3.3), and eventually also to the exact solution — even when the eh interaction is included — via the cumulant ansatz.

The generalisation to extended systems including excitonic effects will be explained following the work of Cudazzo and Reining.[38, 113] According to this work, a dynamical ExcP self–energy Π^{exc} can be constructed with the solution of the static BSE providing the "unperturbed" propagator to be dynamically corrected by ep interactions. We will discuss the shape of this self–energy, and how it should correct the absorption spectrum. Finally, we will present the two final contributions of this thesis: the problem of the mixing of longitudinal and transverse excitons, with its consequences for the absorption spectrum, and our implementation of the exciton–phonon self–energy Π^{exc} in `Yambo`. We will present some preliminary *ab initio* results for bulk hBN.

4.1 Indirect absorption revisited

4.1.1 One–level system coupled to one phonon

Let us consider a filled electron state with energy ϵ, coupled to a boson with frequency Ω (this is called Langreth's model). This model was originally introduced by Lundqvist[138] and then Langreth[139] to study the spectral function of a photoexcited core electron. According to Eqs. (1.25) and (1.109), respectively, the GFs for the electron and the boson at $T = 0$ are

$$\begin{aligned} G^0(\omega) &= \frac{1}{\omega - \epsilon - \mathrm{i}\eta}, \\ D(\omega) &= \frac{1}{\omega - \Omega + \mathrm{i}\eta} - \frac{1}{\omega - \Omega - \mathrm{i}\eta}. \end{aligned} \tag{4.1}$$

Langreth has provided the exact solution to this model in terms of the interacting electron GF. This is given in the time domain by[139, 140]

$$G_{ex}(t_{12}) = G^0(t_{12})\, \mathrm{e}^{-\frac{g^2}{\Omega^2}}\, \mathrm{e}^{\mathrm{i}\frac{g^2}{\Omega}(t_2 - t_1)}\, \mathrm{e}^{\frac{g^2}{\Omega^2}\mathrm{e}^{-\mathrm{i}\Omega(t_2 - t_1)}}, \tag{4.2}$$

where g is the electron–boson coupling constant (with $g^2 \equiv |g|^2$). In the frequency domain, this expression takes the form

$$\begin{aligned} G_{ex}(\omega) &= \mathrm{e}^{-\frac{g^2}{\Omega^2}} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{g^2}{\Omega^2}\right)^n \frac{1}{\omega - (\epsilon + \frac{g^2}{\Omega}) + n\Omega - \mathrm{i}\eta}, \\ G_{ex}^{(N)}(\omega) &= \sum_{n=0}^{N} \sum_{m=0}^{N-m} \frac{(-\frac{g^2}{\Omega^2})^m \frac{g^2}{\Omega^2}^n}{m!n!} \frac{1}{\omega - (\epsilon + \frac{g^2}{\Omega}) + n\Omega - \mathrm{i}\eta}, \end{aligned} \tag{4.3}$$

with the second line describing the truncation of G_{ex} at order N. We can have a look at the resulting spectral function in Fig. 4.1(b) (red shaded region). We see that the energy of the electronic level is dressed by the interaction with the boson, so that $\epsilon_{QP} = \epsilon + g^2/\Omega$, and several satellites are present beyond the quasiparticle peak at ϵ_{QP}. Their energies are $\epsilon_{QP} + n\Omega$ and they correspond to the excitation of n bosons.

We will now try to get this result back by means of perturbation theory, computing the Fan–Migdal self–energy of the system as in Sec. 1.6.6, and as depicted in Fig. 4.1(a). We have

$$\Sigma(\omega) = \mathrm{i}g^2 \int \frac{\mathrm{d}\omega'}{2\pi} G^0(\omega - \omega') D(\omega') = \frac{g^2}{\omega - \epsilon + \Omega - \mathrm{i}\eta}. \tag{4.4}$$

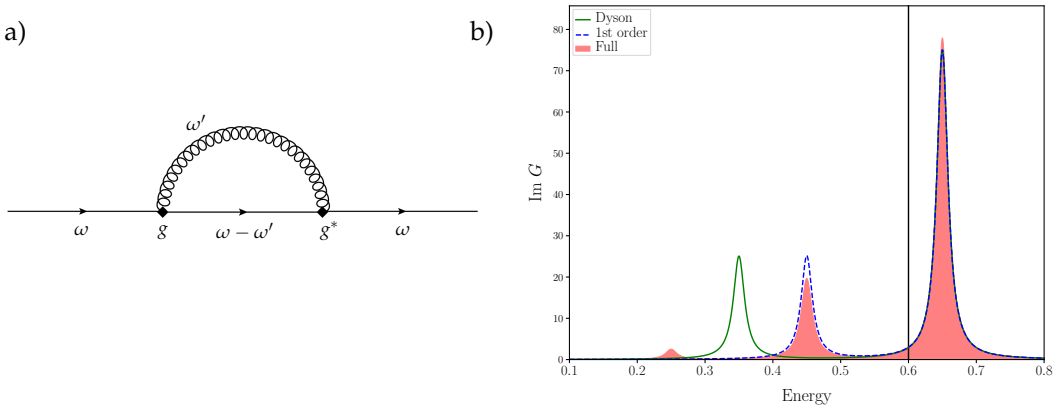


FIGURE 4.1: Langreth's model. (a) First–order self–energy contribution to the Dyson equation of the electron–boson system in the one–level case. (b) Spectral functions for the one–level system. The vertical black line is at the bare electronic energy ϵ. The red shaded area is the exact solution at all orders, Eq. (4.3). The green curve is the result of the Dyson equation (4.5), Im$\widetilde{G}$, which obtains the correct QP peak but misplaces the satellite. The dashed blue curve is the exact solution at first order, Eq. (4.8). The parameters are $\epsilon = 0.6$, $\Omega = 0.2$, $g^2 = 0.01$, $\eta = 0.01$.

From here, we write the Dyson equation $\widetilde{G} = G^0 + G^0\Sigma(\omega)\widetilde{G}$ to include the FM self–energy at all orders. We invert it as $\widetilde{G} = [G^{0,-1} - \Sigma(\omega)]^{-1}$ and we expand the resulting expression for small g^2 since a weak coupling is a requirement for perturbation theory to be valid. We obtain

$$\widetilde{G}(\omega) = \frac{1 - \frac{g^2}{\Omega^2}}{\omega - (\epsilon + \frac{g^2}{\Omega}) - i\eta} + \frac{g^2}{\Omega^2}\frac{1}{\omega - \epsilon + (\Omega + \frac{g^2}{\Omega}) - i\eta}. \tag{4.5}$$

Since the self–energy Σ has a single pole, only one satellite is generated and its energy is $\omega = \epsilon - g^2/\Omega - \Omega$. We then have to compare it with the first–order term ($n = 1$) of the exact solution, $G^{(1)}_{ex}$: this can be seen in Fig. 4.1(b). Although the quasiparticle energy in $\widetilde{G}$ is the same as the exact one ϵ_{QP}, the satellite position is wrong. This might not surprise us since we mentioned already in Sec. 1.3.3 that a GW–type perturbative expansion, though often successful in describing QP corrections, typically gives rise to one satellite in the wrong position. Indeed, the position of the first satellite depends on all the others, which are neglected in a Dyson approach but obviously accounted for in the full solution.

Actually, we can point out that the *full* solution can be recovered from this perturbative result by constructing a series expansion of G in terms of Σ which is different from the Dyson one and is given by (Hedin, Ref. [141])

$$G(t_{12}) = G_{ex}(t_{12}) = G^0(t_{12})e^{C(t_{21})}. \tag{4.6}$$

This is called *cumulant ansatz*, with C being the cumulant (whose expression can be obtained by matching the first–order terms of both the cumulant and Dyson expansions):

$$C(t_{12}) = i\int_{t_1}^{t_2} dt_{1'}\int_{t'_1}^{t_2} dt_{2'}\Sigma(t_{1'2'})e^{i\varepsilon(t'_1 - t'_2)}. \tag{4.7}$$

At first order, we then have

$$G^{(1)}(\omega) = G^{(1)}_{ex}(\omega) = \frac{1 - \frac{g^2}{\Omega^2}}{\omega - \epsilon_{QP} - \mathrm{i}\eta} + \frac{g^2}{\Omega^2}\frac{1}{\omega - \epsilon_{QP} + \Omega - \mathrm{i}\eta}. \tag{4.8}$$

Here, the renormalisation factor for the QP peak (compare with Eq. (1.55)) is given by

$$R(\Omega) = \frac{g^2}{\Omega^2} = -\frac{\partial}{\partial\omega}\Sigma(\omega)\Big|_{\omega=\epsilon_{QP}}. \tag{4.9}$$

Starting from Langreth's model and Hedin's cumulant ansatz, the cumulant expansion has been the object of intense study in the past years.[120, 142–148] It has been rigourously derived and generalised to extended systems with conduction and valence electrons.[149–152] In this case, the cumulant $G(\omega)$ remains an approximation whose validity is at times difficult to estimate (a helpful discussion on this topic is provided in Ref. [153]), but it proved to be very successful in describing satellite structures in various systems, from silicon (where the electrons interact with *plasmons* during photoemission)[149] to transition metal oxides (where phonon–derived sidebands are visible in ARPES measurments).[154] If an extended system is considered, a q–integration must be performed in addition to a sum over band states, which transforms the satellites from simple peaks to DOS–like structures and adds a finite electron–boson lifetime to the electronic states.

Our interest in this model lies in the possibility to apply an analogous approach for the screening function χ, and describe the resulting phonon–derived satellites from the perspective of electronic excitations. For the purposes of this thesis we will stick to the first–order expressions (i.e. we will only study single–phonon processes), although a full "excitonic" cumulant expansion is currently being the object of further studies.[38, 113]

4.1.2 Absorption in a two–level system coupled to one phonon

Let us now consider a two–level system interacting with a phonon and compute its phonon–assisted absorption. We aim to compute the relevant contributions up to first order in the interaction: such contributions are schematically represented on the first line of Fig. 4.2. From this point on, we explicitly take the TDA in order to ensure that the hole and electron propagators have definite time directions; this will result in a retarded response function for absorption, consistent with the way we normally compute it both in the IP and sBSE cases. In the two–level system, the TDA amounts to setting $g_{cv} = 0$ (this is reasonable in materials with a band gap much larger than the characteristic phonon frequency, like hBN). In this case, the first–order corrections to the GFs corresponding to the "valence" v and "conduction" c levels of a two–level system are given by (compare with Eq. (4.8)):

$$\begin{aligned} G^{(1)}_v(\omega) &= \frac{1 - \frac{g^2_{vv}}{\Omega^2}}{\omega - \epsilon^{QP}_v - \mathrm{i}\eta} + \frac{g^2_{vv}}{\Omega^2}\frac{1}{\omega - \epsilon^{QP}_v + \Omega - \mathrm{i}\eta}, \\ G^{(1)}_c(\omega) &= \frac{1 - \frac{g^2_{cc}}{\Omega^2}}{\omega - \epsilon^{QP}_c + \mathrm{i}\eta} + \frac{g^2_{cc}}{\Omega^2}\frac{1}{\omega - \epsilon^{QP}_c - \Omega + \mathrm{i}\eta}. \end{aligned} \tag{4.10}$$

Then, in order to compute the terms (a), (c1) and (c2) of Fig. 4.2[1] for the two–level system, we recognise that they are given by the product (now expressed in the time domain) of the "cumulant" GFs, where the product is kept at first order in g^2:

$$\chi_{cv}^{[c1]+[c2]}(t_{21}) = -\mathrm{i}d_{cv}^2 G_v(t_1 - t_2)G_c(t_2 - t_1)\Big|_{1^{st}\text{ order in } g^2} \tag{4.11}$$

(d_{cv} is the dipole matrix element). We Fourier–transform the expressions Eqs. (4.10) according to Eqs. (1.25) and obtain

$$\begin{aligned}\chi_{cv}^{[c1]+[c2]}(t_{21}) = -\mathrm{i}\theta(t_2 - t_1)\Bigg[d_{cv}^2\left(1 - \frac{g_{vv}^2 + g_{cc}^2}{\Omega^2}\right)\mathrm{e}^{-\mathrm{i}\Delta_{cv}^{QP}(t_2-t_1)} \\ + d_{cv}^2\frac{g_{cc}^2 + g_{vv}^2}{\Omega^2}\mathrm{e}^{-\mathrm{i}[\Delta_{cv}^{QP}+\Omega](t_2-t_1)}\Bigg],\end{aligned} \tag{4.12}$$

with the first term in the sum representing the absorption edge and the second one the satellite due to the emission of the phonon. Here, $\Delta_{cv}^{QP} = \epsilon_c^{QP} - \epsilon_v^{QP}$. We still need to calculate explicitly the "interference" contribution (c3). In order to do that, we first redefine the Green's function so that it already contains the quasiparticle energy instead of the bare one (compare with Eq. (4.2)): $G^{QP}(t_{12}) = G^0(t_{12})\exp\{\mathrm{i}\frac{g^2}{\Omega}(t_2 - t_1)\}$. Now the expression amounts to:[2]

$$\begin{aligned}\chi_{cv}^{[c3]}(t_{21}) &= d_{cv}^2(g_{vv}^* g_{cc} + g_{vv}g_{cc}^*)\int \mathrm{d}t_1' t_2' G_v^{QP}(t_{11'})G_c^{QP}(t_{2'1})\Delta D(t_{2'1'})G_v^{QP}(t_{1'2})G_c^{QP}(t_{22'}) \\ &= -\mathrm{i}\theta(t_2 - t_1)d_{cv}^2\left[\frac{g_{vv}^* g_{cc} + g_{vv}g_{cc}^*}{\Omega^2}\left(\mathrm{e}^{-\mathrm{i}\Delta_{cv}^{QP}(t_2-t_1)} - \mathrm{e}^{-\mathrm{i}[\Delta_{cv}^{QP}+\Omega](t_2-t_1)}\right)\right].\end{aligned} \tag{4.13}$$

Here we have a double contribution due to the possible ordering of the internal times: the phonon can be emitted first either by the hole ($t_1 < t_2' < t_1' < t_2$) or by the electron ($t_1 < t_1' < t_2' < t_2$). Finally, in order to get the absorption spectrum (where now indirect absorption mediated by one phonon is included), we just need to sum Eqs. (4.12) and (4.13) and take the imaginary part (now switching to the frequency domain). The result then takes on the following simple form:

$$\varepsilon_2(\omega) \propto \left[1 - \frac{\mathscr{G}^2}{\Omega^2}\right]d_{cv}^2\delta(\omega - \Delta_{cv}^{QP}) + \frac{\mathscr{G}^2}{\Omega^2}d_{cv}^2\delta(\omega - [\Delta_{cv}^{QP} + \Omega]). \tag{4.14}$$

Here, we have incorporated the ep matrix elements into an "excitation–phonon" coupling given by

$$\mathscr{G}^2 = g_{vv}^2 + g_{cc}^2 - g_{vv}^* g_{cc} - g_{vv}g_{cc}^* = |g_{vv} - g_{cc}|^2. \tag{4.15}$$

Note also that *no dynamical effects are possible* if $g_{cc} = g_{vv}$.

How does this result compare with the textbook time–dependent perturbation theory expression obtained in Sec. 3.3.3 (i.e. Eq. (3.18))? Since we are in the phonon emission case, the energy conservation given by the delta function is the same as

[1] The labels referring to the various contributions are chosen consistently with the general picture presented in Fig. H.1.

[2] Note that in this expression we don't use the full phonon propagator D, but ΔD, which corresponds to D without its static part. ΔD is given by Eq. (I.24). This trick is not necessary if we plan to take the cumulant expansion at the end of the calculation. It is instead necessary to remove spurious QP contributions if we stop at first order, since they are already fully included in G^{QP}.

in Eq. (4.14); furthermore, we have neglected anti–resonant transitions, so the only term that contributes to the transition amplitude is $|T_B|^2 = |S_2^\omega + S_2^\Omega|^2$. If we specialise the corresponding expressions in Eq. (3.14) to the two–level system and we exploit the delta function to replace ω with $\Delta_{cv}^{QP} + \Omega$ in the denominator of S_2^ω, then we have

$$\begin{aligned} S_2^\omega &= \frac{d_{vc}\mathcal{G}_{cc}^*}{\Omega}, \\ S_2^\Omega &= -\frac{\mathcal{G}_{vv}^* d_{vc}}{\Omega}, \\ |T_B|^2 &= \frac{\mathscr{G}^2}{\Omega^2}. \end{aligned} \tag{4.16}$$

Therefore, the indirect part of the absorption is recovered exactly. We see that the (c3) term represents the interference between the two quantum paths of photon absorption, then phonon emission (given by (c1)) and phonon emission, then photon absorption (given by (c2)). Note that a similar "diagrammatic" development of the contributions to indirect absorption was already put forward long ago by Chow in Refs. [155, 156]. In the case of Sec. 3.3.3, the strength of the direct part of the absorption is overestimated, since the first–order perturbation theory Eq. (3.16) does not contain the phononic renormalisation factor equal to $\mathscr{G}^2/\Omega^2$, so it is not possible to take its ratio with the indirect contribution in order to estimate their relative strengths: rather, the two contributions have to be treated on the same footing in order to conserve the spectral weight, as we did in this Section.

"Excitonic" effects can also be trivially included in the two–level system: in fact, we know from App. A that in the TDA the role of the eh interaction is just to reduce the band gap energy to $E_{vc} = \Delta_{cv}^{QP} - K_d = \Delta_{cv}^{QP} - (W_{vc}^{vc} - 2V_{vc}^{vc})$. Therefore, since no change between the excitation and the exciton bases is involved, it is sufficient to replace Δ_{cv}^{QP} by E_{vc} in Eqs. (4.12), (4.13) and (4.15) to obtain the "full" result represented by the second line of Fig. 4.2,

$$\chi_{cv}^{(1)}(t_{21}) = -\mathrm{i}\theta(t_2 - t_1)\left[d_{cv}^2\left(1 - \frac{\mathscr{G}^2}{\Omega^2}\right)\mathrm{e}^{-\mathrm{i}E_{vc}(t_2-t_1)} + d_{cv}^2\frac{\mathscr{G}^2}{\Omega^2}\mathrm{e}^{-\mathrm{i}[E_{vc}+\Omega](t_2-t_1)}\right]. \tag{4.17}$$

We have obtained an expression analytically equivalent to the one for the single–particle GF $G^{(1)}$ in Langreth's model, Eq. (4.8). In fact the analogies run deeper: if we take the cumulant ansatz for $\chi_{cv}^{(1)}(t_{21})$, we actually obtain the *exact solution* for a two–level system interacting with a dynamical boson in the TDA, which is provided in Ref. [148].[3] For reference, in our case this takes the form

$$\begin{aligned} \chi_{cv}^{(1)}(t_{21}) &= \chi_{cv}^{QP}(t_{21})\left[1 + \mathscr{G}^2\left(-1 + \mathrm{e}^{-\mathrm{i}\Omega(t_2-t_1)}\right)\right], \\ \chi(t_{21}) &= \chi_{ex}(t_{21}) = \chi_{cv}^{QP}(t_{21})\mathrm{e}^{\mathscr{G}^2\left(-1+\mathrm{e}^{-\mathrm{i}\Omega(t_2-t_1)}\right)}, \end{aligned} \tag{4.18}$$

where we have defined $\chi_{cv}^{QP}(t_{21}) \equiv -\mathrm{i}\theta(t_2 - t_1)d_{cv}^2\exp[-\mathrm{i}E_{vc}(t_2 - t_1)]$ by analogy with the single–particle case.

This result is a convincing indication that the presented treatment may provide a viable description of phononic satellites in excitonic spectra, and, in the case of realistic systems, a fully *ab initio* way to compute phonon–assisted absorption. In

[3] Section VI.C, equation (34), where we have already explicitly solved the time integrals because our interaction matrix element is just $W_{nnmm}(t - t') = g_{nn}\Delta D(t - t')g_{mm}^*$.

the following we will then consider the realistic case of extended systems with multiple electronic bands and phonon branches, and we will write the corresponding expressions for $\chi^{(1)}$ and $\varepsilon_2(\omega)$.

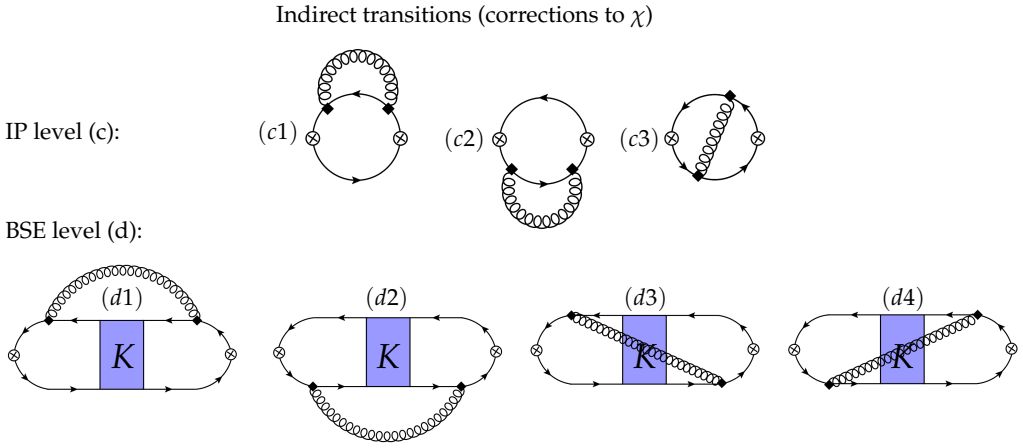


FIGURE 4.2: Contributions to the response function χ due to phonon–assisted transitions (schematic representation). The meaning of the dots and arrows is explained in Chap. 1. The labeling of the diagrams is consistent with the more general Fig. H.1. (c) Phonon–assisted transitions at first order at the IP level. In the case of the two–level system, diagrams (c1) and (c2) are included in Eq. (4.12), while diagram (c3) is given by Eq. (4.13). (d) Phonon–assisted transitions at first order, this time including the static eh interaction kernel K (BSE level). In the case of the two–level system, the corresponding expression is given by Eq. (4.17). In the general case, it is given by Eq. (4.26). Additionally, diagrams (d) are discussed in more detail in App. I (see in particular Fig. I.2).

4.2 Phonon–assisted optical spectra II: the dynamically corrected dielectric function

We summarise below the general theory of the dynamical exciton–phonon interaction based on Refs. [38, 113]. An alternative, more qualitative description of the same theory following a different approach is provided in App. H. Furthermore, Fig. H.1 graphically summarises all the relevant many–body physics included in the treatment.

4.2.1 The exciton–phonon self–energy

According to Ref. [113], the full *dynamical* BSE can be written using the four–times two–particle correlation function as

$$L(1423) = L^{(0)}(1423) + L^{(0)}(12'21')\tilde{\Xi}(1'4'2'3')L(3'44'3). \tag{4.19}$$

Here, $L^{(0)}$ *is the solution of the static BSE* (from now on sBSE), i.e. the two–particle correlation function at zero order in the dynamical part of the interaction, given by Eq. (1.87) with the usual kernel $K = W - 2v$. Then, the kernel $\tilde{\Xi}$, which is in principle a quantity to be computed self–consistently as it depends itself on L, contains the dynamical ep interaction. In principle, the solution of this equation is very complicated, because the dynamical kernel makes it impossible to write it in terms of a two–times L, like in the static sBSE case: therefore Eq. (4.19) is not easily inverted.

However, the kernel $\tilde{\Xi}$ can be considered up to just the first order in the ep interaction, giving the equation

$$L^{(1)}(1423) = L^{(0)}(1423) + L^{(0)}(12'21')\Xi(1'4'2'3')L^{(0)}(3'44'3). \tag{4.20}$$

Now, the kernel Ξ contains the contributions corresponding to Fig. 4.2 (and equivalently to Fig. H.1(d)). Note that these contributions are shown in detail in Fig. I.2, while the explicit expression for Ξ is discussed in App. I. Furthermore, this equation can now be written also for the two–times response function $\chi(13) = -i\hbar L(1313)$ in the excitonic basis, obtaining the one in Fig. H.1(e):

$$\chi^{(1)}_{\alpha\alpha'}(t_{12}) = \chi^{(0)}_{\alpha}(t_{12}) + \int_{t_1}^{t_2} dt_{1'} \int_{t_{1'}}^{t_2} t_{2'}\, \chi^{(0)}_{\alpha}(t_{11'})\Pi^{\text{exc}}_{\alpha\alpha'}(t_{1'2'})\chi^{(0)}_{\alpha'}(t_{2'2}). \tag{4.21}$$

Again, $\chi^{(0)}$ is the usual sBSE response function. Then, $\chi^{(1)}$ represents its first–order dynamical correction. Eq. (4.21) was originally derived in Ref. [38] to include the dynamical part of the screened Coulomb interaction W. However, it can also be applied to the the electron–phonon interaction D while keeping the Coulomb interaction in the static approximation. This can be justified with the argument that retardation effects in the electronic system happen on a much smaller time scale than in the electron–phonon case, since electrons move fast and ions move slowly, and therefore an electronic excitation may still be regarded as instantaneously screened even after one its constituents is scattered by a phonon.

Keeping in mind the analogies with the simpler cases analysed previously, it is possible to derive an *exciton–phonon* self–energy $\Pi^{\text{exc}} = |\mathscr{G}^{\text{exc}}|^2 D\chi^{(0)}$ from the kernel Ξ in the TDA, and then use it to correct the response function and obtain $\chi^{(1)}$ (this is shown in App. I). Π^{exc} has the form of the Fan–Migdal self–energy, Eq. (1.116), where G^0 is replaced by $\chi^{(0)}$ and g by $\mathscr{G}^{\text{exc}}$: this is why we may regard it as an excitonic self–energy. Now, two summations to infinite order are possible. A Dyson equation may be taken, which is invertible, corresponds to a *partial resummation* of the contributions included in the general Eq. (4.19), and yields a correction to the exciton energies while failing to describe the satellites,

$$\chi^{D} = \chi^{(0)} + \chi^{(0)}\Pi^{\text{exc}}\chi^{D}. \tag{4.22}$$

Alternatively, the cumulant expansion may be performed (for a diagonal self–energy), which is able to capture the satellite physics:

$$\chi^{c} = \chi^{(0)}e^{C} \tag{4.23}$$

(with $C_\alpha(t_{12}) = \int_{t_1}^{t_2}\int_{t_1'}^{t_2} dt'_{12}\Pi^{\text{exc}}(t'_{12})e^{iE_\alpha(-t'_{12})}$). For the purposes of this thesis, however, we will just consider the first order of both expressions, that is Eq. (4.21): our previous static approach (Sec. 3.4.5) was also confined to first order. We will also neglect the corrections to the exciton energies and focus on the description of the satellite structures (related at first order to scattering processes mediated by a single phonon). This will already fully clarify the physics involved and its consequences on the absorption spectrum. It will also give us all the ingredients for an implementation in `Yambo` in order to carry out the necessary numerical tests.

