UNIVERSITÀ DEGLI STUDI DI ROMA
"TOR VERGATA"

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI

TESI DI DOTTORATO IN FISICA
XVIII ciclo

Electronic and optical properties of the (111)2×1 diamond surface: an ab-initio study.

Candidata:
Dott. Margherita Marsili

Relatori:
Dott. Olivia Pulci
Prof. Rodolfo Del Sole
Prof. Friedhelm Bechstedt

Coordinatore del dottorato:
Prof. Piergiorgio Picozza

Contents

Contents

Preface

This thesis is mainly devoted to the study of many body effects in the electronic and optical properties of the (111) surface of diamond.

C(111) is the cleavage surface of diamond and the interest on its properties is based on both fundamental and applicative reasons. On one side, in fact, it is one of the growth surfaces for the chemical vapour deposition (CVD) of diamond. CVD allows the fabrication, at a relatively low cost, of diamond films, thus providing the possibility of exploiting the extreme and peculiar properties of diamond and of its surfaces for applications. On the other side, it is a prototype, apparently simple, surface, and for this reason it is quite amazing that up to now no *ab-initio* calculation was able to even qualitatively reproduce its, experimentally found, semiconducting character. The lack of knowledge regarding this surface (in contrast with the other low index surfaces of diamond, silicon and germanium) has to be related to the fact that the state of the art method for electronic properties (namely G_0W_0) could not be applied in the usual scheme.

Experimentally, right after cleavage, the (111) surface of diamond exhibits a 1×1 periodicity, and thus apparently no reconstruction, because hydrogen saturates the dangling bonds. After annealing at more than 1000 K, however, hydrogen desorbs and surface undergoes a 2×1 reconstruction which is the lowest energy configuration. Both experiments and theory agree that the reconstruction geometry is the Pandey chain model, but the details of this reconstruction, i.e. dimerisation or buckling of the surface chains, are still not fully experimentally clarified. All well converged *ab-initio* calculations predict an unbuckled and undimerized geometry, however the surface appears semimetallic in contrast to experimental evidence. Even the introduction of quasi-particle corrections within the usual G_0W_0 scheme does not lead to the opening of the gap between surface states.

As we will see in the next chapters, the peculiarity of the band structure of the surface, and its dependence on the geometry details, do not allow to

understand clearly wether the discrepancy between theory and experiment is due to a wrong equilibrium geometry or to the lack of quasiparticle effects in the DFT band structure. The main task of this thesis is to solve this problem.

To deal with this issue we have used the state of the art methods for the excited state properties within the Green's function theory. First of all, we have confirmed the previously found geometry and DFT semimetallic band structure. Then, having seen that the introduction of the self energy effects within the usual G_0W_0 scheme does not improve the results, we have calculated the quasiparticle energies within an iterative GW scheme. In this scheme the GW calculations were repeated updating the quasiparticle energies until selfconsistency was reached. A gap of about 1 eV was found between the surface states, thus finally solving the discrepancy between theory and experiment.

Unluckly no inverse photoemission data exists for this surface to directly compare with our band gap results. This is why, as shown in the last part of the thesis, we are computing the optical spectra of the surface. Optical experiments, in fact, while not accessing directly the electronic band gap, give important insights on the transitions between occupied and unoccupied surface states, thus yielding relevant informations on the band structure itself, especially when compared to *ab-initio* calculations that are able to include also the effects of the electron-hole attraction, namely the excitonic effects.

This thesis is organised as follows: in the first chapter a brief introduction on diamond (especially on its (111) surface), with an overview on the technological relevancy of this material, is presented; then, in the second chapter, we introduce the results on the ground state properties of the surface and the theory on which these calculations are based, namely Density Functional Theory. In the third chapter we will start dealing with the Many Body Perturbation Theory and see how quasiparticle effects can be introduced in the band structure calculations within the GW approximations of the self energy; we will apply this formalism to the band structure of our surface and we will see how in this case an iterative calculation is needed. Finally, in the last chapter, we will present the results for the optical spectra of this surface computed within the Bethe-Salpeter equation formalism.
Buona lettura!

Chapter 1

Diamond and its (111) surface.

In this chapter we will introduce the main properties of diamond. We will see that diamond is a very attractive material, and we will try to elucidate why properties specific of its surfaces make it now even more appealing for technological applications, ranging from optoelectronic to biosensing. We will then introduce its cleavage surface, namely the (111) surface, and we will look at what it is known up to now about it. We will start from understanding its geometry and reconstruction, finishing with its electronic properties. We will see how the theoretical predictions for this surface are in disagreement with the experimental findings also at a qualitative level; not being able even to qualitatively state if the surface is semiconducting or semimetallic. This problem is the main topic faced during this thesis.

1.1 Diamond

Diamond is an extreme material. Besides being extremely beautiful it is also unique for its physical and chemical properties: it is the hardest material known, it possesses the highest thermal conductivity at room temperature, it is chemically extremely inert, biocompatible and transparent from the deep UV to far infrared.

Diamond is made of a single element, carbon. As it is well known, there is another crystalline form of pure carbon: graphite. If one has ever hold a pencil in his/her hand, he/she has experienced that, besides being informed that the gem on the ring and the tip of the pencils are made of the same atoms, the properties of the two materials are completely different. They are somewhat complementary: graphite is black, and soft, and (this maybe we have not tested holding the pencil in the hand) conductive. The completely different properties stem from the different kind of hybridisation and of the subsequent bonding of the valence electrons of the carbon atoms. The resulting crystal structure

is completely different, as we can see from fig. 1.1. In graphite the carbon

(a)   (b)
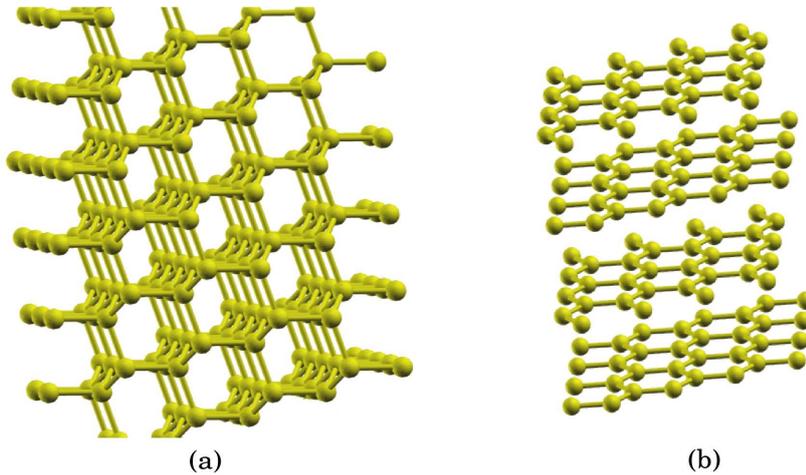
Figure 1.1: Diamond crystal structure (a) compared to the graphite crystal structure (b). The different kind of bondings lead to completely different geometries. Graphite is characterised by weakly bounded layers, whereas diamond is a cubic close packed structure.

atoms are sp^2 hybridised; the lattice is formed by weakly bounded planes of atoms. The atoms within the plane form an hexagonal lattice through the bonds between the three sp^2 electrons; the interaction among the planes is mediated by the remaining p orbitals. Instead, in diamond, the electrons are sp^3 hybridised; as a consequence, each carbon atom is bonded to other four carbon atoms, and all the bonds are equivalent; the resulting lattice is a closed packed cubic lattice.

At normal temperature and pressure, graphite is more stable than diamond and this is the reason why its natural occurance is rare compared to graphite; diamond is created, probably in the earth's mantle, under extremely high pressure and temperatures. However we do not see a conversion of diamond in graphite as soon as it is brought to the normal environment; this is due to the existence of a very large energy barrier for the conversion process. So, even if diamond is not thermodynamically stable at room temperature and pressure, it is kinetically stable. The ones owning diamonds can feel safe, it is highly improbable that they will wake up some day finding the gem turned to a black opaque stone. Diamond is forever, at least on a human time scale.

1.1.1 The role of the diamond and of its surfaces in novel applications

The first problem one has to face wanting to use diamond for any application is its availability and cost. Other problems are specific of the different kind of applications, for example in electronic applications one would like to find a way to dope diamond, or in biological applications it is important to functionalise it; we will rapidly go through each issue, and especially underline the importance of surfaces properties.

Let's first start form the *availability issue*. The first attempts that succeeded in producing diamond artificially are dated back in 1955 [1], the idea was to reproduce the conditions under which diamond forms in nature. The process is called High Pressure High Temperature (HPHT) synthesis: graphite is put under these conditions and with the help of a transition metal catalyst it turns into diamond. Diamond formed in this way has dimensions that range from the micrometer to the millimiter and are used in mechanical application as superabrasive for polishing. However, until recently, HPHT diamond is not enough pure and defect free to be used in more refined applications such as electronics.

Another possibility is to get diamond at lower pressures, synthetizing it from a vapour phase, this technique is called Chemical Vapour Deposition (CVD). The first successful attempt to grow diamond powder at subatmospheric pressure was reported in 1969 [2]; since then, the techniques have developed and now diamond is typically grown using a dilute mix of hydrocarbon (like methane) in hydrogen. Roughly what happens(see [3] and references therein) is that atomic hydrogen is formed for example through an electric discharge in the gas; atomic hydrogen, then, reacts with the hydrocarbon forming some reactive species like CH_3 and C_2H_2 which in turn, transported to the substrate surface, are converted into diamond by a series of still not well understood reactions. Even from this rough description, it is clear that surface chemistry plays a crucial role also in solving the basic problem of low cost availability of diamond. Until very recently, however, CVD diamond could not be used in practical applications because the films were polycrystalline, with different orientation for each crystal. Now, see [4, 5] and references therein, new HPHT methods have been developed that yield ultrapure diamond samples on top of which CVD highly oriented diamond films can grow. The single-crystal diamond synthetized in this way is reaching the quality required for applications.

A significant use of any material in electronic applications, relies strongly on the possibility of controlling its conductivity by providing it of additional carriers through doping. The use of diamond within electronic is desirable in

order to exploit its extreme properties: its great stability yields a *'superior resistance to ionising radiation and high energy particle'* [6] for devices that are able to operate in space; its wide band gap ($E_g = 5.5$ eV) makes it the most promising material for ultraviolet light emission [7]; and its chemical composition, carbon atoms only, make it extremely biocompatible. The inclusion of boron and phosphorus during the CVD growth process allows p-type and n-type doping respectively; the possibility of transforming diamond from insulator to semiconducting opened up the chance of realizing novel electronic and optoelectronic devices such as an ultraviolet diamond emitting diode [8], and a diamond microelectrode array device which is expected to find applications in tools working in paticularly harsh environments [9]. Moreover an enhanced crystalline quality lead to a charge mobility significantly greater than the electron mobility measured for SiC and GaN, two other wide-gap semiconductors currently under study for high-frequency and high-power applications [4, 5, 10]. Despite all of this, the doping mechanism for diamond presents many problems, the main one is the fact that the activation energies for the charge carriers are big (0.38 eV in the case of boron and 0.43 eV in the case of phosphorus) so that at room temperature only a small fraction of the acceptors or donor levels are activated; a possible solution is overdoping, however overdoping can ruin the transport properties of the sample by introducing more defects and scattering centres.
Nevertheless diamond surfaces can be doped without the need of the introduction of foreign atoms in the lattice: as a matter of fact, diamond exhibits a p-type conduction observed when hydrogenated diamond surfaces are exposed to air without the need of introducing any impurities in the lattice [11]. A similar effect has been reported when C_{60} molecules are evaporated on a diamond (111) hydrogenated surface [12]. In this last case the surface conductivity has been explained with a sub-surface hole accumulation layer produced by a charge transfer from the surface to the adsorbate. A similar mechanism could also be responsible for the first case. The discovery of this surface conductivity has lead to the creation of a novel field effect transistor [13].

Diamond is also a very good candidate for developments of new bio-sensors and detectors, especially for its extreme surface inertness which makes it harder to be rejected for both structural and sensor application in the body. Recently an artificial retina coated with diamond ultrananocrystals has been produced [14], the diamond coating is needed to make the device biocompatible and to protect it from the fluids presents in the eye. Another important characteristic of diamond from a biotechnological point of view, is the fact that its surface can become an active substrate, thanks to the surface conductivity and to the functionalisation of its surfaces. Hydrogenated diamond surfaces have in fact

been functionalised attaching on them green fluorescent protein, the enzyme catalase [15] and DNA [16]. Despite of the fact that the molecules were immobilised at the surface, they were still fully functional and active. Moreover the biomolecules are covalently bounded to the diamond surface and the stability of this bonding provides an important advantage with respect to the normally used metal electrodes such as gold.

We have seen how diamond's potentialities stem not only from bulk properties, but also from its surface chemistry and surface properties; carrying out all this potentialities relies strongly on the control and on the understanding of the chemical and physical properties of this surfaces. Moreover, understanding the behaviour of its surfaces is important also from a theoretical point of view, in order to elucidate the principles of surface reconstructions of the other group IV semiconductors, Si and Ge.

1.2 The (111) surface of diamond

C(111) is the cleavage surface of diamond and it is also one of the growth plane of the CVD process; moreover, the incorporation rate of boron in the (111) face is greater than in the other CVD growth plane, namely the (100) surface. Despite all of this, the geometry of the reconstructed C(111)(2 × 1) surface and the electronic band structure that originates from it are still under debate.

1.3 Geometry and reconstruction

Let's consider bulk diamond: one of the four sp^3 orbital of each carbon atom is parallel to the (111) direction. As a consequence, if we cut diamond perpendicularly to the 111 direction, the terminated bulk can present either one or three dangling bonds. The three dangling bonds surface has a much higher formation energy and by cleavage only the single dangling bond surface can be obtained; all the results presented below, and the study of this thesis refer to the single dangling bond surface.
The single dangling bond truncated bulk is shown in fig. 1.2. From the side view we notice that the ideally truncated bulk is formed by a series of atomic layers spaced by a longer and a shorter interval, leading to a bilayer structure. The important thing to notice is that the dangling bonds are not on nearest neighbour; we can thus expect a strong modification of the structure in order to bring the dangling bonds closer and allow them to saturate each other [17].
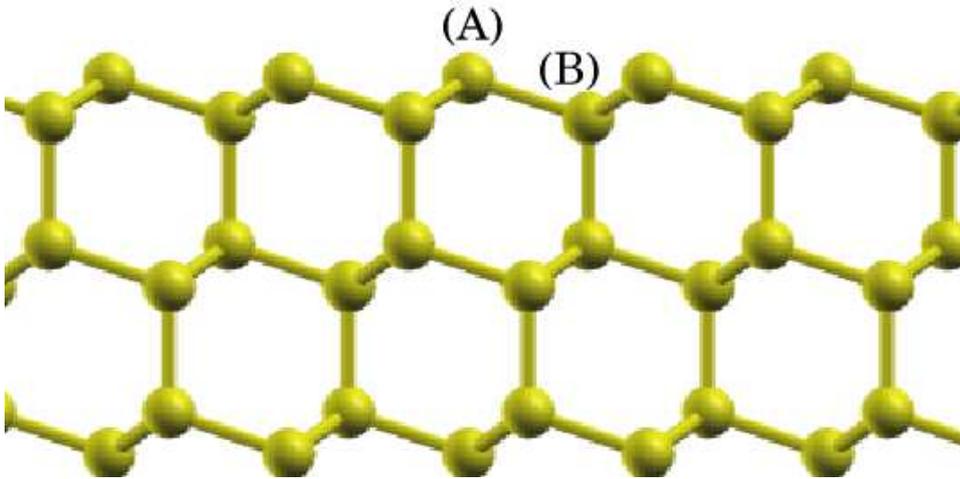
Figure 1.2: Top view of the ideally truncated bulk. The single dangling bond face is presented.

Right after cleavage the surface exhibits a 1×1 reconstruction, which consist of approximately a truncated bulk with each dangling bond saturated by an hydrogen atom. The presence of hydrogen in the polished surface is confirmed by angle resolved photoemission spectra (ARPES) that show that in the 1×1 no occupied surface states are present [18, 19] within the fundamental gap, consistently with the picture of hydrogen saturating the dangling bonds; moreover, infrared-visible sum frequency generation spectroscopy has shown the presence of structures related to C-H bonds [20].

After annealing at more than 1100 K, hydrogen desorbs and the surface undergoes a 2×1 reconstruction. It is now well accepted that the reconstruction geometry of this surface is the Pandey π-chain model [21], shown in fig. 1.3. As already mentioned, the reconstruction involves significant changes in the bonding of the atoms belonging to the first three layers. In each surface unit cell, one atom of the second layer (the atom labeled B in fig. 1.2) breaks its bond with the third layer and rises up to the first layer; at the same time, an atom on the first layer (labeled A) goes down to the second layer bonding to the third layer (compare fig. 1.2 and 1.3). As a result, the dangling bonds
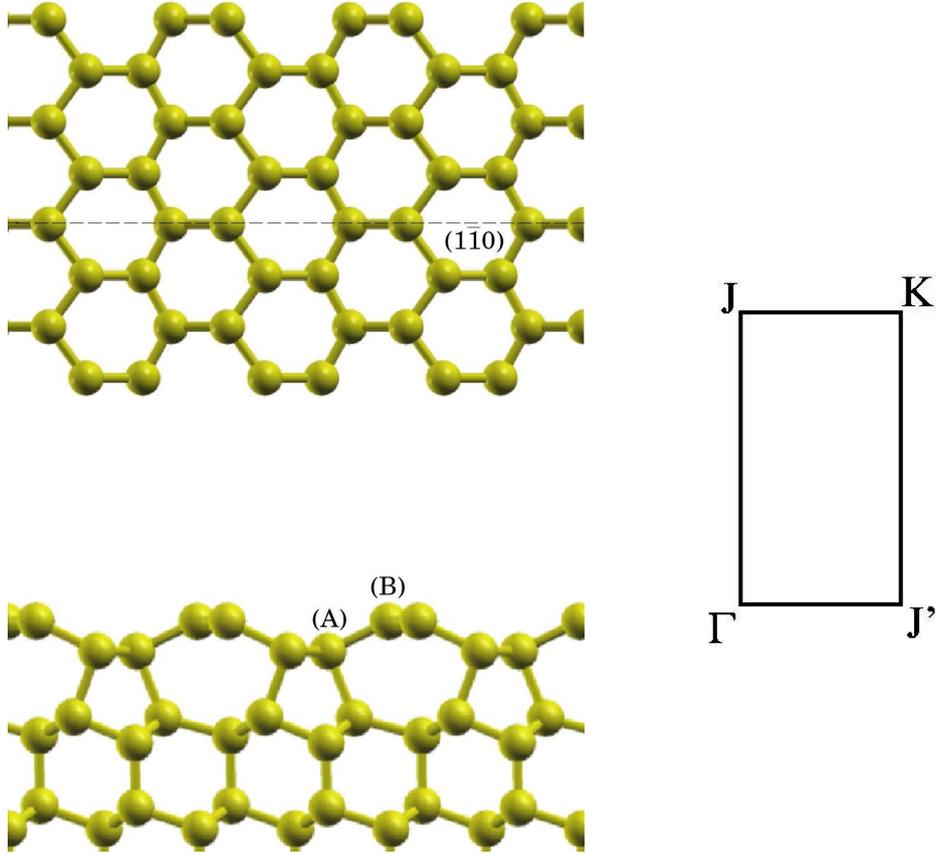
Figure 1.3: Pandey chain model for the (111) 2×1 surface of diamond. The reconstruction involves a significant change in the lattice structure. The irreducible Brillouin zone and its high symmetry points are also shown.

are now on nearest neighbour atoms in the first layer. The electrons on the chain atoms partially dehybridize to three sp^2 and one p_z orbitals, and the p_z like dangling bonds can saturate each other through delocalised π orbitals forming an almost one dimensional chain developing along the y direction. The Pandey chain reconstruction finds the most evident confirmation through the dispersion of the occupied surface states obtained by ARPES measurements [18, 22]. ARPES finds a surface state that is almost dispersionless along the JK and $\Gamma J'$ lines (see inset fig. 1.3), i.e. in the direction perpendicular to the chains; and that at the same time exhibits a strong dispersion along the ΓJ and the $J'K$ line, i.e. parallel to the chain direction. This kind of dispersion is a signature of a one dimensional system, like the Pandey chains.

Besides the dispersion of the occupied states, also the great majority of ex-

perimental results, ranging from medium energy ion scattering [23, 24], low energy electron diffraction (LEED) analysis of the spot intensity vs voltage [25, 26], X-ray diffraction structure analysis [27], confirm the Pandey model for the reconstruction of the (111) surface. In this sense diamond behaves in the same way as silicon and germanium, which exhibit a Pandey chain 2×1 reconstruction.

However, even if there is general consensus regarding the π-bonded chain model, the exact geometry of the unit cell is still controversial; preserving the 2×1 periodicity, there are still two possible degrees of freedom:

- *Intrachain buckling:* the atoms on the chain are not on the same heights. As a result, the two atoms on the chain are not equivalent anymore and there is a charge transfer.

- *Intrachain dimerisation:* the bonds along the chain are not equivalent, one is shorter (almost like a double bond) than the other, and the final structure appears as a sequence of dimers. This kind of distortion breaks the reflection symmetry with respect to the $(1\bar{1}0)$ plane (see fig. 1.3)

While most of the theoretical study within Density Functional Theory (see chap. 2), predict a non buckled non dimerised geometry [17], the situation on the experimental side is rather unclear. Dimerisation seems to be ruled out by ARPES experiment [22], because the mirror symmetry with respect to the $(1\bar{1}0)$ plane is preserved. Medium energy ion scattering data [23], X-ray diffraction structure analysis [27], and infrared-visible sum frequency generation measurements [20] find best agreement for an atomic arrangement featuring tilted chains, but can not completely rule out an unbuckled undimerized geometry. At the same time, core level binding-energy measurements [28], and a recent LEED study [26], see no evidence for buckling.

1.4 Electronic properties, the puzzle of the electronic gap

As already mentioned in the previous section, diamond, silicon, and germanium share the same model for the reconstruction of their (111) 2×1 surface. However, while in the case of germanium and silicon the most stable geometry features tilted chains, diamond chains do not seem buckled [29]. This tiny difference leads to a strong qualitative difference concerning the corresponding band structures: at a DFT level germanium and silicon have a semiconducting surface, whereas diamond appears semimetallic (see fig. 1.4). Experimentally, direct photoemission experiments, which give us the information about the
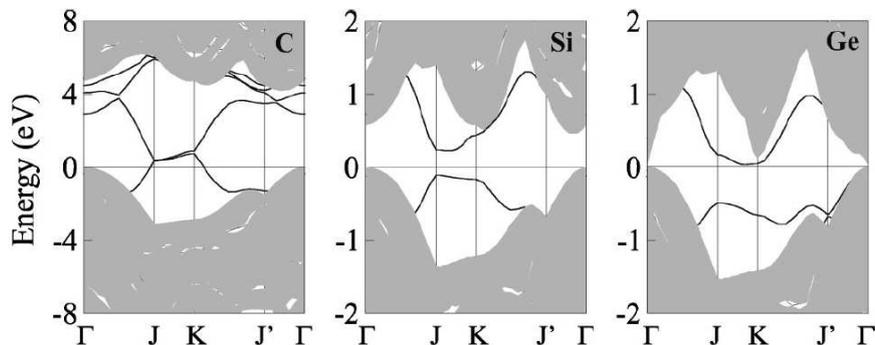
Figure 1.4: DFT band structure for the $(111)2 \times 1$ surface of diamond, silicon and germanium from ref. [29]. As a consequence of the different relaxation details, the band structures differ qualitatively. Germanium and silicon surfaces are semiconducting, while the diamond surface is semimetallic.

occupied surface states, find occupied surface states within the fundamental gap. They show a strong dispersion along the chain direction and are almost flat perpendicularly to it [18, 22], consistently with a reconstruction model exhibiting one dimensional structures like the π bonded chains. However, most importantly, in ref. [22] the surface states are reported to reach their maximum energy 0.5 eV below the Fermi level (see. fig. 1.5). In order to know the experimental band gap we should now know how much low in energy the unoccupied surface band goes. Unfortunately, very little is known about unoccupied surface states: there are evidences of unoccupied surface states from photon photoelectron spectroscopy (2PPES) [30], high resolution soft X-rays absorption spectra [31], and electron energy loss spectroscopy (EELS) [28]. However, all this spectra can not be directly compared to the calculated band structure: excitonic effects (see chap. 4) must be included in the calculations in order to compare to this kind of experiments. So what can be inferred by the experiments is that diamond (111) 2×1 has an electronic gap of at least 0.5 eV.

Where does the disagreement between the calculated band structure and the experimental results come from? In ref. [17] it is shown that the degeneracy of the surface states along the JK line of the Brillouin zone is connected to the equivalence of the two atoms on the chain. In the case of silicon and germanium buckling enhances this equivalence and the degeneracy is lifted. For diamond this seems not likely because the elastic energy loss, due to the distortion of the lattice, prevails with respect to the gain in energy that could be obtained through the electronic band structure. This seems to be in line with a general
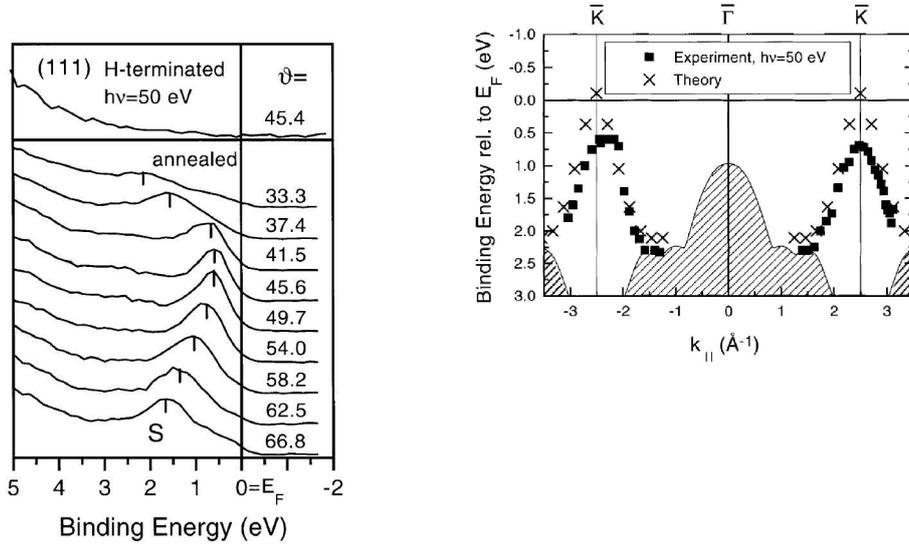
Figure 1.5: ARPES spectra from ref. [22]. The occupied surface state reaches its maximum energy 0.5 eV below the Fermi energy at K.

chemical and physical trend. Within this trend, diamond has an opposite behaviour regarding reconstructions with respect to silicon and germanium. For example, silicon and germanium present very large reconstruction as their most stable geometry (a 7×7 reconstruction for silicon [32], and a $c(2 \times 8)$ for germanium [33]) whereas diamond's most stable geometry has a 2×1 periodicity. The long range reconstructions are not realized in the case of diamond as a consequence of shorter interatomic distances and stronger bonds [29]; and in general for diamond there is the tendency of stabilisation of more symmetric structure.

Despite all this arguments, that justify the presence of an unbuckled and undimerized geometry for diamond it could be possible that the theory is unable to catch some symmetry breaking mechanism yielding a semiconducting band structure. Another possibility is that the discrepancy arises from the fact that, as we will see in more detail in the next chapter, DFT band structures systematically underestimate the electronic gaps of materials. In this case the inclusion of quasiparticle effects could solve the problem.
The main task of this thesis is to try to understand wether the insulating character of the (111) surface of diamond is a consequence of asymmetric geo-

metrical changes, or of many body effects. We will see that many body effects determine the insulating character of this surface.

Chapter 2

Ground state properties

In this chapter we first review the theory that lies behind the ground state calculations performed during this thesis. We illustrate the methods of calculation and show the results for the (111) surface of diamond. We also illustrate the convergency tests done in order to check the impact of the approximations on the results. We will see how, confirming previous results based on the Local Density Approximation (LDA) for the exchange correlation potential, also within the Generalised Gradient Approximation (GGA) we find an undimerized and unbuckled geometry (see previous chapter). The equivalence of the two atoms on the chain is not broken even by the introduction of the spin degrees of freedom within Local Spin Density Approximation (LSDA) that would possibly allow an antiferromagnetic ordering of the electrons' spin along the chain. The resulting band structure is thus semi-metallic in contrast with experimental findings. We finally discuss the meaning of DFT band structure and analyse the link between the geometry and the resulting band structure.