4.2.2 Form of the exciton–phonon coupling

A sketch of the derivation of Π^{exc} and a detailed description of the notation used, together with all of the quantities and expressions involved, is provided in App. I. Here we will just summarise the main results and consider for simplicity the ExcP self–energy to be diagonal in the exciton index (this restriction is necessary if the cumulant expansion is to be taken). Since we are interested in the corrections to the optical absorption spectrum, we start from zero–momentum, optically created excitons (index α). Their scattering with phonons (μq) will involve a sum over finite–q exciton states (βq). Now the self–energy is given by

$$\Pi^{\text{exc}}_{\alpha\alpha}(\omega) = \frac{1}{N_q}\sum_{\mu\beta q}\frac{|\mathscr{G}^{\text{exc}}_{\beta\alpha,\mu q}|^2}{\omega - (E_{\beta q} + \Omega_{q\mu}) + \mathrm{i}\eta}. \tag{4.24}$$

We see that the poles of the self–energy — related to phonon emission only since we are at zero temperature — are at $\omega = E_{\beta q} + \Omega_{q\mu}$. Furthermore, the *exciton–phonon* coupling matrix element $\mathscr{G}^{\text{exc}}$, describing the probability of the scattering of exciton α into exciton βq by phonon μq, appears at the numerator. It is given by

$$\begin{aligned}\mathscr{G}^{\text{exc}}_{\beta\alpha,\mu q} &= g^{\mu q\downarrow}_{\beta\alpha} - g^{\mu q\uparrow}_{\beta\alpha} \\ &= \sum_{\mathcal{K}_1}\left[\sum_{v_2}\left(\mathcal{A}^{c_1k_1,v_2k_1-q}_{\beta q}\right)^* \mathcal{A}^{\mathcal{K}_1}_{\alpha} g^{q\mu}_{v_1k_1,v_2k_1-q} - \sum_{c_2}\left(\mathcal{A}^{c_2k_1+q,v_1}_{\beta q}\right)^* \mathcal{A}^{\mathcal{K}_1}_{\alpha} g^{q\mu}_{c_2k_1+q,c_1k_1}\right].\end{aligned} \tag{4.25}$$

The specific notation used (see Sec. I.2) is chosen with the aim of simplifying the eventual `Yambo` implementation.[4] Note that in the case of a two–level system, $|\mathscr{G}^{\text{exc}}|^2$ reduces correctly to Eq. (4.15) and the final result, Eq. (4.17), is then obtained.

It should be noted that analogous expressions for the exciton–phonon coupling, denoted for brevity as $|\mathscr{G}^{\text{exc}}|^2 = \mathcal{A}^* g \mathcal{A} g^*$, already appeared in the literature, to the best of our knowledge, in Refs. [107] (2017) and [110] (2005). In Ref. [107], more concerned with temperature–dependent corrections to the exciton energies (and possibly with exciton–phonon scattering lifetimes), the excitons were treated as pure bosons from the start (introducing excitonic creation and annihilation operators), and made to couple directly with the electron–phonon interaction written in the exciton basis. A full Dyson equation $\chi^{\text{full}} = \chi^{(0)} + \chi^{(0)}\Xi\chi^{\text{full}}$ was then *assumed*. In Ref. [38], instead, the first–order correction to $\chi^{(0)}$ in the dynamical part of the interaction was derived starting from a general dynamical kernel for the BSE. As for Ref. [110], a model exciton–phonon coupling of the type $|\mathscr{G}^{\text{exc}}|^2 = \mathcal{A}^* g \mathcal{A} g^*$ appears in a 1D tight–binding Hamiltonian including the excitons as linear combinations of electron and hole creation and destruction operators, then individually coupled to the phonon system. The interest here was precisely to describe the phonon–induced sidebands appearing in the absorption spectra of carbon nanotubes in addition to the main absorption peak. In this case, of course, the values of the required physical quantities were chosen and not computed *ab initio*, but the model allows the authors to perform a smooth integration over the q wave vectors in order to obtain (i) the correct renormalisation of the main peak and (ii) a broad structure for the sideband (instead of discrete Lorentzian peaks for each q).

[4] The calligraphic symbol $\mathcal{A}$ for the exciton eigenvectors is used as a reminder that we still did not specify which response function we are correcting: it can be either χ, with $\mathcal{A} = A$, or $\overline{\chi}$, with $\mathcal{A} = \overline{A}$.

Back to our case, by using Eqs. 4.24 and 4.25, we can write $\chi^{(1)} = \chi^{(0)} + \chi^{(0)}\Pi^{\text{exc}}\chi^{(0)}$ in the exciton basis as

$$\chi^{(1)}_{\alpha}(\omega) = \frac{1 - R_{\alpha}}{\omega - E_{\alpha} + \mathrm{i}\eta} + \frac{1}{N_q}\sum_{\mu\beta q}\frac{|\mathscr{G}^{\text{exc}}_{\beta\alpha,\mu q}|^2}{\mathscr{W}^2_{\beta\alpha,\mu q}}\frac{1}{\omega - (E_{\alpha} + \mathscr{W}_{\beta\alpha,\mu q}) + \mathrm{i}\eta}, \tag{4.26}$$

with the renormalisation factor R and energy denominator $\mathscr{W}$ given by

$$\begin{aligned} R_{\alpha} &= \frac{1}{N_q}\sum_{\mu\beta q}\frac{|\mathscr{G}^{\text{exc}}_{\beta\alpha,\mu q}|^2}{\mathscr{W}^2_{\beta\alpha,\mu q}}, \\ \mathscr{W}_{\beta\alpha,\mu q} &= E_{\beta q} - E_{\alpha} + \Omega_{q\mu}. \end{aligned} \tag{4.27}$$

Again, this is a first–order correction in the dynamical part of the interaction.[5] Consequently, we see satellites appear with strength $|\mathscr{G}^{\text{exc}}|^2/\mathscr{W}^2$ (therefore at first order in $|\mathscr{G}^{\text{exc}}|^2$), accompanied by a reduction in the intensity of the main peak — recall that $\chi^{(0)}_{\alpha}(\omega) = [\omega - E_{\alpha} + \mathrm{i}\eta]^{-1}$. The position of the satellite is $\omega = E_{\alpha} + \mathscr{W}_{\beta\alpha,\mu q} = E_{\beta q} + \Omega_{q\mu}$. Writing it with the modified frequency $\mathscr{W}$ emphasises its nature as a satellite of E_{α}; yet, the actual position of the peak is at the energy of the finite–q excitons plus the emitted phonon frequency. This is exactly what we expect from our previous study of indirect absorption with excitons in bulk hBN (see Fig. 3.10). In that case we had to add the phonon frequency $\Omega_{\mu q}$ "by hand" since the exciton–phonon coupling was calculated statically. Here instead, the correct result arises naturally from the dynamical treatment of the interaction in perturbation theory. Finally, we recall from Eq. (1.87) that the optical transition amplitude is $T_{\alpha} = \sum_{\mathcal{K}} \mathcal{A}^{\mathcal{K}}_{\alpha} d_{\mathcal{K}}$, so that the $G = 0$, $G' = 0$ component of the G–space tensor for $\chi^{(1)}$ can be written as

$$\begin{aligned} \chi^{(1)}_{00}(\omega) &= \sum_{\alpha} |T_{\alpha}|^2 \chi^{(1)}_{\alpha}(\omega) \\ &= \sum_{\alpha}\left[\frac{|T_{\alpha}|^2[1 - R_{\alpha}]}{\omega - E_{\alpha} + \mathrm{i}\eta} + \frac{1}{N_q}\sum_{\mu\beta q}\frac{|T_{\alpha}|^2|\mathscr{G}^{\text{exc}}_{\beta\alpha,\mu q}|^2}{\mathscr{W}^2_{\beta\alpha,\mu q}}\frac{1}{\omega - (E_{\alpha} + \mathscr{W}_{\beta\alpha,\mu q}) + \mathrm{i}\eta}\right]. \end{aligned} \tag{4.28}$$

This is the quantity, along with the self–energy Eq. (4.24), that we want to compute by *ab initio* calculations.

4.2.3 Implementation

We implemented the self–energy, in particular Eqs. (4.24) and (4.25), in the `Yambo` code. Notes and details about the implementation and the code structure are provided in App. J for the interested reader. This development is one of the main achievements of this thesis because it enables *ab initio* calculations for a dynamical BSE, and eventually will permit the computation of indirect absorption spectra.

Some test results are shown in Fig. 4.3 for an unconverged bulk hBN calculation (details in the caption). The quantities shown in (a) are $\text{Re}\Pi^{\text{exc}}_{33}(\omega)$ (in red) and $\text{Im}\Pi^{\text{exc}}_{33}(\omega)$ (in blue). The subscript 3 refers to the lowest–lying bright excitonic level (not counting degeneracies) at $q = 0$, whose energy is indicated by the vertical line. In frame (a), Eq. 4.24 is computed using the full response function $\chi^{(0)}$ in the finite–q sBSE calculations (see Eq. (1.62) in Sec. 1.4). This is a recent development in the

[5] Note that we have neglected the the "quasiparticle" correction to the bare exciton energy E_{α}, since we are not interested in that at this point.

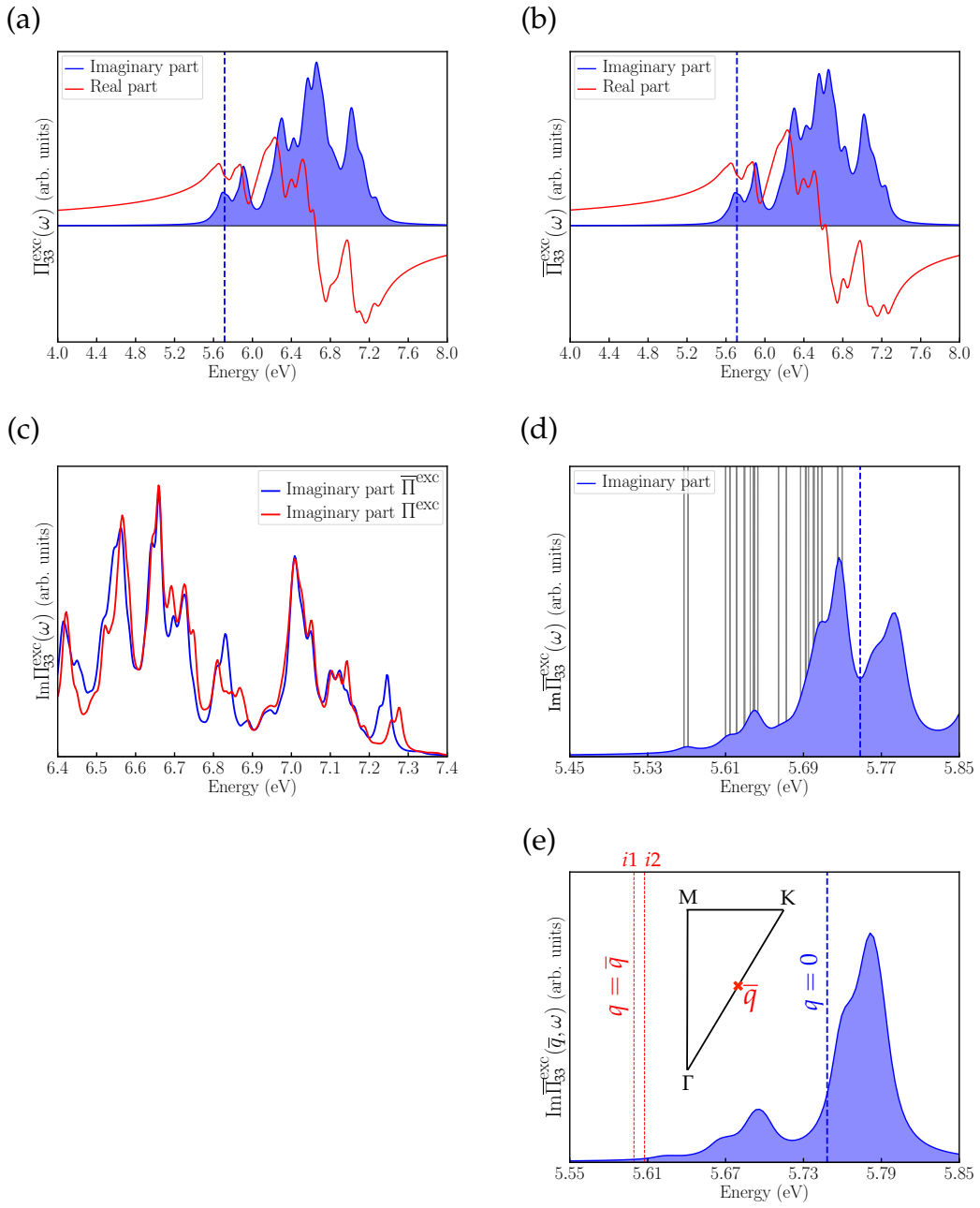


FIGURE 4.3: Exciton–phonon self–energies $\Pi^{exc}_{\alpha\alpha}$, Eq. (4.24), computed for the lowest–lying optically active exciton ($\alpha = 3$) in a test calculation for bulk hBN. We used a $6\times6\times2$ k– and q–grid sampling, two valence and two conduction bands at the DFT level, and eleven excitonic states (note that the exciton energies are rigidly shifted to their converged values). The peak broadenings are set at 0.04 eV. (a) Π^{exc} is computed using the the full response function $\chi^{(0)}$ for the static BSE calculations. (b) $\overline{\Pi}^{exc}$ is computed using the proper response function $\overline{\chi}^{(0)}$ for the static BSE calculations. The imaginary part of Π^{exc} is shown in blue, the real part in red. The dashed blue lines mark the energy of the corresponding exciton E_3. (c) A comparison of $\mathrm{Im}\Pi^{exc}$ (red) and $\mathrm{Im}\overline{\Pi}^{exc}$ (blue) is shown in a narrower energy range. (d) Same as (b), but zoomed in around the energy E_3; the peak broadening has been reduced to 0.01 eV. The black vertical lines indicate the finite–q transitions (as poles of the self–energy) lying below E_3. (e) Same as (d), but this time the self–energy is computed only at momentum $\bar{q}$, whose position within the IBZ is labeled in the inset with a red cross. The dashed red lines represent the two lowest–bound excitons with momentum $\bar{q}$, called $i1$ and $i2$. This plot is meant to be compared to Fig. 3.10(a).

`Yambo` code. Thus, in this case, if we look at Eq. (4.25) we have that $\mathcal{A}^{\mathcal{K}}_{\gamma} = A^{\mathcal{K}}_{\gamma}$. In frame (b), we instead use the *proper* response function $\overline{\chi}^{(0)}$ to compute the self–energy (now denoted $\overline{\Pi}^{\text{exc}}$). This is the response to a transverse external field that is typically evaluated in the context of optical absorption. In this case, as seen in Eq. (1.75), we have that $\mathcal{A}^{\mathcal{K}}_{\gamma} = \overline{A}^{\mathcal{K}}_{\overline{\gamma}}$ (with $\overline{\gamma}$ now explicitly labelling a "transverse" exciton). Additionally, the energies of the poles of $\chi^{(0)}$ and $\overline{\chi}^{(0)}$ may also differ, so we relabel them as E_{α} and $\overline{E}_{\overline{\alpha}}$, respectively, in order to avoid confusion.[6] It can be seen from Fig. 4.3(c) that the self–energy changes depending on the presence or absence of the long–range component of the Coulomb interaction v_0. Indeed, some of the peak structures in $\text{Im}\Pi^{\text{exc}}_{33}$ and $\text{Im}\overline{\Pi}^{\text{exc}}_{33}$ differ both in energy (since $E_{q\beta} + \Omega_{q\mu}$ may be different than $\overline{E}_{q\overline{\beta}} + \Omega_{q\mu}$) and in intensity (since the ExcP coupling will also be different). In particular, we know that in bulk hBN the bright exciton with symmetry E_{1u} is doubly degenerate (therefore $\overline{E}_3 = \overline{E}_4$): if v_0 is included, then this exciton undergoes a LT splitting, with its "longitudinal" part, the part not sensitive to optical absorption, becoming the eleventh state in the series ($E_3 < E_{11}$, see also Ref. [28]). All the other degenerate states in this energy range do not split.

The presence of low–lying indirect transitions is emphasised in Fig. 4.3(d): this is the same as Fig. 4.3(b), but zoomed in around the energy $\overline{E}_3$. We can see the presence of a dense peak structure *below* $\overline{E}_3$, originating from allowed exciton–phonon couplings with $\overline{E}_{\overline{\beta}q} + \Omega_{q\mu} < \overline{E}_3$ and $q \neq 0$ (black vertical lines). In order to provide a loose comparison with the indirect peak structure reported in Fig. 3.10(a) by means of our finite–difference approach (see Chap. 3), we have plotted the self–energy component of q–point $\overline{q}$ only in Fig. 4.3(e). Despite the fact that the two functions compared are technically different[7] and that the exciton and phonon energies are uncoverged in the present calculation, the same characteristic spectral shape is present in both figures: this indicates a likely agreement between the two approaches.

So far, we have discussed results at the self–energy level. As for the correction to the absorption spectrum, we need the imaginary part of the dielectric function $\varepsilon_2(\omega)$, which can be directly obtained from Eq. (4.28). The indirect correction to the absorption associated with the excitonic level α will be similar to its self–energy spectrum, renormalised by the the modified frequencies $\mathscr{W}^2$. However, should we choose to compute $\overline{\chi}^{(1)}(\omega)$, in accordance with standard optical absorption calculations in the direct case, or switch to $\chi^{(1)}(\omega)$? We will discuss this in the next Section.

4.2.4 Longitudinal–transverse exciton mixing

A transverse incoming em field will generate transverse excitations in the system which are properly described by the response function $\overline{\chi}(\mathbf{q}, \omega)$. The longitudinal excitations are absent in $\overline{\chi}$ because the long–range component of the Coulomb interation, v_0, has been removed. We have seen in Sec. 1.4 that this is only valid at $q = 0$ for optical absorption. Therefore, in principle there is no reason to remove the longitudinal poles of the full $\chi(\mathbf{q}, \omega)$ at finite wave vector. Furthermore, in the present treatment the excitations are interacting, and the interaction is mediated by phonon modes that can have both a longitudinal and a transverse component, something

[6] Recall that in our notation $E_{q\alpha} \equiv E_{\alpha}$ if $q = 0$.

[7] The plot in Fig. 3.10(a), based on Eq. (3.40), represents the indirect component of the absorption spectrum at point $\overline{q}$. Therefore, the proper comparison should be with the imaginary part of the second term in Eq. (4.28). In the imaginary part of the self–energy, Eq. (4.24), a scaling factor given by $\mathscr{W}^2$ is missing.

that, as we have mentioned in Sec. 1.6, may be especially relevant for a polar material like hBN. This means that the phonon–mediated mixing of longitudinal and transverse excitons is possible and should be accounted for in our treatment: in order to do that, the exciton–phonon self–energy should always be computed using χ, even when we are interested in the dynamical correction to $\overline{\chi}^{(0)}$, i.e. $\overline{\chi}^{(1)}$. This argument leads us to write Eq. (4.21) in the following way:

$$\overline{\chi}^{(1)}(13) = \overline{\chi}^{(0)}(13) + \overline{\chi}^{(0)}(11'12')\Xi(1'4'2'3')\overline{\chi}^{(0)}(3'34'3), \tag{4.29}$$

where Ξ *is computed using* $\chi^{(0)}$ *and not* $\overline{\chi}^{(0)}$, even though $\overline{\chi}^{(0)}$ is the response function to be dynamically corrected.[8]

This physical consideration can be complemented by a mathematical one. So far we have only considered a first–order correction to $\chi^{(0)}$ or $\overline{\chi}^{(0)}$, but let us consider the full Dyson equation appearing in Eq. (4.22):

$$\chi^{D,-1} = \chi^{(0),-1} - \Pi^{\text{exc}}. \tag{4.30}$$

We know already that in the static case $\chi^{(0)}$ and $\overline{\chi}^{(0)}$ are linked by Eq. (1.62), i.e.

$$\chi^{(0),-1} = \overline{\chi}^{(0),-1} - v_0. \tag{4.31}$$

Therefore, we can replace Eq. (4.31) into Eq. (4.30), and then define $\overline{\chi}^{D,-1} \equiv \overline{\chi}^{(0),-1} - \Pi^{\text{exc}}$. This means, firstly, that we can write an expression relating χ^D and $\overline{\chi}^D$ which is consistent with Eq. (4.31):

$$\chi^{D,-1} = \overline{\chi}^{D,-1} - v_0. \tag{4.32}$$

Secondly, a Dyson equation can now be written also for $\overline{\chi}$, but in terms of Π^{exc}, a quantity that is computed using χ, and not $\overline{\Pi}^{\text{exc}}$, as one might have expected:

$$\overline{\chi}^D = \overline{\chi}^{(0)} + \overline{\chi}^{(0)}\Pi^{\text{exc}}\overline{\chi}^D. \tag{4.33}$$

The validity of this equation guarantees the proper connection of the two response functions χ^D and $\overline{\chi}^D$ to the microscopic dielectric function, concisely written as $\varepsilon = 1 - v_0\overline{\chi}^D$ and $\varepsilon^{-1} = 1 + v_0\chi^D$.

The above considerations mean that we should consider Eq. (4.29), that is, the first order contribution to Eq. (4.32), in order to express correctly the ExcP coupling matrix elements in the absorption case:

$$\begin{aligned}\mathscr{G}^{\text{exc}}_{\beta\bar{\alpha},\mu q} &= g^{\mu q\downarrow}_{\beta\bar{\alpha}} - g^{\mu q\uparrow}_{\beta\bar{\alpha}} \\ &= \sum_{\mathcal{K}_1}\left[\sum_{v_2}\left(A^{c_1k_1,v_2k_1-q}_{\beta q}\right)^* \overline{A}^{\mathcal{K}_1}_{\bar{\alpha}} g^{q\mu}_{v_1k_1,v_2k_1-q} - \sum_{c_2}\left(A^{c_2k_1+q,v_1}_{\beta q}\right)^* \overline{A}^{\mathcal{K}_1}_{\bar{\alpha}} g^{q\mu}_{c_2k_1+q,c_1k_1}\right].\end{aligned} \tag{4.34}$$

Note the appearance of both $\overline{A}^{\mathcal{K}_1}_{\bar{\alpha}}$, referring to the optically created transverse excitons, and $A^{\mathcal{K}}_{\beta q}$, relative to the "internal" scattered excitons. The exciton–phonon coupling strength now includes longitudinal–transverse (LT) mixing and is given

[8] This is clarified by looking at the expressions in App. I: it means that the part in blue brackets of Eq. (I.16) is computed in terms of $\chi^{(0)}$, while the part outside contains $\overline{\chi}^{(0)}$. When the change of basis is made in Eq. (I.17), this will be from the transition basis to the basis of the $\overline{A}^{\mathcal{K}}_{\alpha}$, whereas the internal excitonic indices of Eq. (I.19) will be expressed in terms of the $A^{\mathcal{K}}_{\beta}$.

by

$$\frac{|\mathscr{G}^{\text{exc}}_{\beta\bar{\alpha},\mu q}|^2}{\mathscr{W}^2_{\beta\bar{\alpha},\mu q}} = \frac{|\langle\bar{\alpha}|\hat{\mathscr{G}}^{\text{exc}}_{\mu q}|\beta\rangle|^2}{(E_{\beta q} - \overline{E}_{\bar{\alpha}} + \Omega_{\mu q})^2}. \tag{4.35}$$

Here the energy denominator includes the difference between the transverse and longitudinal exciton energies (including, for example E_α-$\overline{E}_{\bar{\alpha}}$ which might be non–zero).

It is clear then that in the indirect absorption case the knowledge of χ is always required, while for direct absorption only $\overline{\chi}$ is in principle needed. The description of the phonon–mediated LT mixing of excitons represents the final contribution of this thesis. In the future, we are going to investigate the effects of the LT mixing both *ab initio* and with models. We know already that neglecting the mixing might be a reasonable approximation: this is what we implicitly did in our static approach in Sec. 3.4, since we computed the finite–difference derivatives of $\overline{\chi}^{(0)}$ and not $\chi^{(0)}$, still obtaining a good agreement with experiment. However, this might not always be the case. Yet, the LT mixing can be captured also in the context of static approaches, we just need to compute $\chi^{(0)}$ directly: the only additional complication is that the macroscopic dielectric function ε_M must then be obtained using Eq. (1.64) instead of the simpler Eq. (1.65).

4.2.5 Conclusions

In light of these findings, we can now conclude this Chapter by writing our final expression for the microscopic dielectric function, now including exciton–phonon coupling (it is instructive to compare this with Eq. (4.14) and especially with Eq. (1.75) by replacing the index λ with $\bar{\alpha}$):

$$\begin{aligned}\varepsilon_2(\omega) =& \frac{8\pi^2 e^2}{V}\sum_{\bar{\alpha}}|T_{\bar{\alpha}}|^2[1-R_{\bar{\alpha}}]\delta(\omega-\overline{E}_{\bar{\alpha}})+ \\ & \frac{8\pi^2 e^2}{VN_q}\sum_{\bar{\alpha}\mu\beta q}\frac{|T_{\bar{\alpha}}|^2|\mathscr{G}^{\text{exc}}_{\beta\bar{\alpha},\mu q}|^2}{\mathscr{W}^2_{\beta\bar{\alpha},\mu q}}\delta(\omega-(\overline{E}_{\bar{\alpha}}+\mathscr{W}_{\beta\bar{\alpha},\mu q})).\end{aligned} \tag{4.36}$$

This expression paves the way for a significantly improved description of indirect absorption, because it finally gives access to a more complete spectrum, including on the same footing phonon–mediated and direct contributions. The phonon–mediated part is then accurately described in terms of dynamical exciton–phonon interaction (an advance on the theoretical side) and spans the full BZ of the crystal instead of just a single q–point (an advance on the numerical side). Indeed, the self–energy plot in Fig. 4.3(d) already clearly describes an indirect absorption shoulder, whose smooth appearance is due to the fact that we are able to compute the ExcP couplings at many different momenta[9]. The cumulant expansion of Eq. (4.36), given by Eq. (4.33), may also be taken at this point.

As Eq. (4.36) suggests, the self–energies shown in Fig. 4.3 should be computed as $\Pi^{\text{exc}}_{\bar{\alpha}\bar{\alpha}} =$, i.e., including the LT mixing. The necessary ingredients are then: (i) the Kohn–Sham electronic eigenvalues and wave functions — computed in DFT — along with their QP corrections (GW); (ii) the full list of excitonic eigenvalues and eigenvectors associated with the sBSE response function $\chi^{(0)}$; (iii) the excitonic eigenvalues and eigenvectors, only at $q = 0$, associated with the "proper" sBSE response function $\overline{\chi}^{(0)}$; (iv) the full list of phonon frequencies and ep coupling matrix

[9] Note however that in this Figure the couplings are not rescaled by the frequency $\mathscr{W}$.

elements (computed with DFPT). On the one hand, the convergence of the q–sums in Eq. (4.36) is a particularly delicate point: our early results suggest that dense meshes might be needed, especially because the value of $R_{\bar{\alpha}}$ is prone to numerical instabilities (in analogy with standard ep self–energy calculations). Therefore, numerical approaches to accelerate convergence (such as usage of the crystal symmetries to reduce the value of N_q, interpolation schemes, *et cetera*) might be beneficial for complicated systems.

Finally, we point out that Eq. (4.36) constitutes the starting point for a rich variety of additional developments such as an extension to finite temperatures and/or the application of the cumulant ansatz to include multiphonon scatterings.

Conclusion

The research work presented in this thesis has been inspired by two connected scientific problems: (i) on the theoretical side, the development of practical approaches for the *ab initio* description of indirect absorption when strong excitonic effects are present; (ii) on the materials side, the explanation of the fine structures appearing in the optical spectroscopy of hexagonal boron nitride.

We briefly restate below our main results, followed by an assessment of the future research avenues to be explored.

Main results

There is an intense computational and theoretical effort currently ongoing in the study of the optical spectroscopy of 2D semiconductors. Less work has been done for cases where phonon–assisted optical excitations are important to the response of the material because of the difficulties in modeling the combined spectral contributions of both excitons and phonons. We have solved the problem in this thesis, providing a theoretical and computational description of the impact of exciton–phonon coupling on the spectral features of a layered material with strongly bound excitons, namely hBN.

In order to achieve this, we have elucidated the excitonic structure in BN systems. We established that exciton physics is essential to describe reliably this material explaining how layer stacking, nearly–free electron states and the character of the optical gap influence the spectral features. We have also proven that a strong electron–phonon interaction is present in these systems together with a clearly indirect optical gap for systems with more than three layers, including bulk.

We have developed a computationally feasible static approximation for the computation of the exciton–phonon coupling by calculating the variation of the optical response function with respect to the phonon–driven atomic displacements. By applying it to the description of the optical spectra, we were able to reproduce the experimentally observed spectral fine structure: we definitively prove that the much–discussed sidebands in the hBN luminescence spectrum come from the coupling of the two lowest–lying excitons, of wave vector $\bar{q}$, with every in–plane phonon mode of the same wave vector, as expected from our selection rules. The absorption onset, instead, is still dominated by the direct excitons.

After overcoming this first important problem, we have started to tackle the remaining liminations in accuracy such as the limited q–point sampling and the static approximation. We have considered a recent theory about a microscopic description of the exciton–phonon coupling arising from a dynamical correction to the kernel of the Bethe–Salpeter equation. We find that in this view, the phonon–assisted component of the spectrum can be reframed as a collection of satellites of the excitonic spectral function. We have successfully implemented the resulting equations, once adapted to our physical problem, in the `Yambo` code and have obtained promising first results. In particular, we modify our expressions in order to account for the

phonon–mediated longitudinal–transverse mixing of the excitons, an effect that has never been considered before.

Because of the significant steps forward described in this thesis, we believe that it is finally possible to accurately describe indirect optical spectroscopy with first–principles tools.

Outlook

Future theoretical development

The microscopic, dynamical *ab initio* treatment for IP indirect absorption, Ref. [36], is accurate close to the indirect absorption onset but fails away from it. Conversely, the static, supercell–based approach, Ref. [105], is valid up to many order of magnitudes in the absorption coefficient, but fails at the absorption edge.

In the course of this thesis we joined the strong sides of these two approaches in order to find a solution for our specific research problem at hand. However, various enticing questions remain. One is how to build a rigorous theoretical bridge between the two methods, linking for example the exciton–phonon self–energy, Eq. (4.24), to the derivatives of the excitonic oscillator strengths, Eq. (3.36), through a series of controlled approximations. Another is how to construct a microscopic and dynamical theory that allows for the description of the temperature dependence of both exciton energies and exciton–dominated optical spectra on the same footing; on this topic, the "cumulant" approach offers some interesting insight.

Another issue, though far harder to tackle than the previous ones, is the inclusion of the microscopic exciton–phonon coupling into an out–of–equilibrium theoretical framework for the consistent desciption of the luminescence process, including the direct–to–indirect relaxation dynamics of excitations. This will eliminate the need to pass through the van Roosbroeck–Shockley relation in order to describe luminescence (and therefore remove all of its underlying assumptions, while leaving more controlled approximations in their place).

Work in progress

Our first, ongoing, task is to transform our preliminary results obtained with the dynamical Bethe–Salpeter equation into a complete and converged phonon–assisted absorption spectrum for bulk and monolayer BN. The code implementation will be improved and made more efficient with the addition of symmetry operations and parallelisation. It will also be streamlined so that it can be run just as a post–processing tool after the computationally–heavy many–body data have been calculated and stored in databases. Further validation of the implementation will also be performed with the help of simple, semi–analytical models.

This implementation can then be easily extended to cover the temperature dependence of the excitonic optical spectra, and this will be our next step: in fact, solving the exciton–phonon equations for a temperature–dependent phonon propagator leads to the appearance of two satellite structures for each phonon scattering, with the phonon Bose–Einstein occupation factor n_{BE} now appearing. The structure related to phonon emission is now weighted with $n_{BE} + 1$, and the new structure related to phonon absorption is weighted with n_{BE}.

Additionally, our implementation lends itself naturally to the introduction of the cumulant ansatz for the exciton propagator: in this way we hope to be able to explain the overtone peaks in the luminescence spectrum, which are experimentally attributed to multiphonon processes.

So far we have always ignored the possibility to correct the exciton energies themselves using this method, since we focused on the dynamical corrections to the spectral function at first order. However, after working on the excitonic cumulant we will be able to extract complex "quasiparticle" corrections to the exciton energies, with their imaginary part describing exciton–phonon scattering lifetimes. A new development along this line will be helpful in order to tackle the very difficult problem of out–of–equilibrium carrier relaxation in 2D systems, when the correlated motion of electron and holes may strongly impact the dynamics.

Open problems in hBN and beyond

After having established our two approaches to the description of exciton–phonon coupling, we now have the possibility to study indirect absorption in layered materials beyond the state of the art, and yet in a numerically feasible way. We will keep pursuing the answers to the interesting and very relevant questions that still remain on the spectroscopic properties of BN and other 2D materials.

One of the most intriguing findings is related to the hBN single layer: a very recent paper,[157] the very first about luminescence in the single layer (epitaxially grown on a graphite substrate), finds two visible peaks in the spectrum, whereas only one is expected due to its expected direct–band gap nature. It has been suggested that this is due to indirect emission, yet this is still unconfirmed and could be due to other effects such as strain and substrate interactions, something that we would like to investigate.

The spectroscopy of BN multilayers leads as well to another puzzling observation: the luminescence signal, usually very strong, seems to disappear when the number of layers is reduced below 6.[32] We plan to use our results on the exciton dispersion in multilayer BN systems to explore if the signal may be quenched by the excitons related to nearly–free electrons, and/or by changes in the nature of the optical gap.

Furthermore, we are in the process of applying our static, finite–difference approach to other materials possibly also showing an indirect minimum in the exciton dispersion. In particular, we are interested in heterobilayers of transition metal dichalcogenides: their low–lying interlayer excitons allow for photo–induced charge separation, which make them viable for energy applications.

Last but not least, it has been suggested for certain transition metal dichalcogenides that even in the single–layer limit some of them remain indirect band gap materials. This might explain the recently observed rich spectral features.[133] We cannot tackle this with a one q–point approach (the static approach), since a competition between direct and indirect transitions is likely: nonetheless, we can do it with the perturbative approach.

The above examples show that the indirect optical spectroscopy of layered materials is rich with fascinating unanswered questions, that we are now in the position to tackle.

Appendix A

Excitonic Hamiltonian for a two–level system

An exciton in the proper sense of the term cannot exist in a two–level system, because it is not possible to mix transitions arising from different regions in the BZ and/or from different band states, since there are only one occupied and one unoccupied level. We can equivalently say that the exciton is a solid–state concept and cannot be properly applied to an isolated system. However, we can still have mixing between the resonant and anti–resonant character of the same transition as long as we do not employ the TDA.

We start with the excitonic Hamiltonian in the transition basis:

$$H^{n_1 n_2}_{n_3 n_4} = (\epsilon_{n_2} - \epsilon_{n_1})\delta_{n_1 n_3}\delta n_2 n_4 + (f_{n_3} - f_{n_4})\left[2V^{n_1 n_2}_{n_3 n_4} - W^{n_1 n_2}_{n_3 n_4}\right], \tag{A.1}$$

with ϵ_n and f_n being the single–particle energy and occupation of state n, respectively, W being the statically screened Coulomb interaction and V the exchange contribution. We label the hole level as v and the electron level as c. Furthermore, we consider an insulator at zero temperature, so that $f_v - f_c = 1$.

First we define

$$\begin{aligned} K_d &\equiv W^{vc}_{vc} - 2V^{vc}_{vc}, \\ K_{od} &\equiv W^{vc}_{cv} - 2V^{vc}_{cv}, \\ \varepsilon &\equiv \epsilon_c - \epsilon_v - K_d. \end{aligned} \tag{A.2}$$

In a two–level system, W and V are real quantities and are symmetric under exchange of indices (e.g. $W^{vc}_{vc} = W^{cv}_{cv}$). Now we can write

$$H = \begin{bmatrix} \varepsilon & K_{od} \\ -K_{od} & -\varepsilon \end{bmatrix}. \tag{A.3}$$

The solution of the eigenvalue equation $HA_\lambda = E_\lambda A_\lambda$ gives the exciton energies

$$\begin{aligned} E_1 &= +\sqrt{\varepsilon^2 - K^2_{od}}, \\ E_2 &= -\sqrt{\varepsilon^2 - K^2_{od}}, \end{aligned} \tag{A.4}$$

with the plus/minus sign referring to a resonant/anti–resonant excitation. The normalised eigenvectors are

$$A_1 = \begin{pmatrix} -\mathrm{sgn}(K_{od})\sqrt{\frac{\varepsilon+\sqrt{\varepsilon^2-K_{od}^2}}{2\varepsilon}} \\ \frac{|K_{od}|}{\sqrt{2\varepsilon(\varepsilon+\sqrt{\varepsilon^2-K_{od}^2})}} \end{pmatrix},$$
$$A_2 = \begin{pmatrix} -\mathrm{sgn}(K_{od})\sqrt{\frac{\varepsilon-\sqrt{\varepsilon^2-K_{od}^2}}{2\varepsilon}} \\ \frac{|K_{od}|}{\sqrt{2\varepsilon(\varepsilon-\sqrt{\varepsilon^2-K_{od}^2})}} \end{pmatrix}. \tag{A.5}$$

Since the eigenvectors are not orthogonal, we construct the overlap matrix, defined as $N_{\lambda\lambda'} = A_\lambda^{vc\,*} A_{\lambda'}^{vc} + A_\lambda^{cv\,*} A_{\lambda'}^{cv}$:

$$N = \begin{bmatrix} 1 & \frac{|K_{od}|}{\varepsilon} \\ \frac{|K_{od}|}{\varepsilon} & 1 \end{bmatrix},$$
$$N^{-1} = \begin{bmatrix} 1 & -\frac{|K_{od}|}{\varepsilon} \\ -\frac{|K_{od}|}{\varepsilon} & 1 \end{bmatrix} \frac{1}{1-K_{od}^2/\varepsilon^2}. \tag{A.6}$$

Let us now consider the electron–hole correlation function L. According to the BSE, L is given, in frequency space, by the resolvent of H as

$$\begin{aligned} L_{n_3n_4}^{n_1n_2}(\omega) &= [H - \mathbb{1}z]^{-1}{}_{n_3n_4}^{n_1n_2} (f_{n_2} - f_{n_1}) \\ &= (\delta_{n_1v} - \delta_{n_2v}) \sum_{\lambda\lambda'} \frac{A_\lambda^{n_1n_2} N_{\lambda\lambda'}^{-1} A_{\lambda'}^{n_3n_4\,*}}{\omega - E_\lambda + \mathrm{i}\eta_\lambda} \\ &= (\delta_{n_1v} - \delta_{n_2v}) \sum_{\lambda\lambda'} A_\lambda^{n_1n_2} N_{\lambda\lambda'}^{-1} A_{\lambda'}^{n_3n_4\,*} L_\lambda(\omega). \end{aligned} \tag{A.7}$$

In the last step, we defined L_λ as the diagonal propagator in the excitonic basis, in its causal (retarded) $L_1(\omega) = (\omega - E_1 + \mathrm{i}\eta)^{-1}$ and anti–causal (advanced) $L_2(\omega) = (\omega - E_2 - \mathrm{i}\eta)^{-1}$ parts (recall that $E_1 = -E_2$). In our case the propagator $L_{n_3n_4}^{n_1n_2}$ is made up by four components, corresponding to the elements of H: resonant, anti–resonant and off–diagonal. We compute explicitly all $A_\lambda^{n_1n_2} N_{\lambda\lambda'}^{-1} A_{\lambda'}^{n_3n_4\,*}$ terms by making use of Eqs. (A.4), (A.5), (A.6) and we obtain

$$\begin{aligned} L_{vc}^{vc}(\omega) &= \left[\frac{1}{2}\left(\frac{\varepsilon}{E_1}+1\right) L_1(\omega) - \frac{1}{2}\left(\frac{\varepsilon}{E_1}-1\right) L_2(\omega)\right], \\ L_{cv}^{cv}(\omega) &= -\left[\frac{1}{2}\left(\frac{\varepsilon}{E_1}-1\right) L_1(\omega) + \frac{1}{2}\left(\frac{\varepsilon}{E_1}+1\right) L_2(\omega)\right], \\ L_{cv}^{vc}(\omega) &= \frac{\mathrm{sgn}(K_{od})}{2}\sqrt{\frac{\varepsilon^2}{E_1^2}-1}\,[L_1(\omega) - L_2(\omega)] = L_{vc}^{cv}(\omega). \end{aligned} \tag{A.8}$$

Notice that $\varepsilon/E_1 \pm 1 > 0$ if $K_{od} \neq 0$, and that the expressions reduce as expected to the TDA case $L_{vc}^{vc}(\omega) = L_1(\omega)$ when $K_{od} = 0$ (recall that in fact $|K_{od}| = \sqrt{\varepsilon^2 - E_1^2}$). In the TDA case there is no mixing and the transition and "exciton" bases coincide. The energy of the two–level transition is reduced from $\varepsilon_c - \varepsilon_v$ to ε. The same calculation as above also yields the completeness relation $\sum_{\lambda\lambda'} A_\lambda^{n_1n_2} N_{\lambda\lambda'}^{-1} A_{\lambda'}^{n_3n_4\,*} = \delta_{n_1n_3}\delta_{n_2n_4}$.

As a final step, we write the full propagator in space–time coordinates. The

(inverse) Fourier transform of Eq. (A.8) is simply given by the respective transforms of L_1 and L_2, so that we have $L_1(t_{13}) = -\mathrm{i}\theta(t_3 - t_1)\exp(\mathrm{i}E_1(t_1 - t_3))$ and $L_2(t_{24}) = \mathrm{i}\theta(t_2 - t_4)\exp(-\mathrm{i}E_1(t_2 - t_4))$. Now we introduce the single–particle wave functions $\varphi_n(\mathbf{r})$ and obtain the four–point L as

$$\begin{aligned} L(1423) =& L^{vc}_{vc}(t_{1423})\varphi_v(\mathbf{r}_1)\varphi^*_c(\mathbf{r}_2)\varphi^*_v(\mathbf{r}_3)\varphi_c(\mathbf{r}_4)+ \\ & L^{vc}_{cv}(t_{1423})\varphi_v(\mathbf{r}_1)\varphi^*_c(\mathbf{r}_2)\varphi^*_c(\mathbf{r}_3)\varphi_v(\mathbf{r}_4)+ \\ & L^{cv}_{vc}(t_{1423})\varphi_c(\mathbf{r}_1)\varphi^*_v(\mathbf{r}_2)\varphi^*_v(\mathbf{r}_3)\varphi_c(\mathbf{r}_4)+ \\ & L^{cv}_{cv}(t_{1423})\varphi_c(\mathbf{r}_1)\varphi^*_v(\mathbf{r}_2)\varphi^*_c(\mathbf{r}_3)\varphi_v(\mathbf{r}_4). \end{aligned} \tag{A.9}$$

Appendix B

Supercells

We briefly introduce here, for reference, the technique of calculating finite momemntum quantities using supercells. Here by supercell we intend a simulation cell that contains multiple repeated copies of the crystal unit cell along the directions in which the crystal itself is periodic: hence the x and y directions for 2D systems.

Starting with the primitive lattice vectors $\mathbf{a}_i$ that define the unit cell in real space, the reciprocal ones $\mathbf{b}_i$ are defined as $\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij}$, and a wave vector inside the reciprocal unit cell is given by $\mathbf{k} = \sum_i k_i \mathbf{b}_i = \sum_i \frac{n_i}{N_i}\mathbf{b}_i$ $(0 \leq n_i \leq N_i - 1)$. Here, k_i are the *fractional coordinates* and N_i represents the number of real–space unit cells along the i–direction (when the N_i go to infinity, we have a continuum of wave vectors). Therefore, each supercell size $N_x N_y N_z$ can accommodate the real–space periodicity of a perturbation with a corresponding wave vector $\mathbf{k}$ (like the periodicity of a lattice vibration). This means that by performing a simulation at the Γ point of a specific supercell, we also obtain information about the eigenvalues and eigenvectors relative to the reciprocal space k–points with compatible periodicities, which and are then *folded onto* Γ in the supercell. This is exemplified in Fig. B.1 for the energy band of a monoatomic, 1D crystal with lattice parameter a.[1] If we double its lattice constant, $a_2 = 2a$, then the BZ size is halved and the k–point at the zone edge is folded onto the Γ point in the new "supercell" band structure (blue line). If we perform a calculation at Γ on a system four times the original size, $a_4 = 4a$, we then obtain both the state at the zone edge (now at half the periodicity of the new supercell) and a new one coming from the zone center (red lines). This also makes clear that by resorting to a supercell calculation, we are *trading k–points for bands, i.e. decreasing the set spanned by the k–index at the cost of increasing the one spanned by the n–index.*

In order to define the supercell lattice vectors, $\mathbf{a}_{Si}$, and the supercell fractional coordinates, k_{Si}, we use the supercell matrix:[129]

$$\begin{pmatrix} \mathbf{a}_{S1} \\ \mathbf{a}_{S2} \\ \mathbf{a}_{S3} \end{pmatrix} = \begin{pmatrix} N_1 & 0 & 0 \\ 0 & N_2 & 0 \\ 0 & 0 & N_3 \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{pmatrix}, \quad \begin{pmatrix} k_{S1} \\ k_{S2} \\ k_{S3} \end{pmatrix} = \begin{pmatrix} N_1 & 0 & 0 \\ 0 & N_2 & 0 \\ 0 & 0 & N_3 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix}. \tag{B.1}$$

Here, a k–point is correctly folded onto Γ in the supercell if the k_{Si} are integers. For example an M and a K point of an hexagonal BZ have fractional coordinates $(0, 1/2, 0)^T$ and $(1/3, 1/3, 0)^T$ respectively. According to Eq. B.1, we need a $1 \times 2 \times 1$ supercell (i.e. $\mathbf{a}_{M2} = 2\mathbf{a}_2$, $\mathbf{a}_{M1} = \mathbf{a}_1$, $\mathbf{a}_{M3} = \mathbf{a}_3$) to fold M onto Γ, and a $3 \times 3 \times 1$ supercell for K.[2] We performed supercell BSE calculations in order to provide

[1] This is given by the tight–binding expression $E_k = \epsilon + 2\gamma\cos(ka)$ where ϵ and γ are parameters.

[2] A smaller supercell for K can actually be obtained by rotating the lattice vectors to a $\sqrt{3} \times \sqrt{3} \times 1$ configuration.

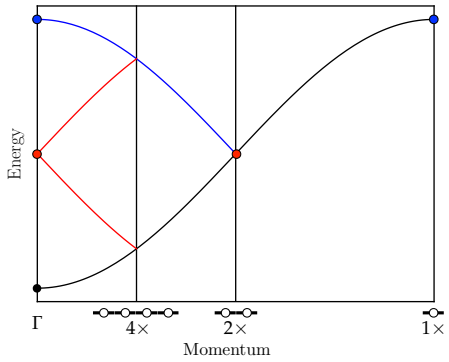


FIGURE B.1: Folding band states onto the Γ point in a one–dimensional crystal. The black line is the dispersion in the primitive BZ, the blue line is the folded part when the real space lattice constant is doubled, and the red lines are the folded parts when the lattice constant becomes four times the original size. A band calculation at the Γ point in the latter supercell BZ will net the folded red and blue points, which are labeled, as additional bands.

support to the TBEM predictions[97] and to test the implementation of the finite–q BSE in the `Yambo` code, providing as a result equivalent finite–momentum wave functions and energies to the ones shown in Fig. 2.10.

However, performing supercell calculations quickly becomes impractical when the size $\mathcal{N} = N_1 N_2 N_3$ of the selected supercell increases beyond the first few integers, since the number of electrons in the system scales accordingly: in order to maintain the same level of convergence as in the unit–cell calculations, the following observations must be made. (i) The total number of k– or q–points N_k and N_q needed for convergence is reduced by a factor $\mathcal{N}$. (ii) The number of states (bands, excitons, phonons) needed for convergence is increased by a factor $\mathcal{N}$. (iii) The number of G–vectors needed for convergence is also increased by a factor $\mathcal{N}$, since the size of a supercell G–vector is a $\mathcal{N}$–submultiple of the corresponding unit–cell one, but they have to achieve the same energy cutoff.

It is possible to show that the minimum supercell size needed to fold a k–point with coordinates $(n_1/N_1, n_2/N_2, n_3/N_3)^T$ is just the least common multiple of the N_i, not their product. This is achieved by constructing a *non–diagonal* supercell, where off–diagonal elements are added to the supercell matrix:[129]

$$\begin{pmatrix} k_{S1} \\ k_{S2} \\ k_{S3} \end{pmatrix} = \begin{pmatrix} S_1 & S_{12} & S_{13} \\ 0 & S_2 & S_{23} \\ 0 & 0 & S_3 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ k_3, \end{pmatrix} \tag{B.2}$$

where the S–elements ($0 \leq S_{12} < S_2$, $0 \leq S_{13}, S_{23} < S_3$) can be found in such a way that the k_{Si} are integers and the new supercell (where the new lattice vectors are not related to the original ones by a simple scaling factor) has the desired minimal size. We make use of non–diagonal supercells in Chapter 3. We have developed a Python script compatible with the QUANTUM ESPRESSO code, to generate input files for any supercells, diagonal or non–diagonal, starting from the input file of a unit cell. Additionally, the supercell atomic positions can be displaced according to a phonon mode with the correct wave vector q if such mode has been previously calculated in the unit cell with DFPT. Such a static displacement is obtained from Eq. (1.99) with a fixed t, e.g.,

$$u_{Li\alpha}^{\lambda q}(t=0) = \frac{c}{\sqrt{M_i}} \mathrm{Re}\left\{ e^{i\mathbf{q}\cdot\boldsymbol{\tau}_L} \xi_{i\alpha}^{\lambda q} \right\}, \tag{B.3}$$

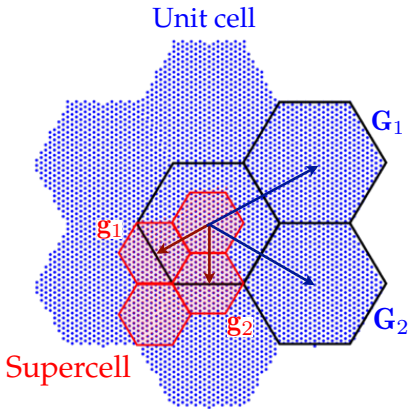


FIGURE B.2: Unit cell versus supercell in reciprocal space. A two–dimensional, hexagonal BZ (blue) is shown alongside its six nearest neighbors. The neighbors are the translation of the original hexagon by a reciprocal lattice vector $\mathbf{G}$. The q–point discrete sampling is shown with the blue dots. In red, the BZ corresponding to a 2×2 real–space supercell is shown, its area being four times smaller. The supercell q–point sampling is shown with red dots and three additional repetitions, two lying on the border of the original unit cell BZ and one completely outside, are displayed. They are connected to the original supercell BZ by a supercell reciprocal lattice vector $\mathbf{g}$.

where c is a custom scaling factor.