2.1 The many body problem

If we do not take into account relativistic effects, which are negligible in the range of energy of our interest, a system of interacting electrons and ions, like for example a solid or a molecule, is ruled by the many body Hamiltonian:

$$\hat{H} = \sum_{I=1}^{N_{ion}} -\frac{1}{2}\nabla_I^2 + \sum_{I=1}^{N_{el}} -\frac{1}{2}\nabla_i^2 + \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J}{|R_I - R_J|} + \frac{1}{2}\sum_{i\neq j}\frac{1}{|r_i - r_j|} + \sum_{i,I}\frac{-Z_i}{|r_i - R_I|}.$$

(2.1)

The first two terms in this Hamiltonian refer respectively to the kinetic energy of the ions (referred to with capital index) and of the electrons, then the ion-ion interaction, the electron-electron interaction and the ion-electron interaction are taken into account. In principle to be able to predict the ground state

properties of this system, the minimisation of the total energy given by the expectation value of 2.1 with respect of both the electronic and ionic degrees of freedom, or else the knowledge of its eigenstate and eigenvalue would be required. Of course this is a very complex task and the solution of it can not be achieved for realistic systems.

However the ions and the electron mass are orders of magnitude different, and since the forces acting on them are the same, we can assume that the electrons instantaneously reach their ground state for each ionic configuration. This fact, or in other words, the fact that the dynamic of ions and electrons involve well separated ranges of energies, allows us to decouple the ionic and the electronic degrees of freedom: this is the well known Born-Oppenheimer approximation [34].

Within the the Born-Oppenheimer approximation the electronic Hamiltonian now becomes:

$$H = \sum_i -\frac{1}{2}\nabla_{\mathbf{r}_i}^2 + V_{ext} + \frac{1}{2}\sum_{i,j}' \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} = T + V_{ext} + V_{e-e}, \qquad (2.2)$$

even though now the ionic coordinates appear as parameters in V_{ext}, eq. 2.2 is still very involved and directly diagonalising it to find its eigenstates is still a formidable task that for complex system is out of the reach of the computational power. So, in order to tackle this problem, it is required to somehow change point of view. We will see, in the next section, that, if one is interested in ground state properties only, maybe the complete solution of the problem set by eq. (2.2) is not necessary. The density functional approach, focusing the attention on the the *simpler* total electronic density $n(r)$ [1] rather than on the more complex many body wavefunction $\Psi(r_1, ..., r_N)$, opens the way to the study of the ground state properties of even very complex systems.

2.2 Minimising the electronic energy: Density Functional Theory

Density Functional Theory (DFT) treats the case of an external, time independent potential, and is based on the seminal paper of Hohenberg and Kohn of 1964[35]. For a review on DFT see for example [36].

---

[1]The electronic density is given by: $n(r) = N \int \Psi^*(r, r_2, ..., r_N)\Psi(r, r_2, ..., r_N)$

2.2.1 Basic theorem: Hohenberg and Kohn

The Hohenberg and Kohn theorem states that once the mutual interaction among the electrons is fixed, in a non time dependent situation and assuming that all the potentials acting on the system are local, the knowledge of the complete Hamiltonian (i.e. the external potential within a constant) and the knowledge of the ground state density of the system represent an equivalent information. This means that, the complete Hamiltonian of the system is univocally determined by the ground state density alone, just like, more obviously, viceversa the ground state density is determined by the full Hamiltonian. In a more concise way what we we want to prove is that:

> *The ground state expectation value of any observable (including the ground state energy) is a unique functional of the exact ground state density*

and also that:

> *In particular, the energy functional, once the external potential is specified, finds it minimum at the ground state density of the system*

We will prove these two statements only for the case of a non degenerate ground state, a generalisation also to degenerate ground state can be found in ref. [36].

We start by considering the Hamiltonian given by eq. (2.2), where, V_{ext} belongs to the set of external potentials $\mathcal{V}$, that lead to a non degenerate ground state. We can associate a ground state wavefunction ψ, to each element $v \in \mathcal{V}$ by the solution of the corresponding Schrödinger equation. We collect all the possible ψ generated in this way in a set, which we will call Ψ. The Schrödinger equation defines a map $\mathbf{C}$ between the sets $\mathcal{V}$ and Ψ:

$$\mathbf{C} : \mathcal{V} \rightarrow \Psi. \tag{2.3}$$

Now, we define a second map $\mathbf{D}$, that associates to each element of Ψ the corresponding density $n(r)$ which will be collected in the set $\mathcal{N}$:

$$\mathbf{D} : \Psi \rightarrow \mathcal{N}. \tag{2.4}$$

We know, by construction, that to each element of Ψ corresponds at least one element of $\mathcal{V}$ and that to each element of $\mathcal{N}$ corresponds at least one element of Ψ, i.e. that the two maps $\mathbf{C}$ and $\mathbf{D}$ are surjective. In their paper, Hohenberg

and Kohn proved with a *reductio ad absurdum* that the two maps are one-to-one maps and thus are fully invertible. This means that the ground state density univocally determines the Hamiltonian of the system. As a first thing we will prove that **C** is fully invertible:

suppose that two different potentials, v and v', that differ by more than a constant, lead to the same ground state $|\psi\rangle$, this means that:

$$\hat{H}|\psi\rangle = (\hat{T} + \hat{V}_{e-e} + \hat{V})|\psi\rangle = E_{gs}|\psi\rangle$$

$$\hat{H}'|\psi\rangle = (\hat{T} + \hat{V}_{e-e} + \hat{V}')|\psi\rangle = E'_{gs}|\psi\rangle. \tag{2.5}$$

$$\tag{2.6}$$

Now, subtracting term by term the two equations, we get:

$$V - V' = E_{gs} - E'_{gs}$$

since the two potentials are local, if the wavefunctions are not null in a set of finite measure [2], the two potentials differ only by a constant, which is against our assumption.

So, what we have just seen is that each ground state belongs only to one possible potential; if now we prove that to each density we can associate only one ground state, we can finally say that the density determines external potential, and so the entire Hamiltonian.

From the variational principle we know that:

$$E_{gs} = \langle\psi|H|\psi\rangle < \langle\psi'|H|\psi'\rangle; \tag{2.7}$$

but

$$\langle\psi'|H|\psi'\rangle = \langle\psi'|H' + V - V'|\psi'\rangle = E'_{gs} + \int n'(r)(v(r) - v'(r))d^3r, \tag{2.8}$$

if now we put (2.8) into eq. (2.7) we get that

$$E_{gs} < E'_{gs} + \int n'(r)(v(r) - v'(r))d^3r. \tag{2.9}$$

Repeating the same argument for the case of $\hat{H}'$, we get:

$$E'_{gs} < E_{gs} + \int n(r)(v'(r) - v(r))d^3r; \tag{2.10}$$

---

[2]This is true if the potentials are sufficiently well behaved, i.e. they don't show infinite barriers etc . . .

if we now assume that the two ground state wavefunctions lead to the same density, i.e. $n(r) = n'(r)$, we can add term by term eq. (2.9) and eq. (2.10), obtaining $E_{gs} + E'_{gs} < E'_{gs} + E_{gs}$, which makes no sense. This means that to each density $n(r)$ corresponds only one possible ground state wavefunction. Since each ground state determines univocally an external potential, the density itself determines the external potential.

So, univocally, the Hamiltonian can be written as a functional of the ground state density $\hat{H} = \hat{H}[n]$; and the same happens for total ground state energy $E = E[n]$.

Now we fix the external potential $v = v_0$, we know that a ground state density n_0 corresponds univocally to this potential. Since the map $\mathbf{D}^{-1}$ that links the densities to the ground state wavefunctions does not depend on the specific choice of v_0, the variational principle assure us that:

$$E_0 = E_{v_0}[n_0] = \min_{n \in \mathcal{N}} E. \tag{2.11}$$

An other important consequence of the independence of $\mathbf{D}^{-1}$ from v_0 is that we can divide the total energy functional into two parts:

$$
\begin{aligned}
E_{v_0}[n] &= \langle \psi[n] | T + V_{e-e} + V_0 | \psi[n] \rangle \\
&= \langle \psi[n] | T + V_{e-e} | \psi[n] \rangle + \int V_0(r) n(r) d^3r \\
&= F_{HK}[n] + \int V_0(r) n(r) d^3r; \tag{2.12}
\end{aligned}
$$

where $F_{HK}[n]$ is a *universal* functional of n, non depending from the choice of v_0

So, in principle, knowing the form of the functional dependence $E[n]$, one could obtain the ground state density and energy of the system implementing its minimisation with respect to the density $n(r)$. However the general form of $E[n]$ is not known, in the next section we will see how the Thomas-Fermi approximation can actually be recovered as a certain form of $E[n]$. The known limits of the Thomas-Fermi approximation can then give an hint on where care should be put to be able to use DFT in practice.

Thomas-Fermi approximation

The Thomas-Fermi equations can be obtained using a certain approximation for the total energy functional $E[n]$. In this approximation the electrons' interaction is accounted to only via the classical electrostatic (Hartree) term:

$$E^{TF}_{e-e}[n] = \frac{1}{2} \int d^3r \int d^3r' \frac{n(r)n(r')}{|r - r'|}, \tag{2.13}$$

and the kinetic energy term is given by:

$$T^{TF}[n] = \frac{2}{(2\pi)^3} \int d^3r t_0(n(r)), \qquad (2.14)$$

where $t_0(n)$ is the kinetic energy density of a free electron gas with density n.

$$t_0(n) = \int_{|k|<k_F} d^3k \frac{k^2}{2m}. \qquad (2.15)$$

With k_F related to n through the expression $k_F = (3\pi^2 n)^{1/3}$. In this way the kinetic energy becomes:

$$T^{TF}[n] = C \int d^3r (n(r))^{5/3}. \qquad (2.16)$$

Where $C = (3\pi^2)^{(5/3)}/2m$. If we now minimise the total energy functional constraining the number of particle through the use of a Lagrange multiplier, we obtain the Thomas-Fermi equations:

$$\frac{5}{3}C(n(r))^{2/3} + \int d^3r' \frac{n(r')}{|r-r'|} + V_{ext}(r) + \lambda = 0. \qquad (2.17)$$

It has been shown that the Thomas-Fermi approximation for the total energy is exact in the limit of an infinite nuclear charge [37], and that this approximation gives a rough description of the charge density. However the Thomas-Fermi atoms do not bind to form molecules or solids nor have a shell structure. So, even if the Thomas-Fermi theory have the merit of having focused the attention on the *correct* variable, i.e. the electronic density, it can not be used, even in its most refined versions, to get reliable predictive calculation.

2.2.2 Kohn and Sham single particle scheme

It is found that the most of the drawbacks of the Thomas Fermi approach are related to the approximation made to the kinetic energy T of the interacting N-electron system. In fact it constitutes a non negligible part of the total energy and an approximation made directly on this term can lead to large errors in the total energies of the system. The breakthrough appeared one year after the Hohenberg and Kohn theorem had been published, with the paper of Kohn and Sham [38], where the authors introduced an ad hoc separation of the terms contributing to the total energy E. The two main advantages of this separation were: first, to provide a single particle scheme to obtain the ground state density and total energy, and second, but still very important, to

have an expression in which the approximations to the unknown part would be, in many cases, not as relevant as before. In the KS scheme the total energy functional of the interacting system is written as:

$$E[n] = T_{KS}[n] + E_H[n] + \int d\mathbf{r} n(\mathbf{r}) v_{ext}(\mathbf{r}) + E_{xc}[n], \qquad (2.18)$$

T_{KS} is the kinetic energy of a non interacting system with density n, E_H is the Hartree contribution to the total energy:

$$E_H[n] = \frac{1}{2} \int d^3r \int d^3r' \frac{n(r)n(r')}{|r - r'|}, \qquad (2.19)$$

and E_{xc} is the remaining part of the total energy which contains exchange-correlation contributions plus the difference between the kinetic energy of the interacting N-electron system T and the kinetic energy T_{KS} of the non-interacting system. The main idea underlying this approach is to map the study of the (complicated) interacting system into the study of a (simpler) non interacting fictitious system whose Hamiltonian is written as:

$$[-\frac{1}{2}\nabla^2 + v_{ext} + v_H + V_{xc}]\phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}), \qquad (2.20)$$

and which has by construction the same ground state density of the interacting system:

$$n(r) = \sum_i f_i |\phi_i(\vec{r})|^3, \qquad (2.21)$$

with f_i being the occupation number of the state i. In Eq. 2.20 $V_{xc} = \frac{\delta E_{xc}[n]}{\delta n}$ is the exchange and correlation potential of the interacting system, and V_{ext} is the same external potential of eq. 2.18.

It is now possible, given an approximation for $E_{xc}[n]$, to solve the Kohn-Sham equations (2.20,2.21) self consistently and calculate the density of the interacting (real) system via Eq. 2.18. Once the density is known, it is possible to calculate the energy of the ground state of the interacting system and hence, by proper minimisation, to find its ground state geometry.

2.2.3 Exchange-correlation energy

If we compare eq. 2.12 and eq. 2.18 we see that the exchange and correlation energy is defined as:

$$\begin{aligned} E_{xc}[n] &= F_{HK}[n] - T_{KS}[n] - E_H[n] = \\ &= T_{HK} - T_{KS} + \langle \Psi[n]|V_{e-e}|\Psi[n]\rangle - E_H[n], \end{aligned} \qquad (2.22)$$

where $\Psi[n]$ is the true ground state wavefunction, ϕ_{KS} is the Kohn and Sham wave function that is obtained through the selfconsistent scheme described in the previous section and that usually is a single slater determinant, T_{HK} is the kinetic energy of the real system, e.g. the expectation value of the kinetic energy operator on the state $\Psi[n]$.

Now, we can divide E_{xc} into two parts:

$$E_{xc}[n] = E_x[n] + E_c[n]. \tag{2.23}$$

The exchange energy E_x is defined, following an Hartree-Fock like definition, as:

$$E_x[n] = \langle \phi_{KS}|V_{e-e}|\phi_{KS}\rangle - E_H[n]. \tag{2.24}$$

It is very important to notice that the exchange energy corrects the spurious self-interaction present in the Hartree energy, just like the Fock term does in the Hartree-Fock scheme. The correlation energy E_c is defined as the difference between E_{xc} and E_x; comparing eq. 2.22 and eq. 2.24, E_C is given by:

$$E_c[n] = \langle \Psi[n]|T + V_{e-e}|\Psi[n]\rangle - \langle \phi_{KS}|T + V_{e-e}|\phi_{KS}\rangle. \tag{2.25}$$

We know that $\Psi[n]$ minimises the total energy of the interacting system, thus $E_c[n] < 0$. At the same time, we know that ϕ_{KS} minimise the total energy of the non interacting system and so we know that $\langle \Psi[n]|T|\Psi[n]\rangle - \langle \phi_{KS}|T|\phi_{KS}\rangle > 0$. So $E_c[n]$ is the sum of a positive kinetic energy contribution and a negative *potential* contribution, which anyhow prevails.

2.2.4 Approximations

Up to now, all the presented derivations are exact. However, to carry out the real calculations, one has to resort to some approximations. For example, as already mentioned, the exact functional dependence of V_{xc} with respect to the density is not known; or, more trivially, the infinite basis set needed to describe the electronic wavefunctions must be truncated at some point. In this section we will go through some of the approximations that have been used within the ground state calculations performed in this thesis.

Local Density approximation

The definitions (2.24), (2.25) while shading some light on the role of the exchange and correlation energy, are actually just formal definition that do not help in practice to build a good functional.

In their original paper [38], Kohn and Sham suggested a first simple approximation for E_{xc} valid in principle only for slow varying density systems: the so called Local Density Approximation (LDA) given by:

$$E_{xc}^{LDA} = \int d\mathbf{r}\, n(\mathbf{r}) \epsilon_{xc}^{\text{heg}}(n(r)), \qquad (2.26)$$

where $\epsilon_{xc}^{\text{heg}}(n)$ is the exchange-correlation energy per electron of a homogeneous electron gas of density n. As we can see from eq. 2.26, within LDA the exchange correlation energy depends only on the local value of the density. Actual calculations use parametrisation of $\epsilon_{xc}^{\text{heg}}$ based on accurate Monte Carlo calculation of the exchange correlation energy of uniform electron gas for different values of the density n [39, 40].

In their original paper Kohn and Sham wrote that they did not " *expect an accurate description of chemical bonding*" within LDA. However, despite the simplicity of the approximation, LDA turned out to be an excellent, very accurate, tool for computing ground state properties and describing the geometry of many, even very complex, systems.

Why LDA works so well?

In this paragraph we will try to get some hints on why LDA is so accurate. In fact this is important to get an idea of how to get better functionals beyond it. To do this we will have first to find another expression for E_{xc} within the coupling constant method, and then connect it to the pair distribution of the system. We will see that LDA fulfils, among others, the sum-rule for the exchange correlation hole, and thus, despite the crude approximation, provides a quite realistic description of the exchange and correlation potential also for non homogenous systems. In what comes next we have followed the very nice derivation of ref. [41].

E_{xc} within the coupling constant method

Let's consider a density $n(r)$ which is the ground state density of a system with Hamiltonian

$$H = T + V_{ext} + V_{e-e}$$

We introduce now a parameter λ that will modulate the strength of the interaction among electrons; at the same time we vary V_{ext} in order to force the system with Hamiltonian $H = T + V_{ext}^{\lambda} + \lambda V_{e-e}$ to have the same ground state

density of the original system. Of course, if $\lambda = 0$, the ground state wavefunction of the system $\psi^{\lambda=0}$ is the Kohn and Sham state ϕ_{KS}; at the same time, for $\lambda = 1$, $\psi^{\lambda=1}$ recovers the ground state wave function Ψ of the real system. If we suppose that the two limiting system are adiabatically connected varying λ from 0 to 1, then we can write the exchange correlation energy as:

$$
\begin{aligned}
E_{xc}[n] &= \langle \Psi[n]|T + V_{e-e}|\Psi[n]\rangle - \langle \phi_{KS}|T|\phi_{KS}\rangle - E_H[n] \\
&= \int_0^1 d\lambda \frac{d}{d\lambda}\langle \psi^\lambda|T + \lambda V_{e-e}|\psi^\lambda\rangle - E_H[n].
\end{aligned}
\tag{2.27}
$$

Using the Hellmann-Feynman theorem [3] to evaluate the derivative in eq. 2.27, we get:

$$
E_{xc}[n] = \int_0^1 d\lambda \frac{d}{d\lambda}\langle \psi^\lambda|V_{e-e}|\psi^\lambda\rangle - E_H[n].
\tag{2.29}
$$

The exchange-correlation hole

Introducing now the two particle reduced density matrix:

$$
\rho_2(r', r)^\lambda = N(N-1)\int dr_3 \cdots \int dr_N |\psi^\lambda(r', r, r_3, \cdots, r_N)|^2,
\tag{2.30}
$$

we can express the expectation value of V_{e-e} in terms of it:

$$
\langle V_{e-e}\rangle_\lambda = \frac{1}{2}\int dr \int dr' \frac{\rho_2(r', r)}{|r - r'|}.
\tag{2.31}
$$

Now, we can interpret $\rho_2(r', r)^\lambda$ as the probability amplitude of finding two electrons in r **and** in r'. So, $\rho_2(r', r)^\lambda$ can be expressed as the product of the probability of finding an electron in r (which is the definition of the density $n(r)$) times the probability of finding an electron in r' given an electron in r, which we will call $n_2^\lambda(r, r')$:

$$
\rho_2(r', r)^\lambda = n(r)n_2^\lambda(r, r').
\tag{2.32}
$$

We know that, since the number of particle is conserved,

$$
\int dr' n_2^\lambda(r, r') = N - 1.
\tag{2.33}
$$

---

[3]The Hellmann-Feynman theorem states that if H_λ is an Hamiltonian depending on a parameter λ, and Ψ_λ is a variational solution (including any eigenstate of H_λ) of energy $E_\lambda = \langle \Psi_\lambda|H_\lambda|\Psi_\lambda\rangle$; then

$$
\frac{dE_\lambda}{d\lambda} = \langle \Psi_\lambda|\frac{dH_\lambda}{d\lambda}|\Psi_\lambda\rangle
\tag{2.28}
$$

We divide now the expression of $n_2^\lambda(r, r')$ into two parts:

$$n_2^\lambda(r, r') = n(r') + n_{xc}^\lambda(r, r'). \qquad (2.34)$$

The $n_{xc}^\lambda(r, r')$ term takes into account the fact that since there is already an electron in r, the probability amplitude of finding an electron in r', can not be anymore the simple $n(r')$, the effect of the interaction and of the exchange among the electrons modify its shape. $n_{xc}^\lambda(r, r')$ is called the exchange correlation hole, for what we will say next. In fact, comparing eq. 2.33 and eq. 2.34, we see that:

$$\int dr' n_{xc}^\lambda(r, r') = -1. \qquad (2.35)$$

This means that the contribution of $n_{xc}^\lambda(r, r')$ to $n_2^\lambda(r, r')$ is negative. Indeed we can think that the presence of an electron in r pushes the other electrons away from it, creating a sort of *hole* around it. The presence of this hole is due to:

- **Exchange effects:** The exchange effects that take into account that electrons of the same spin repel each other via the Pauli exclusion principle, and also corrects the self-interaction term.

- **Correlation effects:** The correlation effects that take into account the Coulomb repulsion, which keeps the electrons of any spin apart.

It is very important to notice that the exchange effects are present in any system (also non interacting) which is ruled by the Fermi-Dirac statistic. Now we try to distinguish the effect of correlation and exchange in n_{xc} by writing:

$$n_{xc}^\lambda(r, r') = n_x(r, r') + n_c^\lambda(r, r'), \qquad (2.36)$$

$n_x(r, r')$, the exchange hole, is defined as the exchange correlation hole of the non interacting system: $n_x(r, r') = n_{xc}^{\lambda=0}(r, r')$; eq. 2.36 implicitely defines $n_c^\lambda(r, r')$. Now, taking into account the definition of the exchange hole, using eq. 2.35 we know that:

$$\int dr' n_x(r, r') = -1, \qquad (2.37)$$

and this automatically implies that:

$$\int dr' n_c^\lambda(r, r') = 0. \qquad (2.38)$$

This fact has an interesting physical interpretation: the amount of charge that is moved away from r by the presence of an electron is only connected to the

exchange effect. In other words the integral in eq. 2.33 must be equal to $N-1$ independently of the kind of interaction between the particles; it just tells that if an electron is in r, then it can not be somewhere else. The kind of statistic of the system, then, gives a first shape to the hole, for example, keeping the electrons of the same spin apart. On top of this, the correlation effects only *polarise* the hole by further modifying its shape. Typically the Coulomb interaction makes the hole deeper and more short ranged.

Now, coming back to the exchange correlation functional, if we take its expression in terms of the integration of the coupling constant λ, i.e. eq. 2.27, and we use eq. 2.31 and eq. 2.32; we obtain:

$$
\begin{aligned}
E_{xc}[n] & = \frac{1}{2} \int dr dr' \frac{n(r)}{|r-r'|} \int_0^1 d\lambda n_{xc}^\lambda(r,r') = \\
& = \frac{1}{2} \int dr dr' \frac{n(r)}{|r-r'|} \bar{n}_{xc}(r,r'),
\end{aligned}
\tag{2.39}
$$

where $\bar{n}_{xc}(r,r')$ is the coupling constant averaged exchange correlation hole. If we compare this expression with the definition of E_{xc}^{LDA} in eq. 2.26, we see that the exchange correlation hole in the case of LDA is given by:

$$
\bar{n}_{xc}^{LDA}(r,r') = \bar{n}_{xc}^{heg}(n(r),|r-r'|),
\tag{2.40}
$$

where $\bar{n}_{xc}^{heg}(n,|r-r'|)$ is the exchange correlation hole of a homogeneous electron gas of density n. Since the homogeneous electron gas is a possible real physical system, $n_{xc}^{LDA}(r,r')$, will obey many of the exact constraint and sum rule of the real exchange correlation hole.

Generalised Gradient Approximation

A possible way to go beyond LDA is to introduce an energy functional that depends also on the derivative of the density. However, it has been seen [41] that the direct, simple second order gradient expansion around the uniform gas limit yields worse results than LDA. This is because the new functionals, do not preserve the constraint for the exchange correlation hole like LDA does. In other words the exchange correlation hole for the gradient expansion functional is somehow less realistic than the LDA one (especially in the long range part). In order to overcome this problem, a second type of functionals was introduced, the so called generalised gradient approximation (GGA) functionals [42, 43].

$$
E_{xc}^{GGA} = \int dr f(n(r), \nabla_r n).
\tag{2.41}
$$

These functionals are built employing real space cutoff for both the exchange and the correlation hole in such a way that they still fulfil the sum rule for the exchange and correlation hole, improving in this way the results.

Pseudopotentials

The form of the energy functional is not the only approximation that is used in typical plane wave *ab-initio* calculations. The pseudopotentials are used to model the interaction between the positive ion and the electrons, in order to get rid of (i)the singularity of the electrostatic potential of the nuclei, (ii) the description of the very localised core electrons, (iii) the nodes of the wave functions of the valence electrons close to the nuclei. The idea arises from the experience that the core electrons are usually not involved in chemical bonds among atoms and their wavefunctions are practically the same in the possible different chemical environment. Their global effect on the valence electrons is to push them away from the nuclei, since the Pauli exclusion principle impose mutual orthogonality and the Coulomb interaction further pushes them apart. The presence of the core electron can thus be assimilated to a repulsive potential, that summed up with the singular attractive potential of the nuclei, provides a much weaker total potential.

So what is done is usually to replace the nuclei and the core electron with an effective potential which is computed by inverting the pseudoatom Schrödinger equation imposing that

- for r greater than a certain 'core' value r_c, the original potential is recovered

- for $r > r_c$, the valence wave function of the pseudoatom are the same as the one of certain reference configuration, typically the neutral free atom

- the energies of the valence states of the pseudoatom are the same as the ones of the original atom

- for $r < r_c$ the pseudoatom wave functions do not have any nodes

- the pseudopotential is transferable, i.e. it is able to describe well the behaviour of the valence electrons in different chemical environments

On top of this requirements the norm conservation of the pseudo wave functions is imposed. This is very important for two reasons: the first is that of course, within a density functional framework the total electronic density should at least fulfil the basic particle number conservation rule; the second is that, in order to carry out transferability, the scattering properties of the ion and the

core electrons for the valence state should be recovered; the logarithmic derivatives of the pseudo wave-functions determines this scattering properties. It is shown, [44], that norm conservation forces the logarithmic derivatives to be the same as the one of the original atom for a certain reference energy and for energies close to it. This is very important because, when the pseudopotential is used to describe chemical bonding, for example in a solid or in a molecule different from the atomic configuration, the energy levels of the valence electrons vary from the atomic eigenvalue used to build the pseudopotential [4].

K-point sampling of the Brillouin zone and kinetic energy cutoff

Exploiting the Bloch theorem [45], in a periodic system, each electronic eigenfunction can be expressed a sum of plane waves:

$$\psi_{n,k}(r) = \sum_G c_{n,k+G} e^{i(k+G)r},\qquad(2.42)$$

where k is a vector in the reciprocal space that spans the first Brillouin zone, and G are the reciprocal lattice vectors. The Kohn and Sham equation now has a simpler form: the Hamiltonian for each k point is a matrix in the reciprocal space (G, G'); and one should diagonalise it in order to get the coefficients of the planewaves.

In principle, both the number of the k points and of the G vectors in (2.42) are infinite. Of course this infinite summations can not be implemented in real calculations and what is actually done is to sample the Brillouin zone with a finite number of k points and to introduce a cutoff in the kinetic energy that limits the summation over the G vectors.

In fact, assuming that the wavefunctions in the k points that are very close to each other are almost identical, we can think of replacing the wavefunctions of the k points in a region of the Brillouin zone with the wavefunction of one k point representative of the entire region. Of course convergency on the number of k points must be achieved; however there are special sets of k points [46] for which, exploiting the symmetry of the system, very accurate calculations can be performed at a reasonable small number of k points.

To limit the number of G vectors, we have to think that usually the most important coefficients for the plane wave basis set are the ones that carry small kinetic energy. The plane wave basis set can then be truncated to include only plane waves with kinetic energy below a certain threshold E_{cut}:

$$\frac{1}{2}|k + G|^2 \leq E_{cut}.\qquad(2.43)$$

---

[4]Nota sui non norm conserving pseudopotential

In all our calculations for diamond, if not else specified, we have used a cutoff energy of 45 Ry.