B.1 Folding of the response function χ

For the reasons explained above, the static RPA screening function $\chi^{RPA}_{GG'}(\mathbf{q})$, Eq. (1.47), which includes sums over G–vectors and band states and must be computed for a sufficiently large number of GG' components, is difficult to compute for large supercells. In order to obviate this problem, we developed a Python script that, taking as input a χ computed in the unit cell (uc), folds it onto a diagonal supercell (sc) of choice, provided the q–point samplings of the uc and sc are commensurate. Let us denote the uc reciprocal space vectors with capital letters and the sc ones with small letters. We know, because the supercell is diagonal, that each uc G–vector must be a multiple of some sc g–vector. We also know, because the momenta grids are commensurate, that each uc Q–point is related to a sc q–point by a specific sc g–vector, so that $\mathbf{Q} = \mathbf{q} + \mathbf{g}_\mathbf{Q}$ ($\mathbf{g}_\mathbf{Q}$ can and will be $\mathbf{0}$; these observations are clarified by looking at Fig. B.2). The next observation is that the sum of $\mathbf{g}_\mathbf{Q}$ with a uc G–vector must be a sc g–vector, so that $\mathbf{g} = \mathbf{G} + \mathbf{g}_\mathbf{Q}$. Therefore, the routine matches the uc and sc extended $Q + G$ grids (see Figure) and finds $\mathbf{g}_\mathbf{Q}$ for each Q–point. Ultimately it converts a previously computed `Yambo` database for χ on the uc $Q + G$ grid to a sc database in the $q + g$ grid that can immediately be read by the code in the context of a supercell BSE calculation. In short, the operation implemented is shown below:

$$\begin{aligned} \text{uc} \rightarrow \chi_{G_1G_2}(\mathbf{Q}) &= \chi(\mathbf{Q}+\mathbf{G}_1, \mathbf{Q}+\mathbf{G}_2) = \chi(\mathbf{q}+\mathbf{g}_\mathbf{Q}+\mathbf{G}_1, \mathbf{q}+\mathbf{g}_\mathbf{Q}+\mathbf{G}_2) \\ &= \chi(\mathbf{q}+\mathbf{g}_1, \mathbf{q}+\mathbf{g}_2) = \chi_{g_1g_2}(\mathbf{q}) \leftarrow \text{sc.} \end{aligned} \tag{B.4}$$

Both of the implementations described here were developed in the form of `yambo-py`[3] classes (respectively called *supercell.py* and *foldvX.py*).

[3] See `https://yambopy.readthedocs.io/en/latest/introduction.html`.

Appendix C

How many single–particle transitions make up a converged exciton?

In order to reduce the excitonic Hamiltonian size and make its diagonalisation more efficient, while performing a BSE calculation, one should only include the single–particle transitions (cvk) that give a finite contribution to the excitonic states of interest. However, the choice of such transitions is somewhat delicate in BN, where the excitonic weights can decay slowly in reciprocal space. Let us consider bilayer hBN. In Fig. C.1(a), The transition energies $\Delta_{cv}(\mathbf{k}) = E_c(\mathbf{k}) - E_v(\mathbf{k})$ obtained from the disentangled GW valence and conduction bands are shown in different colors. In order to obtain converged *ab initio* results for the absorption spectra in multilayer hBN, one might be tempted to only include in the calculations the area around the K point or along the KM region in the BZ (transitions below lines (A), (B) or (C) in the figure). This seems justified by looking at Fig. C.1(b), which shows the weights $\sum_{cv} |\Psi^{\lambda}_{cv}(\mathbf{k})|^2$ of the electronic transitions in the BZ for the lowest-bound bright exciton. However, it can be seen from Fig. C.1(c) that this would produce unconverged spectra. In fact, transitions up to 12 eV and located in the middle of the BZ still give non–negligible contributions to the first exciton, even though the band gap is at 7 eV and the exciton close to 5.5 eV. This is due to the fact that the Fourier intensities decay slowly away from the K point. Therefore, the converged result is obtained by increasing the energy window included in the calculation up to the π^*-σ^* crossing ((E) lines in Fig. C.1(a)-(b)).

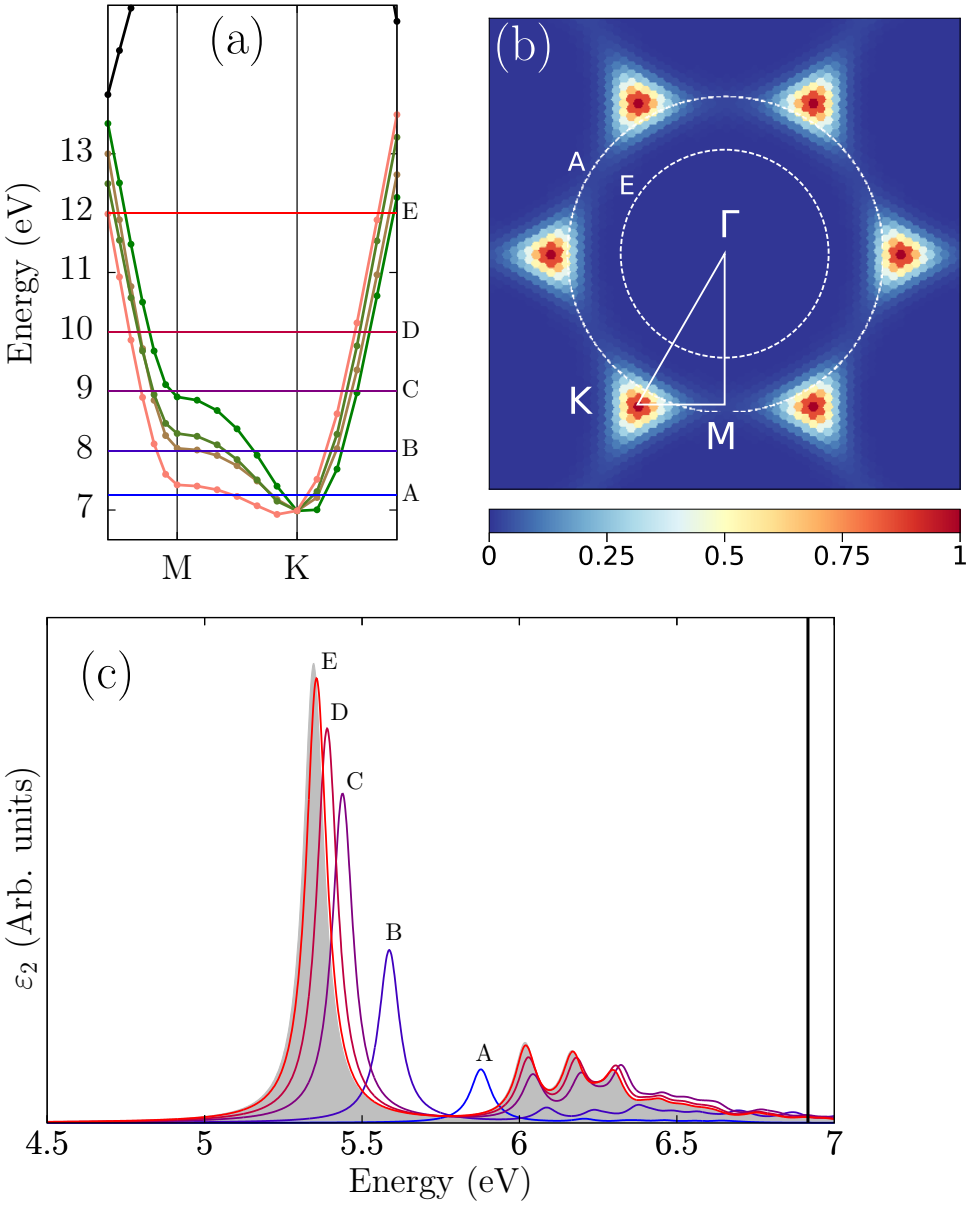


FIGURE C.1: Transition energy region (TER) for absorption, in the case of bilayer hBN. In (a), the transition energies obtained from the disentangled π and π^* bands are shown with different colors in the relevant part of the BZ. The lowest $\sigma \rightarrow \sigma^*$ transition is shown in black. The horizontal lines labeled from A to E represent different TERs. They include (A) only the K point, (B) the lowest transition along KM, (C) all transitions along KM, (D) parts of the ΓM and ΓK directions, (E) all energies of the relevant region. The excitonic weights in k-space for the first bright exciton are shown in (b). Most of the weight comes from the area around the K point. The intersections between the irreducible wedge of the BZ (white triangle) and the white dashed circles labeled A,E represent the fraction of the BZ which is included in the BSE calculations in the two cases. The imaginary part of the dielectric functions obtained using the five TERs from A to E are shown in (c). It can be seen how only a very wide TER – in this case the one labeled (E) – is able to reproduce the fully converged result (gray shadow).

Appendix D

Computational details for all the results presented in the Thesis

In this Appendix we give the computational details regarding all the first–principles calculations presented in this Thesis: converged values, tips for calculations, numerical accuracy and issues, *et cetera*.

D.1 Sections 2.2 and 2.3

These Sections were concerned with optical absorption and excitonic Davydov splitting in multilayer hBN (from monolayer to bulk).

The QE DFT–LDA calculations were performed using norm–conserving von Barth–Car pseudopotentials. The DFT convergence of the energy cutoff (110 Ry being the maximum value used) and of the k–point grid ($12 \times 12 \times (1)/(4)$) is standard procedure. However, if a subsequent `Yambo` calculation is in order, the following additional *caveats* should be considered: (i) if including a high number of conduction bands, make sure that they accurately converged. (ii) Non–symmorphic symmetries should be disabled. (iii) If dealing with a non–bulk system, it is advisable to position the system within the simulation supercell in such a way that it does not break the symmetry of the discrete real space grid used in the Fast Fourier Transform (FFT) routine. For example, if $z = 0$ is the bottom of the supercell, and the interlayer distance is d, the z–coordinates for the layers of a bilayer system should be $-d/2, d/2$; in the trilayer case it should be $-d, 0, d$. This will help `Yambo` identify correctly important symmetries (such as inversion symmetry).

Table D.1 summarizes the most important parameters needed to obtain converged GW π and π^* bands and converged (lowest–lying) excitonic peaks.

The convergence of the internal `Yambo` parameters was carefully checked by regularly increasing each one until differences in band energies (for GW) or excitonic peak positions (for BSE) were less than 0.01 eV each time (except for the pentalayer, where the threshold was increased to 0.02 eV), which is the precision of the GW method. As we are dealing with quasi–2D materials, special attention was paid to the amount of vacuum space introduced between repeated copies of the systems in the vertical direction. The Coulomb cutoff allows convergence with a separation distance of $v_z = 20$ Å. Another important observation is that if h is the thickness of the system and $L_z = h + v_z$ the supercell height, as we increase the number of atomic layers L_z becomes larger, and consequently we might need to use a denser k–point mesh and to sum over more G–vectors. We will also need to some over more unoccupied states since the number of electrons in the system is increased. Table D.1 summarizes the needed parameters to obtain converged GW π and π^* bands and converged (lowest–lying) excitonic peaks.

System	k–point mesh	number of bands	LFE cutoff (Ry)
1L	$24 \times 24 \times 1$	120	10
2L	$36 \times 36 \times 1$	200	10
3L	$42 \times 42 \times 1$	200	9
5L	$48 \times 48 \times 1$	350	12
Bulk	$18 \times 18 \times 6$	280	14

TABLE D.1: Size of the k–point mesh, number of bands and energy cutoff of the local field effects (LFE) in the RPA screening used in the GW and BSE calculations of monolayer (1L), bilayer (2L), trilayer (3L), pentalayer (5L) and bulk hBN. Only the highest values used between the GW and the BSE calculations are reported.

In the GW case, the plasmon–pole approximation was used for the computation of the electronic response function.[158] Its validity was checked, for the monolayer, against the direct integration in frequency space, yielding excellent agreement. Moreover, our GW bandgap value (7.26 eV) for the hBN monolayer is in good agreement with other results obtained with different many–body codes (7.36 eV[95] and 7.37 eV[159]). The numerical shift of 0.1 eV is entirely due to the underlying DFT calculation: the cited results can be obtained exactly by switching to the optimized lattice constant for the monolayer. Our optical spectrum for the monolayer also agrees with the one in Ref. [95].

An additional convergence check was performed on the monolayer, by decreasing the convergence threshold by almost an order of magnitude (using a $48 \times 48 \times 1$ k-point mesh, a vacuum separation of 30 Å, and summing up to 400 states). The results for GW band gap and excitonic peak positions differ by about 0.03 eV (rigid shift) from the ones obtained with the parameters listed in Tab. D.1. We conclude therefore that our results are well converged. Our reference calculations for the bulk system (convergence parameters also listed in the Table) are in agreement with previously established results.[21, 23]

D.2 Section 2.4

In this Section we calculated the exciton dispersion curves for BN systems (bilayer, trilayer, bulk) using the finite–q BSE.

The convergence parameters used are essentially the same as in the previous Section. The DFT–LDA pseudopotentials where changed to the "fhi" ones from Ref. [160]. For bulk, the GW and BSE k/q–grid was increased to $36 \times 36 \times 6$ in order to have better reciprocal–space resolution.

D.3 Section 3.2

Here we computed electron–phonon couplings and the zero–point redshift of the absorption spectrum in monolayer BN.

The prescriptions from the previous two Sections were followed, with the calculations being now performed at the optimised lattice constant for monolayer BN. Checking for the converged values of the total energy and of the forces acting on each ion (respectively for the scf and DFPT calculations) is standard procedure, and the q–point grid sampling for the interatomic force constants was $24 \times 24 \times 1$. In order to guarantee a very accurate convergence of the electron–phonon self–energy,

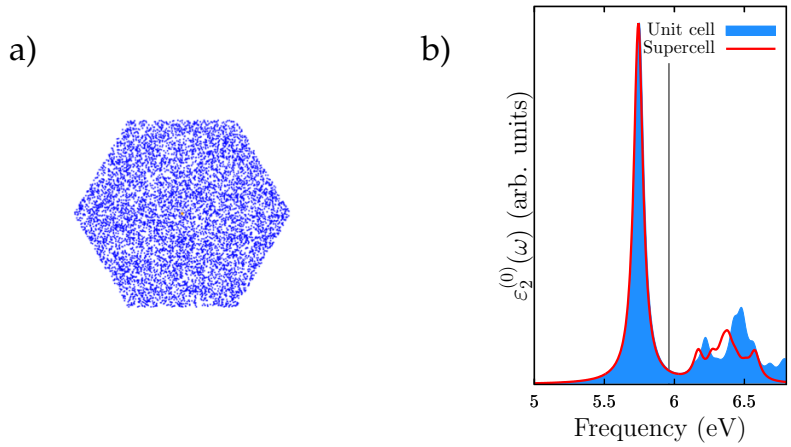


FIGURE D.1: (a) Figure referring to Sec. 3.2. The 8100, randomly generated q–points in the 2D BZ used for the calculation of electron–phonon couplings in monolayer hBN. (b) Figure referring to Sec. 3.4. The imaginary part of the dielectric functions are plotted for the hBN unit cell (blue, DFT–GW–BSE calculation on fine k–point grid) and non–diagonal supercell (red, DFT–scissor–BSE calculation on coarse k–point grid). Only the first 480 states are included in the iterative solution of the BSE in the supercell. The black vertical line is at the energy of the indirect quasiparticle band gap. Both spectra have a broadening parameter of 0.04 eV. Differences in the spectra (due to finite k–point samplings) only occur in the energy region above the band gap. In the relevant energy region below the band gap, both spectra are identical.

40 electronic states were included over a q–grid of 8100 (90×90) *randomly generated* q–points in the 2D BZ (randomized q–samplings can accelerate convergence for low–dimensional systems) — see Fig. D.1(a). The calculations of the required electron–phonon coupling matrix elements $g_{nmk}^{\lambda q}$ were automatised using Python scripts. The random integration method (Sec. 1.5.2) was used to attenuate numerical instabilities in the self–energy q–integrations. The convergence of the electron–phonon self–energy was checked on the full spectral functions for the band edges at the high–symmetry points in the BZ (not just on the band gap energies). It is worth noting that the Newton method (default option) fails to solve properly the quasiparticle equation, Eq. (1.55), and gives wrongly shifted energies. Therefore, we apply the secant method directly on the spectral function $\mathrm{Im}G_{nk}(\omega)$ for each corrected (nk) state with an external Python script.

Supercells.[12] Extreme care was taken in ensuring that every frozen–atoms supercell calculation shared the same level of convergence as the unit–cell one. This was done by properly scaling the internal yambo convergence parameters, and comparing the results of the calculations. The same approach was taken in the construction of the scissor operator for the supercells, in such a way that the resulting scissor+BSE supercell spectra coincided with the GW+BSE unit cell one.

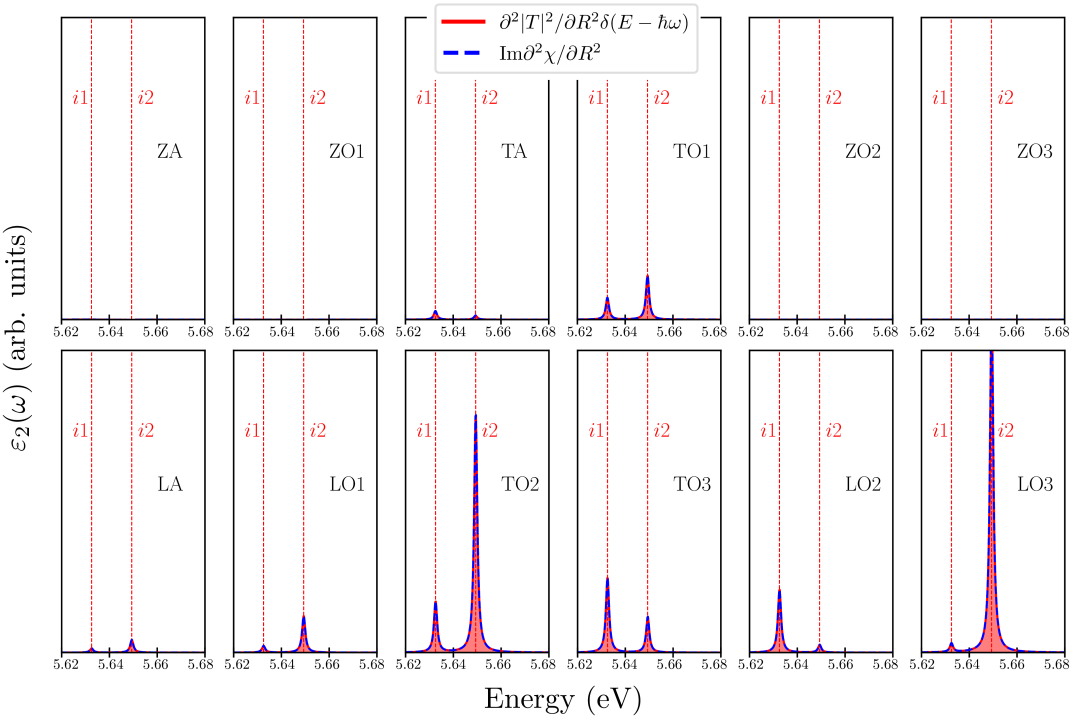


FIGURE D.2: Figure referring to Sec. 3.4: static exciton–phonon couplings between the $i1$ and $i2$ excitons and the 12 different phonons at $\overline{q} = \frac{1}{2}|\Gamma K|$. We compare results obtained from the full second derivative of the response function (dashed blue lines) — which contains also derivatives of the exciton energies — and from the second derivative of just the oscillator strengths (full red lines), see Eq. (3.36). An average over polarization of the incoming light and sum over the 6 equivalent directions of phonons with wave vector $|\overline{q}|$ is performed.

D.4 Section 3.4

The phonon frequencies and eigenmodes were computed with DFPT in the unit cell, using a q–point grid sampling of $18 \times 18 \times 6$. The G_0W_0 and semi–self–consistent GW (sscGW) corrections to the band energies were obtained with `Yambo` for the unit cell, using the plasmon–pole approximation for the dynamical screening. The direct and indirect gaps were converged with the same parameters as in the previous Sections, the QP corrections being computed for the last 4 valence bands and the first 6 conduction bands (the sscGW amounts to an additional opening of the band gap by 0.22 eV with respect to the one–shot G_0W_0 calculation). The fully converged result was subsequently used to construct a k–dependent *scissor operator* (see next Section) in such a way that, when applied to the supercell, it would yield exactly the same optical absorption spectrum as the unit cell (here we neglect the changes in the GW corrections due to lattice displacements). The BSE in the supercell is solved iteratively in the `Yambo` code for the proper response function $\overline{\chi}$ using the SLEPC library[63] for the first 600 eigenvalues and eigenvectors. In the unit cell, a reasonably converged calculation of the static screening can be obtained by considering an $18 \times 18 \times 6$ k–point grid and summing 250 bands. However, the energy window close to the absorption edge is already converged with a $12 \times 12 \times 4$ sampling. As

[1]For a general discussion about how to scale some previously converged parameters in the unit cell in order to obtain an equivalent accuracy level in a supercell, see App. B.

[2]For a general discussion about how to use the scissor operator to replace a GW calculation, and about the specificities of BN systems, see Sec. D.5.

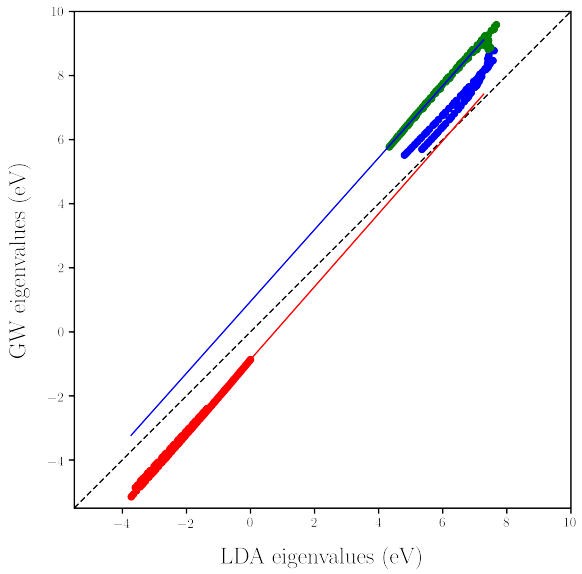


FIGURE D.3: Linear regression to obtain the scissor operator for bilayer hBN. The GW band energies throughout the IBZ are plotted against the DFT–LDA values. Valence bands are in red, conduction bands π^* and σ^* are in green and blue, respectively. The red and blue lines represent the linear fits for valence and conduction states, respectively.

the non–diagonal supercell contains 6 times the atoms of the unit cell, the convergence parameters were changed accordingly, using a $12 \times 2 \times 4$ k–point grid, and including enough states in the Bethe–Salpeter kernel to span the transition energy region relevant for the absorption edge. The static screening was computed summing $1.2 \cdot 6 \cdot 250$ bands. The factor 1.2 is a safety margin to account both for the folded bands from the zone edge and for spurious finite–q bands (see Fig. D.1(b) for a comparison of the optical absorption spectra $\varepsilon_2^{(0)}(\omega)$ between unit cell and supercell). All in all, the spectra of $\varepsilon_{\vec{q}2}^{(2)}(\omega)$ and $R_{\vec{q}}^{\mathrm{sp}}(\omega)$ shown in Figs. 3.10 and 3.11 required ~ 70 DFT–RPA–BSE calculations in the (displaced) 24–atoms non–diagonal supercell. This figure arises when atomic displacements, phonon modes and light polarisation directions are all taken into account. In order to manage the volume of data required both to pre–process and post–process the actual simulations, several Python scripts, interfaced with both QE and `Yambo`, were developed. The main objective of the scripts were: (i) the automatic generation of displaced supercells and relative input files for the two codes; (ii) the direct extraction of the exciton oscillator strengths and energies from the `Yambo` databases (ensuring the highest possible precision), followed by externally performed finite–difference calculation of the second derivatives as well as construction of the required response functions.

Concerning Eq. (3.36), its numerical verification is shown in Fig. D.2.

D.5 Scissor operator

A scissor operator is a correction applied "by hand" to the DFT Kohn–Sham energies in order to avoid computationally expensive GW calculations. This can be useful when the precise value of the QP band gap is not a quantity of interest, or when a BSE calculation is feasible but a GW one is not. The latter case typically applies to

supercell calculations. The scissor energy correction may be just a rigid shift of the conduction bands, but in order to improve accuracy such shift can be made momentum/energy dependent to better reflect the behaviour of the actual GW self–energy. If converged GW calculations are already available, for example in the unit cell, then the scissor operator can be "tailored" so that the optical spectrum resulting from a DFT–scissor–BSE calculation is exactly the same as that of a DFT–GW–BSE one. In this case, the GW energies thoroughout the BZ can be fitted with a linear function, as it is shown in Fig. D.3 for bilayer hBN: the GW band energies are plotted versus the LDA ones, with the valence bands in red and the conduction ones in blue/green. In order to obtain a correct optical absorption spectrum for BN systems, the fitting on the conduction bands should include only the π^* states (in green) and neglect the σ^* ones (blue) which have a different energy shift while not contributing to optical transitions. We call the fitted conduction and valence band slopes s_c and s_v, respectively. The DFT band gap is $\epsilon_g = \epsilon_{cB} - \epsilon_{vT}$ and the intercept of the fits gives the band gap GW correction, $\Delta_g = \Delta_{cB} + |\Delta_{vT}|$. Then, the Kohn–Sham states are corrected as follows:

$$\begin{aligned} E_{vk} &= \epsilon_{vT} - s_v(\epsilon_{vT} - \epsilon_{vk}), \\ E_{ck} &= \epsilon_{cB} + \Delta_g + s_c(\epsilon_{ck} - \epsilon_{cB}). \end{aligned} \tag{D.1}$$

If $s_c = s_v = 1$ we just obtain a rigid shift by Δ_g of the whole band structure.

Appendix E

Reference for expressions appearing in the literature

E.1 Indirect absorption in silicon: Ref. [36]

In this paper, the authors write the following dielectric function for indirect optical transitions mediated by a single phonon:

$$\varepsilon_2(\omega) = C(\omega) \sum_{ijkq\lambda} |S_1 + S_2|^2 P \delta(\epsilon_{ik-q} - \epsilon_{jk} + \hbar\omega \pm \hbar\Omega_{q\lambda}) \tag{E.1}$$

Here, $\Omega_{q\lambda}$ is the phonon frequency and ϵ_{nk} is the energy of an electronic state. The prefactor $C(\omega)$ reads $2\frac{4\pi^2 e^2}{Vm^2 N_k N_q \omega^2}$ with V being the unit cell volume. We also have

$$\begin{aligned} S_1(k,q) &= \sum_m \frac{\mathbf{e}\cdot\mathbf{v}_{imk-q}\, g^{q\lambda}_{mjk}}{\epsilon_{mk-q} - \epsilon_{ik-q} - \hbar\omega + \mathrm{i}\Gamma_{mk-q}} \\ S_2(k,q) &= \sum_m \frac{g^{q\lambda}_{imk}\, \mathbf{e}\cdot\mathbf{v}_{mjk}}{\epsilon_{mk} - \epsilon_{ik-q} \pm \hbar\Omega_{q\lambda} + \mathrm{i}\Gamma_{mk}}, \end{aligned} \tag{E.2}$$

with $\mathbf{v_{ijk}}$ and $g^{q\lambda}_{mjk}$ being the optical matrix element in the dipole approximation and the electron–phonon coupling matrix element, respectively. The light polarisation is $\mathbf{e}$ and $P = (n_{q\lambda} + \frac{1}{2} \pm \frac{1}{2})(f_{ik-q} - f_{jk})$ is the dependence on the occupation factors (Bose–Einstein distribution n for phonons, Fermi–Dirac distribution f for electrons). Finally, Γ_{mk} represents a broadening of the intermediate electronic states. Based on the result obtained in Section 3.3, we can make the following observations on Eq. (E.1):

- Eq. (E.1) is in the "Tamm–Dancoff" approximation (in this context, it means only considering the contributions of resonant transitions $i \to j$), which is only justified if $\epsilon_{jk} - \epsilon_{ik-q} > 0$ and larger than any phonon frequency. Otherwise there are more terms involved than just S_1 and S_2 (their relevance might be negligible anyway).
- The terms in Eq. (E.2) refer to the photon absorption case exclusively. However, if there is competing emission (i.e. phonon–assisted recombination) from the final state, the screening is reduced and one should consider the *net* absorption rate, which is the difference between absorption and emission rates for any transition $i \to j$ and involves more terms than just S_1 and S_2.
- The Fermi functions in the factor P seem to represent the dependence on the occupation factors in the case of the *net* absorption rate for *direct* transitions.[43]

The rest of the P–factor instead seems to refer to the *absolute* absorption rate for *indirect* transitions. Even in the approximation of $f_j << f_i$ where j is a conduction and i a valence band, the occupation dependence has a slightly different form.

- The factors Γ_{nk} are added "by hand" and don't come out of the time–dependent perturbation theory employed in order to derive this formula (they can nonetheless be computed separately with an electron–phonon self–energy calculation).

Appendix F

Magnitude of the Davydov splitting along the wave vector

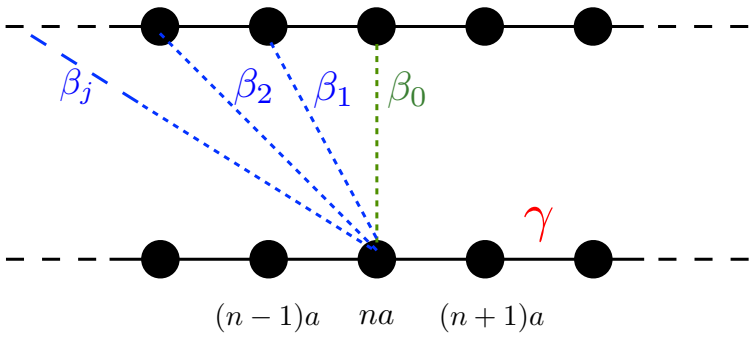


FIGURE F.1: Scheme of the tight–binding model for two interacting linear chains.

Here we construct a simple tight–binding model to reproduce the difference in the magnitude of Davydov splitting for states as one goes from Γ (where the splitting is 6 meV for phonons, 50 meV for excitons in hBN) to large momentum (where both splittings drop below 1 meV). The model consists of two identical simple linear chains with periodic boundary conditions (sketched in Fig. F.1). We consider for simplicity a basis of localised atomic orbitals $|n, X\rangle$, with na being the position along one linear chain (a is the interatomic distance), $X = A, B$ labeling the two chains and the resulting Bloch sum being $|k, X\rangle = 1/\sqrt{N}\sum_n \mathrm{e}^{\mathrm{i}kna}|n, X\rangle$. The intra–chain interaction is not important for our purposes, therefore we limit it to the first nearest neighbours with on–site terms $\langle n, X|\, H\, |n, X\rangle = 0$ and hopping elements $\langle n \pm 1, X|\, H\, |n, X\rangle = c.c. = \gamma$, giving for the two bands of the system the dispersion $E(k) = 2\gamma \cos ka$. We are more interested in the long–range inter–chain interaction, therefore we consider the hopping between sites on different chains up to infinite order: $\langle n, A|\, H\, |n, B\rangle = \beta_0$ and $\langle n \pm j, A|\, H\, |n, B\rangle = \langle n \pm j, B|\, H\, |n, A\rangle = c.c. = \beta_j$ with $j \in [1, \infty)$ (if only the β_0 contribution is retained, the two bands will be parallel and shifted by $2\beta_0$). We obtain the following tight–binding matrix,

$$\mathcal{M} = \begin{pmatrix} \langle k, A|\, H\, |k, A\rangle & \langle k, A|\, H\, |k, B\rangle \\ \langle k, A|\, H\, |k, B\rangle & \langle k, A|\, H\, |k, A\rangle \end{pmatrix}, \tag{F.1}$$

whose elements can be easily computed to give the energies $E_\pm(k) = 2\gamma \cos ka \pm$

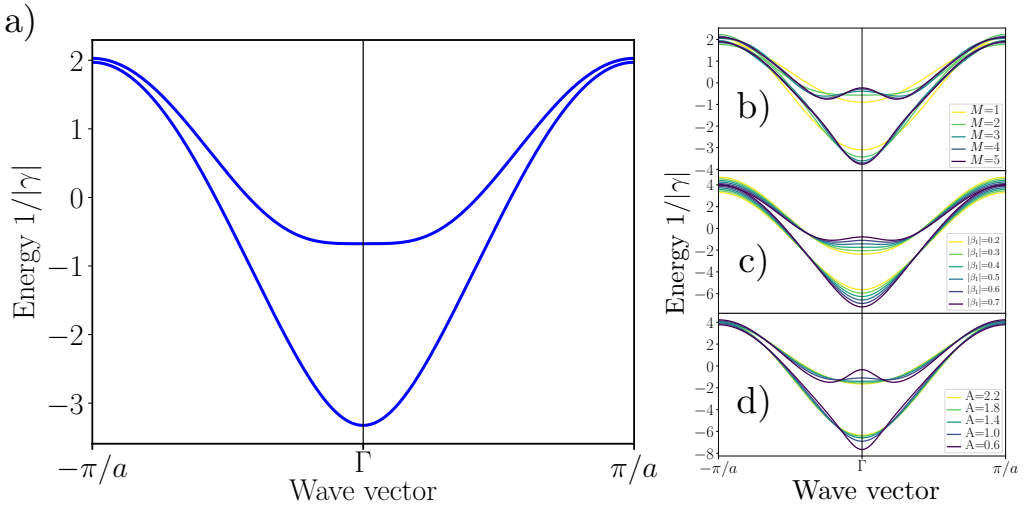


FIGURE F.2: Results of the double linear chain model in the one–dimensional BZ. For all plots $\gamma = -2$, $\beta_0 = -1$, $a = 1$. a) Davydov splitting with $C = 1$, $\beta_1 = -0.6$ and M converged. b) Plot for increasing values of M (darker curve means larger M) with $C = 0.6$ and $\beta_1 = -0.6$. c) Increasing values of $|\beta_1|$ (darker curve means larger $|\beta_1|$) with $C = 1$ and M converged. d) Decreasing values of C (darker curve means smaller C) with $\beta_1 = -0.6$ and M converged.

$[\beta_0 + 2\sum_{m=1}^{\infty}\beta_m \cos mka]$. In order to model the spatial decay of the inter–chain interaction we define $\beta_m \equiv \beta_1 e^{-C(m-1)}$, which gives the β_m coefficients up to infinite order from the parameters β_1 and C (the latter parameter can be made dependent on the inter–chain distance). This allows us to write the analytical results

$$E_{\pm}(k) = 2\gamma\cos ka \pm \left[\beta_0 + \beta_1\frac{-1+e^{C}\cos ka}{\cosh C - \cos ka}\right],$$
$$E_{\pm}^{M}(k) = 2\gamma\cos ka \pm \left[\beta_0 + \beta_1\frac{-1+e^{C}\cos ka + e^{-MC}(\cos Mka - e^{C}\cos(M+1)ka)}{\cosh C - \cos ka}\right], \tag{F.2}$$

where the first expression stands for infinite inter–chain neighbors and the second one for the partial sum to M. Our results, reproducing the strong reduction in the splitting away from Γ, is displayed in Fig. F.2(a). We make the following observations: (i) the inclusion of β_1, which has the periodicity of the BZ, causes the bands to come closer together toward the zone edges; (ii) to reproduce the strong bending observed for the Davydov splitting (see e.g. the phonon dispersion of hBN or MoS_2),[1] it is necessary to have at least a second non–negligible oscillatory contribution (e.g. β_2); this means that (iii) the parameters C and β_1, controlling the long–range decay and the strength of the inter–chain interaction, should take values such that the sum is not converged after just one term and $\beta_1/\beta_0 \gtrsim 0.5$ (Fig. F.2 (b) to (d)).

[1] See http://henriquemiranda.github.io/phononwebsite/phonon.html

Appendix G

Temperature dependence in the finite–difference spectra

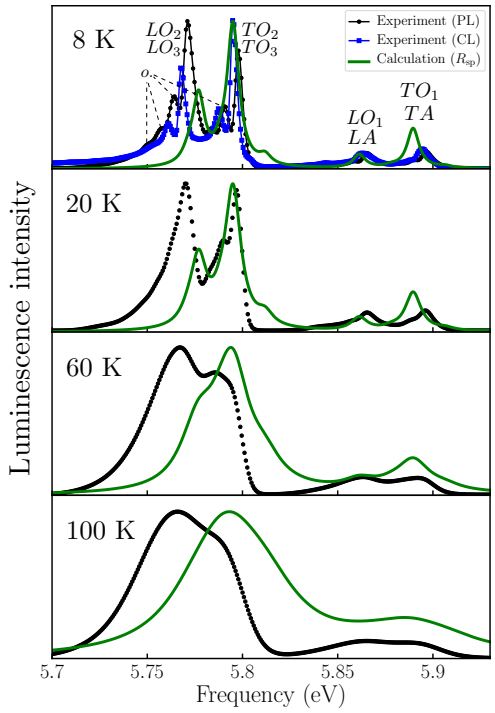


FIGURE G.1: Temperature dependent spectra from Eq. (3.41). The computed spontaneous emission rate (see Sec. 3.4 of the main text) is shown with a green line for various temperatures. The main phonon modes responsible for the peaks are labeled, as well as the overtones ("o."). The peak broadenings as a function of temperature, as well as the effective "excitonic" temperatures, are discussed in the text. Black dots: experimental PL emission spectrum[83] (uncorrected for setup response). Blue squares: experimental CL emission spectrum.[84]

In our calculations from Sec. 3.4, the temperature–dependent exciton lifetime, which is related to the imaginary part of the exciton–phonon self–energy and inversely proportional to the line broadenings η, remains an empirical parameter. As the imaginary part of our response functions intrinsically gives a Lorentzian shape for single peaks ($\mathrm{Im}[\omega - E - i\eta]^{-1}$), we focus on the range of temperatures in which the experimental phonon–assisted peaks can also be reasonably described with a Lorentzian broadening (from 0 to 100 K).[83] We use a linear model where the line

broadening η is given by[109]

$$\eta = \Gamma_0 + aT + bN_{BE}(T), \tag{G.1}$$

where $N_{BE} = [e^{E_O/kT} - 1]^{-1}$ is the Bose-Einstein distribution. The values of the parameters are taken from the experimental fit in Ref. [83]: $\Gamma_0 = 3$ meV, $a = 0.1$ meV/K, $b = 150$ meV, and $E_O = 25$ meV. For completeness, in Fig. G.1 we provide a version of Fig. 3.11(b) of the main text with several temperatures (T= 8, 20, 60 and 100 K). The effective "excitonic" temperatures $T_{\rm exc}$ entering the Boltzmann factor of Eq. (3.41) of the main text are taken from the data points in Fig. 1c of Ref. [19]. At the lowest temperature ($8 - 10$ K, top frame), both the experimental spectra obtained from photoluminescence (PL, black dots, from Ref. [83]) and cathodoluminescence (CL, blue squares, from Ref. [84]) are compared with the computed one (green line). The temperature–dependent data are only available for the PL spectrum. Notice however that, contrary to the CL spectrum, the PL one is uncorrected for the response of the luminescence experimental setup, giving rise to wrong relative peak intensities in the low–energy structure.

Appendix H

The many–body case for phonon–mediated absorption

We have seen with Eq. (1.76) that the response function χ can be generated by the derivatives of the interacting electron density ρ with respect to an external perturbation ϕ_{ext}. We will do this now, only graphically, by considering an electron density where both the electron–electron and the electron–phonon interactions are present at first order. The resulting terms are displayed in Fig. H.1. The zero–order contribution (a) gives the non–interacting eh pairs described by χ^0. To fix the ideas, we can imagine that all the electronic lines in the upper halves of the diagrams refer to conduction bands c, and all those in the lower halves to valence bands v.[1] The first–order contribution in the electron–electron interaction (b) gives rise to the BSE with the kernel of Eq. (1.78). The first–order contribution in the electron–phonon interaction (c) produces terms from (1) to (4) in Fig. H.1. Here, the (c1) and (c2) contributions would be normally already included in the electron–phonon single–particle QP corrections if the single–particle GF is $\overline{G}$: this is why analogous terms are not considered in the electron–electron case. However, the *dynamical* part of the electron–phonon interaction in this diagrams contributes to the satellite structure, therefore we explicitly list these terms here. Contribution (c3) is a new one, that we will call interference term.[2] We will neglect the last phononic term (c4) for two reasons: (i) a contribution of this type is already approximately incorporated in the electron–phonon coupling matrix elements $g^{\mu q}_{nmk}$ if they are computed within DFPT (as mentioned in Sec. 1.6.5; more information in Ref. [49]), and (ii) these terms are zero in the TDA, since for phonons it implies that $g^{\mu q}_{vck} = 0$. Finally, at first order in both the electron–electron and the electron–phonon interactions, we obtain the terms in Fig. H.1(d): these are be the main new contributions that will describe the *exciton*–phonon coupling.

The exciton–phonon scattering cross–section is described by the following processes. First, an electron–hole pair is optically created; then the electron or the hole is scattered by a phonon, changing its energy and momentum; now the scattered particle interacts with the other one via the electronic kernel $K = W - 2v$; later the lost energy and momentum are recovered by the eh pair via another phonon scattering, and ultimately the excitation annihilates emitting optical light. A first–order dynamical correction can then be obtained for the full static BSE (sBSE) by going to all orders in the electronic part: let us imagine inserting an infinite number of kernels $K = W - 2v$ in between the start and end times of the phonon interaction, obtaining the expression in Fig. H.1(e). The kernels that fall outside the phonon

[1] Fixing this actually implies taking the Tamm-Dancoff approximation.
[2] In Ref. [107] this is called "phonon exchange diagram".

interaction part reconstruct the response function χ from the sBSE. This final expression contains the (a), (b), (c) and (d) terms. It is clear that this expression represents a first–order dynamical correction to the sBSE.

Note that a similar development could be made by considering the dynamical part of the electron–electron interaction (and indeed this is the goal in Ref. [38]) in order to describe plasmonic satellites, however this is not our interest here. Therefore, we will maintain the static approximation for the BSE electronic kernel: the full kernel is then made dynamical just by the presence of the electron–phonon interaction mediated by the phonon propagator $D(\omega)$.

The full expression for (e) is discussed in Secs. 4.2.1 and 4.2.2, and explicitly derived in App. I.

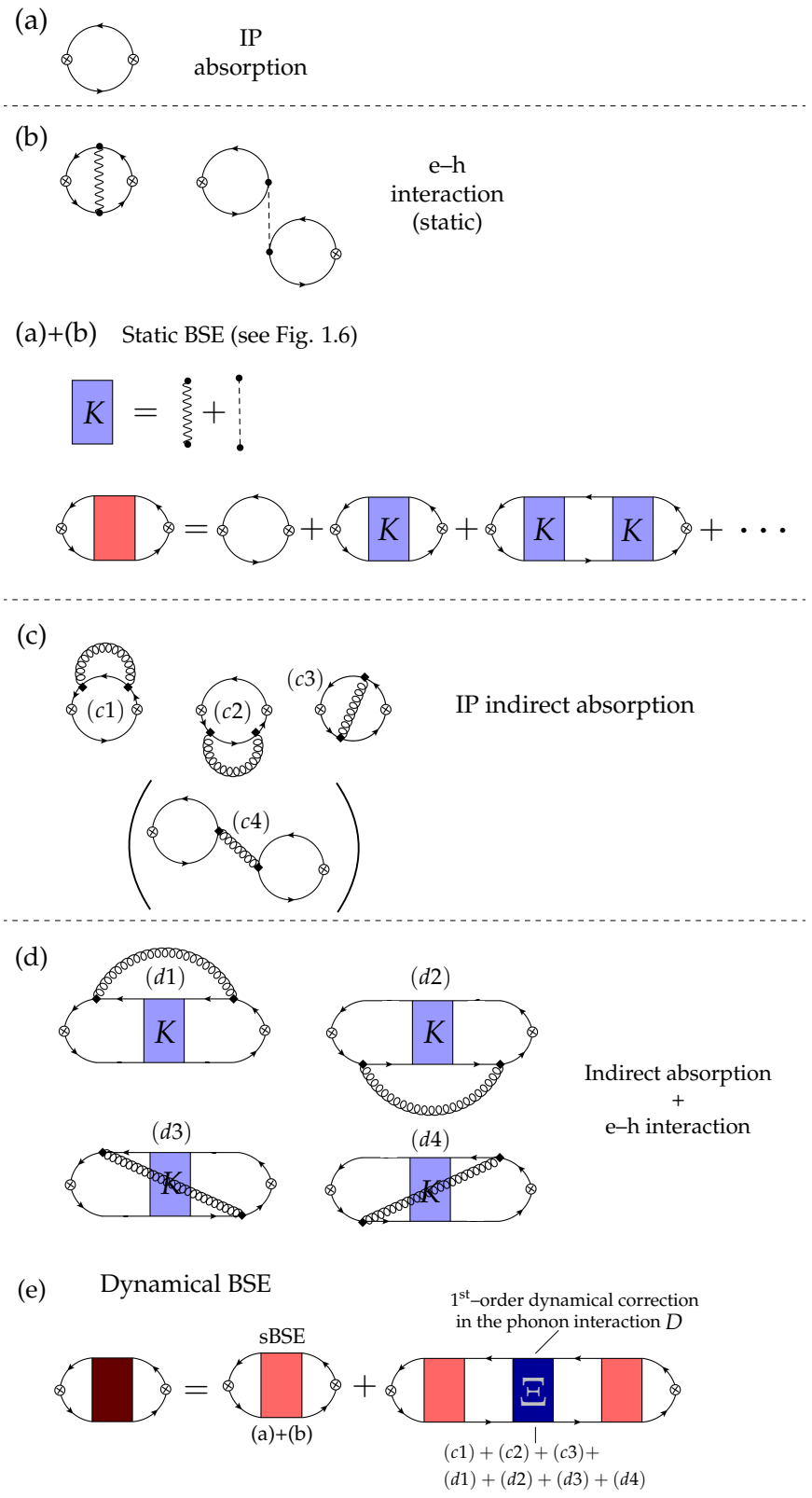


FIGURE H.1: Response function χ at zero and first orders in the electron–electron and ep interaction. (a) Non–interacting eh pairs, Eq. (1.44). (b) Static, electronic eh interaction at first order (i.e. static BSE, sBSE, Eq. (1.87)). We are not interested in its dynamical part here, which is neglected. Note that the single–particle GF $\overline{G}$ (solid lines) already includes quasiparticle corrections, Eq. (1.57). (c) *Dynamical* ep interaction at first order. The phonon QP correction is either neglected or will be summed at a later step to all orders. (d) Both dynamical ep and static eh interaction present at first order (the static phonon contribution to the eh interaction is either neglected or included in the kernel K). (e) By keeping the ep interaction at first order [(c) and (d)] and summing the sBSE kernel at all orders [(a)+(b)], a correction to the sBSE at first order in the ep interaction can be obtained (Eq. (4.21) and Ref. [113]).

Appendix I

Reference for the derivation of the exciton–phonon self–energy

In this Appendix we sketch the derivation of the exciton–phonon self–energy Eq. (4.24) and Fig. H.1(e). We follow Ref. [38] and provide several useful identities for reference. Our conventions for the energy and momentum flows are shown in Fig. I.1(b), as well as a scheme of the notation we employed. These choices will lead to the indexing of the exciton–phonon coupling (ExPC) used in Eq. (4.25), which is particularly favourable in the case of finite–momentum matrix elements computed with QE and `Yambo`. The mathematical expressions will soon become large and complicated; however the schematic representation in Fig. I.2 provides a reference that emphasises the physical meaning of the processes considered.

We consider the two–particle correlation function in the excitation basis, $L_{\mathcal{K}\mathcal{K}'}(t_{13})$ (t_1 and t_3 being the initial and final times, respectively). Let us recall that when the four–times dependence of L is collapsed to just two times, in order to describe simultaneous electron–hole pair generation, we just have $\chi(t_{13}) = -\mathrm{i}L(t_{13})$. Now, $L^{(0)}_{\mathcal{K}\mathcal{K}'}(t_{13})$ is the correlation function obtained in the BSE with a static electronic kernel (sBSE, Sec. 1.5):

$$L^{(0)} = L^{(0)}_0 + L^{(0)}_0 K L^{(0)}, \tag{I.1}$$

where $L^{(0)}_0$ is the independent–particle electron–hole propagation, corresponding to χ^0 elsewhere in the thesis. Notice that in this case we can pass from the excitation to the excitonic basis, where $L^{(0)}$ is diagonal, as

$$L^{(0)}_{\mathcal{K}\mathcal{K}'}(t_{13}) = \sum_\alpha \mathcal{A}^{\mathcal{K}}_\alpha \left[\mathcal{A}^{\mathcal{K}'}_\alpha\right]^* L^{(0)}_\alpha(t_{13}) = \sum_\alpha \mathcal{A}^{\mathcal{K}}_\alpha \left[\mathcal{A}^{\mathcal{K}'}_\alpha\right]^* \theta(t_3 - t_1)\mathrm{e}^{-\mathrm{i}E_\alpha(t_3-t_1)}. \tag{I.2}$$

For now, our $L^{(0)}$ can be equivalently the full one or the one without the long–range components of the Coulomb interaction (elsewhere denoted as $\overline{L}$ or $\overline{\chi}$), as can its eigenvector components $\mathcal{A}$ and exciton energies E_α. We denote as $L^{\mathrm{dyn}}_{\mathcal{K}\mathcal{K}'}(t_{13})$ the correlation function given by the first–order correction to the static kernel in the dynamical phonon propagator $D_{\mu q}$, Eq. (I.24), as in Eq. (4.21). Then, the full correlation function up to first order in the dynamical effects is given by

$$L^{(1)}_{\mathcal{K}\mathcal{K}'}(t_{13}) = L^{(0)}_{\mathcal{K}\mathcal{K}'}(t_{13}) + L^{\mathrm{dyn}}_{\mathcal{K}\mathcal{K}'}(t_{13}). \tag{I.3}$$

I.1 Several identities used in the derivation

T–matrix. We can rework the sBSE in order to cast the electron–hole interaction completely inside a new quantity, called the T–matrix:

$$L^{(0)} = L_0^{(0)} + L_0^{(0)} T L_0^{(0)}. \tag{I.4}$$

This leads, after comparison with the inverted Eq. (I.1), to the following expression for T:

$$T = K + K L^{(0)} K \tag{I.5}$$

(a graphical representation of this equation is shown in Fig. I.1(a)). The usefulness of T consists in the fact that it allows us to rewrite the dynamical correction in Eq. (I.3) explicitly in terms of a modified dynamical kernel Ξ which, as we will see, is in turn written in terms of the T–matrix:

$$L^{(1)} = L^{(0)} + L^{\mathrm{dyn}} = L^{(0)} + L^{(0)} \Xi[T] L^{(0)}. \tag{I.6}$$

The expression of T entering the calculations, with explicit transition indices and time variables, is:

$$\begin{aligned} T_{\mathcal{K}_1\mathcal{K}_2}(t_{1423}) &= \mathrm{i} K_{\mathcal{K}_1\mathcal{K}_2} \delta(t_{13})\delta(t_{34})\delta(t_{12}) - \sum_{\mathcal{K}_3\mathcal{K}_4} K_{\mathcal{K}_1\mathcal{K}_3} L^{(0)}_{\mathcal{K}_3\mathcal{K}_4}(t_{31}) K_{\mathcal{K}_4\mathcal{K}_2} \delta(t_{12})\delta(t_{34}) \\ &= \mathrm{i} K_{\mathcal{K}_1\mathcal{K}_2} \delta(t_{13})\delta(t_{34})\delta(t_{12}) - S_{\mathcal{K}_1\mathcal{K}_2}(t_{31})\delta(t_{12})\delta(t_{34}). \end{aligned} \tag{I.7}$$

Replacing the static electronic kernel with a difference of single–particle transitions and exciton energies. By reworking the eigenvalue equation for the excitonic Hamiltonian, Eq. (1.84), we obtain the following identities:

$$\begin{aligned} \sum_{\mathcal{K}'} K_{\mathcal{K}\mathcal{K}'} \mathcal{A}^{\mathcal{K}'}_\alpha &= [\epsilon_{ck} - \epsilon_{vk} - E_\alpha] \mathcal{A}^{\mathcal{K}}_\alpha, \\ \sum_{\mathcal{K}} K_{\mathcal{K}\mathcal{K}'} \left[\mathcal{A}^{\mathcal{K}}_\alpha\right]^* &= [\epsilon_{c'k'} - \epsilon_{v'k'} - E_\alpha] \left[\mathcal{A}^{\mathcal{K}'}_\alpha\right]^*. \end{aligned} \tag{I.8}$$

Notice that here we indicate the single–particle energy associated to $\overline{G}_{nk}$ as ϵ_{nk} and not as E_{nk} like in the main text: this is to avoid confusion with the exciton energies. By using Eqs. (I.2) and (I.8) we can rewrite the S function in the T–matrix expression, Eq. (I.7), in terms of single–particle energies and the excitonic quantities:

$$S_{\mathcal{K}_1\mathcal{K}_2}(t_{31}) = \sum_\alpha [\epsilon_{c_1k_1} - \epsilon_{v_1k_1} - E_\alpha][\epsilon_{c_2k_2} - \epsilon_{v_2k_2} - E_\alpha] \mathcal{A}^{\mathcal{K}_1}_\alpha \left[\mathcal{A}^{\mathcal{K}_2}_\alpha\right]^* L^{(0)}_\alpha(t_{31}). \tag{I.9}$$

Expressing the three–times $L^{(0)}$ in terms of time–ordered, two–times $L^{(0)}$. As can be seen from the diagrams in Fig. I.2, L^{dyn} contains, to the left and to the right, two sBSE correlation functions depending on three time variables: we have $L^{(0)}_{(cvk)(c_1v_1k_1)}(t_{12'11'})$ and $L^{(0)}_{(c_2v_2k_2)(c'v'k')}(t_{3'34'3})$, respectively. In order to evaluate the time integrals involving these functions, we notice that they can be reduced to two–times L^0 functions

via the following identity:

$$\begin{aligned} L^{(0)}_{(cvk)(c_1v_1k_1)}(t_{12'11'}) &= \theta(t_{2'}-t_{1'})\mathrm{e}^{\mathrm{i}\epsilon_{c_1k_1}(t_{1'}-t_{2'})}L^{(0)}_{(cvk)(c_1v_1k_1)}(t_{11'}) \\ &+ \theta(t_{1'}-t_{2'})\mathrm{e}^{\mathrm{i}\epsilon_{v_1k_1}(t_{1'}-t_{2'})}L^{(0)}_{(cvk)(c_1v_1k_1)}(t_{12'}), \\ L^{(0)}_{(c_2v_2k_2)(c'v'k')}(t_{3'34'3}) &= \theta(t_{3'}-t_{4'})\mathrm{e}^{\mathrm{i}\epsilon_{c_2k_2}(t_{4'}-t_{3'})}L^{(0)}_{(c_2v_2k_2)(c'v'k')}(t_{3'3}) \\ &+ \theta(t_{4'}-t_{3'})\mathrm{e}^{\mathrm{i}\epsilon_{v_2k_2}(t_{4'}-t_{3'})}L^{(0)}_{(c_2v_2k_2)(c'v'k')}(t_{4'3}) \end{aligned} \tag{I.10}$$

(this can be proven relatively easily for $L^{(0)}_0 = \overline{G}\overline{G}$; the proof for $L^{(0)}$ is more cumbersome and is done using the T–matrix expression, Eq. (I.4)).