2.3 Application of the theory to C(111)2×1

For the DFT calculations presented in this paragraph, and also where is not else specified, we used the FHI98MD code [47] which is a plane-waves basis set code. In this code the minimisation of the total energy functional, with respect to both the electronic $\{\psi_{k,n}\}$ and ionic degrees of freedom $\{R_I\}$, is achieved within the subsequent self-consistent scheme:

1. Input geometry and a first guess of the electronic wavefunction are provided.

2. The electronic density is computed from the wavefunctions; this density is used to build the Hartree term and the exchange correlation potential within the chosen approximation

3. The Kohn and Sham equations are solved by an iterative minimisation technique, until the system is close enough to the Born-Oppenheimer surface, i.e. close enough to the electronic ground state for the given geometry; the result are new electronic wavefunctions

4. The forces on the ions are computed and their classical equation of motion is integrated for one time step

5. Steps from 2 to 4 are repeated until the forces on the ion are below a certain threshold (0.025 eV/Åand the electronic ground state is reached

In our calculations, the exchange-correlation functional is treated within the generalised gradient approximations (GGA) of Perdew et al. [42, 43]. The ion potential is represented by a norm-conserving, fully separable pseudopotentials.

2.4 Equilibrium geometry of the (111) surface of diamond

The convergency test on the DFT parameters, which we will not describe in detail, determines, for 45 Ry energy cutoff, a lattice constant of 3.54Å, 0.8% less than the experimental value. The Car-Parrinello optimisation of the clean 2×1 surface was performed on a slab made of 12 atomic layers of carbon atoms plus 5.4 Å of vacuum. The initial ionic positions were changed from the ideal

position breaking all the possible symmetries of the ideal lattice except for the inversion of the z axis; all the the atomic layers were allowed to move except for the central 2 layers. Convergency test on the slab thickness reaching 24 atomic layers have been performed. A set of 16 $\vec{k}$ points was used to sample the BZ.

As already introduced in chapter 1, it is now well accepted that the reconstruction of clean C(111) is the Pandey chain model. This model involves significant changes from the ideally truncated bulk. In fact in the ideally truncated bulk the dangling bonds belong to atoms which are second nearest neighbors; the atoms rearrange themselves, in order to have dangling bonds on nearest neighbours forming, chains along the y directions. Indeed, as we can see form fig. 2.1, in the relaxation, per each 2x1 unit cell, one atom on the second layer breaks its bond with the third layer and rises toward the first layer, at the same time a bond is formed between an atom belonging to the first layer that lowers toward the second layer. As a result, looking at the surface from the side, we have a sequence of five fold and seven fold rings, while the ideal surface has only sixfold rings. The dangling bonds have mostly a p-like character that allows π bonding between the surface atoms along the chain [17]

As already mentioned, there is strong debate about the exact details of the relaxation, mainly because, as we will see in the next paragraph, the DFT band structure of the surface depends significantly on those details. The main discussion concerns the presence of buckling and the degree of dimerisation of the surface chains. The buckling Δz is defined as the difference between the surface atomic coordinates in the (111) direction, while the degree of dimerisation Δ is defined as $\Delta = \frac{d_1 - d_2}{d_1 + d_2}$, d_1 and d_2 being the two bond lengths within the surface chain.

In the optimisation of the geometry we started from several configurations, involving Pandey chains with slight negative and positive buckling [48] and dimerisation. The final relaxed geometry, common to all the starting points, is shown in Fig.(2.1). We find that the buckling and the dimerisation of the chain vanish. In particular we find for the buckling $\Delta z = 0.004 \mathring{A}$ and for the degree of dimerisation $\Delta = 0$. The bond length of the atom in the chain is $1.427 \mathring{A}$ in agreement with previous DFT-LDA calculations [49, 50]. The Pandey chain reconstruction affects also the deeper layers of the slab. Conformly to the previous DFT-LDA calculations we find that the length of the bonds d_{12}, d'_{12} that connects the chain to the subsequent layer is almost (within $0 - 0.1\%$) equal to the bulk bond length, while the distance d' between the atoms in the second layer chain is stretched by 0.9%. Even in the absence of buckling and dimerisation, the atoms on the top chain are not equivalent, this is due to the rising and the lowering of the atoms in the 3rd and 4th layer. In fact,
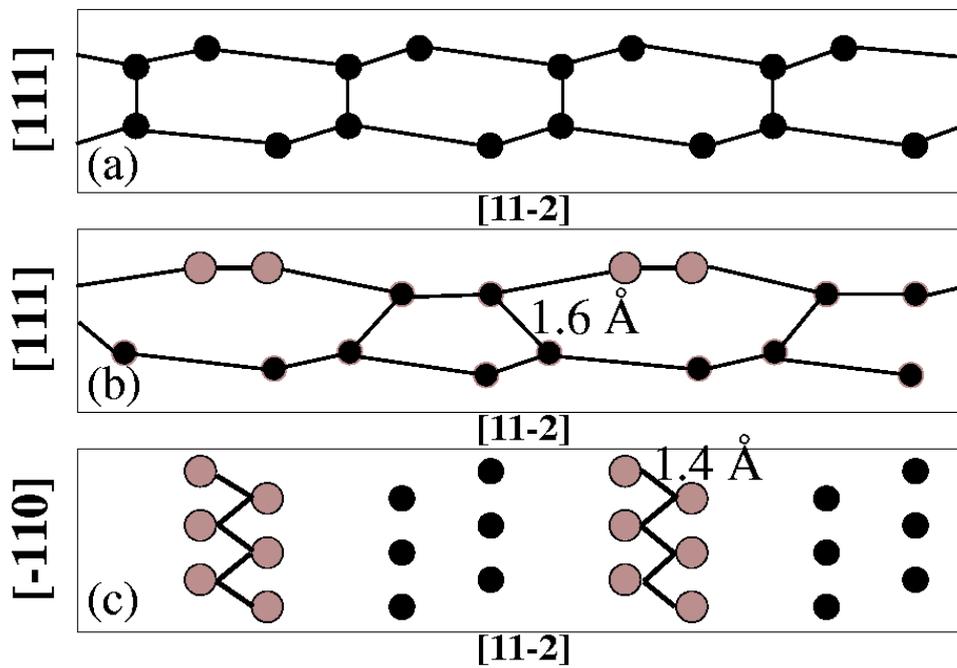
Figure 2.1: (a)Ideally truncated and (b)(c) fully reconstructed C(111) surface. (a) and (b) are side views, (c) is the top view of the surface. In agreement with previous DFT calculations, the surface relaxes to a non buckled non dimerised Pandey chain geometry.

confirming previous DFT-LDA calculations [50], we find stronger buckling in the 4th layer ($\Delta z = 0.1659 \overset{\circ}{A}$). The values of the buckling on the deeper layers are reported on Fig. 2.1b. The bonds between the 2nd and 3rd layer d_{23} and d'_{23} differ significantly: d_{23} is 6.7% larger than the average bulk length and d'_{23} is 4.1% larger. All this geometry parameter are listed in the table (2.1), where we also list for a comparison the results of previous *ab-initio* calculations and the result of the latest LEED spot intensity vs voltage experiment [26].

Method	d	Δ	Δz	d_{12}	d'_{12}	d'	d_{23}	d'_{23}
DFT-LDA [51]	1.44	1.4	0.00	< 2.0	< 2.0	< 2.0	8.0	8.0
DFT-LDA [49]	1.43	0.0	< 0.01	0.1	-0.1	0.9	6.6	4.5
DFT-LDA [50]	1.425	0.0	0.006	0.2	-0.2	0.9	6.5	4.6
DFT-GGA	1.427	0.00	0.004	0	0.1	0.9	6.7	4.1

Table 2.1: DFT-LDA predictions of structural parameters of the C(111) 2×1 pandey-chain reconstructions: -d is the average intrachain bond length -Δ is the degree of dimerisation- Δz is the first layer buckling - d_{12}, d'_{12}, d_{23} and d'_{23} are the changes in the length of the bonds connecting the 1st to the 2nd layer and the 2nd to the 3rd -d' is the average bond length within the second layer atomic chain. GGA calculations: present work. All lengths are in Å.

The geometrical details obtained for our slab within GGA are in very good agreement with all the previous DFT-LDA plane waves calculations [49, 50] except for ref. [51]. In particular in [51] a chain dimerisation of 1.4% is found which leads to a correct band structure, despite of the fact that most experimental predictions suggests an undimerized geometry. It has been argued [50] that a possible reason for the found dimerised surface is that the slabs used in those calculation were not thick enough (8 layers) to make possible a complete relaxation involving many deeper layers as was instead found by the other *ab-initio* calculations.

2.5 DFT electronic structure and symmetries of the reconstruction

Even if there is no formal justification, one can try to interpret the eigenvalues of the Kohn and Sham structure as single particle excitation energies of the system, building the DFT band structure. The DFT electronic band structure of the clean C(111)(2×1) surface is shown in Fig.(2.2).

The surface bands that lie in the gap are originated by the bonding and anti-bonding combinations of the p_z orbitals of the atoms in the chain. As
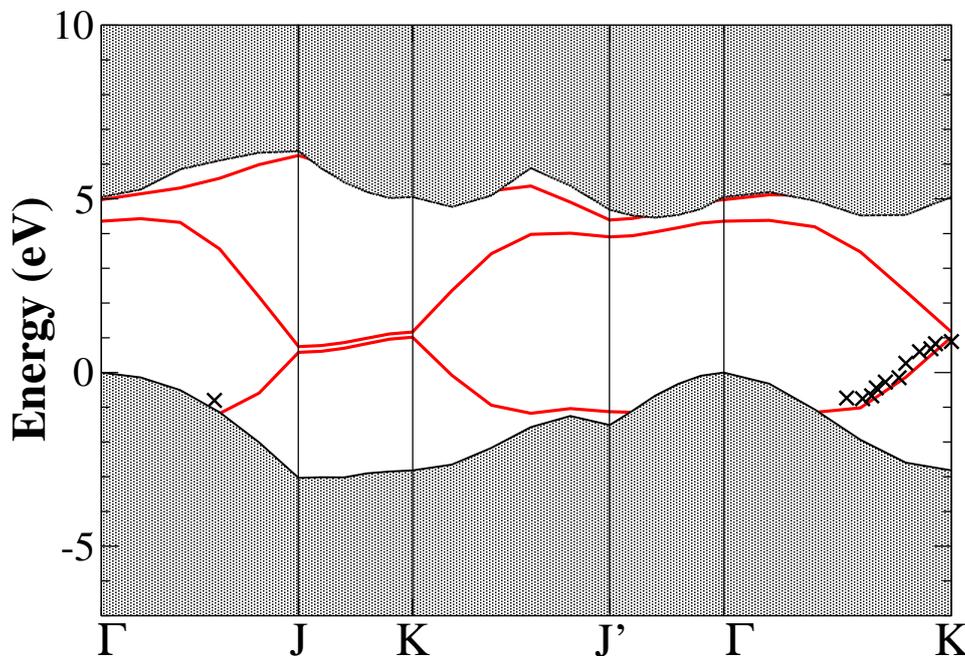
Figure 2.2: Electronic band structure of the 2×1 surface. In agreement with previous DFT calculation the surface is semimetallic. Between J and K the upward dispersion of the surface bands, due to the interaction between different chains, makes both the bands cross the Fermi level.

shown in Fig.(2.1), the chains develop in the ($\bar{1}$10) direction and are an almost one dimensional structure, this characteristic should lead to a strong dispersion of the surface bands along the chain direction that correspond to the ΓJ and KJ' parts of the IBZ boundaries and, viceversa, to a less dispersive part of the band structure along the directions that are perpendicular to the chains, that correspond to the JK and J'Γ directions. The calculated band structure seeeems in good agreement with these thoughts and the small dispersion along the JK and J'Γ is possibly due to the interaction between the chains through the substrate [49]. The dispersion of the calculated bands seems to be in qualitatively good agreement with experimental data points: experiments for occupied states obtained from angle resolved photoemission show a surface state dispersing 1eV going from 2/3ΓJ to a point close to J [18] and 1.9eV going from 1/2ΓK to K [22]. In our case the dispersions are of 1.3eV and 2.3eV respectively. However in the JK directions the calculated two surface bands are nearly degenerate and their upward dispersion (~ 0.4 eV) makes them cross

the Fermi level producing a semimetallic character of the slab. This, as already mentioned, is in contrast with experimental data that predict a gap of at least 0.5 eV [22].

Symmetry of the chain and DFT band structure

Tight binding, one dimensional models of the π-chains; help to get insights about the physical origin of the almost degenerate surface states present along the JK line of the band structure. In fact what is shown, [17], is that the degeneracy of the states along the JK line is strictly connected to the equivalence of the two atoms on the chain; if this equivalence is removed then the gap between surface states may open.

As we have already mentioned, the two atoms on the chain are not really equivalent if we take into account the presence of the third layers. As a matter of fact, there is a little splitting ($\sim 0.1eV$) between surface states, but this does not prevent the surface to be semimetallic thanks to the dispersion of both of the surface bands along the JK line. There are different ways to gain more inequivalency: a chance is given by a distortion of the lattice, through buckling or dimerisation of the chain for example; another possibility is to polarise the spin of the atoms of the chains forming an antiferromagnetic order. For the same surface, Si and Ge follow the first way and the chains, in their case, exhibit buckling [52]. The mechanism is regulated by the delicate balance between the gain of band structure energy and the loss of elastic one while distorting the system. Probably, since the carbon bonds are stronger than Si and Ge bond, the balance is unfavourable in the case of diamond, and the system seems to unlike distortions. This may also justify the fact that the most stable reconstruction for the (111) surface of diamond is the (2x1); while in the case of Si and Ge we have more fancy (7x7), c(2x2) reconstructions.

Dimerisation is excluded experimentally for symmetry reasons [22], while, as already discussed in chap. 1, there is still debate concerning the buckling of the chains; so, starting from the fully relaxed geometry, we have studied the total energy and the band structures of systems with increasing buckling. The results are listed in tab. (2.2) and fig. (2.3). As we can see comparing tab. 2.2 and fig. (2.3), only a strong (in terms of energy cost) buckling could induce an opening of the gap.

As a last try we tested the possibility of antiferromagnetic ordering along the chain within the local spin density approximation (LSDA). LSDA is a generalisation of LDA in order to include the spin degrees of freedom. The energy functional is now dependent (still locally) also on the magnetisation
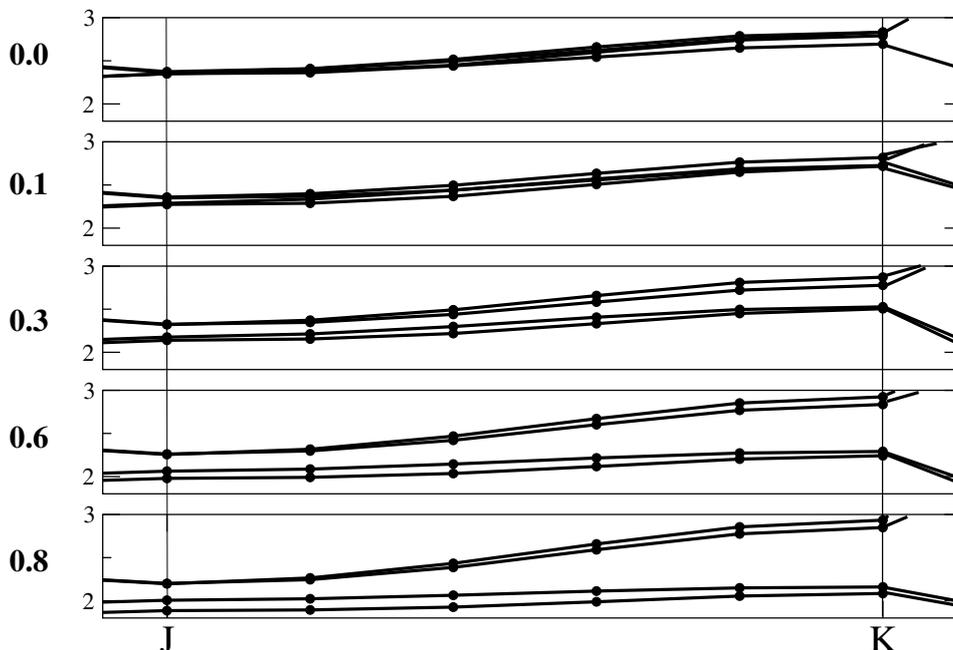
Figure 2.3: Electronic band structure of the 2×1 surface, along the JK directions for different values of the buckling of the surface chains. The degeneracy of the surface bands along the JK line is lifted by the presence of buckling. The upward dispersion of the bands, however, makes the surface become semiconducting only in the case of a 0.8 Åbuckling. Such a big value of buckling is very unlikely: most experiments (see chap. 1) do not find any evidence of buckling or estimate it to be 0.3 Å; moreover total energy calculations (see tab. chap2energybuckl) show that such a buckling is energetically unfavourable.

Buckling (Å)	Total Energy (a.u.)
0.0	-137.528651
0.1	-137.526757
0.3	-137.511821
0.6	-137.464567
0.8	-137.418744

Table 2.2: Total energy of slabs vs buckling.

of the system. Also in this case we started from different initial configuration involving different values of the magnetisation of the slab. In all cases we found a non magnetic ground state, essentially the LDA one.

The discrepancy of the DFT band structure, however, can not be only addressed to the possible, unlikely, incapability of DFT to capture some inequivalence of the system. In general, as we will see in the next paragraph, DFT band structures underestimate the electronic gap of semiconductors; the main part of this thesis is devoted to the study of the possibility that also in this case such a low value of the gap is not a consequence of asymmetric geometrical changes, rather of many body effects.

2.6 The band gap problem

The Kohn and Sham eigenvalues, introduced in the theory as Lagrange multipliers, in principle do not have any physical meaning; not even in the sense that Hartree-Fock eigenvalues do within Koopman's theorem. However they are often interpreted as addition and removal energies; and the agreement with experiments even if only qualitative, is already remarkable in view of the absence of theoretical foundations. Thus it is important to understand whether by getting better and better functionals, one could hope to get to a more quantitative description of the band structure within the Kohn Sham scheme. We will see that this is not the case, however it is still important to get insights on the physical content of the Kohn Sham eigenvalues.

In ref. [53], Perdew *et al.* extended the DFT formalism to fractional particle number states by defining proper energy functionals for statistical mixtures. In this framework, they proved that the ground state energy of $(N + \omega)$-particle state [5] (where N is an integer and $0 \leq \omega \leq 1$) is given by:

$$E_{gs}^{N+\omega} = (1 - \omega)E_{gs}^N + \omega E_{gs}^{N+1}. \tag{2.44}$$

From eq. 2.44, we can see that the ground state energy versus the total number of particles is given by a series of straight line segments. While E_{gs}^N [6] is a continuous function of N, its derivative, namely the chemical potential $\mu = \frac{\partial E}{\partial N}$ might be discontinuous as the variable N passes through an integer value. In ref. [53], it was also shown that:

$$\frac{\delta E^N[n]}{\delta n(r)} = \frac{\partial E}{\partial N}\Big|_{N-\delta} = \mu^-, \tag{2.45}$$

where δ is a positive infinitesimal quantity, and by μ^- we mean the left limit of μ. Eq. 2.45 states that, in principle, the left limit of the chemical potential

---

[5]Described by the corresponding statistical mixture

[6]In this case we are dropping the condition that N is an integer number.

is correctly given by the functional derivative of the total energy functional of an N-particle calculation.

At the same time, Janak's theorem [54] states that the partial derivative of the total energy functional with respect to the occupation of a certain Kohn Sham state is given by the corresponding Kohn-Sham eigenvalues:

$$\frac{\partial E}{\partial f_i} = \epsilon_i. \tag{2.46}$$

So comparing eq. 2.45 with eq. 2.46, it is easy to see that within DFT the left limit of the chemical potential is given correctly by the energy of the highest Kohn Sham occupied state:

$$\mu^-(N) = \epsilon_N^{(N)}. \tag{2.47}$$

From eq. 2.46, it is also straightforward [55] to see that the ionisation potential $I(N)$, and the electron affinity $A(N)$ are given by:

$$I(N) = E_{N-1} - E_N = -\int_{N-1+\delta}^{N-\delta} \frac{\partial E}{\partial N} dN = -\mu^-(N) = -\epsilon_N^{(N)}$$

$$A(N) = E_N - E_{N+1} = -\int_{N+\delta}^{N+1-\delta} \frac{\partial E}{\partial N} dN = -\mu^+(N) = -\mu^-(N+1) = -\epsilon_{N+1}^{(N+1)}. \tag{2.48}$$

The electronic gap of the system is defined as the difference between the ionisation potential and the electron affinity, and so:

$$E_g = I - A = \epsilon_{N+1}^{(N+1)} - \epsilon_N^{(N)}. \tag{2.49}$$

Eq. 2.49 is very important since it states that the true gap of a system is correctly given by the difference of the energies of the Kohn Sham highest occupied states of the $N+1$ and N particle system. How is it connected to the Kohn and Sham gap, namely $E_g^{KS} = \epsilon_{N+1}^{(N)} - \epsilon_N^{(N)}$? Let's consider again the expression of the *true* gap:

$$E_g = \mu^+(N) - \mu^-(N) = \left(\frac{\delta E[n]}{\delta n}\bigg|_{N+\delta} - \frac{\delta E[n]}{\delta n}\bigg|_{N-\delta}\right); \tag{2.50}$$

only terms arising from the kinetic and exchange correlation part of the total energy functional might contribute to this expression. The remaining part of the total energy functional will give a null contribution since, in extended systems, the addition of one electron is an infinitesimal perturbation, and the change in the density is null being the volume infinite. Therefore all the terms

(i.e. the Hartree potential, the external potential) of the Hamiltonian which are analytic with respect to the density can not yield a finite, of order one, effect on the eigenvalues. Taking all this into account,

$$E_g = \frac{\delta T}{\delta n}\bigg|_{N+\delta} - \frac{\delta T}{\delta n}\bigg|_{N-\delta} + v_{xc}^+ - v_{xc}^-. \tag{2.51}$$

Now let's consider the Kohn Sham system: it is a system of free particle moving in an effective potential v_{xc}. Therefore the electronic gap of this system is given only by the discontinuity of the kinetic part of the energy functional, so that:

$$E_g^{KS} = \frac{\delta T}{\delta n}\bigg|_{N+\delta} - \cdot \frac{\delta T}{\delta n}\bigg|_{N-\delta} \tag{2.52}$$

Comparing eq. 2.51 and eq. 2.52, we can see that the DFT Kohn Sham gap E_g^{DFT} of the N particle system, is related to the *true* quasiparticle gap through the discontinuity Δ of the exchange and correlation potential when an electron is added to the system.

$$E_g = E_g^{KS} + \Delta = \epsilon_{N+1}^{(N)} - \epsilon_N^{(N)} + v_{xc}^+ - v_{xc}^-. \tag{2.53}$$

Now, if Δ is very small, a good enough functional would provide the correct gap; viceversa, if Δ is a big quantity, no matter how good the functional is, the Kohn Sham gap does not provide a good approximation for the quasiparticle gap. The discontinuity of the exchange and correlation potential was found [56, 57] to be a consistent (80%) part of the error; the problem of quasiparticle energies should then be addressed by other theories.

Chapter 3

Electronic properties

In this chapter we will see how a proper description of electronic spectra may be achieved through many body perturbation theory, within the GW approximation for the electronic self energy. In the case of the (111) surface of diamond, however, the usual implementation of the GW calculation can not be successfully applied and we need to go beyond it to face the problem of semimetallicity of the surface; hence an iterative scheme is implemented. In this scheme the energies present in the expression of the Green function and of the microscopic dielectric function are updated until selfconsistency is reached. A finite direct (∼ 1 eV) gap between surface states is obtained, in better agreement with the available experiments. We will discuss briefly the issue of selfconsistency within GW. Finally, we introduce the GWΓ approximation for the self energy, where vertex correction are included; we will apply this method to the calculation of the electronic affinity of the surface.

In the final part of the previous chapter, we have seen how there is strong evidence [56, 57] that within a Kohn-Sham scheme the quantitative determination of the electronic gap of semiconductors is beyond reach, even if the exact exchange correlation energy functional is known. However, the DFT band structure may represent a first-order approximation and a good (the *best available* even) starting point for the implementation of other theories more suited to tackle the problem of the description of electronic spectra of materials.

Let's first simply try to understand what are we comparing to when we calculate band structures, and introduce the main ingredients required to describe these phenomena properly. Typically the electronic band structure is obtained by (angle resolved) photoemission spectroscopy (ARPES) experiments, both direct and inverse. Roughly speaking, as schematically depicted in fig. 3.1, in the direct PES, light is shone on the sample and, due the absorption of a photon, an electron is excited from the valence band and leaves the sample. In the inverse photoemission setup, on the other hand, an electron beam
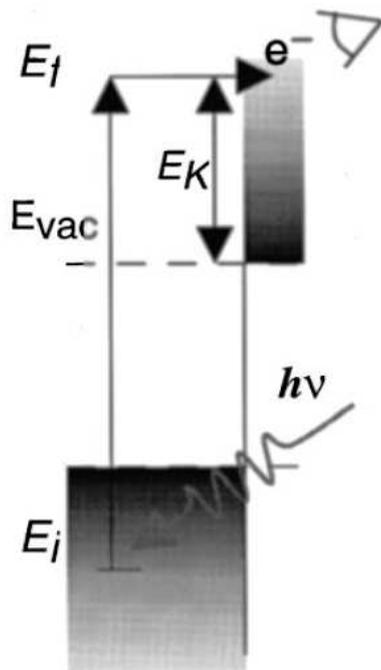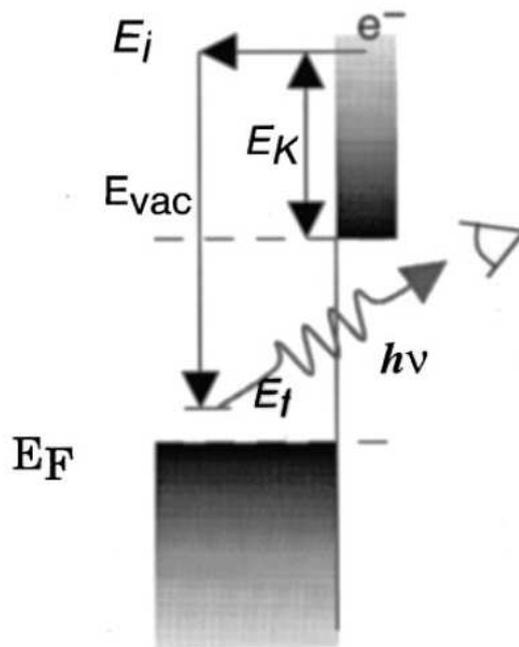
Figure 3.1: Schematic picture of direct and inverse photoemission experiments. From ref. [58].

impinges on the sample and as an electron is captured, occupying an empty conduction state, energy is released through the emitted photon. If, in the case of direct photoemission, for example, one considers the final electron as being completely decoupled from the system, the conservation of the total energy and momentum, yields information about the amount of energy and momentum transferred to the system in order to bring it from the neutral ground state to the **charged** excited state, i.e.:

$$E_\gamma + E_0^N = E_{kin}^{e^-} + E_n^{N-1} \Rightarrow E_n^{N-1} - E_0^N = E_\gamma - E_{kin}^{e^-}, \qquad (3.1)$$

, where E_γ and $E_{kin}^{e^-}$ are the energies of the incoming photon and of the ejected electron respectively; E_0^N is the energy of the N-particle ground state, and E_n^{N-1} is the energy of the n-th excited state of the $(N+1)$-particle system. In an independent-electron framework, considering for istance Hartree-Fock within Koopman's theorem, the energy levels of the electrons are calculated once and for all through the selfconsistent solution of the Hartree-Fock equations. In the ground state, the electrons will fill up the lowest $N/2$ levels; the energy of the system is given by the sum of the energy of the occupied levels calculated with respect to a common reference value, for example, the vacuum level. Now, if a photon is absorbed and an electron is ejected from one of these states, such as from the s-th level, the energy of the system is still the the sum of the energy of the occupied levels, so we get:

$$E_n^{N-1} - E_0^N = E_\gamma - E_{kin}^{e^-} = -\epsilon_s. \qquad (3.2)$$

The energy required to excite the system is exactly the energy level of the ejected electron measured with respect to the vacuum level, the binding energy of the electron. In the same way, the electron, added to the system in an inverse photoemission experiments, will occupy one of the *fixed* empty levels above the Fermi energy. Hartree Fock calculations performed in this way, don't give a good description of the band structure, often overestimating the band gap of semiconductors by more than 100 percent.

What is missing in this kind of calculation is the correlation between the electrons. The electrons, in fact, respond selfconsistently to the addition or removal of an electron because they 'feel' the modification of the system, which, for example, becomes polarised as it tries to screen the added charge. This response causes a change in the potential felt by the the new electron and all the energy levels are modified with respect to the ones of the neutral system. So, in order to describe more conveniently what is happening in a photoemission experiment, starting from a Hartree Fock picture, we need to introduce correlation for example by adding to the equation for each single particle state the potential induced by the added (removed) particle, acting on the particle itself.

As a matter of fact, we will come up to an equation, namely the quasiparticle equation, in which the electronic self energy will play the role of this potential induced by the presence of the particle through the response of the others. This requires to use Green's function theory.

3.1 The Green function

The single particle Green's function is defined as:

$$G(x, t; x', t') = \frac{\langle \Psi_0 | \mathcal{T}\{\widehat{\psi}(x, t)\widehat{\psi}\dagger(x', t')\} | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle}. \tag{3.3}$$

Where $|\Psi_0\rangle$ is the N-particle ground state, $\widehat{\psi}(x, t)$ ($\widehat{\psi}\dagger(x, t)\}$) is the Heisenberg creation (annihilation) field operator and we assume that the spin degrees of freedom are included in the variable x. $\mathcal{T}$ is the time ordered product, meaning that:

$$\mathcal{T}\{\widehat{\psi}(x, t)\widehat{\psi}\dagger(x', t')\} = \begin{cases} \widehat{\psi}(x, t)\widehat{\psi}\dagger(x, t) & t > t' \\ -\widehat{\psi}\dagger(x', t')\widehat{\psi}(x, t) & t < t' \end{cases}. \tag{3.4}$$

From the definition of time-ordered product we see that, for $t > t'$, the Green's function describes the propagation of a particle from the point (x', t') to the point (x, t); viceversa, for $t < t'$, it describes the propagation of a hole.

The knowledge of the Green's function enables the computation of many properties of the system:

- The expectation value of any single particle observable in the ground state

- The ground state energy of the system

- The excitation spectra of the system

From now on we will focus our attention on this third point; very good introductions about the Green's function formalism and about the Many Body Perturbation Theory can be found in ref. [59, 60, 61].

3.1.1 The Lehman representation

After a temporal Fourier transform, the Green's function G can be formally written, in what is called its Lehman representation, as

$$G(x, x'; \omega) = \sum_s \frac{f_s(x)f_s^*(x')}{\omega - \epsilon_s - i\eta\mathrm{sgn}(\mu - \epsilon_s)}, \tag{3.5}$$

where s and s' run over the N+1 and N-1 electrons excited states, η is positive and infinitesimal, μ is the Fermi energy of the system, and

$$f_s(x) = \begin{cases} \langle \Psi_0 | \widehat{\psi}(x) | \Psi_{N+1}^s \rangle & \epsilon_s = E_{N+1}^s - E_N^0 & \text{when} & \epsilon_s > \mu \\ \langle \Psi_{N-1}^s | \widehat{\psi}(x) | \Psi_0 \rangle & \epsilon_s = E_N^0 - E_{N-1}^s & \text{when} & \epsilon_s < \mu \end{cases} . \qquad (3.6)$$

In the last formula $\langle \Psi_{N\pm1}^n |$ is the n-th excited state of the $N\pm1$-particle system, and $E_{N\pm1}$ its corresponding energy. By looking at equation 3.5 and 3.6, we see that the poles of the Green's function are at the energy differences we are interested in when we compare to ARPES spectra, namely the addition and removal energy of one electron.

If we take the imaginary part of eq. 3.5, we get the so called spectral function:

$$\frac{1}{\pi} \mathcal{I}m(G(x, x'; \omega)) = \pm A(x, x'; \omega) = \pm \sum_s f_s(x) f_s^*(x') \delta(\omega - \epsilon_s). \qquad (3.7)$$

The spectral function is a sort of density of available states weighted by the Lehman amplitudes which give the probability, after the creation (destruction) of a particle at point (x, t), to find the system in the eigenstate Ψ_{N+1}^s (Ψ_{N-1}^s). If now we take the spectral function $A(x, x'; \omega)$, and we 'project' it onto a chosen suitable one-electron state $|\phi_i\rangle$, we compute:

$$A_{ii}(\omega) = \int dx dx' A(x, x'; \omega) \phi_i^*(x) \phi_i(x')$$

$$= \sum_s \delta(\omega - E_s) \int dx dx' f_s(x) \phi_i^*(x) f_s^*(x') \phi_i(x')$$

$$= \sum_s \delta(\omega - E_s) \langle \Psi_0 | \widehat{c}_i | \Psi_{N+1}^s \rangle \langle \Psi_{N+1}^s | \widehat{c}_i \dagger | \Psi_0 \rangle. \qquad (3.8)$$

(The last line of this equation is valid only for $\omega > 0$, and $\widehat{c}_i\dagger$, $(\widehat{c}_i)$ is the creation (annihilation) operator of an electron in i-th single particle state). Eq. 3.8 says that if we take a system at its ground state and add (remove) a particle in the state ϕ_i, the amplitude of the oscillations at frequency ω will be given by $A_{ii}(\omega)$.

Suppose that we have a system of non interacting electrons with Hamiltonian $\hat{H}_0$. If ϕ_i is an eigenstate of the single particle Hamiltonian, then:

$$A_{ii}(\omega) = \delta(\omega - \epsilon_i). \qquad (3.9)$$

In fact after the creation of the electron in the state ϕ_i, the system is still in an eigenstate of the Hamiltonian and will oscillate only with the proper frequency
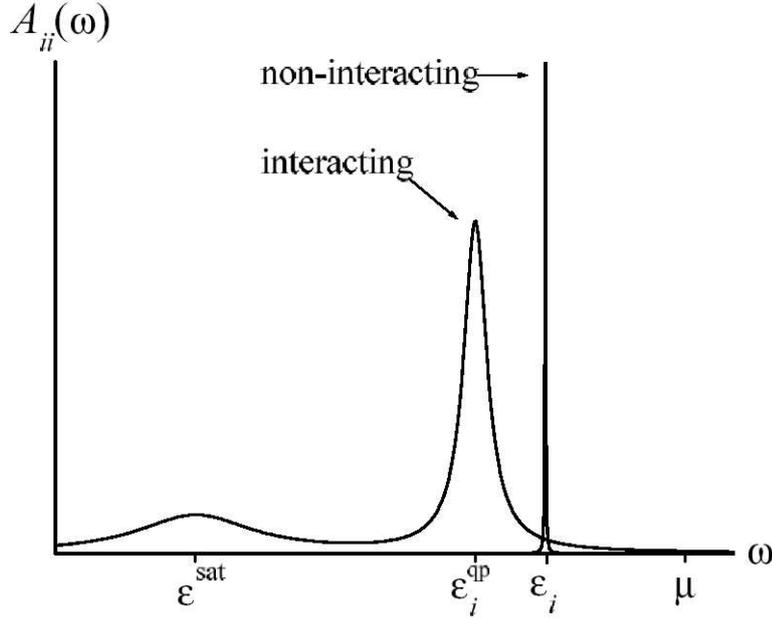
Figure 3.2: Schematic picture of the spectral function $A_{ii}(\omega)$. In the non interacting case $A_{ii}(\omega)$ is a delta function; when the interaction is switched on, the position of the peak changes, the structures now have a finite width and also higher energy satellite structures might appear. From ref. [62].

ϵ_i. As a matter of fact, also in the case of interacting system $A_{ii}(\omega)$ can exhibit some structures that may reflect that particular single particle state is 'quite' stable; this kind of states that oscillate for a reasonable number of times before being damped are the famous *quasiparticles*. Generally the A_{ii} might have a form similar to the one presented in fig. 3.2: close to the Fermi level there is a strong quasiparticle peak whose width is connected to its life time, satellite structures at higher energies might also appear, usually they are interpreted as excited state made by the quasiparticle plus a plasmon.

3.1.2 The quasiparticle equation

The Lehman representation (eq. 3.5) sheds light on the physical information contained in the Green's function but actually it does not provide any insight on how to get it.

It can be shown [59, 60, 61] that the Green's function of the interacting system is connected to the Green's function of the non interacting system, G^0, via an

integral equation, the so called Dyson equation, which takes the form:

$$G(1,2) = G_0(1,2) + \int d3\,d4\, G_0(1,3)\Sigma(3,4)G(4,2). \tag{3.10}$$

. Each index stands for space, time and spin coordinates for example $1 \equiv \{r_1, \sigma_1, t_1\}$. Σ is the proper self energy operator which is non hermitian, non local, and, when turning to Fourier space, energy dependent. We consider the Hartree Hamiltonian as the non interacting Hamiltonian $\hat{H}_0$. In principle, knowing the expression for the self-energy one could use eq. 3.10 to obtain the Green's function. However, here we want to show that eq. 3.10 is equivalent to a single particle Schrödinger equation.

From eq. 3.10 it follows that:

$$G^{-1}(x, x'; \omega) = G_0^{-1}(x, x'; \omega) - \Sigma(x, x'; \omega); \tag{3.11}$$

we know, see ref. [59, 60, 61] ,that

$$G_0^{-1}(x, x'; \omega) = \delta(x - x')[\omega - h_0(x)]. \tag{3.12}$$

Inserting eq. 3.12 in eq. 3.11, we get:

$$G^{-1}(x, x'; \omega) = \delta(x-x')[\omega-h_0(x)]-\Sigma(x, x'; \omega) = \delta(x-x')\omega-H(x, x'; \omega) = \omega-\hat{H}, \tag{3.13}$$

where we have defined an effective Hamiltonian given by:

$$\hat{H}(x, x'; \omega) = h_0(x)\delta(x - x') + \Sigma(x, x'; \omega) \tag{3.14}$$

which is not hermitian and energy dependent due to Σ. It is easy to prove that we can write the Green's function G as:

$$G(x, x'; \omega) = \sum_n \frac{\psi_n(r, \omega)\psi_n^*(r', \omega)}{\omega - E_n(\omega)} \tag{3.15}$$

where the $\psi_n(r, \omega)$ and $E_n(\omega)$ obey the equation:

$$\left(-\frac{\nabla_r^2}{2} + V_{ext}(r) + V_H(r)\right)\psi_n(r, \omega) + \int dr' \Sigma(r, r', \omega)\psi_n(r', \omega) = E_n(\omega)\psi_n(r, \omega). \tag{3.16}$$

The $\psi_n(r, \omega)$, while being a complete set, are not necessary orthonormal and the $E_n(\omega)$ are not necessarily real, since Σ is not hermitian.[1] The poles of the Green's function are now at the complex energies that fulfil the equation

---

[1]In fact G should be really expressed in terms of the left and right eigenvalues of $\hat{H}$ in a biorthogonal form, see ref. [63].

$\omega = E_n(\omega)$. Given a suitable expression for Σ, the poles of the Green function can be obtained solving the single particle like equation given in 3.16. It is worth noticing that these equations reduce to the Hartree equations when $\Sigma=0$, to the Hartree-Fock ones when $\Sigma = iGV$, and to the Kohn-Sham equations when, instead of Σ, a local, hermitian and energy independent operator is taken: $\Sigma \rightarrow V_{xc}^{KS}(r)$.

This observation gives an *a-posteriori* justification of the use of the Kohn-Sham equations to calculate the excitation energies; the qualitative agreement between DFT and experimental band structure results comes from the fact that the Kohn-Sham equations already buries a lot of physics, containing, at least approximately, exchange and correlation effects. In other words, $V_{xc}^{KS}(r)$ is often a quite good approximation for the 'true' self-energy Σ.

3.1.3 Complex or real poles?

In the previous section we have seen that, paying the price of dealing with a complicate non local, non hermitian, energy dependent operator Σ, we have regained a single quasiparticle picture. The poles of the Green's function now lie in the whole complex plane, where the Green's function is analytically continued; for the subtle mathematics that stands beyond this, refer to [63].

However we have seen that, in the Lehman representation (eq. 3.5), the poles of the Green's function are at the exact eigenenergies of the system upon addition and removal of a particle; in fact they lie infinitesimally close to the real axis. [2] Are the two expression for the Green's function, eq. 3.5 and eq. 3.15, contradictory? In this section we will try to shed light to this point following an illustrative example from ref. [58].

First of all we want to keep in mind that, if we take the thermodynamic limit, the set $\{\epsilon_s\}$ becomes a continuous set of excitations. So, in the thermodynamic limit, the Green's function definition in eq. 3.5 will have, instead of a sum, an integral over the variable s; as a consequence rather than isolated poles, it will present a branch cut on the real axis.

Now we assume a specific form for the Green's function in the Lehman representation; we then compute the spectral function, and finally we will see how the same spectral function can be obtained with a much simpler formulation, by allowing the Green's function frequency domain to expand in the whole complex plane. The important thing is that the spectral function is the quantity that is directly connected to the experiments.

Let us assume that a generic matrix element of the Green's function in the

---

[2]Note that for energy below (above) the Fermi level the analytic continuation of the Green's function is performed in the lower (upper) imaginary plane.

Lehman representation has the form:

$$G(\omega) = \sum_s \frac{g(s)}{\omega - s + i\eta}, \tag{3.17}$$

where $g(s)$ has the form:

$$g(s) = \frac{E_2}{(s - E_1)^2 + E_2^2}. \tag{3.18}$$

If now we take the thermodynamic limit, s becomes a continuous variable, the sum present in eq. 3.17 becomes an integral and $G(\omega)$ has a branch cut along **R**. If now we take the imaginary part of the Green's function we end up with a spectral function given by:

$$A(\omega) = \frac{1}{\pi} \frac{E_2}{(\omega - E_1)^2 + E_2^2}. \tag{3.19}$$

It is straightforward that the same spectral function could be obtained taking the imaginary part of a Green's function given by:

$$G(\omega) = \frac{1}{\omega - (E_1 + iE_2)}. \tag{3.20}$$

So in this case, in principle, it would have been sufficient to find only a *simple* but complex pole to describe a spectral structure given by a continuos density of excited states. The real part of the pole gives the position at which the peak in the spectral function is located, whereas the imaginary parts gives its width. It is as if one is in front of a choice between:

1. Describing the dynamic of the system, upon the addition or removal of the particle, referring exactly to the eigenstate of the system. Each component will oscillate at its proper frequency infinitely in time (the energies are real) and the damping is just a matter of interference. This is a possible procedure, but we have to keep in mind that we are dealing with very complicated spectra (for solids they are continuous) and eigenstates.

2. Describing the dynamic of the system, upon the addition or removal of the particle, through the evolution of elementary excitations called single quasiparticles, which are not the exact eigenstate of the system: they will have a finite lifetime, will not be orthogonal etc. but still they yield a good description of the dynamic of the excited state [3]. Obviously this

---

[3]The quasiparticles follow the Fermi statistic, in fact they might be thought as the single particle excitations of the non interacting system which are adiabatically connected to excitations of the real system.

kind of picture makes sense only when the lifetimes are sufficiently high, i.e. until the poles are close enough to real axis. In fact it would be useless to introduce some kind of excitations that don't even have the time to oscillate for a reasonable time before being completely damped.

Both the approaches are, in principle exact, however in practice it is always simpler to resort to some expression of the selfenergy and to solve the quasi-particle equation searching for a set of simple (but complex!) poles. This is the very advantage in using Green's function theory.

3.2 Hedin's equations

In the same way in which we use an approximation for V_{xc}^{KS}, we have now to face the problem of finding a suitable expression for Σ. Within the Many-Body theory we have a set of exact equations, the Hedin's equations [64, 65], which define implicitly the self-energy Σ.
This set of equations is obtained (see Appendix A of ref. [64]) starting from the equation of motion for the single particle Green's function. It is known [59, 60, 61], that this equation of motion involves the two particle Green's function; the two particle Green's function equation of motion in turn will depend on the three particle Green's function and so on, yielding an infinite hierarchy of equations. Hedin has shown that this chain of equations can be broken by introducing in the Hamiltonian an external source that couples to the density, and that at the end of the derivation is set to zero. The new thing now is that the two particle Green's function can be obtained as a functional derivative of the Green's function with respect to this external source, and so it is formally removed from the equation of motion of the single particle Green's function. Introducing other characteristic quantities of the system like the screened Coulomb interaction $W(1,2)$, the time ordered polarisation $P(1,2)$, and the vertex function $\Gamma(1,2;3)$, a close set of equation can be derived.

$$\Sigma(1,2) = i \int W(1^+,3)G(1,4)\Gamma(4,2;3)d(34);$$

$$W(1,2) = V(1,2) + \int W(1,3)P(3,4)V(4,2)d(34);$$

$$P(1,2) = -i \int G(2,3)G(4,2^+)\Gamma(34;1)d(34),$$

$$\Gamma(1,2;3) = \delta(1,2)\delta(1,3) + \int \frac{\delta\Sigma(1,2)}{\delta G(4,5)}G(4,6)G(7,5)\Gamma(6,7;3)d(4567). \quad (3.21)$$

The fifth equation, closing the set, is the Dyson equation given by eq. 3.10. It is very important to stress that, not only the Hedin equations form a set of closed selfconsistent equation that in principle could be solved using brute force (actually this procedure can not be applied for realistic systems); but most importantly they lead directly to the possibility of a perturbative expansion in terms of the screened Coulomb interaction W. This is very important and physically sound, in fact a direct expansion in terms of the bare Coulomb interaction would be unsuitable due to the strenght of this kind of interaction; whereas in most solid the Coulomb interaction is usually screened via the polarisation of the system, and the effective potential, by which the particles interact, is much weaker.

The Hedin equation offer also the possibility to analyse the physical meaning of the involved quantity: we have already seen how the self energy can be thought of as the effective potential that a particle, added to the system, feels due to the change in the charge distribution induced by its own presence. We can see from the third equation (3.21) that the polarisation P, which represents exactly this mechanism, is made by the propagation of particles and holes, that are not independent due to the presence of the vertex Γ. Γ contains the information on how the particle and hole interact with each other, and it can be seen, looking at the last equation, that it is connected to the variation of the induced potential during the polarisation process. The entanglement of this set of equations reflects the selfconsistency of the response of the system.

3.3 The G_0W_0 approach for the self-energy

The full solution of the Hedin equations can not be found for realistic systems and what is usually performed is an iterative procedure starting from some expression of the self-energy. If, for example, we start from $\Sigma = 0$ we immediately find that $G = G_0$, i.e. the non interacting Green function, and $\Gamma(1,2;3) = \delta(1,2)\delta(1,3)$; inserting this values in the expression of the self-energy we get $\Sigma(1,2) = iG(1,2)W(2,1)$, this is the so called GW approximation [65]. Now starting from $\Sigma = 0$, in principle one should use the Hartree Green's function and screening in the expression of Σ, which is a very poor approximation leading to bad wavefunctions and band structure. State of the art calculations for quasiparticle corrections use instead the *best G, best W* approach [66] in which the wavefunction and energies used to build the Green's function and the dielectric function are the best available, namely the DFT ones. So, in principle we want to calculate:

$$\Sigma(r, r'; \omega) = \frac{i}{4\pi} \int_{-\infty}^{\infty} e^{i\omega'\delta} W(r, r'; \omega) G(r, r'; \omega + \omega') d\omega'. \qquad (3.22)$$

The Green's function in terms of the DFT wavefunctions and energy looks like:

$$G(r, r'; \omega) = \sum_{n,k} \frac{\phi_{nk}^{DFT}(r)\phi_{nk}^{*DFT}(r')}{\omega - \varepsilon_{nk}^{DFT} + i\eta \text{sgn}(\varepsilon_{nk}^{DFT} - \mu^{DFT})}. \tag{3.23}$$

And the screened Coulomb interaction, given in term of the microscopic dielectric function $\epsilon(r, r'; \omega)$ and of the bare Coulomb interaction $V(r, r')$ is:

$$W(r, r'; \omega) = \int dr'' \epsilon^{-1}(r, r''; \omega) V(r'', r'). \tag{3.24}$$

If we compare this expression with the Hedin's equation defining W, we see that the microscopic dielectric function is connected to the irreducible polarisability P by the equation:

$$\epsilon(r, r'; \omega) = \delta(r - r') - \int dr'' P(r, r''; \omega) V(r'', r'), \tag{3.25}$$

if now we take the zero-th order approximation for P and exploit the 3d translational symmetry, we get for $P_0 = \chi$:

$$\chi_{G,G'}(k, \omega) = \frac{2}{N_k} \sum_{n_1 k_1, n_2 k_2} (f_{n_1 k_1} - f_{n_2 k_2}) \frac{\langle n_1 k_1 | e^{-i(k+G)r} | n_2 k_2 \rangle \langle n_2 k_2 | e^{i(k+G')r'} | n_1 k_1 \rangle}{\varepsilon_{n_1 k_1} - \varepsilon_{n_2 k_2} - \omega - i\delta} \tag{3.26}$$

where G and G' are reciprocal lattice vectors, k belongs to the first Brillouin zone, $|n_i k_i\rangle$ are Kohn-Sham states labeled by their band and wave vectors indices and $E_{n_i k_i}$ are their corresponding energies. From eq. 3.26, we see that only transitions between Kohn-Sham states with k vectors that differ by $k + G_0$ (here G_0 is a generic reciprocal lattice vector) contribute to the response function.

In terms of $\chi_{G,G'}(k, \omega)$, we can get via eq. 3.25, the dielectric matrix in reciprocal space:

$$\epsilon_{G,G'}(q, \omega) = \delta_{G,G'} - V(q + G)\chi_{G,G'}(q, \omega). \tag{3.27}$$

From this, we can easily write the screened interaction in reciprocal space as:

$$W_{G,G'}(q, \omega) = \epsilon_{G,G'}^{-1}(q, \omega)v(q + G') \tag{3.28}$$

which is connected to the real space by the Fourier transform given by:

$$W(r, r'; \omega) = \sum_{q,G,G'} e^{i(q+G)r} W_{G,G'}(q, \omega) e^{-i(q+G')r'}. \tag{3.29}$$

3.3.1 The plasmon pole approximation for the dielectric function

Looking at the equations 3.27, 3.28 and 3.29, we can easily see that in order to get the screened Coulomb interaction for all the frequency range, in principle the inversion of the microscopic dielectric matrix for each frequency is needed; of course this is a huge bottleneck. To avoid this procedure, a model of the inverse of the dielectric function, within what is called the *plasmon pole approximation* is typically made, and is performed all through the calculations presented in this thesis.

This kind of model [67], is based on the observation that, typically, the imaginary part of $\epsilon_{GG'}^{-1}(q, \omega)$ is a peaked function of ω. It is then assumed that the for each q point, $\epsilon_{GG'}^{-1}(q, \omega)$ has the form given by:

$$\epsilon_{GG'}^{-1}(q, \omega) = \delta_{GG'} + \frac{\Omega_{GG'}^2(q)}{\omega^2 - \tilde{\omega}_{GG'}^2(q)};\tag{3.30}$$

the unknown quantities $\Omega_{GG'}(q)$ and $\tilde{\omega}_{GG'}(q)$ can be fixed by imposing the fulfilment of some specific sum rules, like in ref. [67]; or by calculating the exact value of $\epsilon_{GG'}^{-1}(q, \omega)$ for two chosen frequencies on the imaginary axis, typically for $\omega = 0$ and ω equal to some value above the plasma frequency, like in ref. [68]. In the code [69], that we have used for the calculation presented in this thesis, the latter procedure is implemented. We may see that, even if the plasmon pole approximation reduces the computational effort needed to compute $\epsilon_{GG'}^{-1}(q, \omega)$, it still remains a formidable task because it requires the inversion of the dielectric matrix two times for each q point, while the dielectric matrix can easily be very large.

Typically the plasmon pole approximation is known to work very well for semiconductors, the reason for this can be ascribed to the fact that in the calculation of the self energy, an integral over the frequencies is required; thus, if the model is capable to catch the average features of the true dielectric function, the calculation should remain very accurate. In addition, the plasmon pole approximation is often a quite good approximation of the true dielectric matrix like it has been shown in ref. [67] for the case of bulk silicon or in ref. [70] for the case of a sodium nanocluster. Moreover the quasiparticle energies seem to be quite insensitive to the different plasmon pole model used [71, 72]. The drawback of the plasmon pole model is that it can not be used if one is interested in calculating the lifetimes of the quasiparticles, this is because, in this case, a much more accurate description of the imaginary part of the self energy is required [73, 74].

3.3.2 Actual implementation: perturbative approach

Now, within the plasmon pole approximation, we have an expression for $\epsilon_{GG'}^{-1}(q,\omega)$; we can therefore calculate the screened Coulomb interaction W. From eq. 3.23 we have also the Green's function; putting all this ingredients together we can build the self energy operator. What do we do next?

We could remember we are trying to solve the Hedin's equation within an iterative scheme; as a consequence, we should continue the iterations using the new Green's function and self energy in the expression of the vertex, then update the polarisability and so on so forth until selfconsistency has been reached; this would mean a complete solution of the Hedin's equation which is out of reach of computational power for any realistic system.

However, there is an other kind of self consistency which might be thought desirable: the self energy depends on the Green's function and one could desire to obtain a Green's function and a self energy which are consistent with each other, while keeping the vertex equal to the identity. The issue of selfconsistency within GW is still very controversial, we will devote a whole paragraph (3.4.2) to its discussion; for the moment we want to describe the commonly applied method which relies on a *'single shot GW'* (here named G_0W_0), in which the selfenergy operator is built using the DFT eigenvalues and eigenfunctions, and the quasiparticle energies are computed using first order correction to the DFT energies.

In fact let's look again at the quasiparticle equation 3.16, computed at the quasiparticle energy ϵ_{nk}^{QP},

$$\left(-\frac{\nabla_r^2}{2} + V_{ext}(r) + V_H(r) \right)\psi_{nk}^{QP}(r)) + \int dr' \Sigma(r,r',\epsilon_{nk}^{QP})\psi_{nk}^{QP}(r') = \epsilon_{nk}^{QP}\psi_{nk}^{QP}(r)$$

and at the Kohn and Sham equation 2.20

$$\left(-\frac{\nabla_r^2}{2} + V_{ext}(r) + V_H(r) \right)\phi_{nk}^{KS} + V_{xc}(r)\phi_{nk}^{KS} = \varepsilon_{nk}^{KS}\phi_{nk}^{KS},$$

as already mentioned the two equations look very similar. If we assume now that $\psi_{nk}^{QP}(r) \sim \phi_{nk}^{KS}$ [4], multiplying both equations per ϕ_{nk}^{KS} and integrating over

---

[4]The DFT wave functions are not only the best available, they are also, usually, relatively close to the real quasiparticle wave functions, as it has been seen for the case of bulk silicon by Hybertsen and Louie in [67] and as it has been discussed by Hedin in [75]. However there are cases, like the (110) surface of GaAs [76], or SiH$_4$ [77], in which the two kind of wave function may differ more significantly. It has also been shown recently [78], that while reaching a self consistent solution of the quasiparticle equation the wave function may deviate from the original DFT ones. In general, [79], care should be put also in the case when the original DFT band gap is much smaller than the GW correction (which is the case of our surface along the JK line).

r, we would get:

$$\langle nk|\Sigma(\epsilon_{nk}^{QP})|nk\rangle - \langle nk|V_{xc}|nk\rangle = \epsilon_{nk}^{QP} - \varepsilon_{nk}^{KS}. \qquad (3.31)$$

We can see that the self energy operator is computed here at the quasiparticle energy which is unknown, so our task is still not completed and we have to expand the selfenergy operator around the DFT energy and keep only the linear term getting:

$$\Sigma(\epsilon_{nk}^{QP}) \sim \varepsilon_{nk}^{KS} + \left.\frac{\partial\Sigma(\omega)}{\partial\omega}\right|_{\omega=\varepsilon_{nk}^{KS}} (\varepsilon_{nk}^{KS} - \epsilon_{nk}^{QP}). \qquad (3.32)$$

If we use eq. 3.32 in eq. 3.31, we finally obtain for the self energy corrections to the DFT band structure:

$$\epsilon_{nk}^{QP} - \varepsilon_{nk}^{KS} = \frac{\langle nk|\Sigma(\epsilon_{nk}^{KS})|nk\rangle - \langle nk|V_{xc}|nk\rangle}{1 - \langle nk|\frac{\partial\Sigma(\omega)}{\partial\omega}|_{\omega=\varepsilon_{nk}^{KS}}|nk\rangle}. \qquad (3.33)$$

To summarise, following the flow chart presented in fig. 3.3, we see that the G_0W_0 calculation performed during this thesis followed roughly this scheme: an initial ground state calculation was done to get the input wavefunctions and energies; using this wavefunction and energies, the RPA response function χ was built, following eq. 3.26, for two frequencies for each q vector; the microscopic dielectric matrix was then computed and inverted for the two frequencies and then, via the plasmon pole approximation, the inverse macroscopic dielectric matrix was obtained for all the frequency range; finally the expectation value of the selfenergy calculated at the DFT energies and its first derivative was computed; all these pieces are then put together to get the quasiparticle correction.