I.2 Specific notation for finite momentum quantities

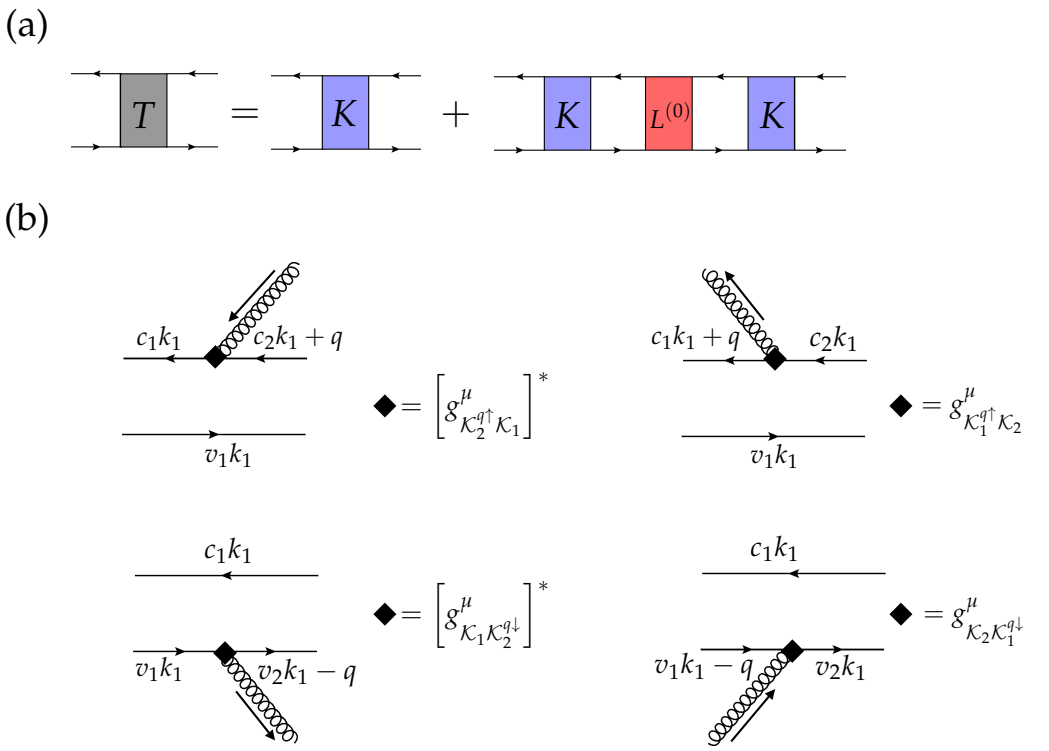


FIGURE I.1: (a) Graphical expression for the T–matrix in terms of the sBSE kernel K and correlation function $L^{(0)}$. (b) Choice of notation and momentum flows for electron–phonon couplings with simultaneous electron–hole propagation (see also text).

Valence bands are denoted with the letter v and conduction ones with c. If the phonon scattering happens on the valence band, we denote the momentum transfer as $q\downarrow$, while if it happens on the conduction band we write $q\uparrow$. The momentum flows in both cases are chosen as shown in Fig. I.1(b), so that in terms of transition indices $\mathcal{K}$ we may write

$$\begin{aligned} \mathcal{K}_1^{q\downarrow} &= (c_1k_1)(v_1k_1-q), \\ \mathcal{K}_1^{q\uparrow} &= (c_1k_1+q)(v_1k_1). \end{aligned} \tag{I.11}$$

In particular, when expressing the electron–phonon coupling matrix element $g^{q\mu}_{nk_n,mk_m}$ using transition indices, we write

$$\begin{aligned}
\left[g^{\mu}_{\mathcal{K}_1\mathcal{K}_2^{q\downarrow}}\right]^* &= \left[g^{q\mu}_{v_1k_1,v_2k_1-q}\right]^* \delta_{c_1c_2}\delta_{k_1k_2}, \\
g^{\mu}_{\mathcal{K}_2\mathcal{K}_1^{q\downarrow}} &= g^{q\mu}_{v_2k_1,v_1k_1-q}\delta_{c_1c_2}\delta_{k_1k_2}, \\
g^{\mu}_{\mathcal{K}_1^{q\uparrow}\mathcal{K}_2} &= g^{q\mu}_{c_1k_1+q,c_2k_1}\delta_{v_1v_2}\delta_{k_1k_2}, \\
\left[g^{\mu}_{\mathcal{K}_2^{q\uparrow}\mathcal{K}_1}\right]^* &= \left[g^{q\mu}_{c_2k_1+q,c_1k_1}\right]^* \delta_{v_1v_2}\delta_{k_1k_2}.
\end{aligned} \tag{I.12}$$

I.3 Calculation of the first–order dynamical correction to the response function

With reference to Fig. H.1, L^{dyn} is given by the sum of all the contributions previously discussed,

$$L^{\mathrm{dyn}}_{\mathcal{K}\mathcal{K}'}(t_{13}) = L^{[c1]+[c2]+[c3]+[e1]+[e2]+[e3]+[e4]}_{\mathcal{K}\mathcal{K}'}(t_{13}). \tag{I.13}$$

The terms $[e1]$ to $[e4]$ contain the T–matrix, i.e. the full summation of the kernel K to infinite order, and at first order in K reduce to the $[d1]$–$[d4]$ terms in Fig. H.1.

Now, we will proceed to write explicitly the $[e2]$ term to emphasise some crucial aspects of the derivation. The remaining terms can be obtained in a similar way: all the $[e]$ terms are explicitly shown diagrammatically in Fig. I.2, while the $[c]$ terms are sketched in Fig. I.3. We have

$$L^{[e2]}_{\mathcal{K}\mathcal{K}'}(t_{13}) = \sum_{\mathcal{K}_1,\mathcal{K}_2}\int \mathrm{d}t_{1'2'3'4'} L^{(0)}_{\mathcal{K}\mathcal{K}_1}(t_{12'11'})\Xi^{[e2]}_{\mathcal{K}_1\mathcal{K}_2}(t_{1'4'2'3'})L^{(0)}_{\mathcal{K}_2\mathcal{K}'}(t_{3'34'3}). \tag{I.14}$$

Here, the two $L^{(0)}$ functions refer to the propagation of the exciton via a static electronic kernel before and after the phonon scattering event. The kernel $\Xi^{[e2]}$, which contains the dynamical phonon scattering at first order (and the static electronic kernel at all orders), refers in particular to the case of the hole component of the exciton being scattered. The $[e1]$ term refers to the conduction electron being scattered, and $[e3]$, $[e4]$ are the interference terms.

We can write $\Xi^{[e2]}$ explicitly by looking at the corresponding diagram in Fig. I.2:

$$\begin{aligned}
\Xi^{[e2]}_{\mathcal{K}_1\mathcal{K}_2}(t_{1'4'2'3'}) = \sum_{\mathcal{K}_1\mathcal{K}_2}\sum_{\mu q}&\left[g^{\mu}_{\mathcal{K}_1\mathcal{K}_3^{q\downarrow}}\right]^* g^{\mu}_{\mathcal{K}_2\mathcal{K}_4^{q\downarrow}}\mathrm{i}D_{\mu q}(t_{1'3'}) \\
&\int \mathrm{d}t_{1''t_{3''}}\overline{G}_{\mathcal{K}_3^{q\downarrow}}(t_{1'1''})T_{\mathcal{K}_3^{q\downarrow}\mathcal{K}_4^{q\downarrow}}(t_{1''4'2'3''})\overline{G}_{\mathcal{K}_4^{q\downarrow}}(t_{3''3'}).
\end{aligned} \tag{I.15}$$

Here, we already explicitly introduced the T–matrix and the single–particle GFs $\overline{G}$ are given by Eq. (1.57). They may contain GW and/or the electron–phonon quasiparticle corrections. The first task is the evaluation of the time integrals in Eq. (I.15) for all the diagrammatic contributions: this will lead us to the exciton–phonon self–energy. Secondly, thanks to Eq. (I.14), we will obtain the dynamical correction to $L^{(0)}$ as in Eq. (I.3). This will in turn provide us with an expression for the phonon–assisted absorption spectrum as $\chi^{(1)} = -\mathrm{i}L^{(1)}$, $\mathrm{Im}\chi^{(1)}(\omega) \to \varepsilon_2(\omega)$.

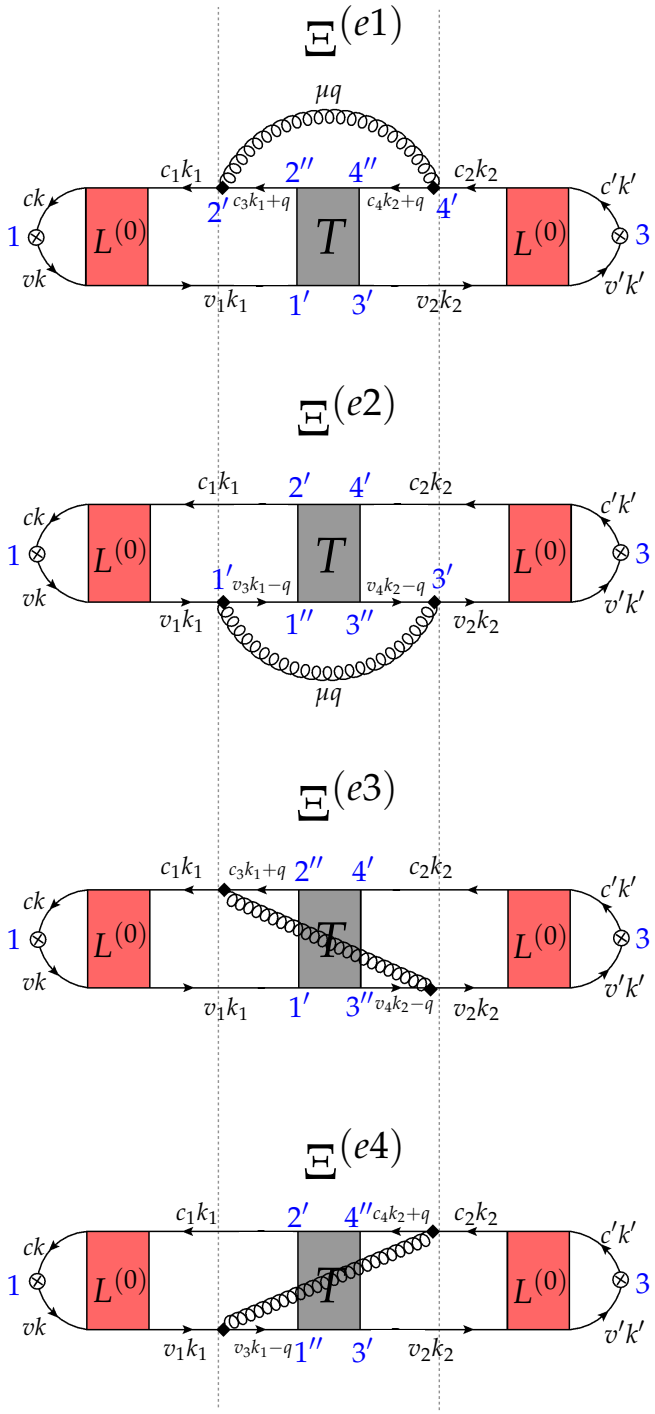


FIGURE I.2: Diagrammatic representation of the terms entering the first–order dynamical correction to the two–particle response function, L^{dyn}. The time labels are written in blue. Only the terms corresponding to Fig. H.1(d) are shown (with the T–matrix replacing the sBSE kernel: therefore they are denoted with the letter (e).)

The evaluation proceeds as follows. First, we use Eq. (I.10) to break down Eq. (I.14) in two integrals involving only two–times $L^{(0)}$ functions with different internal time orderings. Second, we evaluate the internal time integrals appearing in Eq. (I.15) using Eq. (I.7) for the T–matrix. We also use Eq. (I.8), along with the completeness relation $\delta_{\mathcal{K}\mathcal{K}'} = \sum_\alpha A_\alpha^{\mathcal{K}} \left[A_\alpha^{\mathcal{K}'}\right]^*$, to simplify the resulting expressions in terms

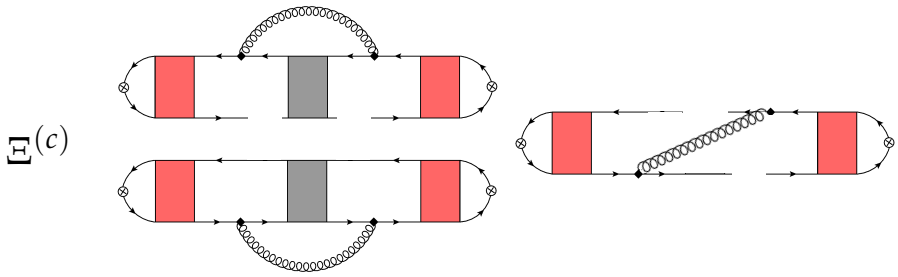


FIGURE I.3: Graphical representation of the terms corresponding to Fig. H.1(c), where the T–matrix is not present. The sum of all the $\Xi^{(e)}$ and $\Xi^{(c)}$ terms adds up to the full dynamical kernel Ξ shown in Fig. H.1(e) and Eqs. (I.6) and (4.21).

of excitonic quantities. In the end, the calculation for $[e2]$ produces two terms: one, only involving the phonon propagator D, is exactly cancelled by the result for $[c2]$, while the other remains. Therefore, the resulting contribution from $[e2]$ reads

$$L^{[e2]}_{\mathcal{K}\mathcal{K}'}(t_{13}) = -\mathrm{i}\sum_{\mathcal{K}_1,\mathcal{K}_2}\int_{t_1}^{t_3}\mathrm{d}t_{1'}\int_{t_{1'}}^{t_3}t_{3'}\,L^{(0)}_{\mathcal{K}\mathcal{K}_1}(t_{11'})D_{\mu q}(t_{1'3'}) \\ \sum_{q\mu}\left[\sum_{\mathcal{K}_3\mathcal{K}_4}\sum_{\beta}\left(g^{\mu}_{\mathcal{K}_1\mathcal{K}_3^{q\downarrow}}\right)^{*}\mathcal{A}^{\mathcal{K}_3^{q\downarrow}}_{\beta}\,g^{\mu}_{\mathcal{K}_2\mathcal{K}_4^{q\downarrow}}\left(\mathcal{A}^{\mathcal{K}_4^{q\downarrow}}_{\beta}\right)^{*}\right]\mathrm{e}^{-\mathrm{i}E_{\beta q}(t_{3'}-t_{1'})}L^{(0)}_{\mathcal{K}_2\mathcal{K}'}(t_{3'3}). \tag{I.16}$$

The quantity between blue square brackets represents a contribution to the *exciton–phonon coupling* coming from this term. The other $[e]$ terms are evaluated in a similar way and after cancellations with the corresponding $[c]$ terms produce analogous results.

In particular, we can now write the complete L^{dyn} in the exciton basis, obtaining the expression

$$\begin{aligned} L^{\mathrm{dyn}}_{\alpha\alpha'}(t_{13}) &= \sum_{\mathcal{K}\mathcal{K}'}\left(\mathcal{A}^{\mathcal{K}}_{\alpha}\right)^{*}\mathcal{A}^{\mathcal{K}'}_{\alpha'}L^{\mathrm{dyn}}_{\mathcal{K}\mathcal{K}'}(t_{13}) \\ &= \int_{t_1}^{t_3}\mathrm{d}t_{1'}\int_{t_{1'}}^{t_3}t_{3'}\,L^{(0)}_{\alpha}(t_{11'})\Pi^{\mathrm{exc}}_{\alpha\alpha'}(t_{1'3'})L^{(0)}_{\alpha'}(t_{3'3}), \end{aligned} \tag{I.17}$$

where all the complicated ExcP interaction has been hidden inside an *exciton–phonon self–energy* $\Pi^{\mathrm{exc}}_{\alpha\alpha'}$. Notice that in the case of optical absorption, the excitons labeled as α represent the optical excitations induced by the external laser field: therefore, α excitons always have $q = 0$. The internal sum over excitons β, instead, includes every possible exciton level $E_{\beta q}$ that can be connected to $E_{\alpha 0} \equiv E_{\alpha}$ by emitting one phonon with frequency $\Omega_{q\mu}$. The self–energy can be written in a "GW"–type form as

$$\Pi^{\mathrm{exc}}_{\alpha\alpha'}(t_{1'3'}) = \sum_{\mu\beta q}\mathcal{D}^{\mu}_{(\alpha,\beta q)(\beta q,\alpha')}(t_{1'3'})L^{(0)}_{\beta q}(t_{1'3'}). \tag{I.18}$$

Here, the exciton–phonon "interaction" $\mathcal{D}$ contains the dynamical phonon propagator D and the quantities between blue square brackets in Eq. (I.16) and analogous terms, which represent the coupling.

I.4 Form of the exciton–phonon coupling

We can write the interaction $\mathcal{D}$ by making the phonon propagator appear explicitly:

$$\begin{aligned}
\mathcal{D}^{\mu}_{(\alpha,\beta q)(\beta q,\alpha')}(t_{1'3'}) = &-\mathrm{i}D_{\mu q}(t_{3'1'})\sum_{\mathcal{K}_1\mathcal{K}_2}\left(\mathcal{A}^{\mathcal{K}_1}_{\alpha}\right)^{*}\Xi^{q\mu,\uparrow\uparrow}_{\mathcal{K}_1\mathcal{K}_2}\mathcal{A}^{\mathcal{K}_2}_{\alpha'} \quad [e1]\\
&-\mathrm{i}D_{\mu q}(t_{1'3'})\sum_{\mathcal{K}_1\mathcal{K}_2}\left(\mathcal{A}^{\mathcal{K}_1}_{\alpha}\right)^{*}\Xi^{q\mu,\downarrow\downarrow}_{\mathcal{K}_1\mathcal{K}_2}\mathcal{A}^{\mathcal{K}_2}_{\alpha'} \quad [e2]\\
&+\mathrm{i}D_{\mu q}(t_{3'1'})\sum_{\mathcal{K}_1\mathcal{K}_2}\left(\mathcal{A}^{\mathcal{K}_1}_{\alpha}\right)^{*}\Xi^{q\mu,\uparrow\downarrow}_{\mathcal{K}_1\mathcal{K}_2}\mathcal{A}^{\mathcal{K}_2}_{\alpha'} \quad [e3]\\
&+\mathrm{i}D_{\mu q}(t_{1'3'})\sum_{\mathcal{K}_1\mathcal{K}_2}\left(\mathcal{A}^{\mathcal{K}_1}_{\alpha}\right)^{*}\Xi^{q\mu,\downarrow\uparrow}_{\mathcal{K}_1\mathcal{K}_2}\mathcal{A}^{\mathcal{K}_2}_{\alpha'}. \quad [e4]
\end{aligned} \tag{I.19}$$

The Ξ–factors are the ExcP coupling contributions coming from the various terms in Fig. I.2, the one on the second line being given by the quantity in blue square brackets in Eq. (I.16). Their explicit expression reads:

$$\begin{aligned}
\Xi^{q\mu,\uparrow\uparrow}_{\mathcal{K}_1\mathcal{K}_2} &= \sum_{\mathcal{K}_3\mathcal{K}_4}\sum_{\beta}\left(g^{\mu}_{\mathcal{K}_3^{q\uparrow}\mathcal{K}_1}\right)^{*}\mathcal{A}^{\mathcal{K}_3^{q\uparrow}}_{\beta}g^{\mu}_{\mathcal{K}_4^{q\uparrow}\mathcal{K}_2}\left(\mathcal{A}^{\mathcal{K}_4^{q\uparrow}}_{\beta}\right)^{*},\\
\Xi^{q\mu,\downarrow\downarrow}_{\mathcal{K}_1\mathcal{K}_2} &= \sum_{\mathcal{K}_3\mathcal{K}_4}\sum_{\beta}\left(g^{\mu}_{\mathcal{K}_1\mathcal{K}_3^{q\downarrow}}\right)^{*}\mathcal{A}^{\mathcal{K}_3^{q\downarrow}}_{\beta}g^{\mu}_{\mathcal{K}_2\mathcal{K}_4^{q\downarrow}}\left(\mathcal{A}^{\mathcal{K}_4^{q\downarrow}}_{\beta}\right)^{*},\\
\Xi^{q\mu,\uparrow\downarrow}_{\mathcal{K}_1\mathcal{K}_2} &= \sum_{\mathcal{K}_3\mathcal{K}_4}\sum_{\beta}\left(g^{\mu}_{\mathcal{K}_3^{q\uparrow}\mathcal{K}_1}\right)^{*}\mathcal{A}^{\mathcal{K}_3^{q\uparrow}}_{\beta}g^{\mu}_{\mathcal{K}_2\mathcal{K}_4^{q\downarrow}}\left(\mathcal{A}^{\mathcal{K}_4^{q\downarrow}}_{\beta}\right)^{*},\\
\Xi^{q\mu,\downarrow\uparrow}_{\mathcal{K}_1\mathcal{K}_2} &= \sum_{\mathcal{K}_3\mathcal{K}_4}\sum_{\beta}\left(g^{\mu}_{\mathcal{K}_1\mathcal{K}_3^{q\downarrow}}\right)^{*}\mathcal{A}^{\mathcal{K}_3^{q\downarrow}}_{\beta}g^{\mu}_{\mathcal{K}_4^{q\uparrow}\mathcal{K}_2}\left(\mathcal{A}^{\mathcal{K}_4^{q\uparrow}}_{\beta}\right)^{*}.
\end{aligned} \tag{I.20}$$

This can be written in a considerably more compact form with the additional definitions

$$\begin{aligned}
g^{\mu q\downarrow}_{\beta\alpha} &\equiv \sum_{\mathcal{K}_1\mathcal{K}_2}\left(\mathcal{A}^{\mathcal{K}_2^{q\downarrow}}_{\beta}\right)^{*}\mathcal{A}^{\mathcal{K}_1}_{\alpha}g^{\mu}_{\mathcal{K}_1\mathcal{K}_2^{q\downarrow}},\\
g^{\mu q\uparrow}_{\beta\alpha} &\equiv \sum_{\mathcal{K}_1\mathcal{K}_2}\left(\mathcal{A}^{\mathcal{K}_2^{q\uparrow}}_{\beta}\right)^{*}\mathcal{A}^{\mathcal{K}_1}_{\alpha}g^{\mu}_{\mathcal{K}_2^{q\uparrow}\mathcal{K}_1}.
\end{aligned} \tag{I.21}$$

Thus, we obtain the exciton–coupling matrix element as

$$\mathscr{G}^{\mathrm{exc}}_{\beta\alpha,\mu q} = g^{\mu q\downarrow}_{\beta\alpha} - g^{\mu q\uparrow}_{\beta\alpha}. \tag{I.22}$$

These definitions are chosen in such a way that now the interaction $\mathcal{D}$ will be given by $D|\mathscr{G}^{\mathrm{exc}}|^2$ if $\alpha = \alpha'$. Now, the full non–diagonal exciton–phonon self–energy, Eq. (I.18), is written as

$$\Pi^{\mathrm{exc}}_{\alpha\alpha'}(t_{1'3'}) = -\mathrm{i}\sum_{\mu\beta q}D_{\mu q}(t_{1'3'})\mathscr{G}^{\mathrm{exc}}_{\beta\alpha',\mu q}\left(\mathscr{G}^{\mathrm{exc}}_{\beta\alpha,\mu q}\right)^{*}L^{(0)}_{\beta q}(t_{1'3'}). \tag{I.23}$$

Armed with this result, we may now take the cumulant expansion, $\chi^{c} = \chi^{(0)}\exp^{\Pi^{\mathrm{exc}}\chi^{(0)}}$. This will automatically include approximated descriptions of multi-phonon scattering processes and renormalisations of the exciton energies. If instead we want to stick to the first order, where cumulant and Dyson approaches coincide,

then the final result for $\chi^{(1)} = -\mathrm{i}L^{(1)}$, that is Eq. (I.3), can be obtained by solving Eq. (I.17) explicitly, using $L^{(0)}(t_{1'3'}) = \theta(t_{3'} - t_{1'})\exp[-\mathrm{i}E_{\beta q}(t_{3'} - t_{1'})]$. In this case we need to neglect completely the spurious contribution to the renormalisation of the exciton energy (because it is only treated correctly at all orders) and focus on the single–phonon satellite. Therefore, we remove the static part of the phonon propagator and only consider its purely dynamical part, given by (see Eqs. (1.109) and (1.110))

$$\begin{aligned}\Delta D_{\lambda q}(t_{21}) &= D_{\lambda q}(t_{21}) - D^{st}_{q\lambda}(t_{21}) = \\ &- \mathrm{i}\left[\theta(t_1 - t_2)\mathrm{e}^{-\mathrm{i}\omega_{q\mu}(t_1-t_2)} + \theta(t_2 - t_1)\mathrm{e}^{\mathrm{i}\omega_{q\mu}(t_1-t_2)}\right] + \frac{2}{\omega_{q\mu}}\delta(t_2 - t_1).\end{aligned} \tag{I.24}$$

I.5 Dynamical response function

Let us consider first the case of a one–exciton system (α) interacting with a single phonon at frequency Ω, with a diagonal self–energy. The result is (in Fourier space)

$$\chi^{(1)}_{\alpha\alpha}(\omega) = \frac{1 - \frac{|\mathscr{G}^{\mathrm{exc}}_{\alpha\alpha}|^2}{\Omega^2}}{\omega - E_\alpha + \mathrm{i}\eta} + \frac{|\mathscr{G}^{\mathrm{exc}}_{\alpha\alpha}|^2}{\Omega^2}\frac{1}{\omega - (E_\alpha + \Omega) + \mathrm{i}\eta}. \tag{I.25}$$

One satellite due to phonon emission appears in the spectrum. Notice that this result, a part from the quantities involved, is analytically equivalent to the first–order dynamical GF in Langreth's model (one–level fermion interacting with one boson, Eq. (4.8)). Notice that here we have the *bare* exciton energies instead of the *quasiparticle*–corrected ones, since we neglected this part of the theory and focused only on the dynamical structure of χ. Moreover, if we take the limit of a two–level system (one valence state v, one conduction state c), the result correctly reduces to the expression found in Sec. 4.1.2, Eq. (4.18), with the exciton–phonon coupling in particular reducing from Eq. (I.22) to eq. (4.15).