The calculation of the selfenergy expectation value is usually performed dividing the self energy into two parts, the exchange selfenergy Σ_x, which arises from the contribution of the δ part of the dielectric matrix, and the correlation self-energy Σ_c that comes from the other term of the dielectric matrix.

$$\langle nk|\Sigma|nk\rangle = \langle nk|\Sigma_x|nk\rangle + \langle nk|\Sigma_c|nk\rangle \qquad (3.34)$$

where

$$\langle \Sigma_x \rangle = -\frac{4\pi}{N\Omega} \sum_{Gk'n'} \frac{1}{|q+G|^2} \langle nk|e^{i(q+G)r}|n'k'\rangle \times$$

$$\langle n'k'|e^{-i(q+G)r}|nk\rangle f_{n'k'} \qquad (3.35)$$

and

$$\langle \Sigma_c \rangle = -\frac{2\pi}{N\Omega} \sum_q \sum_{GG'} \sum_{n'k'} \frac{1}{|q+G||q+G'|} \langle nk|e^{i(q+G)r}|n'k'\rangle\langle n'k'|e^{-i(q+G')r}|nk\rangle \times$$
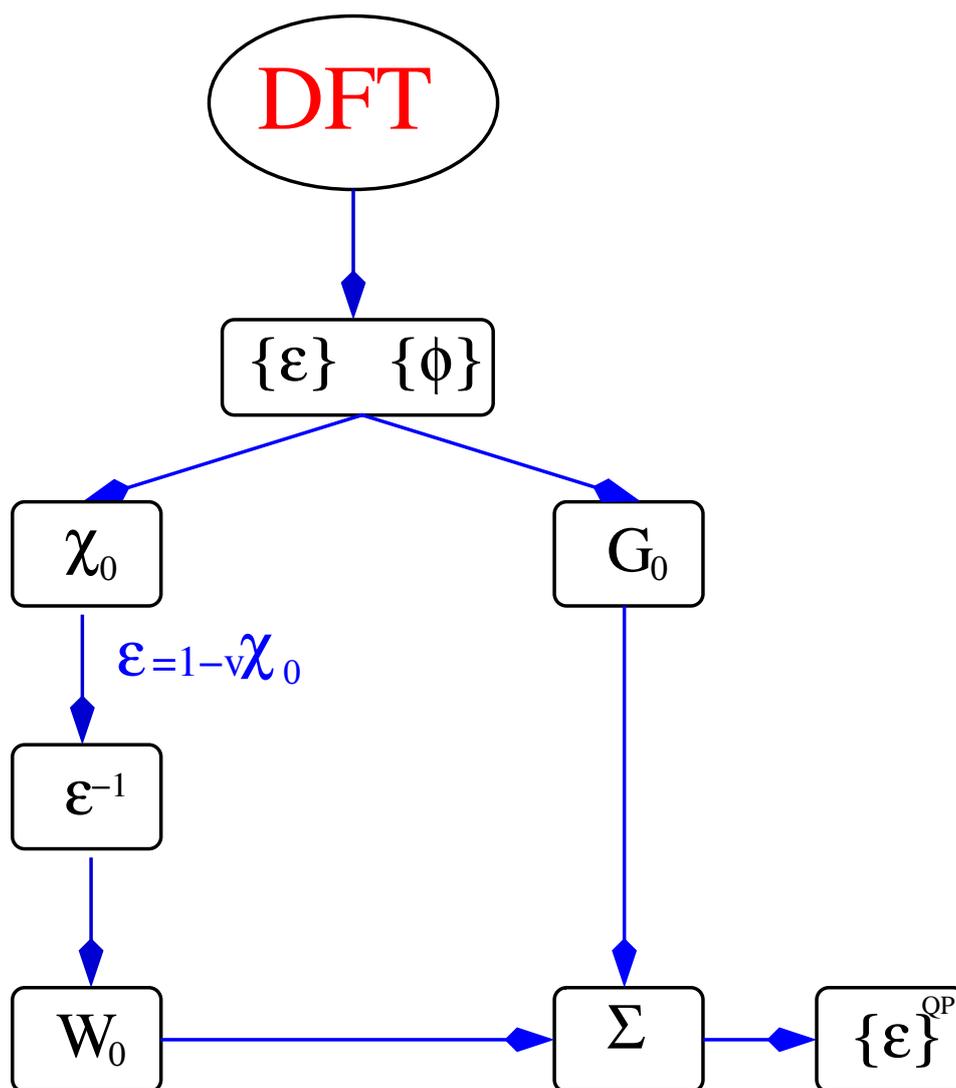
Figure 3.3: Schematic flow chart of the steps needed to perform a G_0W_0 calculation.
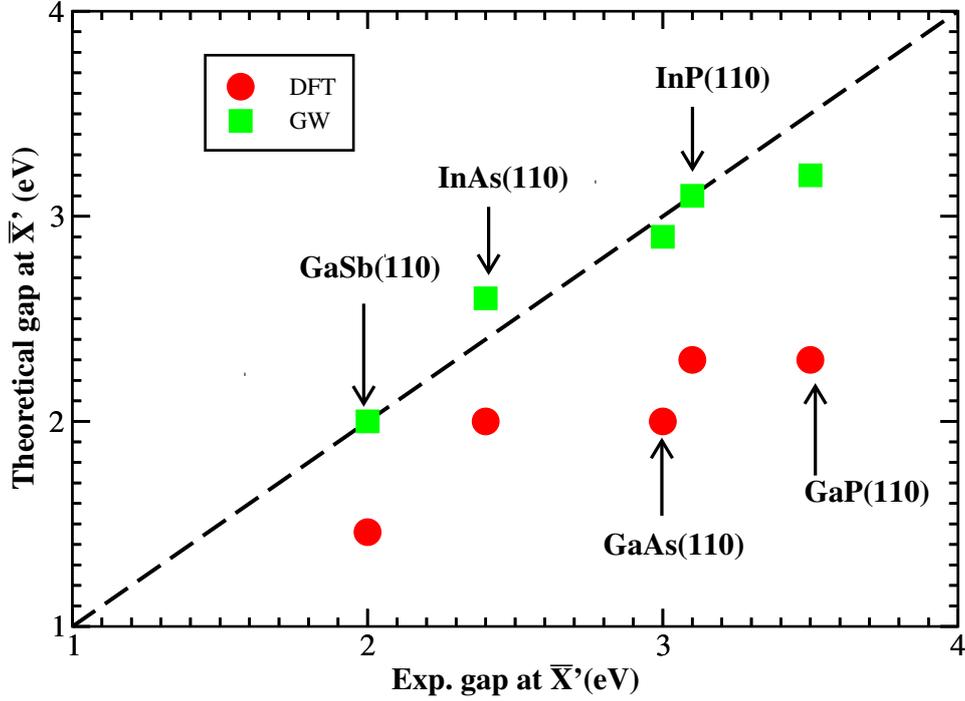
Figure 3.4: DFT (circles) and GW (squares) band gap calculation for different surfaces, in units of experimental gap. In most of the case the DFT underestimation of the gap is removed by the GW corrections which bring the theoretical value much closer to the experimental one. From ref [80].

$$\Big[\frac{\Omega^2_{GG'}(q)}{\tilde{\omega}_{GG'}(q)(\omega - \tilde{\omega}_{GG'}(q) - \epsilon^{KS}_{n'k'})} - \frac{2f_{n'k'}\Omega^2_{GG'}(q)}{(\omega - \epsilon^{KS}_{n'k'})^2 - \tilde{\omega}_{GG'}(q)^2}\Big]. \qquad (3.36)$$

Where $f_{n'k'}$ is the occupancy of the state $(n'k')$ It is very important to notice that the exchange selfenergy is energy independent and moreover it is constructed through a summation over occupied states. Its calculation is hence much less expensive than the calculation of the correlation part of the self energy, which instead involves a sum over both occupied and empty states. In fact, the convergency on the number of bands is a real bottleneck for GW calculations, and sometimes, due to the high number of states needed, it makes even the input ground state calculation really heavy. This and similar G_0W_0 schemes are the state of the art method for computing band structures and in most cases leads to a band gap in excellent agreement with experiments, solving generally the DFT band gap problem. For an example of the success of this method see fig. 3.4.

3.3.3 G_0W_0 for diamond bulk

Also in diamond bulk, DFT shows the typical underestimation of the electronic gap. We have performed G_0W_0 calculation using as input wavefunctions and energies the ones coming out from a DFT calculation employing a GGA exchange correlation functional, 45 Rydberg cutoff, at the theoretical equilibrium lattice constant of 5.7eV. A 19 k point mesh was used in the GW calculations, it included high symmetry points of the Brillouin zone; the wavefunctions were obtained in a non-selfconsistent DFT calculation using a well converged density calculated on a 10 Chadi-Cohen k point mesh.

Typically GW corrections converge at lower cutoff energies with respect to the DFT calculation, meaning that a minor number of G vectors is needed. In all our calculation we have used the same number of G vectors for the wavefunctions and for the dielectric matrix. The results of convergency test on the number of G vectors, the number of empty states and the position of the second frequency chosen for the plasmon pole approximation fit, for a generic k point, are shown on table 3.1.

G vectors	Empty Bands	ω_2 (Hartree)	$\epsilon_{n=5}^{GW}(k) - \epsilon_{n=4}^{GW}(k)$ (eV)
229	100	0.5	16.04
229	100	0.7	16.07
229	120	0.7	16.07
307	100	0.7	16.04

Table 3.1: In this table we list the results of the convergency test performed for the k point $k = (-0.375, 0.250, 0.000)$ (reciprocal lattice units), with respect to the number of G vectors, the number of empty states and the position of the second frequency ω_2 chosen for the plasmon pole approximation fit.

3.3.4 G_0W_0 for the (111) surface of diamond

We have seen in section 2.5 that the DFT band structure associated to the equilibrium geometry of the (111) surface of diamond is semimetallic in contrast with experiments and common wisdom. The absence of the gap may be related to the incapability of the ground state theory to catch some inequivalence on the atoms belonging to the surface chains (see fig. 2.1); however, this seems unlikely because dimerisation is excluded by experiments [22], only a strong unphysical buckling could open the gap, and, moreover, the surface seems not to like an antiferromagnetic ordering of the spins along the chain. Nevertheless,

there is also the possibility that the absence of the gap, at the DFT level, is *just* due to its underestimation, which, as we have been discussing, is typical of Kohn-Sham band structure. This and the next sections are devoted to the analysis of this possibility: wether or not the absence if the gap can be related to quasiparticle effects.

As a first thing, we have calculated quasiparticle corrections to the DFT energies, by evaluating the diagonal elements of $(\Sigma - V_{xc}^{DFT})$ between DFT states, using the G_0W_0 method shown in the previous section.

We have used 1995 G vectors for both the dielectric matrix and the electronic wavefunction, 800 empty bands in the correlation part of the self energy, and we have chosen a plasmon pole frequency of 0.7 Hartree.
Convergency test, on the number of k points employed, were performed on a thin slab made of 6 layers of Carbon atoms with one of the surface saturated by hydrogen. We have tested a 5 and a 13 k point mesh for the irreducible Brillouin zone. Results for the gaps on the high symmetry points of the Brillouin zone are listed in tab. 3.2

k **point**	**DFT (5k)**	**DFT (13k)**	G_0W_0 **(5k)**	G_0W_0 **(13k)**
J	0.0	0.0	0.0	0.0
K	0.0	0.1	0.1	0.0
Γ	3.2	3.2	4.1	4.2
J'	5.0	5.0	6.3	6.2

Table 3.2: Convergency test for the sampling of the Brillouin zone: the G_0W_0 gaps for the high symmetry points are reported in comparison with their corresponding DFT values. All energies are expressed in eV.

The results presented in tab. 3.2 show that, by using the 13 k point mesh, we should be converged within 0.1 eV. We have then carried out the same calculation on a slab of 12 Carbon atomic layer, without any Hydrogen. The results are presented in tab. 3.3. Comparing tab. 3.2 and 3.3, the main difference between the two slabs can be found at Γ. This is explained knowing that, close to Γ, states associated to hydrogen appear within the gap (of course they do not appear in the symmetric hydrogen free slab).
In fig. 3.5 the results on the convergency with respect to the number of G vectors in the description of the wavefunctions is presented, for the case of the 12 layers slab. This can be done by testing the value of the exchange self energy varying the number of G vectors. In fact, the computation of Σ_x,

k **point**	**DFT**	G_0W_0
J	0.2	0.2
K	0.2	0.2
Γ	4.5	6.1
J'	5.0	6.3

Table 3.3: Results for the G_0W_0 gaps for the high symmetry points are reported in comparison with the DFT value, for the 12 layers slab. We can see already that the G_0W_0 method still leads to a semimetallic surface. All energies are expressed in eV.
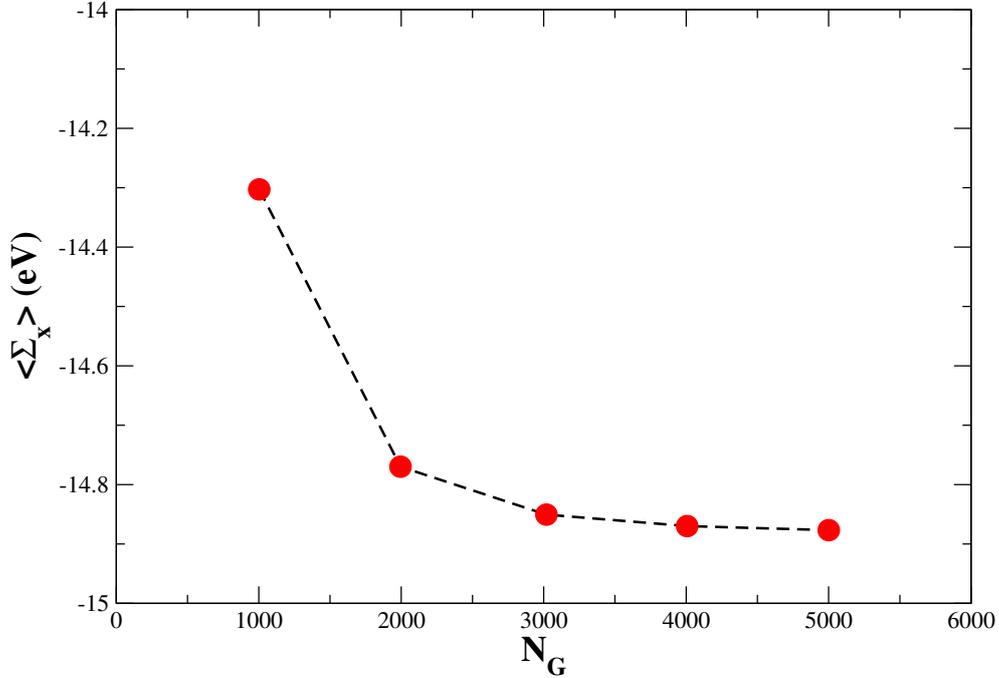
Figure 3.5: Expectation value of the exchange part of the self energy for the state belonging to the 47th band of the k point (.125000, .484375, .000000) of a 2×16 mesh. varying the number of the G vectors used in the description of the wavefunctions. 1995, the number of G vectors used for all the GW calculations presented in this thesis, seems to provide an already well converged value of Σ_x.

which involves a summation on occupied states only, is much less expensive, and calculations can be carried out also for many G vectors. As it can be seen

from fig. 3.5, 1995 is already a good number of G vectors for the degree of convergency that we aim to.

To see the effect of the G_0W_0 on the surface band structure, we present in fig. 3.6 a zoom of the band structure in the area of interest, i.e. along the JK line. As it is clear from the figure, the G_0W_0 corrections do not change the
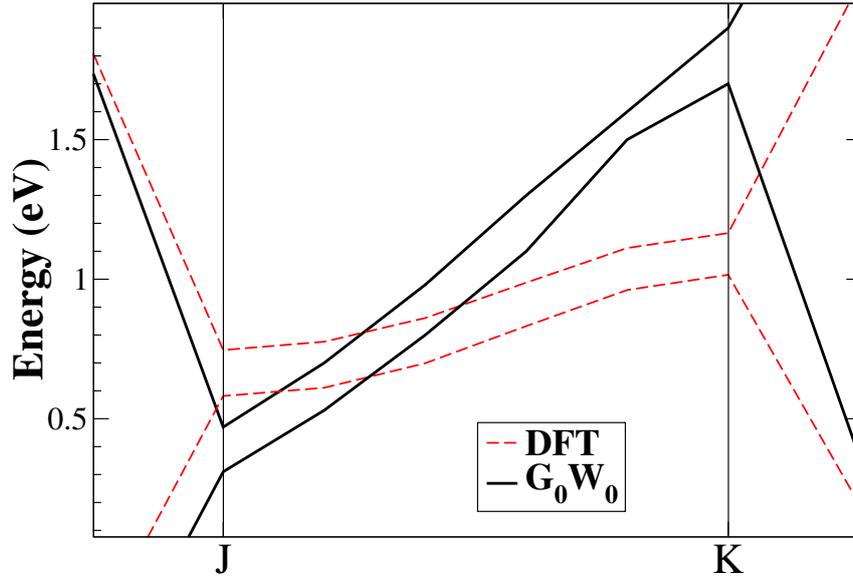
Figure 3.6: Surface states within the DFT-GGA (dashed lines) and within the G_0W_0 scheme (solid lines), along the JK direction.

situation: the upward dispersion from J to K of the nearly degenerate bands, which caused the metalicity of the surface, is still present. This is due to the fact that, at the DFT level, the two surface states at J are lower in energy than the two surface states at K. As a consequence of such energy band ordering, the states at J result occupied and those at K empty. GW corrections, which strongly depend on the occupancy of the states, can not open a gap in this case since the starting DFT band structure has a wrong occupation of the bands. Moreover, even with the wrong occupation, if the orbitals of the upper and lower bands were different one from the other (and this could happen for example in the presence of some distortion like buckling) then the exchange and correlation effects would be different for one orbital with respect to the other, and by chance a gap could open. Instead, at is shown in fig. 3.7 for the high symmetry point K, the orbitals of the upper and lower bands are almost identical: they both exhibit a dangling bond character, the only difference is on which atom on the chain they are localised around; there is no way, apart from having a different occupation, by which exchange and correlation effects should distinguish between them. The same thing happens all along the JK

line. To summarise, for this system, the usual G_0W_0 scheme does not work,
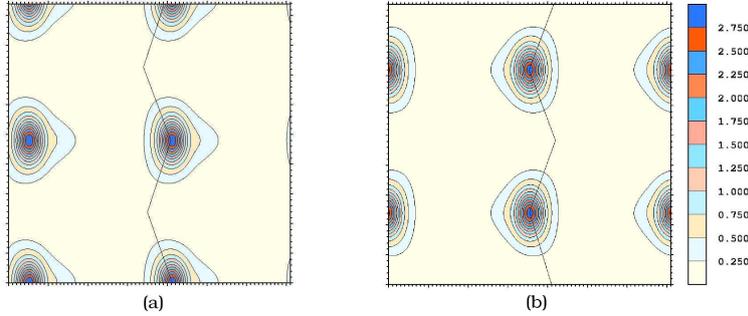
Figure 3.7: Top view of $|\phi(r)|^2$ for the upper (panel a) and lower (panel b) surface bands at K presented in figure 3.6. We can see that the two orbitals are almost indistinguishable apart from the fact that one orbital is localised on one atom of the chain and the other orbital is localised on the other atoms. The charge density is intergrated for each state under consideration for the upper half of the symmetric slab.

this is due to the specific details of the surface; in fact the dispersion along the JK line does not allow a differentiation of the upper and lower surface bands with respect to the occupation so that exchange and correlation effects can not significantly alter the gap between the two bands along the JK line. We have to look for more refined schemes.

Slater exchange as a different starting point

We have just seen how, a possible way to make exchange and correlation effect act differently for the upper and lower surface bands is to have a different occupation for the two bands. We have tested the possibility of modifying the

	J	K	Γ	J'
DFT (GGA)	0.02	0.03	3.15	4.95
DFT(Slater)	0.02	0.07	4.92	5.65

Table 3.4: Gaps between surface states at the high symmetry k points, computed using GGA and Slater approximation to the exchange and correlation potential. All energies are expressed in eV.

exchange correlation potential from the beginning (during the DFT calculation) in order to slightly open the gap and then use the G_0W_0 scheme. We have gone through this possibility by using a Slater exchange potential, but, as it can be seen from fig. 3.8, the surface stays semimetallic, thus not providing a better
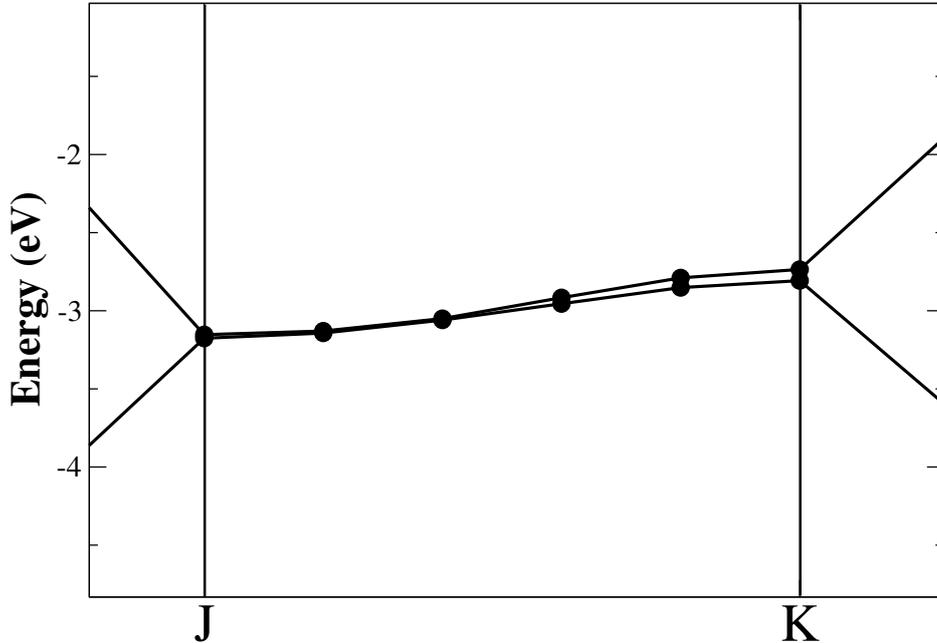
Figure 3.8: Surface bands along the JK line obtained using Slater exchange and correlation potential. The surface stays semimetallic.

starting point for the GW computation. In tab. 3.4, we see that comparing to the GGA results, the gaps in Γ and K' are increased, whereas in J and K they are still almost null. Still, exchange can not see any difference between the two bands.

3.4 Beyond G_0W_0

3.4.1 Non perturbative GW

As we have already discussed, generally, the Kohn-Sham wavefunctions are thought to be an excellent approximation to the true quasiparticle wavefunction. If this is true, the selfenergy operator should be almost diagonal in the basis of Kohn-Sham orbitals and the off diagonal elements $\Sigma_{(nk,n'k')}$ should be very small when compared to the difference of the Kohn-Sham energies of the states among which they are calculated $\epsilon_{nk}^{KS} - \epsilon_{n'k'}^{KS}$. In this case, the usual

G_0W_0 approach, that takes in consideration only the diagonal part of the self energy, works typically very well. There are some cases however in which there is the need of changing also the wavefunctions; a possible way to go beyond the G_0W_0 approach trying to include the effects of the off diagonal terms of Σ, consist in building the whole selfenergy matrix, and then diagonalise it. This approach is called *non perturbative GW* and is described in ref. [76].

Results for the (111) surface of diamond

In the specific case of our surface, the condition $|\Sigma_{(nk,n'k')}| \ll |\epsilon_{nk}^{KS} - \epsilon_{n'k'}^{KS}|$ might not hold along the JK line, because the DFT gap is vanishing. A test calculation within the non-perturbative GW scheme described in [76], including the off diagonal elements of Σ, did not give any improvement to the band gap, since the off diagonal elements are very small (two order of magnitude smaller than the diagonal ones). This is a consequence of the different symmetry of the Bloch sums at J built with the dangling bonds of the two surface atoms in the (2×1) cell: one is even under reflection with respect to the $(1\bar{1}0)$ plane, while the other one is odd (see fig 2.1). Hence they can not be mixed by the Hamiltonian. In a two-dimensional model, this is true all along the JK line; hence no gap opening can be generated by a mixing of DFT wavefunctions of dangling-bond-like surface states.
In other words, due to symmetry reason, along the JK line the self energy is already almost perfectly diagonal, and the Kohn-Sham wavefunctions are almost identical to the quasiparticle ones. This is very important to hold in mind, especially when we will be computing optical properties; in fact the Kohn-Sham wavefunctions are of a really good quality at the JK line but this might be not true anymore for k points that are very close to the JK line (thus having a very small gap) but not exactly on it.

3.4.2   Iterative *GW*

In the presence of a correct band ordering, and occupancy of the states already at the DFT-GGA level, typically G_0W_0 works really well. As we have seen, in the case of degeneracy of DFT eigenvalues, care should be put in considering also the off diagonal elements of the self energy, in order to remove the degeneracy by diagonalising the self energy operator within the degenerate space. In our case this is not working for symmetry reason. Here we have, within DFT-GGA or LDA, a wrong band ordering with respect to the energy; and the states' occupancy of the DFT band structure will not be the same as the one in the GW band structure once the degeneracy has been removed; hence we have to search for a new procedure. We have carried out GW calculations

using Kohn-Sham eigenvalues and eigenfunctions starting from an artificial electronic configuration in which the states with the correct symmetry were occupied. This approach finds a justification only if a self-consistent scheme is subsequently applied, and if the converged result does not depend on the starting point. In this spirit, we have performed a GW calculation updating the quasiparticle energies till self-consistency was obtained. While reaching self-consistency, we could choose to update the eigenvalues only in the Green function part of Σ, or in both the Green function and the screening part, which, of course, is a much more time consuming procedure.

Description of the scheme implemented

The iterative scheme followed roughly this steps:

1. An input configuration, i.e. a band structure and the corresponding wavefunctions, is provided

2. The GW corrections are calculated for a certain number of bands for each k points. We chose to evaluate the GW corrections for 10 bands above the Fermi level and 10 bands below it.

3. The GW corrections to the remaining states are evaluate through a linear extrapolation

4. The expression of the selfenergy is updated with the new energies in the Green's function and, eventually, in the screened Coulomb interaction.

5. Steps 2 to 5 are repeated until selfconsistency has been reach within a certain threshold.

In each iteration we have to find the new quasiparticle energies with respect to the old one; considering, like in the case of G_0W_0 that during this procedure the wavefunctions stay the same, we get:

$$E_{QP}^{N+1} \sim E_{QP}^N + (\langle \Sigma^{N+1}(E_{QP}^{N+1}) \rangle - \langle \Sigma^N(E_{QP}^N) \rangle); \qquad (3.37)$$

again we need the self energy calculated at the new quasiparticle energy. In the same spirit as in the G_0W_0 case, we linearise its value around the old quasiparticle energy:

$$\Sigma^{N+1}(E_{QP}^{N+1}) \sim \Sigma^{N+1}(E_{QP}^N) + \left.\frac{\partial \Sigma^{N+1}}{\partial E}\right|_{E=E_{QP}^N} (E_{QP}^{N+1} - E_{QP}^N). \qquad (3.38)$$

Taking into account that in E_{QP}^N there is the contribution of the self energy calculated linearising around the $N-1$ value and so on, we obtain that

$$E_{QP}^{N+1} \sim E_{QP}^N + \frac{E_{DFT}^0 - E_{QP}^N - \langle V_x c \rangle + \langle \Sigma^{N+1}(E_{QP}^N) \rangle}{\langle 1 - \frac{\partial \Sigma^{N+1}}{\partial E}\Big|_{E=E_{QP}^N} \rangle}. \qquad (3.39)$$

This formula has been used throughout all the calculations that will be illustrated below.

Test on a 5 k point grid

Due to the smaller computational cost, we have performed most of the tests for the thin hydrogen saturated slab with 5 k points. First of all, we simply tested the procedure to check if there was any chance that in a reasonable number of iterations we could reach self consistency. In fact we have to keep in mind that each GW calculation is very heavy, and for each iteration many GW calculations are required, since we wish to compute the corrections for as many bands as possible (actually we did it for 20 bands). We started from the artificial configuration in which the lower surface band was occupied (see fig. 3.6), we computed in this way the self energy using the old screening W_0 (and this is what we called G_1W_0), or we completely recalculated the self energy (G_1W_1), and went on until selfconsistency was reached, within 0.1 eV. As we can see

K points	DFT	G_0W_0	G_1W_0	G_1W_1	G_2W_2
J	0.02	0.03	1.07	0.57	0.62
K	0.03	0.06	1.10	0.58	0.63

Table 3.5: Results for the gaps at J and K within the iterative GW procedure for the 5 k-point grid. All energies are expressed in eV.

from the table 3.5, convergency is achieved within few iterations. Apparently the effect of updating the screening is very strong, but as we will show in the next section, this is a spurious effect of the k point sampling, in fact for the 13 k point mesh the update of the screening changes much more slightly the results.

We have also performed a test GW calculation starting from an arbitrary large (1.6 eV) and small (0.6 eV) gap between J and K. Already at the first step of the computation of the quasiparticle energies, we found for the gap a value that stayed within 0.1 eV from the converged one for the updated and the non updated screening procedure, making sure in this way that the final converged

value was independent of the starting point. The results are reported in Tab. 3.6.

	Fixed $W = W_0$ $E_g^{G_1 W_0}$			Updated W $E_g^{G_2 W_2}$	
	$E_g^{in} = 0.6$	$E_g^{in} = 1.6$	$E_g^{in} = E_g^{DFT}$	$E_g = 1.6$	$E_g^{in} = E_g^{DFT}$
J	1.08	1.12	1.07	0.72	0.62
K	1.09	1.14	1.10	0.78	0.63

Table 3.6: Calculated gaps for a test calculation starting from an arbitrary small and big gap, compared with the results of the iterative calculation starting from the DFT energies and artificial occupation of the gaps. We can see that both the updated and the non updated screening calculations lead to gap values very close to the converged ones.

Final results

Taking this result into account, we performed a much more elaborated calculation using 13 $\vec{k}$ points mesh in the IBZ for the 12 layers slab. The results are listed in Tab.(3.7): a minimum gap of about 1 eV opens between the surface states at J. In fig. 3.9 is the final band structure. Both procedures (updating

Gap	J	K	Γ	J'
DFT	0.16	0.15	4.52	5.03
$G_0 W_0$	0.19	0.18	6.07	6.30
$G_1 W_0$	1.06	1.06	5.82	6.27
$G_1 W_1$	0.93	0.95	5.71	6.17

Table 3.7: The gaps between surface states within DFT and for each GW self-consistency cycle. The subscripts of the GW gaps refer to the iteration number: $G_0 W_0$, $G_1 W_0$, $G_1 W_1$ values of the gap are related, respectively to the GW calculation obtained directly using the DFT eigenvalues, updating the eigenvalues only in G and updating the eigenvalues both in G and in W. All energies are expressed in eV.

only G, or updating G and W) lead to a semiconducting surface, and a good agreement with the available photoemission experiments is found. We have hence found another piece of the the puzzle concerning the electronic structure of the diamond (111)(2×1) surface; its insulating character might not be a consequence of asymmetric geometrical changes, rather of many body effects. Using this self-consistent quasiparticle approach we have found a surface-state gap of about 1 eV. As we have seen, the procedure should lead to a value

independent from the starting configuration, being it either an artificial *ad hoc* occupied configuration, or a configuration in which the conduction band was shifted rigidly by a small or big amount. Unluckily there are no inverse photoemission experiments to directly compare to; but now optical experiments are being carried out [81]; this experiments, even if not accessing directly to the electronic gap, yield information about transitions between the occupied and unoccupied surface states, as we will see in the next chapter.
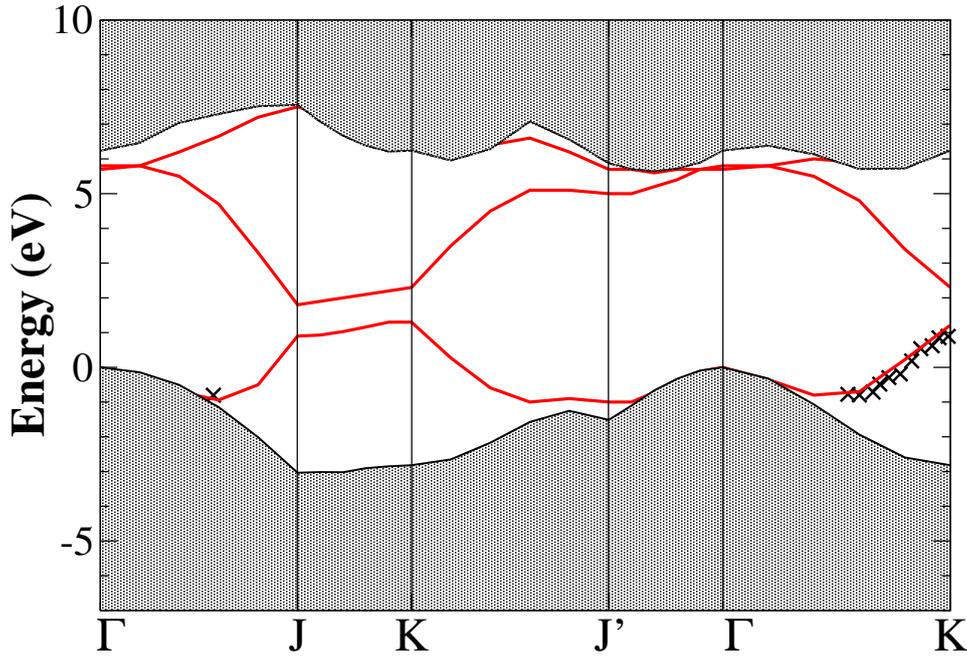
Figure 3.9: Electronic band structure of the C(111)2×1 surface within the self-consistent GW scheme. A gap of about 1 eV opens along the JK direction. Crosses: experimental results from [18, 22].

Selfconsistency and GW

We have seen in the previous section that an iterative procedure is needed in the case of the (111) surface of diamond in order to get some agreement with the experiments. As we have already mentioned, the issue of self-consistency within GW is still very controversial: there are many possible levels of self consistency (i.e. updating also the wavefunctions, including the quasiparticle renormalisation factor or not, etc.) and the different methods yield different results depending also on the implementation of an all-electron or a pseudopotential scheme [82, 83, 84, 85, 86, 87, 88]. The simple update of the quasiparticle energies has already been implemented in previous calculations

by other groups. In systems with semimetallic band structures at LDA level [89, 90], and in systems for which G_0W_0 does not give good agreement with experiments [91, 92], the update of the quasiparticle energies yields improved results. This latter approach has been also substantiated by a recent study of vertex corrections within self consistent GW: vertex corrections in W and in the self energy cancel each other to a good extent, while the energy update yields relevant changes of quasiparticle energies [93]. Also in our calculations, carried out according to this procedure, the self consistent scheme leads to a better agreement with experiments.

3.4.3 $GW\Gamma$

In this section we will focus on the possibility of going beyond the G_0W_0 approach, considering also some part of the vertex corrections within what is called the $GW\Gamma$ approximation [66]. As we will see, this approximation corrects the the quasiparticle energies, but the electronic gaps, which are differences of quasiparticle energies will not be significantly affected by this kind of corrections; there are however, other quantities, like the electron affinity which do depend on this absolute value. In their case it might be important to include this kind of vertex corrections.

In the next paragraph we will first analyse what is the electron affinity, we will try to understand how we can compute it, we will then look a little bit more in detail on the $GW\Gamma$ approximation, and finally we will give some results concerning the (111) surface of diamond.

The electron affinity

The electron affinity χ of a system is defined as the amount of energy needed to get an electron from the conduction band out into the vacuum, it is thus the difference between the vacuum level E_{vac} and the conduction band minimum E_{CBM}.

$$\chi = E_{vac} - E_{CBM}. \tag{3.40}$$

Written like this the electron affinity of a material seems a true bulk quantity which should not depend on surface orientation. However what is actually found is that χ depends on the orientation of the surface from which the electron is extracted. In fact, besides band bending, the vacuum level which can be experimentally measured is actually not the *vacuum level at infinity* $E_{vac}(\infty)$, defined as the energy of an electron at rest at an infinite distance from the sample which is of course surface-invariant [94]; but what is been measured is the *vacuum level at the surface* $E_{vac}(S)$, defined as the energy of an electron at rest close to the surface, i.e. (following ref. [94]) at a distance larger

than the interatomic distances but smaller than the size of the sample. This quantity is strongly dependent on the specificity of the surface geometry and on the presence of different adsorbates. In other words what we can access is not the energy difference between the conduction band minimum and the vacuum level but the energy barrier that the electron must overcome in order to get into the vacuum. The existence of this energy barrier can be explained in terms of the presence of the interface alone: in fact, in an ideally truncated bulk, some of the surface electron will *spill out* into the vacuum creating a negative charge region just outside the solid, at the same time the region just inside is left with a net positive charge. The total effect is thus the presence of an electric dipole at a microscopic level which gives rise to a potential which prevents other electrons to leave the solid [95]. Of course reconstructions and adsorbates strongly influence this microscopic dipole changing the potential barrier and thus the electron affinity. In particular the changes in χ due to the presence of adsorbates can be addressed very intuitively to differences in the electronegativity of the adsorbate's atoms with respect to the substrate's ones. The ability of the adsorbate atom to attract more (less) the electron, induces a polarisation of the bond, which creates a negative (positive) charged region at the surface raising (lowering) χ. In this sense we can easily understand that the presence of hydrogen lowers the energy barrier at the surface of diamond, and of course oxygen behaves in the exact opposite way. It is very important to notice that the electron affinity of a system is not affected by the presence of band bending; band bending at a surface is in fact generated by the presence of the mesoscopic surface charge region and affects in the same way both $E_{vac}(s)$ and E_{CBM} [94, 95].

Supercell, slab calculations and alignment with bulk

In order to compute the value of the electron affinity, we see from eq. (3.40), that we need to know both the energy level of the conduction band minimum and the energy level of the vacuum with respect to a generic common reference value. However E_{CBM} is obtained from a bulk calculation in which, exploiting also the periodicity of the plane wave basis set, the primitive cell is repeated an infinite time in all directions. In this calculation the vacuum level is ill defined since there is actually no vacuum. The level of the conduction band minimum can anyway be computed with respect to the bottom of the macroscopically averaged [96, 97] total electrostatic potential. To include the presence of a surface we use the supercell approach in which the surface is rapresented by a slab of finite thickness; also in this case the periodicity of the basis set is exploited and the interaction between the replicas of the slab is avoided if enough vacuum is included in the geometry. Now within the slab geometry, if

the slab is thick enough and the vacuum is deep enough, we are able to know the value of the vacuum level with respect to the bottom of the macroscopically averaged total electrostatic potential, which we finally compare with E_{CBM} to get χ.

Vertex correction to Σ

If we compare eq. (3.16) and (2.20) we see that using DFT wavefunctions and energies actually means starting from the expression of the self-energy $\Sigma(1,2) = \delta(1,2)V_{xc}(1)$. It has been shown [66] that starting from this local expression for the self-energy, the vertex function is a two-point function and the self-energy itself in the next iteration has still a "GW" form but with an effective interaction $\tilde{W}$:

$$\Sigma(1,2) = iG(1,2)\tilde{W} \tag{3.41}$$

where

$$\tilde{W} = v\left[1 - \chi^0(v + K_{xc})\right]^{-1} \tag{3.42}$$

and

$$K_{xc}(1,2) = \frac{\delta V_{xc}(1)}{\delta n(2)}. \tag{3.43}$$

It has been shown [66] that inserting in this way vertex corrections, i.e. both in the screening and in the self-energy, leads to some sort of cancellations and the effects on the band gap are negligible. However in the same paper it has been shown that the absolute values of the energies change, and this shift, while not affecting the band gaps could strongly influence the electron affinity, which depends just on one quasiparticle energy at the time.

Electron affinity calculations for diamond surfaces

In all the calculations we have carried out the following scheme: the GW and $GW\Gamma$ band gaps (E_{GAP}^{GW}, $E_{GAP}^{GW\Gamma}$) and the correction for the energy of the valence band maximum (ΔE_{VBM}^{GW}, $\Delta E_{VBM}^{GW\Gamma}$) for the bulk were computed; at this point, after the Car-Parrinello optimisation of the slab geometry was performed, the vacuum level E_{vac} of the slab was computed. E_{vac} was aligned with the energy levels of the bulk calculation by computing the difference, ΔE_{all}, between the macroscopically averaged total electrostatic potential of the bulk and the one of the bulk region in the slab calculation, and by shifting E_{vac} of this amount. The electron affinity χ was finally calculated as:

$$\chi_{GW} = E_{vac} + \Delta E_{all} - (E_{VBM} + \Delta E_{VBM}^{GW} + E_{GAP}^{GW}). \tag{3.44}$$

	Γ	X	L	E_{GAP}	Top Valence (Γ)	Bottom Conduction (Γ)
LDA	5.65	11.27	11,39	4.26	4.32	9.97
GW	7.43	12.96	13.42	5.7	4.32-0.64	9.97+1.14
GWΓ	7.39	13.0	13.42	5.7	4.32-0.12	9.97+1.62

Table 3.8: DFT, GW and GWΓ gaps of bulk diamond at high symmetry points, and values of the top of the valence and bottom of the conduction bands at Γ point. All energies are expressed in eV.

In this case we have used for the exchange correlation potential the Local Density Approximation (LDA) in the parametrisation of Perdew and Zunger [40].

GW and GWΓ on diamond bulk

The ground state density was computed using a 10 Chadi-Cohen [46] k-point mesh in the irreducible Brillouin zone. With this density, the LDA energies for the high symmetry points Γ, X and L were calculated. GW and GWΓ corrections to LDA energies were computed within the plasmon pole model [68] for the screened Coulomb interaction using 331 G-vectors and 120 empty bands. The results for the electronic gaps are shown in Tab. 3.8, where the absolute shifts of the bands at Γ (the valence band maximum) are also shown. As we can see, the GWΓ and GW gaps are practically the same, whereas the absolute shift of the valence and conduction bands at Γ point is quite different. As we will see this difference leads to a significative effect for the electron affinity.

C (111) and (100) surfaces

The Car-Parrinello optimisation was performed using a 16 layer slab, plus 6 layers of vacuum, with a cutoff energy of 40 Ryd, in the case of the (111) surface; for the (100), instead, we relaxed the geometry for a slab formed by 12 layers of carbon and 18 layers of vacuum, with a cutoff of 40 Ryd and a 16 k-point mesh. After aligning the slab calculation to the bulk energies, following

	χ_{DFT}	χ_{GW}	$\chi_{GW\Gamma}$	Exp
C(111)2×1	1.6	0.8	0.3	0.38 [98]
C(001)2×1	1.94	1.14	0.62	0.5 [99]

Table 3.9: Electron affinity for the diamond (111) and (100) surface computed at the DFT, GW and GWΓ level. All energies are in eV.

(3.44), we computed the electron affinity. The results are summarised in Tab. 3.9. As we can see, in both cases, the inclusion of the vertex correction seems to be imporatnt to gain a good agreement with experiments.

Chapter 4

Optical properties

In this chapter we will go through the theory that lies behind the calculations of the optical properties of diamond (111) presented in this thesis. This formalism develops within the framework of many body perturbation theory, which allows the computation of the microscopic dielectric function at different levels of sophistication. These different levels are connected to the possibility of including in the spectra different physical effects such as local fields effect, which take into account the inhomogeneity of the system, and the electron hole interaction which takes into account the correlation of the electrons and the holes present in a polarised medium. First, we will try to get an idea about the ingredients needed to describe the neutral excitations in solids. We will then see how the macroscopic dielectric function, which is the quantity directly involved in the experiments, can be connected to the microscopic dielectric function that comes out from the calculations. Then, within an iteration scheme of Hedin's equations, the Bethe Salpeter equation for a generalised four-point polarisability will be derived, and we will see how the Bethe Salpeter equation can be cast into a two body Hamiltonian problem. We will then see how it is typically solved in practice, and apply the method for the computation of a diamond bulk absorption spectra. Then, we will look at Reflectance Anisotropy Spectroscopy (RAS) which is an optical spectroscopy sensitive to surfaces. We will see that the RAS spectra are connected to the slab polarisability α and we will see how α can be computed. We will finally compute the RAS spectra for diamond (111); we will see how the transitions between surface bands give strong signal at low energies. This signals, if detected, could give another proof of the existence of a finite gap between surface state and, moreover, give an indirect information about its magnitude.

4.1 Neutral excitations in solids

In order to describe optical or electron energy loss (EEL) spectra, one has to take into account excited states of a completely different nature with respect to the ones involved in photoemission spectroscopy. In the latter the final state of the system is charged since one electron has been removed or added to it; in the first, instead, the system, exchanging energy with the probe, is left in what is called a neutral excited state. Excitons and plasmons are examples of neutral excited states that are mostly detected by optical and EEL spectroscopy respectively; in a very schematic way excitons can be thought as bounded electron-hole states, while plasmons can be seen as coherent oscillations of electronic charge [100]. The structures associated with excitons are usually located in the low energy part of the spectra, while plasmons' structures are located at higher energies. The main news about this kind of excitations is that the excited electrons stay in the solid, and one might imagine an absorption experiment as a sequence of a direct and an inverse photoemission experiment. If that was the case, one could think that in order to get a good absorption spectra what is needed is actually 'just' a very good band structure. In terms of calculations this would mean that an accurate GW calculation, i.e. the solution of an effective one body problem, would do the main job in obtaining a good optical spectra. Indeed this is not the case, and this can be easily understood noticing, for example, that the conduction band levels represent addition energies, meaning the energy required to add an electron to a N particle state. Instead, in the case of the absorption, thought as a sequence of two photoemission experiments, the electron would be added to a N-1 particle state, and it is obvious to think that the addition energies in this case would be different. So, one might think that, even if a single GW calculation is not enough, maybe two GW calculations could do the job: one for the N and one for the N-1 particle system. However, this is still not the case because, especially in solids, the fact that the hole is formed at the same time that the electron is excited, in other words the correlation between the electron-hole pair, are effects that can not be neglected. In order to take the electron-hole correlation into account, one has to consider the presence of the hole and the electron simultaneously; it is clear, then, that, at least within the Green's function formalism, any effective single particle problem could not describe this phenomena accurately [1], and that one has to resort to an effective two body problem. Indeed this is the case: we will see that the solution of the Bethe-Salpeter equation from which very

---

[1]Actually, optical properties can also be calculated exploiting an effective single particle formalism within the Time Dependent Density Functional Theory (TDDFT) [58, 101] which is a generalisation of DFT to time dependent external potentials. Within TDDFT the polarisability of the system χ can be expressed in terms of the independent particle polarisability

accurate optical spectra can be obtained, can be actually cast into an effective two body Hamiltonian problem.

4.2 Microscopic and macroscopic dielectric functions

In this section we will try to understand the link between the macroscopic quantities involved in experiments and the microscopic ones that come out from ab-initio calculations. We will see that care must be put into the averaging process and that not all the macroscopic quantities are just simply connected to their microscopic counterparts by a macroscopic average. The most relevant example of this fact is given by the macroscopic dielectric function itself.

The section is organised as follows: first we will just review some basic relations relevant in the macroscopic world connecting the absorption of an electromagentic field in a medium and its macroscopic dielectric function. Then we will turn to the microscopic world and see how the density density response function is connected to the reducible polarisability and to the microscopic dielectric function. We will then put things together and get the relation between the microscopic and macroscopic worlds. At last we will see that the macroscopic dielectric function can be cast in terms of a modified reducible polarisability in a form suitable to treat the local fields effects and the electron hole interaction on the same footing.

χ_0 through a Dyson like equation:

$$\chi(r, r'; \omega) = \chi_0(r, r'; \omega) + \int \int dr'' dr''' \Big[\chi_0(r, r''; \omega) \Big(v(r'', r''') + f_{xc}(r'', r''') \Big) \chi(r''', r'; \omega) \Big]$$

where f_{xc} is called the exchange-correlation kernel, and it is the functional derivative of the exchange correlation time dependent potential with respect to the time dependent density, namely:

$$f_{xc}(r', r') = \frac{\delta V_{xc}([n], r, t)}{\delta n(r', t')}.$$

. Unfortunately the natural generalisation of LDA to time dependent problems, namely the adiabatic local density approximation (ADLDA) for the kernel, does not yield accurate spectra in the case of extended systems. More complicated kernels, whose expression is derived by a direct comparison with the MBPT formalism, are nowdays available [102, 103]. In the case of MBPT-derived kernel the quality of the spectra is excellent also for extended systems; however the actual implementation are computationally as heavy as a full MBPT calculation.

4.2.1   Macroscopic electrodynamics

Within classical electrodynamics in CGS units, the electromagnetic field in the presence of a dielectric medium is given by the Maxwell equations [104]:

$$\nabla \wedge H = \frac{1}{c}\frac{\partial D}{\partial t} + \frac{4\pi}{c}j_{ext}; \qquad \nabla \wedge E = -\frac{1}{c}\frac{\partial B}{\partial t}$$

$$\nabla D = 4\pi\rho_{ext}; \qquad \nabla B = 0 \qquad (4.1)$$

here E and H are the total electric and magnetic field, D is the electric displacement, B is the magnetic induction, and ρ_{ext} and j_{ext} are the external charge density and current. It is very important to notice that this equation refers to quantity which are averaged within volumes that are *'physically infinitesimal'* [104] but still contain many atoms and lattice sites.

Up to know this equations are completely general and apparently no specific properties of the material under consideration show up. The peculiarity of each material comes out when one considers the link between the various quantity present in eqns. 4.1: the total electric field E in a dielectric will be different from the applied external field D, and this difference is due to the specific capability of the material to polarise. As a matter of fact, $D = E + 4\pi P$, where P is the polarisation of the system. Neglecting non linear effects, the polarisation is related to the electric field through the electric susceptibility tensor χ_{ij}:

$$P_i(r,t) = \int dr'dt'\chi_{ij}(r,r';t-t')E_j(r',t') = \int dr'dt'\chi_{ij}(r-r';t-t')E_j(r',t').$$
$$(4.2)$$

The second equation holds since on the macroscopic scale the system is homogeneous. In Fourier space eq. 4.2 becomes:

$$P_i(q,\omega) = \chi_{ij}(q,\omega)E_j(q,\omega). \qquad (4.3)$$

The dielectric tensor ϵ, which connects the electric displacement to the electric field, namely in Fourier space $D = \epsilon E$, is then related to the electric susceptibility tensor by:

$$\epsilon_{ij}(q,\omega) = \delta_{ij} + 4\pi\chi_{ij}(q,\omega). \qquad (4.4)$$

For the frequency range we are interested, the wavelength of the light is much bigger than the typical lattice scale, and one is actually interested in knowing only $\epsilon_{ij}(0,\omega)$ which we will more briefly call $\epsilon(\omega)$ [2]. In other words, for this range of frequency, the wavelengths of the field are such that in eq. 4.2, the electric field can be considered constant in the range in which the electric

---

[2]For cubic systems, moreover, the tensor ϵ_{ij} is diagonal and there actually there is only one independent component, i.e. $\epsilon_{i,i} = \epsilon_{j,j} = \epsilon$; $i,j = x,y,z$

susceptibility is different from zero; so that E can be taken out of the integral. Eq. 4.2 then becomes:

$$P_i(r,t) = \int dt' (\int dr' \chi_{ij}(r-r', t-t')) E_j(r,t') = \int dt' \chi_{ij}(q=0, t-t') E_j(r,t').$$

We have to note that the electronic characteristic frequencies are within the frequency range of our interest, and, in fact, we can not neglect the frequency dependency of the dielectric tensor.

The macroscopic optical properties of a material can also be expressed in terms of a complex refractive index $\tilde{n} = n + i\kappa$. In terms of $\tilde{n}$ the propagation of the electric field in the medium in the absence of external sources is given by:

$$E(x,t) = E_0 e^{i\frac{\omega}{c}x\tilde{n}} e^{-i\omega t} = E_0 e^{i\frac{\omega}{c}nx} e^{-\frac{\omega}{c}\kappa x} e^{-i\omega t}. \tag{4.5}$$

The complex refractive index is connected to the dielectric function by $\tilde{n} = n + i\kappa = \sqrt{\epsilon}$; which give the real and imaginary part of the dielectric function in terms of the refractive index n and of the extinction coefficients κ:

$$\epsilon_1 = n^2 - \kappa^2$$

$$\epsilon_2 = 2n\kappa. \tag{4.6}$$

From eq. 4.5 we can see that, when light travels from the point x to the point x' in a medium of complex refractive index $\tilde{n}$, its intensity is reduced following:

$$I(x') = I(x)e^{-\frac{2\omega}{c}\kappa(x'-x)} = I(x)e^{-\alpha(x'-x)}. \tag{4.7}$$

This relation defines the absorption coefficient α as:

$$\alpha = \frac{2\omega}{c}\kappa = \frac{\omega\epsilon_2}{nc} \tag{4.8}$$

from which we see that the imaginary part of the dielectric function is linearly related to the absorption coefficient.

Experimentally the optical functions are measured by different techniques, in relation to the interesting frequency range and to the type of materials: absorption experiments give a direct access to the absorption coefficients, but since one has to measure the intensity of the light *transmitted* through the sample, which decreases exponentially with the sample thickness, thinner and thinner sample would be required. Another possibility is to perform reflection measurements, since the reflectance of the sample is related to the complex refractive index by the Fresnel formulae [105]. The drawback of this technique is the sensitivity of the spectra to the sample surface contaminants. Last

but not least is the determination of the complex dielectric function through the technique of optical ellipsometry, In this technique a linear polarised light (neither s- nor p- polarised [3]) impinges the sample at an angle ϕ and the ratio $\sigma = \frac{r_p}{r_s}$ between the p and s components reflectivities is measured. The complex dielectric function is then given by [106]:

$$\epsilon = \sin^2 \phi + \sin^2 \phi \tan^2 \phi \left(\frac{1 - \sigma}{1 + \sigma} \right)^2.$$

4.2.2 Response functions and time ordered Green's function

In this section we want to derive some relations between microscopic quantities such as the microscopically induced charge density, the microscopic dielectric tensor etc... that will turn out useful in the next paragraph where we will derive the connection between the macroscopic world and the microscopic one. The theory of the linear response to an external perturbation allows us to express the charge density $n_{ind}(x)$ induced by the presence of an external scalar potential $\phi_{ext}(x,t)$ in terms of a response function $\chi(x,t;x't')$, so that:

$$n_{ind}(x,t) = \int_{-\infty}^{\infty} \int d^3x' dt' \chi(x,t;x't') \phi_{ext}(x',t'). \tag{4.9}$$

Within the linear response theory it can be shown (see ref.[59]) that $\chi(x,t;x't')$ is the retarded density-density correlation function which in turn, for $\omega > 0$, is equal to the time ordered reducible polarisability P_{red} [4]. The induced charge in turns creates a potential $\phi_{ind}(x,t)$ given by

$$\phi_{ind}(x,t) = \int d^3x' v(x - x') n_{ind}(x',t) \tag{4.10}$$

so that the total potential $\phi_{tot}(x,t)$ can be written as [5]

$$\phi_{tot} = (1 + vP_{red})\phi_{ext} = \epsilon^{-1}\phi_{ext}. \tag{4.11}$$

---

[3]p-polarisation corresponds to a polarisation parallel to the plane of incidence, while s-polarisation to a polarisation perpendicular to it.

[4]In the previous chapter we have introduced the irreducible polarisability P and we have seen how it is connected to the screened Coulomb interaction and to the microscopic dielectric function, namely $W = v + WPv$, and $\epsilon = 1 - Pv$. The reducible polarisability, see ref. [59], is connected to the irreducible polarisability via a Dyson like equation $P_{red} = P + PvP_{red}$. As a consequence the screened Coulomb interaction now is $W = v + vP_{red}v$, and the inverse microscopic dielectric function is $\epsilon^{-1} = 1 + vP_{red}$

[5]From now on we imply that v, P_{red}, ϵ^{-1}, act as integral operators, and that all the relations are valid for $\omega > 0$; else we replace the polarisabilities with their corresponding retarded version.

This allows us [6] to identify the role played at the microscopic scale by the inverse microscopic dielectric function with the role of the inverse macroscopic dielectric function present in classical electrodynamics. Now, which is the role of the irreducible polarisability P? Let's suppose to have an integral operator X that connects the induced charge density to the total potential of the solid so that $n_{ind} = X\phi_{tot}$ then

$$\phi_{ind} = vX\phi_{tot} = \phi_{tot} - \phi_{ext} \Rightarrow \phi_{tot} = (1 - vX)^{-1}\phi_{ext}$$

but we know that $\phi_{tot} = \epsilon^{-1}\phi_{ext}$ so that $\epsilon = (1 - vX)$. If we remember that the microscopic dielectric function is given, in terms of P, by the relation $\epsilon = 1 - vP$, we can identify P and X. So the irreducible polarisation, for positive frequencies, is the functional derivative of the charge density with respect to the *total* potential:

$$n_{ind} = P\phi_{tot}. \tag{4.12}$$

4.2.3 Macroscopic dielectric function, local fields effects

In the last two paragraphs we have seen how the inverse of the microscopic dielectric function and the inverse of the macroscopic one play the same role in connecting the total potential to the external perturbation. Here, following the work of Adler [107], we will get the macroscopic dielectric function in terms of the microscopic one.