Below, we will write the full expressions for Π^{exc} and $\chi^{(1)}$ in frequency space. We consider both the case of a non–diagonal (upper formulas) and diagonal (lower formulas) self–energy, and for the sake of generality we let the exciton α have any momentum, labeled as $\bar{q}$. In the main text, we always discuss the diagonal case with $\bar{q} = 0$ for optical absorption. The self–energy is

$$\begin{aligned}\Pi^{\mathrm{exc}}_{\alpha\alpha'}(\bar{q},\omega) &= \frac{1}{N_q}\sum_{\mu\beta q}\frac{\mathscr{G}^{\mathrm{exc}}_{\beta\bar{q}+q\ \alpha'\bar{q},\ \mu q}\left(\mathscr{G}^{\mathrm{exc}}_{\beta\bar{q}+q\ \alpha\bar{q},\mu q}\right)^*}{\omega - (E_{\beta\bar{q}+q} + \Omega_{q\mu}) + \mathrm{i}\eta}, \\ \Pi^{\mathrm{exc}}_{\alpha\alpha}(\bar{q},\omega) &= \frac{1}{N_q}\sum_{\mu\beta q}\frac{|\mathscr{G}^{\mathrm{exc}}_{\beta\bar{q}+q\ \alpha\bar{q},\ \mu q}|^2}{\omega - (E_{\beta\bar{q}+q} + \Omega_{q\mu}) + \mathrm{i}\eta},\end{aligned} \tag{I.26}$$

and for $\chi^{(1)}$ we have:

$$\begin{aligned}\chi^{(1)}_{\alpha\alpha'}(\bar{q},\omega) &= \frac{1 - R_{\alpha\alpha'\bar{q}}}{\omega - E_{\alpha\bar{q}} + \mathrm{i}\eta} + \frac{1}{N_q}\sum_{\mu\beta q}\frac{\mathscr{G}^{\mathrm{exc}}_{\beta\bar{q}+q\ \alpha'\bar{q},\ \mu q}\left(\mathscr{G}^{\mathrm{exc}}_{\beta\bar{q}+q\ \alpha\bar{q},\mu q}\right)^*}{\mathscr{W}_{\beta\alpha\bar{q},\mu q}(\mathscr{W}_{\beta\alpha\bar{q},\mu q} - \Delta_{\alpha'\alpha\bar{q}})}\frac{1}{\omega - (E_{\alpha\bar{q}} + \mathscr{W}_{\beta\alpha\bar{q},\mu q}) + \mathrm{i}\eta}, \\ \chi^{(1)}_{\alpha\alpha}(\bar{q},\omega) &= \frac{1 - R_{\alpha\bar{q}}}{\omega - E_{\alpha\bar{q}} + \mathrm{i}\eta} + \frac{1}{N_q}\sum_{\mu\beta q}\frac{|\mathscr{G}^{\mathrm{exc}}_{\beta\bar{q}+q\ \alpha\bar{q},\ \mu q}|^2}{\mathscr{W}^2_{\beta\alpha\bar{q},\mu q}}\frac{1}{\omega - (E_{\alpha\bar{q}} + \mathscr{W}_{\beta\alpha\bar{q},\mu q}) + \mathrm{i}\eta}.\end{aligned} \tag{I.27}$$

Like in Sec. 4.1.1, the renormalisation factors are given by

$$\begin{aligned} R_{\alpha\alpha'\bar{q}} &= \frac{1}{N_q}\sum_{\mu\beta q}\frac{\mathscr{G}^{\rm exc}_{\beta\bar{q}+q\ \alpha'\bar{q},\ \mu q}\left(\mathscr{G}^{\rm exc}_{\beta\bar{q}+q\ \alpha\bar{q},\mu q}\right)^*}{\mathscr{W}_{\beta\alpha\bar{q},\mu q}(\mathscr{W}_{\beta\alpha\bar{q},\mu q}-\Delta_{\alpha'\alpha\bar{q}})} = -\frac{\partial\Pi^{\rm exc}_{\alpha\alpha'}(\bar{q},\omega)}{\partial\omega}\bigg|_{\omega=E_{\alpha\bar{q}}}, \\ R_{\alpha\bar{q}} &= \frac{1}{N_q}\sum_{\mu\beta q}\frac{|\mathscr{G}^{\rm exc}_{\beta\bar{q}+q\ \alpha\bar{q},\ \mu q}|^2}{\mathscr{W}^2_{\beta\alpha\bar{q},\mu q}} = -\frac{\partial\Pi^{\rm exc}_{\alpha\alpha}(\bar{q},\omega)}{\partial\omega}\bigg|_{\omega=E_{\alpha\bar{q}}}, \end{aligned} \tag{I.28}$$

with the energy denominators being

$$\begin{aligned} \mathscr{W}_{\beta\alpha\bar{q},\mu q} &= E_{\beta\bar{q}+q} - E_{\alpha\bar{q}} + \Omega_{q\mu}, \\ \Delta_{\alpha'\alpha\bar{q}} &= E_{\alpha'\bar{q}} - E_{\alpha\bar{q}}. \end{aligned} \tag{I.29}$$

Finally, in the diagonal case, we go in the basis of the G–vectors, Eq. (1.87), so that we can write the $\chi^{(1)}$ corresponding to a periodic crystal that may be computed with first–principles calculations:

$$\begin{aligned} \chi^{(1)}_{GG'}(\bar{q},\omega) &= \sum_\alpha\sum_{\mathcal{K}}\left(\mathcal{A}^{\mathcal{K}}_{\alpha\bar{q}}\right)^*\varrho_{\mathcal{K}}(\bar{\mathbf{q}}+\mathbf{G})\sum_{\mathcal{K}'}\mathcal{A}^{\mathcal{K}'}_{\alpha\bar{q}}\varrho^*_{\mathcal{K}'}(\bar{\mathbf{q}}+\mathbf{G}')\,\chi^{(1)}_{\alpha\alpha}(\bar{q},\omega), \\ \chi^{(1)}_{00}(\bar{q},\omega) &= \sum_\alpha |T_\alpha|^2\chi^{(1)}_{\alpha\alpha}(\bar{q},\omega). \end{aligned} \tag{I.30}$$

Appendix J

Notes on the implementation of the exciton–phonon self–energy

We give here a brief outline of the steps taken so far for the implementation of the exciton–phonon coupling and self–energy in the `Yambo` code.

Symmetries. Ideally, an exciton–phonon calculation would be performed in a q–vector grid spanning just the irreducible part of the BZ. However, this task is made difficult by the fact that the phases of both the exciton eigenvectors $\mathcal{A}$ and ep coupling matrix elements g are relevant to construct the coupling $\mathcal{A}g^*\mathcal{A}^*g$. Since the BSE is solved in the full BZ, and the g matrix elements are computed in DFPT using the symmetries of the crystal, in order to match the correct $\mathcal{A}$ with the correct g in Eq. (4.25) it is necessary to know how these quantities transform under crystal symmetry operations. This is not a trivial task, especially when we are dealing with degenerate band and exciton states. As a consequence, for the moment, all the exciton–phonon calculations have to be performed in the full BZ *with the symmetries of the system turned off*. Clearly, this represents a major hindrance for a calculation that is already computationally very demanding on its own: therefore we will pursue the solution of this issue in the future.

Electron–phonon coupling matrix elements. Within the TDA, we only have two types of ep coupling matrix elements, as defined in Sec. 1.6 and discussed in App. H:

$$\begin{aligned} g^{\mu q}_{vv'k} &= \langle \overline{vk} | \Delta_{q\mu} V_{eff} | \overline{v'k-q} \rangle \,/\, (2\Omega_{\mu q}), \\ g^{\mu q}_{c'ck+q} &= \langle \overline{c'k+q} | \Delta_{q\mu} V_{eff} | \overline{ck} \rangle \,/\, (2\Omega_{\mu q}). \end{aligned} \tag{J.1}$$

In `Yambo`, the default convention for the momentum transfer q is the one on the first line (i.e. $k - k' = q$). In order to obtain the second line (i.e. $k' - k = q$), we need to introduce a mapping function M that associates the index k in the momentum grid and the index q in the momentum *transfer* grid to a third momentum p in the same grid as k:

$$\begin{aligned} &M : \{k+q\} \to \{p\}, \\ &p = M(k+q) = p|_{k+q}, \\ &k = M^{-1}(p) = p|_{k+q} - q. \end{aligned} \tag{J.2}$$

In this way, the required matrix element in the $k' - k$ grid can be recovered from one already computed in the $k - k'$ grid:

$$g^{\mu q}_{c'ck+q} = g^{\mu q}_{c'cp|_{k+q}}. \tag{J.3}$$

As for the phonon frequencies at the denominators, it is implied that the three acoustic modes at $q = 0$, where $\Omega_{q\mu} = 0$, are not included in the $q\mu$ sums.

Exciton eigenvector components. The last basic ingredient is $\mathcal{A}_\alpha^{\mathcal{K}q}$, coming from the diagonalisation of the excitonic Hamiltonian. Again, the default convention in `Yambo` for finite–q calculations is

$$|\alpha q\rangle = \sum_{\mathcal{K}} \mathcal{A}_\alpha^{\mathcal{K}q\downarrow} |ck\rangle \otimes |vk-q\rangle \tag{J.4}$$

(for the meaning of the $(q\downarrow)$ notation, the reader is referred to App. H). For the construction of the ExcP coupling we also need:

$$\mathcal{A}_\alpha^{\mathcal{K}q\uparrow} = \{\langle ck+q| \otimes \langle vk|\} |\alpha q\rangle . \tag{J.5}$$

This can be obtained from the $(q\downarrow)$ components by again using the mapping Eq. (J.2):

$$\begin{aligned} \mathcal{P} &\to (cvp|_{k+q}) , \\ \mathcal{A}_\alpha^{\mathcal{K}q\uparrow} &= \mathcal{A}_\alpha^{\mathcal{P}q\downarrow} . \end{aligned} \tag{J.6}$$

The self–energy loop. Out objective is to compute Eq. (I.26). The exciton and phonon energies, as well as the EPC matrix elements and exciton eigenvector components, are read externally from databases produced by (i) the `Yambo` finite–q BSE calculations and (ii) the `Yambo` interface with QE (for the EPC). We start from Eq. (I.20) for the Ξ–components of the exciton–phonon coupling. We break the four terms down using the following definitions (where $S = \downarrow, \uparrow$):

$$\begin{aligned} \xi_{\beta\mathcal{K}}^{\mu q\downarrow} &\equiv \sum_{\mathcal{K}'} g_{\mathcal{K}\mathcal{K}'q\downarrow}^{q\mu} \left(\mathcal{A}_\beta^{\mathcal{K}'q\downarrow}\right)^* , \\ \xi_{\beta\mathcal{K}}^{\mu q\uparrow} &\equiv \sum_{\mathcal{K}'} g_{\mathcal{K}'q\uparrow\mathcal{K}}^{q\mu} \left(\mathcal{A}_\beta^{\mathcal{K}'q\uparrow}\right)^* , \\ \Xi_{\mathcal{K}_1\mathcal{K}_2}^{q\mu,SS'} &= \sum_\alpha \overline{\xi_{\alpha\mathcal{K}_1}^{\mu qS}} \xi_{\alpha\mathcal{K}_2}^{\mu qS'} \end{aligned} \tag{J.7}$$

(here the overline indicates the complex conjugate). In particular, switching back to band indices and using the previously defined mappings, the ξ–components are written as

$$\begin{aligned} \xi_{\beta\mathcal{K}}^{\mu q\downarrow} &\equiv \sum_{v'} g_{vv'k}^{q\mu} \left(\mathcal{A}_\beta^{cv'k}\right)^* , \\ \xi_{\beta\mathcal{K}}^{\mu q\uparrow} &\equiv \sum_{c'} g_{c'cp|_{k+q}}^{q\mu} \left(\mathcal{A}_\beta^{c'vp|_{k+q}}\right)^* . \end{aligned} \tag{J.8}$$

Equation (J.8) represents the first quantity computed in the innermost loops of the "exciton–phonon" Fortran routine. In a subsequent step, the quantity T_L is computed from the ξ–terms and the exciton eigenvectors $\mathcal{A}_\alpha^{\mathcal{K}}$ at $q = 0$:

$$T_{L\,\alpha\beta}^{q\mu} = \sum_{\mathcal{K}} \mathcal{A}_\alpha^{\mathcal{K}} (\xi_{\beta\mathcal{K}}^{\mu q\downarrow} - \xi_{\beta\mathcal{K}}^{\mu q\uparrow}), \tag{J.9}$$

and by defining $T_R = \overline{T}_L$ we reconstruct the ExcP coupling matrix element as

$$T^{q\mu}_{\alpha\alpha'\beta} = T_{L\,\alpha\beta}^{\;\;q\mu}\; T_{R\,\alpha'\beta}^{\;\;q\mu}. \tag{J.10}$$

Now, loops over β and μ are performed. Here we add the frequency denominator in order to obtain the self–energy at each q,

$$\Pi^{q}_{\alpha\alpha'}(\omega) = \sum_{\mu\beta} \frac{1}{N_q} \frac{T^{q\mu}_{\alpha\alpha'\beta}}{\omega - (E_{\beta q} + \Omega q\mu) + \mathrm{i}\eta}. \tag{J.11}$$

Every time a BSE calculation is performed at a certain q, the code calls the exciton–phonon routine to compute $\Pi^{q}_{\alpha\alpha'}$. The routine returns as output both $\Pi^{q}_{\alpha\alpha'}$ and $T^{q\mu}_{\alpha\alpha'\beta}$. Finally, the full self–energy is given by

$$\Pi^{\mathrm{exc}}_{\alpha\alpha'}(\omega) = \sum_{q} \Pi^{q}_{\alpha\alpha'}(\omega). \tag{J.12}$$

Bibliography

[1] H. Kuzmany, *Solid–State Spectroscopy: An Introduction* (Springer, 2009).

[2] R. M. Martin, *Electronic Structure – Basic Theory and Practical Methods* (Cambridge University Press, 2004).

[3] R. M. Martin, L. Reining, and D. M. Ceperley, *Interacting electrons – Theory and Computational Approaches* (Cambridge University Press, 2016).

[4] K. Lejaeghere, G. Bihlmayer, T. Björkman, P. Blaha, S. Blügel, V. Blum, D. Caliste, I. E. Castelli, S. J. Clark, A. Dal Corso, S. de Gironcoli, T. Deutsch, J. K. Dewhurst, I. Di Marco, C. Draxl, M. Dułak, O. Eriksson, J. A. Flores-Livas, K. F. Garrity, L. Genovese, P. Giannozzi, M. Giantomassi, S. Goedecker, X. Gonze, O. Grånäs, E. K. U. Gross, A. Gulans, F. Gygi, D. R. Hamann, P. J. Hasnip, N. A. W. Holzwarth, D. Iuşan, D. B. Jochym, F. Jollet, D. Jones, G. Kresse, K. Koepernik, E. Küçükbenli, Y. O. Kvashnin, I. L. M. Locht, S. Lubeck, M. Marsman, N. Marzari, U. Nitzsche, L. Nordström, T. Ozaki, L. Paulatto, C. J. Pickard, W. Poelmans, M. I. J. Probert, K. Refson, M. Richter, G.-M. Rignanese, S. Saha, M. Scheffler, M. Schlipf, K. Schwarz, S. Sharma, F. Tavazza, P. Thunström, A. Tkatchenko, M. Torrent, D. Vanderbilt, M. J. van Setten, V. Van Speybroeck, J. M. Wills, J. R. Yates, G.-X. Zhang, and S. Cottenier, "Reproducibility in density functional theory calculations of solids", Science **351**, 1415 (2016).

[5] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, "Phonons and related crystal properties from density-functional perturbation theory", Rev. Mod. Phys. **73**, 515–562 (2001).

[6] X. Gonze and C. Lee, "Dynamical matrices, Born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory", Phys. Rev. B **55**, 10355–10368 (1997).

[7] M. S. Hybertsen and S. G. Louie, "Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies", Phys. Rev. B **34**, 5390–5413 (1986).

[8] M. Rohlfing and S. G. Louie, "Electron-hole excitations and optical spectra from first principles", Phys. Rev. B **62**, 4927–4944 (2000).

[9] A. Castellanos–Gomez, "Why all the fuss about 2D semiconductors?", Nat. Photonics **10**, 202 EP (2016).

[10] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, "Electric Field Effect in Atomically Thin Carbon Films", Science **306**, 666–669 (2004).

[11] C. N. R. Rao, H. S. S. Ramakrishna Matte, and U. Maitra, "Graphene Analogues of Inorganic Layered Materials", Angew. Chem. Int. Edit. **52**, 13162–13185 (2013).

[12] L. Wirtz and A. Rubio, "Optical and Vibrational Properties of Boron Nitride Nanotubes", in *B-C-N Nanotubes and Related Nanostructures* (Springer New York, New York, NY, 2009), pp. 105–148.

[13] S. Galambosi, L. Wirtz, J. A. Soininen, J. Serrano, A. Marini, K. Watanabe, T. Taniguchi, S. Huotari, A. Rubio, and K. Hämäläinen, "Anisotropic excitonic effects in the energy loss function of hexagonal boron nitride", Phys. Rev. B **83**, 081413 (2011).

[14] K. Watanabe, T. Taniguchi, and H. Kanda, "Direct-bandgap properties and evidence for ultraviolet lasing of hexagonal boron nitride single crystal", Nat. Mater. **3**, 404–409 (2004).

[15] G. Watanabe, T. Taniguchi, T. Niiyama, K. Miya, and M. Taniguchi, "Far-ultraviolet plane-emission handheld device based on hexagonal boron nitride", Nat. Photonics **3**, 591–594 (2009).

[16] Y. Kubota, K. Watanabe, O. Tsuda, and T. Taniguchi, "Deep Ultraviolet Light-Emitting Hexagonal Boron Nitride Synthesized at Atmospheric Pressure", Science **317**, 932–934 (2007).

[17] J. D. Caldwell, I. Aharonovich, G. Cassabois, J. H. Edgar, B. Gil, and D. N. Basov, "Photonics with hexagonal boron nitride", Nat. Rev. Mater. **4**, 552–567 (2019).

[18] A. Pierret, J. Loayza, B. Berini, A. Betz, B. Plaçais, F. Ducastelle, J. Barjon, and A. Loiseau, "Excitonic recombinations in $h-$BN: From bulk to exfoliated layers", Phys. Rev. B **89**, 035414 (2014).

[19] G. Cassabois, P. Valvin, and B. Gil, "Hexagonal boron nitride is an indirect bandgap semiconductor", Nat. Photonics **10**, 262–266 (2016).

[20] R. V. Gorbachev, I. Riaz, R. R. Nair, R. Jalil, L. Britnell, B. D. Belle, E. W. Hill, K. S. Novoselov, K. Watanabe, T. Taniguchi, A. K. Geim, and P. Blake, "Hunting for Monolayer Boron Nitride: Optical and Raman Signatures", Small **7**, 465–468 (2011).

[21] B. Arnaud, S. Lebègue, P. Rabiller, and M. Alouani, "Huge Excitonic Effects in Layered Hexagonal Boron Nitride", Phys. Rev. Lett. **96**, 026402 (2006).

[22] S. Wang, Y. Li, J. Yip, and J. Wang, "The excitonic effects in single and double-walled boron nitride nanotubes", J. Chem. Phys. **140**, 244701 (2014).

[23] L. Wirtz, A. Marini, M. Grüning, C. Attaccalite, G. Kresse, and A. Rubio, "Comment on "Huge Excitonic Effects in Layered Hexagonal Boron Nitride"", Phys. Rev. Lett. **100**, 189701 (2008).

[24] K. Watanabe and T. Taniguchi, "Jahn-Teller effect on exciton states in hexagonal boron nitride single crystal", Phys. Rev. B **79**, 193104 (2009).

[25] G. Cassabois, P. Valvin, and B. Gil, "Intervalley scattering in hexagonal boron nitride", Phys. Rev. B **93**, 035207 (2016).

[26] L. Wirtz, A. Marini, and A. Rubio, "Excitons in Boron Nitride Nanotubes: Dimensionality Effects", Phys. Rev. Lett. **96**, 126104 (2006).

[27] C. Attaccalite, M. Bockstedte, A. Marini, A. Rubio, and L. Wirtz, "Coupling of excitons and defect states in boron-nitride nanostructures", Phys. Rev. B **83**, 144115 (2011).

[28] J. Koskelo, G. Fugallo, M. Hakala, M. Gatti, F. Sottile, and P. Cudazzo, "Excitons in van der Waals materials: From monolayer to bulk hexagonal boron nitride", Phys. Rev. B **95**, 035125 (2017).

[29] G. Fugallo, M. Aramini, J. Koskelo, K. Watanabe, T. Taniguchi, M. Hakala, S. Huotari, M. Gatti, and F. Sottile, "Exciton energy-momentum map of hexagonal boron nitride", Phys. Rev. B **92**, 165122 (2015).

[30] E. Torun, H. P. C. Miranda, A. Molina-Sánchez, and L. Wirtz, "Interlayer and intralayer excitons in MoS_2/WS_2 and $MoSe_2/WSe_2$ heterobilayers", Phys. Rev. B **97**, 245427 (2018).

[31] M. Xu, T. Liang, M. Shi, and H. Chen, "Graphene-Like Two-Dimensional Materials", Chem. Rev. **113**, 3766–3798 (2013).

[32] L. Schue, B. Berini, A. C. Betz, B. Placais, F. Ducastelle, J. Barjon, and A. Loiseau, "Dimensionality effects on the luminescence properties of hBN", Nanoscale **8**, 6986–6993 (2016).

[33] Y. Feng, J. A. Soininen, A. L. Ankudinov, J. O. Cross, G. T. Seidler, A. T. Macrander, J. J. Rehr, and E. L. Shirley, "Exciton spectroscopy of hexagonal boron nitride using nonresonant x-ray Raman scattering", Phys. Rev. B **77**, 165202 (2008).

[34] M. Gatti and F. Sottile, "Exciton dispersion from first principles", Phys. Rev. B **88**, 155113 (2013).

[35] X. Blase, A. Rubio, S. G. Louie, and M. L. Cohen, "Quasiparticle band structure of bulk hexagonal boron nitride and related systems", Phys. Rev. B **51**, 6868–6875 (1995).

[36] J. Noffsinger, E. Kioupakis, C. G. Van de Walle, S. G. Louie, and M. L. Cohen, "Phonon-Assisted Optical Absorption in Silicon from First Principles", Phys. Rev. Lett. **108**, 167402 (2012).

[37] M. Zacharias, C. E. Patrick, and F. Giustino, "Stochastic Approach to Phonon-Assisted Optical Absorption", Phys. Rev. Lett. **115**, 177401 (2015).

[38] P. Cudazzo and L. Reining, "Correlation satellites in optical and loss spectra", Submitted. (Private communication with P. Cudazzo) (2019).

[39] T. Galvani, F. Paleari, H. P. C. Miranda, A. Molina-Sánchez, L. Wirtz, S. Latil, H. Amara, and F. Ducastelle, "Excitons in boron nitride single layer", Phys. Rev. B **94**, 125303 (2016).

[40] F. Paleari, T. Galvani, H. Amara, F. Ducastelle, A. Molina-Sánchez, and L. Wirtz, "Excitons in few-layer hexagonal boron nitride: Davydov splitting and surface localization", 2D Materials **5**, 045017 (2018).

[41] F. Paleari, H. P. C. Miranda, A. Molina-Sánchez, and L. Wirtz, "Exciton-Phonon Coupling in the Ultraviolet Absorption and Emission Spectra of Bulk Hexagonal Boron Nitride", Phys. Rev. Lett. **122**, 187401 (2019).

[42] J. D. Martin, "What's in a Name Change?", Physics in Perspective **17**, 3–32 (2015).

[43] G. Grosso and G. Pastori Parravicini, *Solid State Physics* (Academic Press, 2000).

[44] G. Onida, L. Reining, and A. Rubio, "Electronic excitations: density-functional versus many-body Green's-function approaches", Rev. Mod. Phys. **74**, 601–659 (2002).

[45] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors* (Springer, 2010).

[46] G. Strinati, "Application of the Green's functions method to the study of the optical properties of semiconductors", Riv. Nuovo Cimento **11**, 1–86 (1988).

[47] W. Schäfer and M. Wegener, *Semiconductor Optics and Transport Phenomena* (Springer, 2002).

[48] F. Bechstedt, *Many-Body Approach to Electronic Excitations* (Springer, 2015).

[49] F. Giustino, "Electron-phonon interactions from first principles", Rev. Mod. Phys. **89**, 015003 (2017).

[50] G. D. Mahan, *Many-Particle Physics* (Plenum Press, 1990).

[51] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many–Particle Systems* (Dover Publications, 2003).

[52] H. Bruus and K. Flensberg, *Many-Body Quantum Theory in Condensed Matter Physics: an Introduction* (Oxford University Press, 2004).

[53] L. Reining, "The GW approximation: content, successes and limitations", Wiley Interdiscip. Rev. Comput. Mol. Sci. **8**, e1344 (2018).

[54] P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas", Phys. Rev. **136**, B864–B871 (1964).