Let's suppose that the perturbing external potential is given by a monochromatic, slowly varying potential:

$$\phi^{ext} = \phi^{ext}(q, \omega)e^{i(qr-\omega t)}. \tag{4.13}$$

The total potential, given by the sum of the external potential and of the one that is induced by the polarisation of the medium, will present $q + G$ wavevector components (here G is a generic reciprocal lattice vector) and thus will vary much faster in space, so that in a general way we can write [7]:

$$\phi^{tot}(r, t) = \sum_G \phi(q, G; \omega)e^{i(q+G)r}e^{-i\omega t} \tag{4.14}$$

---

[6]We assume from now on that we are talking about positive frequencies, or about retarded functions.

[7]We can see how the total potential gains the $q+G$ components considering its expression in terms of the microscopic dielectric function:

$$\phi^{tot}(r, t) = \int dr'dt'\epsilon^{-1}(r, r'; t - t')\phi^{ext}(r', t') == \phi^{ext}(q, \omega)\int dr'\epsilon^{-1}(r, r'; \omega)e^{i(qr'-\omega t)} =$$

The macroscopic potential, equal to the space average of eq. 4.14, will then be given by:

$$\langle \phi^{tot} \rangle = \phi(q, 0; \omega)^{i(qr - \omega t)}.$$ (4.15)

This is because in eq. 4.14, the space dependence is only contained in the exponential $e^{i(q+G)r}$, and the $G \neq 0$ components are rapidly oscillating in real space and are thus averaged to 0.

In the same way, we can think that the induced potential is gaining $q + G$ components, and we can write it as:

$$\phi^{ind}(r, t) = \sum_G \phi^{ind}(q, G; \omega) e^{i(q+G)r} e^{-i\omega t}.$$ (4.16)

Also in this case we can perform a space average getting:

$$\langle \phi^{ind} \rangle = \phi^{ind}(q, 0; \omega)^{i(qr - \omega t)}.$$ (4.17)

The longitudinal[8] macroscopic dielectric function is then given by:

$$\epsilon^L(q, \omega) = 1 - \frac{\langle \phi^{ind} \rangle}{\langle \phi^{tot} \rangle} = 1 - \frac{\phi^{ind}(q, 0; \omega)}{\phi(q, 0; \omega)} = \frac{\phi^{ext}}{\phi(q, 0; \omega)}.$$ (4.18)

So we want to find the ratio $\phi^{ext} / \phi(q, 0; \omega)$ in terms of microscopic quantities. To do this, we remember from the previous section, that the induced charge density, in terms of the irreducible polarisability P, is given by $n^{ind} = P\phi^{tot}$; so that we can relate ϕ^{ind} and ϕ^{tot} with:

$$\phi^{ind}(r, t) = \int dr' V(r - r') n^{ind}(r', t) =$$

$$= \int dr' dr'' V(r - r') P(r', r''; \omega) \phi^{tot}(r'', \omega) e^{-i\omega t}.$$

$$= \phi^{ext}(q, \omega) e^{-i\omega t} \sum_{G,G'} \int_{1BZ} d^3k \epsilon_{GG'}^{-1}(k) \int dr' e^{-i(k+G)r} e^{i(k+G+q)r'} =$$

$$= \phi^{ext}(q, \omega) e^{-i\omega t} \sum_{G,G'} \epsilon_{GG'}^{-1}(q) e^{-i(q+G)r} = \sum_G \phi(q, G; \omega) e^{i(q+G)r} e^{-i\omega t}$$

where $\phi(q, G; \omega) = \sum_{G'} \phi^{ext}(q, \omega) \epsilon_{GG'}^{-1}(q)$.
Here we have used the Fourier expansion for the inverse microscopic dielectric function defined by: $\epsilon^{-1}(r, r'; \omega) = \sum_{G,G'} \int_{1BZ} d^3k \epsilon_{GG'}^{-1}(k) e^{-i(k+G)r} e^{i(k+G')r'}$, and $\epsilon_{GG'}^{-1}(k) = \epsilon^{-1}(k + G, k + G')$.

[8]Here we are assuming that the electric field is generated by a scalar potential, and thus are talking about the *longitudinal* dielectric function.

Which, exploiting the Fourier representation of the operators, becomes:

$$\phi^{ind}(r,t) = \sum_{G,G'} \frac{1}{|q+G|^2} P_{G,G'}(q,\omega)\phi^{tot}(q,G';\omega)e^{i(q+G)r}e^{-i\omega t}. \quad (4.19)$$

Comparing eq. 4.16 and eq. 4.19, we can express the Fourier component of the induced potential as:

$$\phi^{ind}(q,G;\omega) = \frac{1}{|q+G|^2} \sum_{G'} P_{G,G'}(q,\omega)\phi^{tot}(q,G';\omega). \quad (4.20)$$

However, referring to eq. 4.18, we still want to obtain a relation between the Fourier component of the total potential and the external potential. In order to do this, we remember that the microscopic dielectric function is $\epsilon = 1 - Pv$. From this equation, expliciting P in Fourier space we get:

$$P_{G,G'}(q,\omega) = \delta_{G,G'}|q+G'|^2 - |q+G'|^2\epsilon_{G,G'}(q,\omega). \quad (4.21)$$

Inserting this relation in eq. 4.20, we get:

$$\phi^{ind}(q,G;\omega) = \phi^{tot}(q,G;\omega) - \frac{1}{|q+G|^2} \sum_{G'} |q+G'|^2\epsilon_{G,G'}(q,\omega)\phi^{tot}(q,G';\omega)$$

which gives

$$|q+G|^2 \underbrace{\left[\phi^{tot}(q,G;\omega) - \phi^{ind}(q,G;\omega)\right]}_{\phi^{ext}(q,\omega)\delta_{G,0}} = \sum_{G'} |q+G'|^2\epsilon_{G,G'}(q,\omega)\phi^{tot}(q,G';\omega). \quad (4.22)$$

Now, if we multiply each term by $\epsilon^{-1}_{G'',G}(q,\omega)$ and we sum up over G,

$$\epsilon^{-1}_{G'',0}(q,\omega)|q|^2\phi^{ext}(q,\omega) = \sum_{G'} \delta_{G'',G'}|q+G'|^2\phi^{tot}(q,G';\omega) \quad (4.23)$$

and expliciting ϕ^{tot}, we finally get:

$$\phi^{tot}(q,G';\omega) = |q+G''|^{-2}|q|^2\phi^{ext}(q,\omega)\epsilon^{-1}_{G'',0}(q,\omega). \quad (4.24)$$

Let's go back to the expression for the longitudinal macroscopic dielectric function given by 4.18:

$$\epsilon^L(q,\omega) = 1 - \frac{\phi^{ind}(q,0;\omega)}{\phi(q,0;\omega)} = \frac{\phi^{ext}(q;\omega)}{\phi(q,0;\omega)} =$$

$$= \frac{\phi^{ext}(q;\omega)}{|q+G''|^{-2}|q|^2\phi^{ext}(q,\omega)\epsilon^{-1}_{G'',0}(q,\omega)}\bigg|_{G''=0} = \frac{1}{\epsilon^{-1}_{0,0}(q,\omega)}. \quad (4.25)$$

In the optical limit, the transverse and longitudinal dielectric function are equal [108], so that we can generically write the macroscopic dielectric function as:

$$\epsilon_M(\omega) = \lim_{q \to 0} \frac{1}{\epsilon_{0,0}^{-1}(q,\omega)}. \tag{4.26}$$

The EEL spectra, in the limit of vanishing exchanged wave vector, is proportional to the imaginary part of the inverse of the macroscopic dielectric function; while the absorption spectra is proportional to the imaginary part of the macroscopic dielectric function (see eq.(4.27)).

$$EEL \propto \left\{ \frac{1}{\epsilon_M} \right\} \qquad Abs \propto \left\{ \epsilon_M \right\}. \tag{4.27}$$

In an inhomogeneous system, the off-diagonal terms of the dielectric matrix are non-zero and so, obviously, $\epsilon_{G=0,G'=0}$ is different from $\frac{1}{\epsilon_{G=0,G'=0}^{-1}}$. Computing the macroscopic dielectric function as prescripted by eq. (4.26) takes into account the fact that an external spatially homogeneous electric potential can induce, by mean of the polarisation of the system, an electric potential that fluctuates in space with Fourier components equal to the reciprocal lattice vectors. This are the so called local field effect. So looking at eq. (4.26) we can see that in order to take into account the local field effects we have to compute the inverse of the microscopic dielectric function $\epsilon_{G,G'}^{-1}$ and then take the macroscopic limit. The inverse of the microscopic dielectric matrix can be computed knowing the reducible polarisability P_{red} remembering that $\epsilon^{-1} = 1 + v P_{red}$, where v is the bare coulomb interaction. The reducible polarisability satisfies a Dyson-like equation that connects it to the irreducible polarisability P (see note in pg. 80).

4.2.4 Macroscopic dielectric function and modified polarisability

However there is another way to directly obtain the macroscopic dielectric function, which is more convenient because it allows to treat the excitonic and local field effects on the same footing: it can be shown (see appendix B of ref. [58]) that the macroscopic dielectric function can be rewritten in terms of a modified reducible polarisability $\bar{P}$ as:

$$\epsilon_M(\omega) = 1 - \lim_{q \to 0} v(q)_0 \bar{P}_{G=G'=0}(q,\omega). \tag{4.28}$$

The modified polarisability satisfies a Dyson like equation in terms of the irreducible polarisability P and a modified coulomb potential $\bar{v}$ very similar to

the equation for the reducible polarisability:

$$\bar{P} = P + P\bar{v}\bar{P}. \tag{4.29}$$

The modified Coulomb interaction is defined by subtracting the $G = 0$ Fourier component to the bare coulomb interaction. It's interesting to notice that the $\bar{v}$ term is entirely responsible of the local field effect. In fact, following ref. [102], if we suppose to neglect this term the macroscopic dielectric function would be given by:

$$\epsilon_M(\omega)^{no\bar{v}} = 1 - \lim_{q \to 0} v(q)_0 P_{G=G'=0}(q,\omega) \tag{4.30}$$

if we remember the definition of the microscopic dielectric function, namely $\epsilon = 1 - vP$, looking at eq. 4.30 we see that in this case

$$\epsilon_M(\omega)^{no\bar{v}} = \lim_{q \to 0} \epsilon_{G=G'=0}(q,\omega)$$

which is equivalent in neglecting the off diagonal elements of $\epsilon_{G,G'}$ in eq. 4.26 meaning that the local field effects are neglected as already discussed.

To summarise we have seen how the optical and EEL spectra are connected to the macroscopic dielectric function, the macroscopic dielectric function can be obtained in terms of the reducible polarisability P_{red} or in terms of a modified reducible polarisability $\bar{P}$. In both cases the knowledge of the irreducible polarisability P of the system is required and P can be computed within the Hedin's equations scheme 3.21 at different levels of approximation. In the next sections we will see the quality of the different spectra that we can obtain with different approximation of P.

4.3   Independent quasiparticle approximation

We have seen that the Hedin's equations can not be solved exactly but only using a self consistency scheme (a full self consistency can not be obtained for real system due to numerical limits). Adopting the same level of approximation that we have introduced to obtain GW quasiparticle energies, i.e. neglecting the vertex correction, we get the so called random phase approximation (RPA) for the dielectric matrix. Within this approximation the irreducible polarisability is given by:

$$P_{IQP}(x, x'; \omega) = 2 \sum_{i,j} (f_i - f_j) \frac{\phi_i(x)\phi_j^*(x)\phi_i^*(x')\phi_j(x')}{\epsilon_i - \epsilon_j - \omega - i\eta} \tag{4.31}$$

here (i, j) are single quasiparticle state labels, f_i and ϕ_i are the corresponding occupation and wave function respectively. From this equation, neglecting the local field effects, in the length gauge we have that the response to a longitudinal field for $q \to 0$

$$\epsilon_M(\omega) = 1 - \lim_{q \to 0} v(q) \sum_{vc} \frac{|\langle v| \exp^{-iqr} |c\rangle|^2}{\omega - (\epsilon_c - \epsilon_v) - i\eta}. \tag{4.32}$$

Here v is the coulomb interaction, $|v\rangle$ $(|c\rangle)$ represent valence (conduction) single particle state and ϵ_v (ϵ_c) is its corresponding energy. From equation 4.32 we see that the sum of all the possible independent transitions gives rise to the spectra. In this picture the structure of the spectra are located at quasiparticle energy differences and the line shape is determined by the matrix elements of e^{iqr} between a valence state and a conduction state independently for each transition. In fact in a single quasiparticle scheme the neutral excited state can be seen as a sum of independent excitation of one electron from the valence to the conduction band. Comparing the RPA spectra with experiments usually both the peak positions and the line shape are wrong. In particular, depending on which single particle scheme is used, whether DFT or GW, the position of the peaks will be shifted toward lower or higher energy. The spectra is also strongly dependent on local fields effects, which must be included for anisotropic systems like for example surfaces and nanowires. However even including local fields effects the optical spectra of solids obtained within RPA are often not even in qualitative agreement with experiments.

4.4 Bethe-Salpeter equation

In the previous section we have seen that the vertex corrections can not be completely neglected in order to obtain optical spectra in agreement with experiments. Neglecting vertex corrections means to neglect the interaction between the holes and the electrons that are formed within the polarisation process of the system. In fact, we can see from the Hedin's equation that defines the irreducible polarisability 3.21, that, setting $\Gamma = \delta\delta$, leads to an expression for P, namely $P = -iGG$, in which the hole and the electron propagators are independent one to the other. In the description of absorption experiments, in which the excited electron remains in the sample, evidently the effects of the electron-hole interaction, i.e. the excitonic effects can not be neglected. Within manybody perturbation theory it is shown that, in order to include the effects of the electron-hole interaction in the spectra, one has to introduce an effective two body Hamiltonian,the so called excitonic Hamiltonian. In a very

schematic way we will see how the excitonic Hamiltonian can be obtained.
In order to include vertex corrections, we can think to iterate once more the selfconsistency cycle of the Hedin's equations. If we start from the GW approximation for the selfenergy Σ and iterate, we obtain for the vertex function the following expression:

$$\Gamma(123) = \delta(12)\delta(13) + iW(1^+, 2) \int G(16)G(72)\Gamma(673)d(67) \qquad (4.33)$$

where we have used $\frac{\delta\Sigma}{\delta G} = iW$ and neglected the term $iG\frac{\delta W}{\delta G}$. Multiplying equation (4.33) by $G(4,1)G(2,5)$ and integrating it with respect to the variables $(1,2)$ we can obtain a Dyson like equation for a generalised three-point polarisability defined as:

$$^3P(312) = -i \int G(16)G(72)\Gamma(673)d(67). \qquad (4.34)$$

The Dyson like equation for this three-point polarisability takes the form:

$$^3P(345) = -iG(43)G(35) + \int G(41)G(25)W(1^+2)^3P(312)d(12). \qquad (4.35)$$

We can generalise this equation for a four-point polarisability introducing the four-point screened coulomb interaction $^4W(1234) = W(12)\delta(13)\delta(24)$, getting a four-point integral equation for the generalised irreducible polarisability ^4P:

$$^4P = \quad ^4P_{IQP} +^4 P^4W^4P_{IQP}. \qquad (4.36)$$

We can now generalise the equation (4.29) in terms of four-point quantities getting:

$$^4\bar{P}(1234) =^4 P(1234) + \int \quad ^4P(1256)\delta(56)\delta(78)\bar{v}(57)^4\bar{P}(7834)d(5678). \qquad (4.37)$$

Using the expression for ^4P given by eq. (4.36) in eq. (4.37) we obtain the *Bethe-Salpeter equation* for $^4\bar{P}$:

$$^4\bar{P} =^4 P_{IQP} +^4 P_{IQP}K^4\bar{P}. \qquad (4.38)$$

The kernel K of this integral equation contains an electron-hole exchange contribution $\bar{v}$ and an electron hole attraction $-W$ term:

$$K(1234) = \delta(12)\delta(34)\bar{v}(13) - \delta(13)\delta(24)W(12). \qquad (4.39)$$

In principle to obtain $^4\bar{P}$, one could invert the integral equation 4.38, contract the indices of the four point polarisability and get the macroscopic dielectric

function. However, this is not the procedure performed in the actual calculations for two reasons: the first one is that it would require the inversion of the four point integral equation for each frequency, which is quite time consuming; the second reason is that, as we will see, an effective two body Hamiltonian formalism carries much more information, especially regarding the excitonic eigenstates and eigenvalues. In the next section we will see how to get to this effective two body formalism.

4.4.1 Effective two particle Hamiltonian

In order to be able to pass to a two particle Hamiltonian formulation of the problem, we need to define the projection of a four-point quantity into the transition space, i.e. into a basis made of couples of single quasiparticle states, usually labeled by a band and a wave vector indices which we will summarise in the index n_i; two of this indices rapresent a *transition* which, when needed, we will call t. If A(1,2,3,4) is a generic four-point operator, its projection is given by

$$A_{tt'} = A_{(n_1,n_2)(n_3,n_4)} = \int dr_1 dr_2 dr_3 dr_4 A(1,2,3,4)\phi_{n_1}(r_1)\phi_{n_2}^*(r_2)\phi_{n_3}^*(r_3)\phi_{n_4}(r_4).$$

(4.40)

It is important to notice that in the transition space the unity operator is given by $\delta_{tt'} = \delta_{n_1 n_3}\delta_{n_2 n_4}$. And that the independent particle polarisability P^0 is diagonal:

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

(4.41)

From eq. 4.28, the macroscopic dielectric matrix in terms of the transition space matrix elements of $\bar{P}$ is given by:

$$\epsilon_M(\omega) = 1 - \lim_{q \to 0} v_0(q) \sum_{n_1,n_2,n_3,n_4} \bar{P}_{(n_1,n_2)(n_3,n_4)}(\omega)\langle n_1|e^{-iqr}|n_2\rangle\langle n_3|e^{iqr'}|n_4\rangle.$$

(4.42)

So, to get ϵ_M we need the matrix elements $\bar{P}_{(n_1,n_2)(n_3,n_4)}(\omega)$; in order to obtain them let's start by projecting the Bethe Salpeter equation 4.38 in the transition space:

$$\bar{P}_{(n_1,n_2)(n_3,n_4)}(\omega) = P^0_{(n_1,n_2)(n_3,n_4)}(\omega) + P^0_{(n_1,n_2)(n_5,n_6)}(\omega)K_{(n_5,n_6)(n_7,n_8)}(\omega)\bar{P}_{(n_7,n_8)(n_3,n_4)}.$$

(4.43)

Formally, we get $\bar{P}$ by inverting this last equation, so that we can write:

$$\bar{P}_{(n_1,n_2)(n_3,n_4)}(\omega) = (1 - P^0 K)^{-1}_{(n_1,n_2)(n_5,n_6)} P^0_{(n_5,n_6)(n_3,n_4)}.$$

(4.44)

Let's look more carefully at the operator $(1 - P^0 K)$, a generic matrix element can be written as:

$$(1 - P^0 K)_{(m_1,m_2)(m_3,m_4)} = \delta_{m_1 m_3} \delta_{m_2 m_4} - P^0_{(m_1,m_2)(m_5,m_6)} K_{(m_5,m_6)(m_3,m_4)} =$$

$$= \delta_{m_1 m_3} \delta_{m_2 m_4} - \frac{(f_{m_2} - f_{m_1}) \delta_{m_1 m_5} \delta_{m_2 m_6}}{\epsilon_{m_2}^{QP} - \epsilon_{m_1}^{QP} - \omega} K_{(m_5,m_6)(m_3,m_4)} =$$

$$= \delta_{m_1 m_3} \delta_{m_2 m_4} - \frac{f_{m_2} - f_{m_1}}{\epsilon_{m_2}^{QP} - \epsilon_{m_1}^{QP} - \omega} K_{(m_1,m_2)(m_3,m_4)} =$$

$$= \frac{1}{\epsilon_{m_2}^{QP} - \epsilon_{m_1}^{QP} - \omega} [(\epsilon_{m_2}^{QP} - \epsilon_{m_1}^{QP} - \omega) \delta_{m_1 m_3} \delta_{m_2 m_4} - (f_{m_2} - f_{m_1}) K_{(m_1,m_2)(m_3,m_4)}].$$

We define the excitonic Hamiltonian as:

$$H^{2p,exc}_{(m_1,m_2)(m_3,m_4)} = (\epsilon_{m_2}^{QP} - \epsilon_{m_1}^{QP}) \delta_{m_1 m_3} \delta_{m_2 m_4} - (f_{m_2} - f_{m_1}) K_{(m_1,m_2)(m_3,m_4)} \quad (4.45)$$

so that

$$(1 - P^0 K)_{(m_1,m_2)(m_3,m_4)} = \frac{1}{\epsilon_{m_2}^{QP} - \epsilon_{m_1}^{QP} - \omega} [H^{2p,exc}_{(m_1,m_2)(m_3,m_4)} - \omega \delta_{m_1 m_3} \delta_{m_2 m_4}].$$

$$(4.46)$$

Now, summarising the transition indix as:

$$(m_1 m_2) \equiv t \qquad (m_3, m_4) \equiv t'$$

and writing:

$$E_{m_2} - E_{m_1} - \omega = \Delta E_t; \qquad H^{2p,exc}_{(m_1,m_2)(m_3,m_4)} = H^{2p,exc}_{tt'}; \qquad (f_{m_2} - f_{m_1}) = f_t$$

we can write

$$(1 - P^0 K)_{tt'} = \frac{1}{\Delta E_t} [H^{2p,exc}_{tt'} - \omega \delta_{tt'}]$$

and

$$(P^0)_{tt'} = \frac{f_t}{\Delta E_t} \delta_{tt'}.$$

From eq. 4.44, we get:[9]

$$\bar{P}_{tt'} = (1 - P^0 K)^{-1}_{tt''} P^0_{t''t'} = [H^{2p,exc} - I\omega]^{-1}_{tt''} \Delta E_{t''} \frac{f''_t}{\Delta E''_t} \delta_{t''t'} = [H^{2p,exc} - I\omega]^{-1}_{tt'} f_{t'}.$$

$$(4.47)$$

---

[9]Here we are using the general property that, given a matrix B, defined by $B_{ij} = a_i A_{ij}$, its inverse is given by $(B)^{-1}_{ij} = (A)^{-1}_{ij} \frac{1}{a_j}$

In principle, we still have to invert a four point operator for each frequency, however we can use the spectral representation of the inverse of an operator getting:

$$[H^{2p,exc} - I\omega]^{-1}_{tt'} = [H^{2p,exc} - I\omega]^{-1}_{(n_1,n_2)(n_3,n_4)} = \sum_{\lambda\lambda'} \frac{A^{n_1 n_2}_\lambda S^{-1}_{\lambda\lambda'} A^{*n_3 n_4}_{\lambda'}}{E^{exc}_\lambda - \omega} \quad (4.48)$$

where $\{A^{n_1 n_2}_\lambda\}$ and $\{E_\lambda\}$ are solutions of the eigenvalue problem:

$$H^{2p,exc}_{(n_1,n_2)(n_3,n_4)} A^{n_3 n_4}_\lambda = E_\lambda A^{n_1 n_2}_\lambda \quad (4.49)$$

and S is the overlap matrix given by:

$$S_{\lambda\lambda'} = \sum_{n_1 n_2} A^{*n_1 n_2}_\lambda A^{n_1 n_2}_{\lambda'}. \quad (4.50)$$

If the excitonic Hamiltonian were hermitian, then $\{E_\lambda\}$ would be real and the overlap matrix would be a delta function, $S_{\lambda\lambda'} = \delta_{\lambda\lambda'}$; however this is not always the case.
Inserting eq. 4.48 into eq. 4.47, we get for $\bar{P}$:

$$\bar{P}_{(n_1,n_2)(n_3,n_4)} = \sum_{\lambda\lambda'} \frac{A^{n_1 n_2}_\lambda S^{-1}_{\lambda\lambda'} A^{*(n_3 n_4)}_{\lambda'}}{E^{exc}_\lambda - \omega} (f_{n_4} - f_{n_3}). \quad (4.51)$$

The presence of the factor $(f_{n_4} - f_{n_3})$ implies that only the matrix elements of the excitonic Hamiltonian $H^{2p,exc}_{(n_1,n_2)(n_3,n_4)}$ in which the state labeled by n_3 and the state labeled by n_4 have different occupation contribute. Within this subspace the two particle excitonic Hamiltonian will have the form:[10]

$$H^{2p,exc}_{(n_1,n_2)(n_3,n_4)} = \begin{pmatrix} H^{2p,reso}_{(n_1=v,n_2=c)(n_3=v',n_4=c')} & \bar{K}_{(n_1=v,n_2=c)(n_3=c',n_4=v'} \\ -\bar{K}_{(n_1=c,n_2=v)(n_3=v',n_4=c')} & -H^{2p,reso}_{(n_1=c,n_2=v)(n_3=c',n_4=v')} \end{pmatrix}. \quad (4.52)$$

Where $H^{2p,reso}$, the resonant part of the excitonic hamiltonian is given by:

$$H^{2p,reso}_{(vck)(v'c'k')} = (\epsilon^{QP}_{ck} - \epsilon^{QP}_{vk'})\delta_{vv'}\delta_{cc'}\delta_{kk'} + 2\bar{v}^{v'c'k'}_{vck} - W^{v'c'k'}_{vck}. \quad (4.53)$$

Here $(\epsilon^{QP}_{ck}, \epsilon^{QP}_{vk'})$ are the quasiparticle energies, calculated within the GW approximation, of the states (ck) and (vk'). As already mentioned, the excitonic

---

[10]Since in this subspace the single particle states belonging to each couple $n_i n_j$ have different occupation we will refer to them as v and c, namely valence and conduction bands.

Hamiltonian is not necessary hermitian, however, if we take into account only its resonant part, it is hermitian, and the macroscopic dielectric function is:

$$\epsilon_M(\omega) = 1 - \lim_{q \to 0} v(q) \sum_\lambda \frac{\left| \sum_{(n_1 n_2)} \langle n_1 | \exp^{-iq\dot{r}} | n_2 \rangle A_\lambda^{(n_1 n_2)} \right|^2}{E_\lambda^{exc} - \omega - i\eta} \qquad (4.54)$$

where each label n_i represents a couple of band and wave vector indices. Comparing this formula with the RPA dielectric function 4.32, we can see that in this case the dielectric function has poles at the eigenvalues of the excitonic Hamiltonian which are not anymore differences of occupied and unoccupied quasiparticle states energies. This fact moves the position of the structures in the spectra usually to lower energies. As we can see from eq. 4.49, each eigenstate of the excitonic Hamiltonian 4.49 is made by a sum of independent quasiparticle transitions; each transition contributes to it with a weight equal to $A_\lambda^{n_1 n_2}$. This mixing of independent transitions is present in the numerator of the dielectric function 4.54; and its effect is to strongly modify the lineshape of the spectra.

From eq.(4.54), we can see that to obtain the macroscopic dielectric function we need to diagonalise the excitonic Hamiltonian and to know the quasiparticle wavefunctions (in order to compute the dipole matrix elements present in the numerator). As in most *ab initio* calculation, we assume that Kohn-Sham wave functions are a very good approximation to the quasiparticle's ones. The computation of the matrix elements of the excitonic Hamiltonian, which is also a very time consuming part of the calculation, requires the knowledge of the quasiparticle energies and of the screened Coulomb interaction and also relies on the assumption that the quasiparticle wave functions are very close to the Kohn-Sham ones. At this level we limit the calculation to a static RPA screening, and the quasiparticle energies are computed within the GW approximation of the self energy.

To summarise, in order to compute the excitonic spectra we need a well converged geometrical structure calculation, Kohn-Sham wave functions, quasiparticle energies and the statically screened Coulomb interaction present in the excitonic Hamiltonian.