[55] W. Kohn and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects", Phys. Rev. **140**, A1133–A1138 (1965).

[56] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials", J. Phys. Condens. Matter **21**, 395502 (2009).

[57] A. Damascelli, "Probing the Electronic Structure of Complex Systems by ARPES", Phys. Scripta **T109**, 61 (2004).

[58] A. Marini, C. Hogan, M. Grüning, and D. Varsano, "yambo: An ab initio tool for excited state calculations", Comput. Phys. Commun. **180**, 1392–1403 (2009).

[59] D. Sangalli, A. Ferretti, H. Miranda, C. Attaccalite, I. Marri, E. Cannuccia, P. Melo, M. Marsili, F. Paleari, A. Marrazzo, G. Prandini, P. Bonfà, M. O. Atambo, F. Affinito, M. Palummo, A. Molina-Sánchez, C. Hogan, M. Grüning, D. Varsano, and A. Marini, "Many-body perturbation theory calculations using the yambo code", J. Phys. Condens. Matter **31**, 325902 (2019).

[60] V. Ambegaokar and W. Kohn, "Electromagnetic Properties of Insulators. I", Phys. Rev. **117**, 423–431 (1960).

[61] R. Del Sole and E. Fiorino, "Macroscopic dielectric tensor at crystal surfaces", Phys. Rev. B **29**, 4631–4645 (1984).

[62] R. Haydock, "The Recursive Solution of the Schrodinger Equation", in , Vol. 35, edited by H. Ehrenreich, F. Seitz, and D. Turnbull, Solid State Physics (Academic Press, 1980), pp. 215–294.

[63] V. Hernandez, J. E. Roman, and V. Vidal, "SLEPc, A scalable and flexible toolkit for the solution of eigenvalue problems", ACM Trans. Math. Softw. **31**, 351–362 (2005).

[64] D. Y. Qiu, F. H. da Jornada, and S. G. Louie, "Screening and many-body effects in two-dimensional crystals: Monolayer MoS_2", Phys. Rev. B **93**, 235435 (2016).

[65] C. A. Rozzi, D. Varsano, A. Marini, E. K. U. Gross, and A. Rubio, "Exact Coulomb cutoff technique for supercell calculations", Phys. Rev. B **73**, 205119 (2006).

[66] S. Reichardt, "Many-Body Perturbation Theory Approach to Raman Spectroscopy and Its Application to 2D Materials", PhD thesis (Physics and Materials Science Research Unit, University of Luxembourg, 2018).

[67] H. Miranda, "Ab initio approaches to Resonant Raman Spectroscopy of Transition Metal Dichalcogenides", PhD thesis (Physics and Materials Science Research Unit, University of Luxembourg, 2017).

[68] P. Brüesch, *Phonons: Theory and Experiments I* (Springer, 1982).

[69] S. Poncé, Y. Gillet, J. Laflamme Janssen, A. Marini, M. Verstraete, and X. Gonze, "Erratum: "Temperature dependence of the electronic structure of semiconductors and insulators" [J. Chem. Phys. 143, 102813 (2015)]", J. Chem. Phys. **146**, 099901 (2017).

[70] H. Y. Fan, "Temperature Dependence of the Energy Gap in Monatomic Semiconductors", Phys. Rev. **78**, 808–809 (1950).

[71] H. Y. Fan, "Temperature Dependence of the Energy Gap in Semiconductors", Phys. Rev. **82**, 900–905 (1951).

[72] E. Antončík, "On the theory of temperature shift of the absorption curve in non-polar crystals", Czech. J. Phys. **5**, 449–461 (1955).

[73] P. B. Allen and V. Heine, "Theory of the temperature dependence of electronic band structures", J. Phys. C Solid State **9**, 2305 (1976).

[74] P. B. Allen and M. Cardona, "Theory of the temperature dependence of the direct gap of germanium", Phys. Rev. B **23**, 1495–1505 (1981).

[75] P. B. Allen and M. Cardona, "Temperature dependence of the direct gap of Si and Ge", Phys. Rev. B **27**, 4760–4769 (1983).

[76] S. Poncé, Y. Gillet, J. Laflamme Janssen, A. Marini, M. Verstraete, and X. Gonze, "Temperature dependence of the electronic structure of semiconductors and insulators", J. Chem. Phys. **143**, 102813 (2015).

[77] S. Poncé, E. R. Margine, C. Verdi, and F. Giustino, "EPW: Electron-phonon coupling, transport and superconducting properties using maximally localized Wannier functions", Comput. Phys. Commun. **209**, 116–133 (2016).

[78] J. Serrano, A. Bosak, R. Arenal, M. Krisch, K. Watanabe, T. Taniguchi, H. Kanda, A. Rubio, and L. Wirtz, "Vibrational Properties of Hexagonal Boron Nitride: Inelastic X-Ray Scattering and *Ab Initio* Calculations", Phys. Rev. Lett. **98**, 095503 (2007).

[79] W. Aggoune, C. Cocchi, D. Nabok, K. Rezouali, M. A. Belkhir, and C. Draxl, "Dimensionality of excitons in stacked van der Waals materials: The example of hexagonal boron nitride", Phys. Rev. B **97**, 241114 (2018).

[80] C. E. Dreyer, J. L. Lyons, A. Janotti, and C. G. V. de Walle, "Band alignments and polarization properties of BN polymorphs", Appl. Phys. Express **7**, 031001 (2014).

[81] D. Wickramaratne, L. Weston, and C. G. V. de Walle, "Monolayer to Bulk Properties of Hexagonal Boron Nitride", J. Phys. Chem. C **122**, 25524–25529 (2018).

[82] A. Marini, "Ab Initio Finite-Temperature Excitons", Phys. Rev. Lett. **101**, 106405 (2008).

[83] T. Q. P. Vuong, G. Cassabois, P. Valvin, S. Liu, J. H. Edgar, and B. Gil, "Exciton-phonon interaction in the strong-coupling regime in hexagonal boron nitride", Phys. Rev. B **95**, 201202 (2017).

[84] L. Schué, L. Sponza, A. Plaud, H. Bensalah, K. Watanabe, T. Taniguchi, F. ˜. Ducastelle, A. Loiseau, and J. Barjon, "Bright Luminescence from Indirect and Strongly Bound Excitons in h-BN", Phys. Rev. Lett. **122**, 067401 (2019).

[85] L. V. Keldysh, "Coulomb interaction in thin semiconductor and semimetal films", J. Exp. Theor. Phys. **29**, 658 (1979).

[86] P. Cudazzo, I. V. Tokatly, and A. Rubio, "Dielectric screening in two-dimensional insulators: Implications for excitonic and impurity states in graphane", Phys. Rev. B **84**, 085406 (2011).

[87] A. Davydov, *Theory of Molecular Excitons* (MacGraw-Hill, 1969).

[88] P. Dawson, "Dipole dipole interactions and Davydov splitting in crystals", J. Phys. Chem. Solids **36**, 1401–1403 (1975).

[89] X. Luo, Y. Zhao, J. Zhang, Q. Xiong, and S. Y. Quek, "Anomalous frequency trends in MoS_2 thin films attributed to surface effects", Phys. Rev. B **88**, 075320 (2013).

[90] M. Staiger, R. Gillen, N. Scheuschner, O. Ochedowski, F. Kampmann, M. Schleberger, C. Thomsen, and J. Maultzsch, "Splitting of monolayer out-of-plane $A_1^{'}$ Raman mode in few-layer WS_2", Phys. Rev. B **91**, 195419 (2015).

[91] A. Molina-Sánchez, K. Hummer, and L. Wirtz, "Vibrational and optical properties of MoS2: From monolayer to bulk", Surf. Sci. Rep. **70**, 554–586 (2015).

[92] H. P. C. Miranda, S. Reichardt, G. Froehlicher, A. Molina-Sánchez, S. Berciaud, and L. Wirtz, "Quantum Interference Effects in Resonant Raman Spectroscopy of Single- and Triple-Layer MoTe2 from First-Principles", Nano Lett. **17**, 2381–2388 (2017).

[93] L. Wirtz, A. Marini, M. Grüning, and A. Rubio, "Excitonic effects in optical absorption and electron-energy loss spectra of hexagonal boron nitride", 2005.

[94] F. Caruso, P. Rinke, X. Ren, A. Rubio, and M. Scheffler, "Self-consistent *GW*: All-electron implementation with localized basis functions", Phys. Rev. B **88**, 075105 (2013).

[95] P. Cudazzo, L. Sponza, C. Giorgetti, L. Reining, F. Sottile, and M. Gatti, "Exciton Band Structure in Two-Dimensional Materials", Phys. Rev. Lett. **116**, 066803 (2016).

[96] K. A. Mengle and E. Kioupakis, "Impact of the stacking sequence on the bandgap and luminescence properties of bulk, bilayer, and monolayer hexagonal boron nitride", APL Materials **7**, 021106 (2019).

[97] L. Sponza, H. Amara, C. Attaccalite, S. Latil, T. Galvani, F. Paleari, L. Wirtz, and F. Ducastelle, "Direct and indirect excitons in boron nitride polymorphs: A story of atomic configuration and electronic correlation", Phys. Rev. B **98**, 125206 (2018).

[98] M. Posternak, A. Baldereschi, A. J. Freeman, E. Wimmer, and M. Weinert, "Prediction of Electronic Interlayer States in Graphite and Reinterpretation of Alkali Bands in Graphite Intercalation Compounds", Phys. Rev. Lett. **50**, 761–764 (1983).

[99] M. Posternak, A. Baldereschi, A. J. Freeman, and E. Wimmer, "Prediction of Electronic Surface States in Layered Materials: Graphite", Phys. Rev. Lett. **52**, 863–866 (1984).

[100] X. Blase, L. X. Benedict, E. L. Shirley, and S. G. Louie, "Hybridization effects and metallicity in small radius carbon nanotubes", Phys. Rev. Lett. **72**, 1878–1881 (1994).

[101] X. Blase, A. Rubio, S. G. Louie, and M. L. Cohen, "Stability and Band Gap Constancy of Boron Nitride Nanotubes", Europhys. Lett. **28**, 335–340 (1994).

[102] S. Hu, J. Zhao, Y. Jin, J. Yang, P. Hrvoje, and J. G. Hou, "Nearly Free Electron Superatom States of Carbon and Boron Nitride Nanotubes", Nano Lett. **10**, 4830–4838 (2010).

[103] E. Kioupakis, P. Rinke, A. Schleife, F. Bechstedt, and C. G. Van de Walle, "Free-carrier absorption in nitrides from first principles", Phys. Rev. B **81**, 241201 (2010).

[104] C. E. Patrick and F. Giustino, "Unified theory of electron–phonon renormalization and phonon-assisted optical absorption", J. Phys. Condens. Matter **26**, 365503 (2014).

[105] M. Zacharias and F. Giustino, "One-shot calculation of temperature-dependent optical spectra and phonon-induced band-gap renormalization", Phys. Rev. B **94**, 075125 (2016).

[106] Y. Toyozawa, "Theory of Line-Shapes of the Exciton Absorption Bands", Progr. Theor. Phys. **20**, 53–81 (1958).

[107] G. Antonius and S. G. Louie, "Theory of the Exciton-Phonon Coupling", ArXiv e-prints (2017).

[108] B. Segall and G. D. Mahan, "Phonon-Assisted Recombination of Free Excitons in Compound Semiconductors", Phys. Rev. **171**, 935–948 (1968).

[109] S. Rudin, T. L. Reinecke, and B. Segall, "Temperature-dependent exciton linewidths in semiconductors", Phys. Rev. B **42**, 11218–11231 (1990).

[110] V. Perebeinos, J. Tersoff, and P. Avouris, "Effect of Exciton-Phonon Coupling in the Calculated Optical Absorption of Carbon Nanotubes", Phys. Rev. Lett. **94**, 027402 (2005).

[111] D. Christiansen, M. Selig, G. Berghäuser, R. Schmidt, I. Niehues, R. Schneider, A. Arora, S. M. de Vasconcellos, R. Bratschitsch, E. Malic, and A. Knorr, "Phonon Sidebands in Monolayer Transition Metal Dichalcogenides", Phys. Rev. Lett. **119**, 187402 (2017).

[112] A. Chernikov, V. Bornwasser, M. Koch, S. Chatterjee, C. N. Böttge, T. Feldtmann, M. Kira, S. W. Koch, T. Wassner, S. Lautenschläger, B. K. Meyer, and M. Eickhoff, "Phonon-assisted luminescence of polar semiconductors: Fröhlich coupling versus deformation-potential scattering", Phys. Rev. B **85**, 035201 (2012).

[113] P. Cudazzo, "Cumulant approach to the exciton–phonon interaction", To be published. (Private communication with P. Cudazzo) (2019).

[114] L. H. Hall, J. Bardeen, and F. J. Blatt, "Infrared Absorption Spectrum of Germanium", Phys. Rev. **95**, 559–560 (1954).

[115] K. H. Michel and B. Verberck, "Theory of elastic and piezoelectric effects in two-dimensional hexagonal boron nitride", Phys. Rev. B **80**, 224301 (2009).

[116] T. Sohier, M. Gibertini, M. Calandra, F. Mauri, and N. Marzari, "Breakdown of Optical Phonons' Splitting in Two-Dimensional Materials", Nano Lett. **17**, 3758–3763 (2017).

[117] L. Wirtz, A. Rubio, R. A. de la Concha, and A. Loiseau, "Ab initio calculations of the lattice dynamics of boron nitride nanotubes", Phys. Rev. B **68**, 045425 (2003).

[118] F. Giustino, S. G. Louie, and M. L. Cohen, "Electron-Phonon Renormalization of the Direct Band Gap of Diamond", Phys. Rev. Lett. **105**, 265501 (2010).

[119] H. Mishra and S. Bhattacharya, "Giant exciton-phonon coupling and zero-point renormalization in hexagonal monolayer boron nitride", Phys. Rev. B **99**, 165201 (2019).

[120] F. Bechstedt, K. Tenelsen, B. Adolph, and R. Del Sole, "Compensation of Dynamical Quasiparticle and Vertex Corrections in Optical Spectra", Phys. Rev. Lett. **78**, 1528–1531 (1997).

[121] C. Verdi and F. Giustino, "Fröhlich Electron-Phonon Vertex from First Principles", Phys. Rev. Lett. **115**, 176401 (2015).

[122] M. F. Pereira and K. Henneberger, "Microscopic theory for the influence of Coulomb correlations in the light-emission properties of semiconductor quantum wells", Phys. Rev. B **58**, 2064–2076 (1998).

[123] K. Hannewald, S. Glutsch, and F. Bechstedt, "Theory of photoluminescence in semiconductors", Phys. Rev. B **62**, 4519–4525 (2000).

[124] P. M. M. C. de Melo and A. Marini, "Unified theory of quantized electrons, phonons, and photons out of equilibrium: A simplified ab initio approach based on the generalized Baym-Kadanoff ansatz", Phys. Rev. B **93**, 155102 (2016).

[125] P. T. Landsberg, *Recombination in semiconductors* (Cambridge University Press, 1991) Chap. Radiative recombination (mainly for bands).

[126] W. van Roosbroeck and W. Shockley, "Photon-Radiative Recombination of Electrons and Holes in Germanium", Phys. Rev. **94**, 1558–1560 (1954).

[127] H. B. Bebb and E. W. Williams, *Semiconductors and semimetals* (Elsevier, 1972) Chap. Photoluminescence I: Theory.

[128] B. Laikhtman, "Are excitons really bosons?", J. Phys. Condens. Matter **19**, 295214 (2007).

[129] J. H. Lloyd-Williams and B. Monserrat, "Lattice dynamics and electron-phonon coupling calculations using nondiagonal supercells", Phys. Rev. B **92**, 184301 (2015).

[130] M. S. Dresselhaus, G. Dresselhaus, and A. Jorio, *Group theory: Application to the Physics of Condensed Matter* (Springer, 2008).

[131] T. Q. P. Vuong, G. Cassabois, P. Valvin, V. Jacques, A. V. D. Lee, A. Zobelli, K. Watanabe, T. Taniguchi, and B. Gil, "Phonon symmetries in hexagonal boron nitride probed by incoherent light emission", 2D Materials **4**, 011004 (2017).

[132] T. Q. P. Vuong, G. Cassabois, P. Valvin, V. Jacques, R. Cuscó, L. Artús, and B. Gil, "Overtones of interlayer shear modes in the phonon-assisted emission spectrum of hexagonal boron nitride", Phys. Rev. B **95**, 045207 (2017).

[133] W.-T. Hsu, L.-S. Lu, D. Wang, J.-K. Huang, M.-Y. Li, T.-R. Chang, Y.-C. Chou, Z.-Y. Juang, H.-T. Jeng, L.-J. Li, and W.-H. Chang, "Evidence of indirect gap in monolayer WSe2", Nat. Commun. **8**, 929 (2017).

[134] K. Shindo, "Effective Electron-Hole Interaction in Shallow Excitons", Journal of the Physical Society of Japan **29**, 287–296 (1970).

[135] G. Strinati, "Dynamical Shift and Broadening of Core Excitons in Semiconductors", Phys. Rev. Lett. **49**, 1519–1522 (1982).

[136] G. Strinati, "Effects of dynamical screening on resonances at inner-shell thresholds in semiconductors", Phys. Rev. B **29**, 5718–5726 (1984).

[137] A. Marini and R. Del Sole, "Dynamical Excitonic Effects in Metals and Semiconductors", Phys. Rev. Lett. **91**, 176402 (2003).

[138] B. I. Lundqvist, "Single-particle spectrum of the degenerate electron gas", Phys. Kondens. Mater. **6**, 193–205 (1967).

[139] D. C. Langreth, "Singularities in the X-Ray Spectra of Metals", Phys. Rev. B **1**, 471–477 (1970).

[140] M. Cini, "Theory of the Auger effect in solids: Plasmon effects in electron spectroscopies of valence states", Phys. Rev. B **17**, 2486–2493 (1978).

[141] L. Hedin, "Effects of Recoil on Shake-Up Spectra in Metals", Phys. Scripta **21**, 477–480 (1980).

[142] O. Gunnarsson, V. Meden, and K. Schönhammer, "Corrections to Migdal's theorem for spectral functions: A cumulant treatment of the time-dependent Green's function", Phys. Rev. B **50**, 10462–10473 (1994).

[143] F. Bechstedt, M. Fiedler, C. Kress, and R. Del Sole, "Dynamical screening and quasiparticle spectral functions for nonmetals", Phys. Rev. B **49**, 7357–7362 (1994).

[144] F. Aryasetiawan, L. Hedin, and K. Karlsson, "Multiple Plasmon Satellites in Na and Al Spectral Functions from Ab Initio Cumulant Expansion", Phys. Rev. Lett. **77**, 2268–2271 (1996).

[145] B. Holm and F. Aryasetiawan, "Self-consistent cumulant expansion for the electron gas", Phys. Rev. B **56**, 12825–12831 (1997).

[146] J. S. Zhou, "Theory of electron spectroscopy beyond the state-of-the-art. An improved description of fermion-plasmon coupling in Green's function calculations", PhD thesis (Theoretical Spectroscopy Groupe, École Polytechnique, 2016).

[147] L. Hedin, "On correlation effects in electron spectroscopies and theGWapproximation", J. Phys. Condens. Matter **11**, R489–R528 (1999).

[148] J. S. Zhou, J. J. Kas, L. Sponza, I. Reshetnyak, M. Guzzo, C. Giorgetti, M. Gatti, F. Sottile, J. J. Rehr, and L. Reining, "Dynamical effects in electron spectroscopy", J. Chem. Phys. **143**, 184109 (2015).

[149] M. Guzzo, G. Lani, F. Sottile, P. Romaniello, M. Gatti, J. J. Kas, J. J. Rehr, M. G. Silly, F. Sirotti, and L. Reining, "Valence Electron Photoemission Spectrum of Semiconductors: *Ab Initio* Description of Multiple Satellites", Phys. Rev. Lett. **107**, 166401 (2011).

[150] J. J. Kas, J. J. Rehr, and L. Reining, "Cumulant expansion of the retarded one-electron Green function", Phys. Rev. B **90**, 085112 (2014).

[151] S. M. Story, J. J. Kas, F. D. Vila, M. J. Verstraete, and J. J. Rehr, "Cumulant expansion for phonon contributions to the electron spectral function", Phys. Rev. B **90**, 195135 (2014).

[152] J. P. Nery, P. B. Allen, G. Antonius, L. Reining, A. Miglio, and X. Gonze, "Quasiparticles and phonon satellites in spectral functions of semiconductors and insulators: Cumulants applied to the full first-principles theory and the Fröhlich polaron", Phys. Rev. B **97**, 115145 (2018).

[153] B. Gumhalter, V. Kovač, F. Caruso, H. Lambert, and F. Giustino, "On the combined use of GW approximation and cumulant expansion in the calculations of quasiparticle spectra: The paradigm of Si valence bands", Phys. Rev. B **94**, 035103 (2016).

[154] C. Verdi, F. Caruso, and F. Giustino, "Origin of the crossover from polarons to Fermi liquids in transition metal oxides", Nat. Commun. **8**, 15769 (2017).

[155] W. S. Chow, "Theory of Phonon-Assisted Optical Absorption in Semiconductors. I.", Phys. Rev. **185**, 1056–1061 (1969).

[156] W. S. Chow, "Theory of Phonon-Assisted Optical Absorption in Semiconductors. II", Phys. Rev. **185**, 1062–1063 (1969).

[157] C. Elias, P. Valvin, T. Pelini, A. Summerfield, C. J. Mellor, T. S. Cheng, L. Eaves, C. T. Foxon, P. H. Beton, S. V. Novikov, B. Gil, and G. Cassabois, "Direct band-gap crossover in epitaxial monolayer boron nitride", Nat. Commun. **10**, 2639 (2019).

[158] P. Larson, M. Dvorak, and Z. Wu, "Role of the plasmon-pole model in the *GW* approximation", Phys. Rev. B **88**, 125205 (2013).

[159] F. Hüser, T. Olsen, and K. S. Thygesen, "Quasiparticle GW calculations for solids, molecules, and two-dimensional materials", Phys. Rev. B **87**, 235132 (2013).

[160] M. Fuchs and M. Scheffler, "Ab initio pseudopotentials for electronic structure calculations of poly-atomic systems using density-functional theory", Comput. Phys. Commun. **119**, 67–98 (1999).

Acknowledgements

*This PhD project was funded by the Luxembourg National Research Fund (FNR) through the AFR PhD individual grant scheme (project EXCPHON/*11280304*: "Phonon-assisted optical absorption in layered materials").*

If I were to choose the most charitable and optimistic viewpoint on the status of today's scientific research, I would describe it as a collective and collaborative endeavour. In this spirit, I will list below in random order the names of those researchers whose direct or indirect contributions helped shape this thesis.

Henrique Miranda for developing the *yambopy* pre/postprocessing tool for `Yambo`, lending me his expertise in scientific computing and code development early in my PhD, and crucial discussions on excitons, phonons, electron–phonon coupling, numerical approaches and data analysis. Thomas Galvani for developing the tight–binding exciton model, invaluable discussions about boron nitride excitons, 2D screening and many other topics, as well as for his keen mathematical insight, which helped me more than once. Hakim Amara for his help with the tight–binding exciton model. François Ducastelle for his kindness and willingness to share his deep knowledge on the topics of excitons, phonons, and solid–state physics in general. Julien Barjon for his kindness and for discussing with me his experimental results on cathodoluminescence in BN systems. Phuong Vuong, Guillaume Cassabois, and Bernard Gil for discussing with me their photoluminescence results on bulk hBN. I especially thank G. Cassabois for pointing out a crucial inconsistency in my preliminary calculations and P. Vuong for her kindness. Alejandro Molina–Sánchez for guiding me into the world of *ab initio* many–body simulations, helping me devise the framework for my workflows, and innumerable discussions and suggestions about the computational and physical issues I stumbled upon, especially regarding GW, BSE and electron–phonon coupling. I thank him as well for hosting me at the University of Valencia for three weeks as a visiting researcher. Sven Reichardt for extremely helpful discussions about many–body physics, Feynman diagrams, and possible approaches to solve conceptual issues. Marios Zacharias for discussing the Williams–Lax approach with me in several occasions, as well as for providing input files for boron nitride supercells with optimally displaced atomic positions. Feliciano Giustino for his kindness and for crucially pointing out the existence of the van Roosbroeck–Shockley relation. Lorenzo Sponza for the many discussions on the topics of BN excitons and finite–momentum BSE. Engin Torun for tackling with me computational and conceptual issues related to excitons in 2D materials, many–body simulations, HPC usage, automatised workflows and data analysis. Matteo Barborini for sharing his deep expertise on computational issues regarding complex simulations of materials. Pedro Melo for his expertise on luminescence and many–body physics in general. Andrea Marini for his kindness and invaluable theoretical and code development support: from reviewing my exciton–phonon notes, discussing the physics (e.g. the consequences of the longitudinal–transverse splitting of excitons), and helping me find my way through the `Yambo` Fortran routines. I thank him as well for hosting me at the CNR in Rome for one week, early in my PhD, and then again for two months near the end as a visiting researcher. Davide Sangalli for technical support around `Yambo`'s issues, discussions about excitons and help with code development. On this note, I also wish to acknowledge the entire `Yambo` developers team. Claudio Attaccalite and Elena Cannuccia for kindly inviting me to the University of Roma–Tor Vergata to

give a talk. Guido Fratesi and Gian Paolo Brivio for kindly inviting me to the University of Milano–Bicocca to give a talk, as well as for introducing me to the world of condensed matter physics and first–principles calculations during my Master's degree. Pierluigi Cudazzo for teaching me about his theory for a dynamical BSE and the cumulant expansion, letting me learn all the derivations and expand upon his work. I benefited as well from the many insighful discussions about both conceptual and practical issues in many–body physics, and from his deep expertise in the topic. Valentin Plugaru for his invaluable help as system administrator of the HPC facilities at the University of Luxembourg. His kindness, readiness to assist, accurate assessment of problems, pedagogical discussions, and willingness to go the extra mile — even when my cluster usage was bordering on insane — were instrumental in ensuring the feasibility of my heaviest and most numerous simulations. Lucia Reining for her kindness, for the careful reading of the manuscript and for pointing out various errors and flaws in my presentation of the theory in Chapter 1, which I have tried to correct.

Finally, I wish to extend my most sincere gratitude to my supervisor, Ludger Wirtz. The relevance of his support, counsel, and mentorship during these four years can hardly be overstated. I am grateful to have been part of the Theoretical Solid–State Physics Group at the University of Luxembourg.