4.5 Example of excitonic calculation: Diamond bulk

4.5.1 Actual implementation

All the excitonic calculations presented in this thesis are performed using the *EXC* code [109], which is a plane wave basis set code. As already mentioned, one of the main task of any excitonic calculation is building the excitonic Hamiltonian. Computed in the reciprocal space, considering only spin-singlet class of solution, a generic matrix element of the resonant part of the Hamiltonian will look like [102]:

$$H^{reso}_{(vck)(v'c'k')} = (\epsilon^{QP}_{ck} - \epsilon^{QP}_{vk})\delta_{vv'}\delta_{cc'}\delta_{kk'} + 2\bar{v}_{(vck)(v'c'k')} - W_{(vck)(v'c'k')} \qquad (4.55)$$

where

$$\bar{v}_{(vck)(v'c'k')} = \frac{4\pi}{\Omega}\sum_{G\neq 0}\frac{1}{|G|^2}\langle ck|e^{iGr}|vk\rangle\langle v'k'|e^{-iGr}|c'k'\rangle \qquad (4.56)$$

and

$$W_{(vck)(v'c'k')} = \frac{4\pi}{\Omega}\sum_{GG'}\frac{\epsilon^{-1}_{GG'}(q)}{|q+G|^2}\langle ck|e^{i(G+q)r}|c'k'\rangle\langle v'k'|e^{-i(G'+q)r'}|vk\rangle\delta_{(k-k'),q} \qquad (4.57)$$

v is the exchange term, responsible for the local fields effects and for the propagation of the electron-hole excitation[11], while W rapresents the screened Coulomb attraction between the electron and the hole. [12] Besides the fact that the wavefunctions used to build the excitonic Hamiltonian are not the true quasiparticle wavefunctions, and that the dielectric matrix is calculated within RPA and only its static part is retained, there are many more approximations that are, or might be, performed solving the Bethe Salpeter equation. Let's look more carefully at them:

- The excitonic Hamiltonian should have infinite dimensions: the number of k vectors and of conduction states should be infinite, and one should

---

[11]Looking at the formulation of the exchange term in real space, namely $\int dr\int dr'\phi^*_{ck}(r)\phi_{vk}(r)v(r-r')\phi_{c'k'}(r')\phi^*_{v'k'}(r')$, we can see that it is made of terms in which an electron-hole pair is annihilated in r' and an other electron hole pair is created in r, thus propagating in the medium. The exchange term is repulsive and it may be seen as the repulsion between holes and electrons belonging to different electron-hole pairs [110].

[12]Also in this case from the real space formulation of W we can read more easily its meaning. Calculated in real space W is $W = \int dr\int dr'\phi^*_{ck}(r)\phi_{c'k'}(r)W(r,r')\phi_{vk}(r')\phi^*_{v'k'}(r')$. We can see that W is made up of terms in which an electron present in r interacts, and scatters with a hole in r'.

also take into account all the valence bands. Of course this is not done and one uses a discrete sampling of the Brillouin zone and a finite number of conduction bands. Typically also the number of valence states is limited. Convergency should be reached with respect to all of this three quantities.

- Also the number of the G vectors present in 4.56 and 4.57 is in principle infinite, and the sums must be truncated at some point. Usually convergency is reached with much fewer G vectors than the one needed in ground state calculations.

- If only the resonant part of the excitonic Hamiltonian is used, then the positive energy transitions are completely decoupled from the negative energy ones. Typically, in optical spectra the effects of the coupling terms can be neglected, however they might become important when considering quantities based also on the real part of ϵ, like for example EEL spectra [58].

- The screened Coulomb interaction could be considered taking into account only the diagonal terms in G, G'. Typically the effects of this approximation are small in homogeneous systems but can become significant in strongly inhomogeneous ones like in the case of clusters and surfaces.

In most of the calculations that will be presented all of this approximations are used. When possible the effects of such approximation are analysed.

4.5.2   Absorption spectra of diamond

As an example of excitonic calculation we show the spectra of bulk diamond. Within LDA approximation for the v_{xc} we find a lattice parameter of $6.71a.u.$. The first Brillouin zone has been sampled with 2048 shifted k-points. To obtain this set we started with 60 special k-points in the irreducible Brillouin zone, and by applying all the lattice symmetry operations to this set, a second uniform grid of 2048 k-points was obtained within the first Brillouin zone. A small shift along a non high symmetry direction was then applied to obtain the final grid of crystallographically inequivalent k-points. The quasi particle corrections have been computed for the 2048 shifted k-points by extrapolating from the corrections for the non-shifted grid calculated using 349 plane waves and 100 empty bands. In the calculation of the macroscopic dielectric functions the contributions of the transitions coming from the two valence and three conduction bands around the gap are included. It is important to notice that in this case the spectra has not been obtained by direct diagonalisation of the

excitonic Hamiltonian, but using an iterative algorithm, namely the Haydock recursive algorithm [111], presented in more detail in appendix A to reduce the computer main memory required for the calculation. The In fig. 4.1 we see that the Bethe-Salpeter spectra is in very good agreement with experimental data [112]: in particular it corrects the blue shift of the GW-RPA spectra.
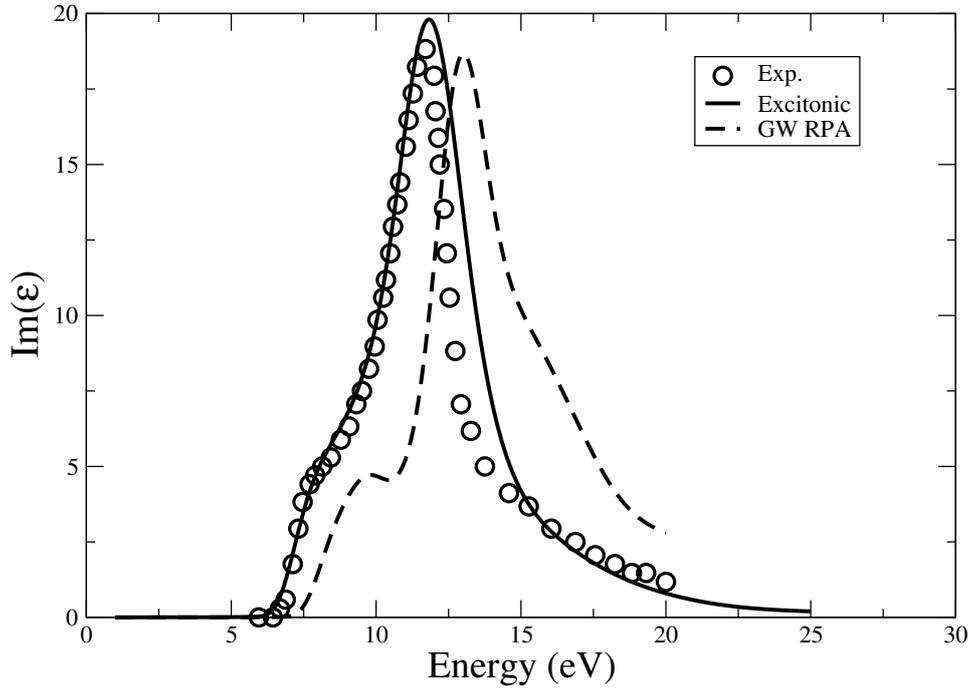
Figure 4.1: Diamond optical spectra. Dashed line: RPA with GW energies; solid line: Bethe-Salpeter with GW energies; circles: experiment (from [112]).

4.6 Reflectance Anisotropy Spectroscopy

Reflectance Anisotropy Spectroscopy is a typical optical probe for surfaces (for a recent review see [113]). It measures the difference in reflectance Δr between two normal incident light beam polarised linearly along two orthogonal directions in the surface plane (xy), normalised to the mean reflectance:

$$\frac{\delta r}{r} = \frac{2(r_x - r_y)}{r_x + r_y}. \tag{4.58}$$

This technique, although maintaining the properties of being highly non destructive, since light doesn't have a very strong interaction with matter, is still surface sensitive, because, in cubic materials and neglecting non linear effects, the isotropic contribution from the bulk cancels out.

In the next section we will briefly see how we can take into account the presence of a surface in describing the the propagation and reflection of light; since all our calculations are performed using the repeated slab geometry we will introduce the half slab polarisability, which is the quantity that connects calculations and experiments; we will finally show the results for the (111) surface at the RPA, GW and Bethe Salpeter level.

4.6.1   RAS calculations

In sec. 4.2.1, we have seen how, for cubic crystals, a simple, non local, isotropic relation holds between the macroscopic electric field E, and electric displacement vectors D, namely:

$$D_i(r,\omega) = \epsilon(\omega)E_i(r,\omega). \tag{4.59}$$

This relation can be obtained assuming that at the macroscopic scale, for long enough wavelengths, the system is homogeneous. In the simplest picture we can think that a surface is an abrupt interface between two regions in space were 4.59 still holds but with two different dielectric constants. In this scheme we can modelize the presence of a surface along the (xy) plane, with the crystal occupying the upper $(z > 0)$ volume, using a dielectric function of the form:

$$\epsilon(\omega,z) = \theta(-z) + \epsilon_b(\omega)\theta(z) \tag{4.60}$$

taking this form into account and solving the Maxwell equations for the propagation of light, one gets to the well known Fresnel formulae which for the normal incidence reflectivity give:

$$R = \left|\frac{\tilde{n}-1}{\tilde{n}+1}\right|^2. \tag{4.61}$$

It is clear that in this framework all the surface effects are neglected, no contribution of the modification of the response function of the material due to surface induced changes in the electronic structure are present; no anysotropy, and no non-locality are there. However anysotropy, non-locality and inhomogeneity are basic features in the response functions of a surface, and can not be neglected if one wishes to compare the results of calculations with surface sensitive techniques such as RAS.

In order to go beyond the Fresnel picture, a perturbative scheme is followed.

This scheme was implemented for a vacuum-jellium interface by Bagchi et al. [114], and generalised to a crystal surface by Del Sole [115]. The unperturbed system is rapresented by the fictitious abrupt Fresnel interface with dielectric constant given by eq. 4.60, and the reflection coefficients are calculated to the first order in the expansion parameter $d/\lambda = (\omega/c)d$, where d is the surface thickness. It is important to notice that the perturbation is actually not small, and the the electric field is not continuous at the interface; the perturbative expansion of ref. [114, 115] works because only continuous quantities (the electric field component parallel to the surface, the electric displacement component perpendicular to the surface) are taken into account in the derivation. Within the perturbative scheme the first order correction to the Fresnel reflectivity R_0 is given by:

$$\frac{\Delta R_i}{R_0} = \frac{R_i - R_0}{R_0} = \frac{4\pi\omega}{c} Im \frac{\langle \Delta\epsilon_{ii} \rangle}{\epsilon_b - 1}; \qquad i = x, y \qquad (4.62)$$

where

$$\langle \Delta\epsilon_{ii} \rangle = \int dz \int dz' [\epsilon_i^{-1} j(z, z') - \delta_{ij}\delta(z - z')\epsilon_b(z)]. \qquad (4.63)$$

In all our calculations, however we have a repeated slab geometry, in this case it is more convenient to express $\langle \Delta\epsilon_{ii} \rangle$, in the terms of the so called halfslab polarisability ϵ_{ii}^{hs} given by [116]:

$$\epsilon_{ii}^{hs}(\omega) = \frac{1}{4\pi d} \int_{-\infty}^{d} dz \int_{-\infty}^{+\infty} dz' [\epsilon_{ii}(\omega; z, z') - \delta(z - z')] \qquad (4.64)$$

where d is usually half of the slab thickness. In practice the imaginary part of the half slab polarisability, is given by the imaginary part of macroscopic dielectric function of the supercell, multiplied by d; it has in fact the dimension of a length; and, in parallel with the level of approximation for the calculation of the macroscopic dielectric function, it can be computed in a single particle framework, or including excitonic effects.

4.7 RAS on the (111) surface of diamond

In the previous two chapters we have seen how the most stable reconstruction for the (111) surface of diamond is the 2×1 Pandey Chain model. Within this model almost one dimensional chains of atoms are present on the surface along the y direction. We have seen how the presence of a buckled or dimerised geometry of the chains is strictly connected to the possibility of opening a gap at the DFT level. However such buckling or dimerisation seem very unlikely

and the most stable geometry, featuring undimerized and unbuckled chain, is associated to a semimetallic DFT band structure. Finally, we have seen how the inclusion of quasiparticle corrections within an iterative GW scheme leads to a semimetallic surface with a minimum direct gap of ~ 1 eV between surface states. Unfortunately, as already mentioned, experimentally not much is known about the empty surface states of C(111)2 × 1. A direct information of the magnitude of the gap could be given by inverse photoemission spectra which, however, are not available. Optical spectroscopy, like RAS, could give insights on the transitions between occupied and unoccupied surface states, and, as a matter of fact, RAS experiments are currently being set up [81]; it is thus extremely interesting to compute the optical spectra of this surface.

The optical properties of this surface at the RPA level have been calculated in 1996 by Noguez and Ulloa [117] using a tight binding approach. As expected, they found a large anysotropy in the surface dielectric response as a consequence of the presence of the one dimensional Pandey chains in the surface reconstruction. However, no fully ab-initio calculation has ever been performed, nor calculations that included excitonic effects, which, on other diamond surfaces, [118] are very important.

In this section we will present the results regarding the optical spectra of C(111)2 × 1, at both the RPA and the Bethe-Salpeter level. We will see how, due to the extremely large dispersion of the surface bands moving away from the JK line (the electronic gap opens of about 4 eV, going from J to Γ, see fig. 3.9 and tab. 3.7), a very dense k-point grid is needed to converge the low energy side of the spectra ($0 \div 1.5$ eV), while the high energy side of the spectra ($1.5 \div 6$ eV) is converged at a much broader k-point grid. We will present the results of the excitonic calculation for a 2×16 k-points grid in the irreducible Brillouin zone. For this k-point mesh the high energy side of the spectra is well converged up to 6 eV and the low energy side is converged up to 0.3 eV (the RPA peak position is 0.3 eV away from its converged position). At last, we will see how an interpolation scheme for the screened Coulomb interaction must be implemented in order to use a much denser k-point mesh. Test calculations on bulk systems are presented, and the scheme is subsequently applied to the calculation of the optical response of the surface for a 4×64 k-point mesh.

4.7.1 RPA

The RPA RAS spectra obtained for different grid of k points is shown in fig. 4.2. The number of empty bands used in each calculation is shown in the legend of the graph. First of all, let's compare this result with the tight binding result of Noguez et al. [117]: also in our case there is a strong negative peak at low energies, and some structures appear around 4 eV. The low energy peak comes

97

(a)

(b)
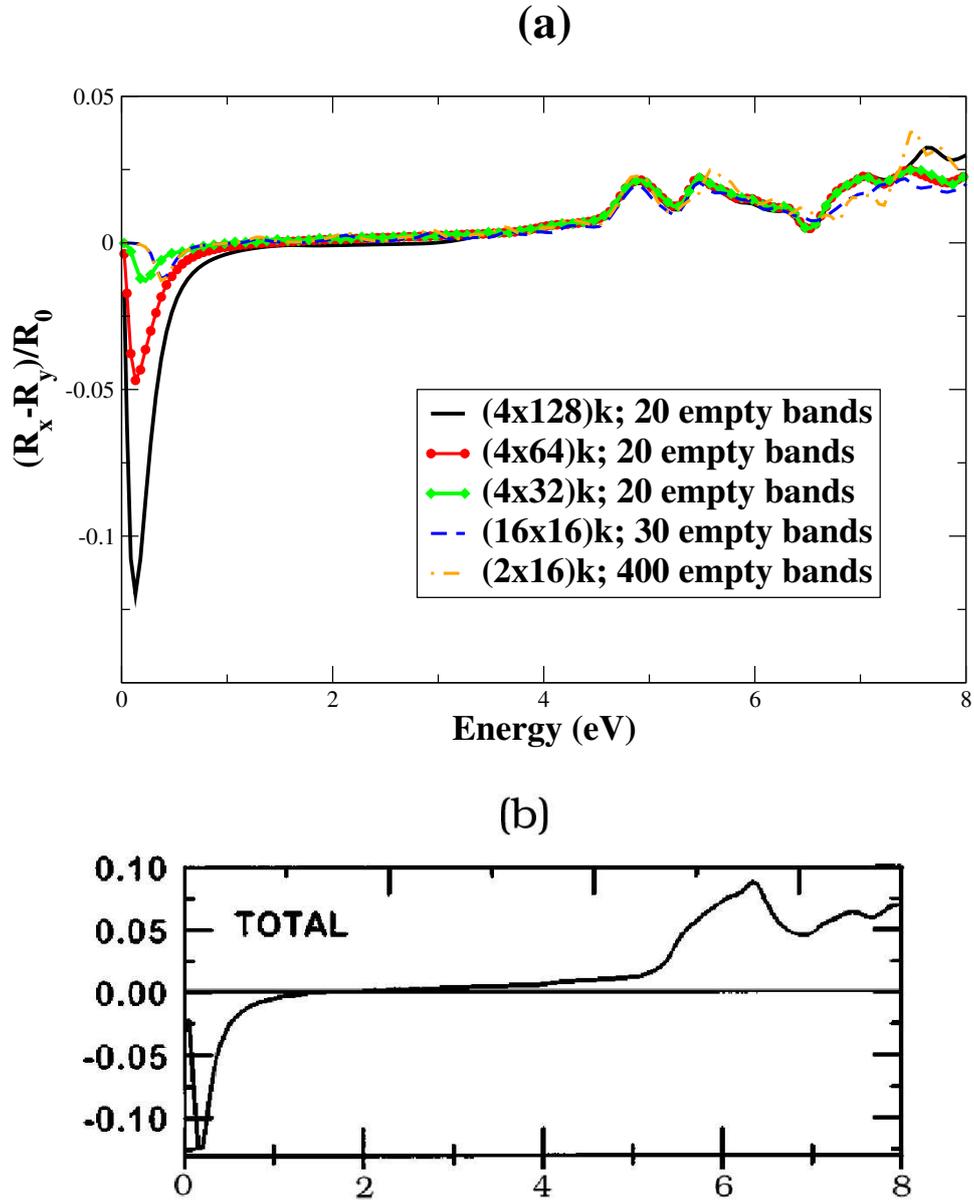
Figure 4.2: RPA RAS computed for different k-point grids (a), compared with the tight binding results (b) of Noguez et al. [117]. Qualitatively the DFT and tight binding spectra are similar. It is important to notice that the low energy peak position is converged using at least a 4×64 k-point mesh, while the high energy side of the spectra is converged already using a 2×16 grid. (Surface chains are along the y direction)

from transitions between surface states for k points close to the JK line. The sign of the low energy peak is clearly determined by the fact that the optical response is much stronger for light polarised along the chain direction (the chains are parallel to the y direction in our case, see fig. 2.1), than for light polarised perpendicular to it. Moreover, if we think that exactly at the JK line the surface state for each k point have dangling bond character with different parity with respect to the mirror symmetry $y \rightarrow -y$; then it is clear that right on the JK line the response for light polarised along x will be exactly zero, since p_x, has vanishing matrix elements with respect to the two surface states, while it will be maximum for light polarised along y

From fig. 4.2 we can also see that in the region of energy ranging from 1.5 to 6 eV, a 2×16 k point mesh and the inclusion of ten empty bands give already a quite good description of the spectra. (When we will introduce excitonic effects, however, we must remember that transitions of different energies might mix so that convergency on the number of bands used must be checked again). The convergency of the position of the low energy peak is very delicate. In fact its position is very sensitive to the sampling of the Brillouin zone, especially to how much close to the JK line the k-point set can go. In fig. 4.3 we show the dispersion of the DFT gap between surface states for a fixed values of K_x, varying K_y: we can see that, for every value of K_x, the gap varies very quickly so that going from $K_y = 0.25$ to $K_y = 0.5$ the gap value decreases from 4 to ~ 0 eV. As a consequence, it is clear that, in order to catch transitions of the minimum energy, one should get really close to the JK line. For example the red crosses in fig. 4.3 show the sampling along the K_y direction that would be performed by a 2×16 k-point mesh: the first transition found is already $0.3 eV$! Looking at the results for the 2×16 k-point mesh we must keep in mind that the accuracy for the low energy peak is at most $0.2 - 0.3 eV$.

4.7.2 Excitonic effects (1)

In this section we show the results of the excitonic calculations using a 2×16 k-point grid. First of all, the convergency with respect to the parameters of the calculations was checked. The convergency test followed roughly this scheme:

- The number of iterations of the Haydock algorithm (see app. A) needed for convergency of the spectra in the $0 \div 6$ eV energy range is obtained.

- The number of G vectors needed in the description of the wave function and of the exchange term (eq. 4.56) is obtained. The calculations were performed using only one G vector for the screened interaction term (eq. 4.57), and considering 10 valence and 10 conduction bands (the number of bands needed to converge the RPA spectra up to 6 eV).

(a)
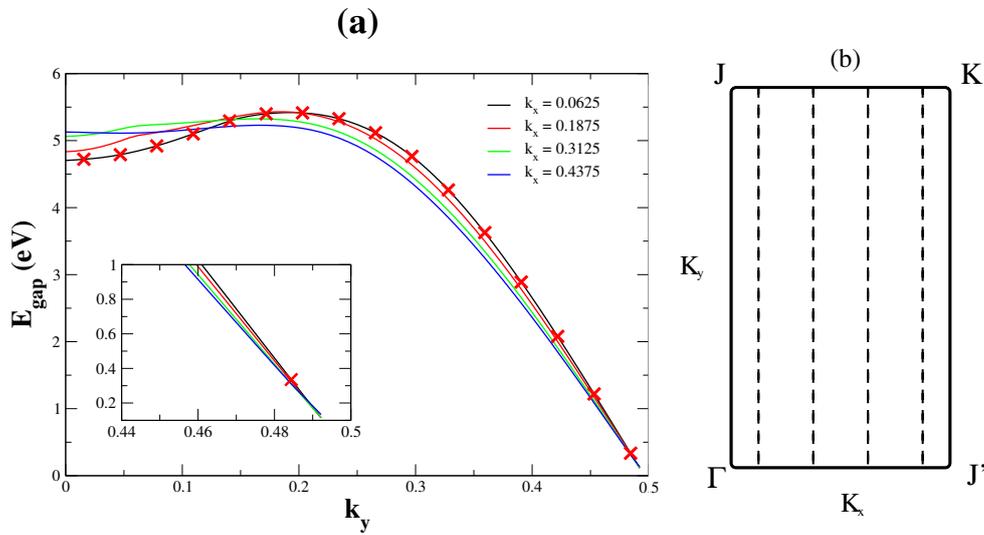
Figure 4.3: (a) Dispersion of the gap between surface states varying K_y at four fixed values of K_x, namely $K_x = (0.0625, 0.1875, 0.3125, 0.4375)$. In all the four cases the gap exhibits a very strong dispersion. The red crosses correspond to the sampling along the K_y direction in the case of a 2×16 mesh, while the continous line to the case of a 4×64 mesh. (All energies are expressed in eV and the k-vectors components are expressed in relative units).(b) The four lines in the irreducible Brillouin zone along which the gap has been calculated. (Surface chains are along the y direction)

- The number of G vectors in the screened Coulomb interaction is also brought to convergency.

- Convergency on the number of bands is achieved
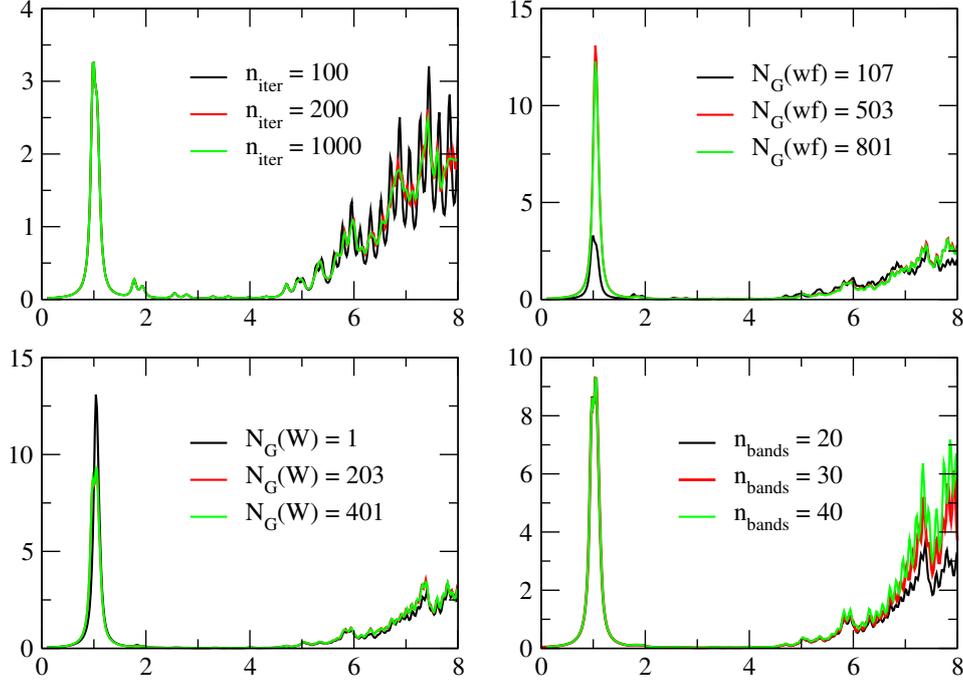
Figure 4.4: Results of the convergency test for the supercell macroscopic dielectric function, in the case of light polarised along the y directions. All calculations have been performed including excitonic effects and using a 2×16 sampling of the IBZ. (Surface chains are along the y direction)

In fig. 4.4 the results of the convergency test are presented for the supercell dielectric function in the case of light polarised along the y direction. We can see that the spectra are well converged, in the energy range of our interest, using 200 iterations of the Haydock algorithm, 503 G-vectors in the description of the wavefunction and of the exchange part, 203 G-vectors for the screened Coulomb interaction term, and taking into account 15 valence and conduction bands.

We also checked the impact on the spectra of the quality of the RPA dielectric matrix used in the computation of the screened Coulomb interaction term. In fact, remembering the expression of the RPA irreducible polarisability (eq. 3.26), and knowing that the inversion of the dielectric matrix must be performed, the quality of the screening might depend on the number of empty

bands considered in the computation of eq. 3.26 and on the dimension of the matrix to be inverted. In fig. 4.5, two spectra are presented, using in one case an RPA screening obtained considering 600 empty bands and inverting a 1995×1995 matrix, in the other case 200 bands and a 1003×1003 matrix were used. As we can see the two spectra are almost identical.
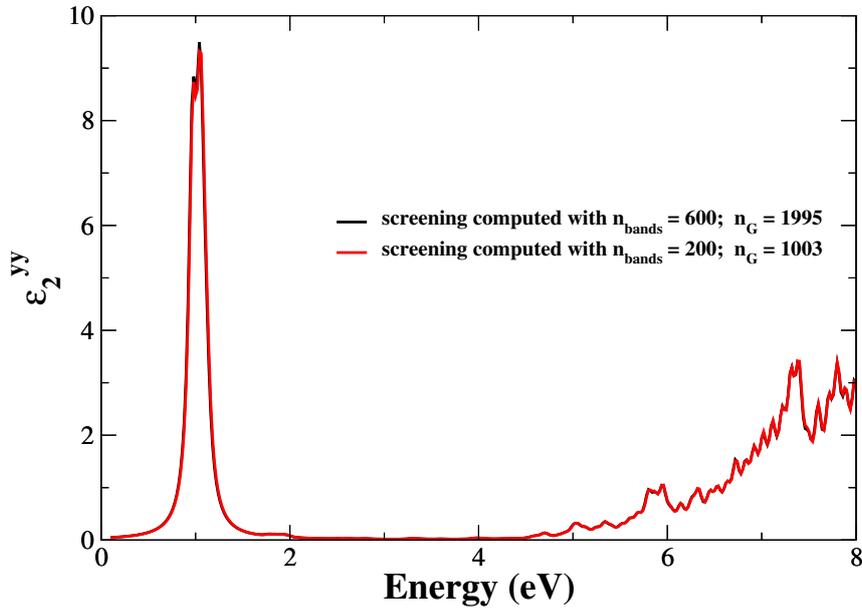
Figure 4.5: Excitonic spectra obtained using two different RPA screening, using 600 bands and a 1995×1995 matrix inversion, or 200 bands and a 1003×1003 matrix inversion. Both calculations were performed using a 2×16 sampling of the IBZ. (Surface chains are along the y direction)

The effect of the computation of the excitonic Hamiltonian taking into account also the off-diagonal terms of the screened Coulomb interaction, and the coupling to the negative frequency transitions was analysed. The results for the low energy peak are shown in fig. 4.6. We can see that in the case of the (111) surface of diamond their effect is negligible. This result is not surprising, because the off-diagonal terms of the screened Coulomb interaction are important only in the case of very inhomogeneous systems like clusters, while, in this case, we are perturbing the system with an electric field parallel to the surface, i.e. we are probing it along the direction in which the system is more homogenous; concerning the coupling term, it is typically more important for the study of EEL spectra.
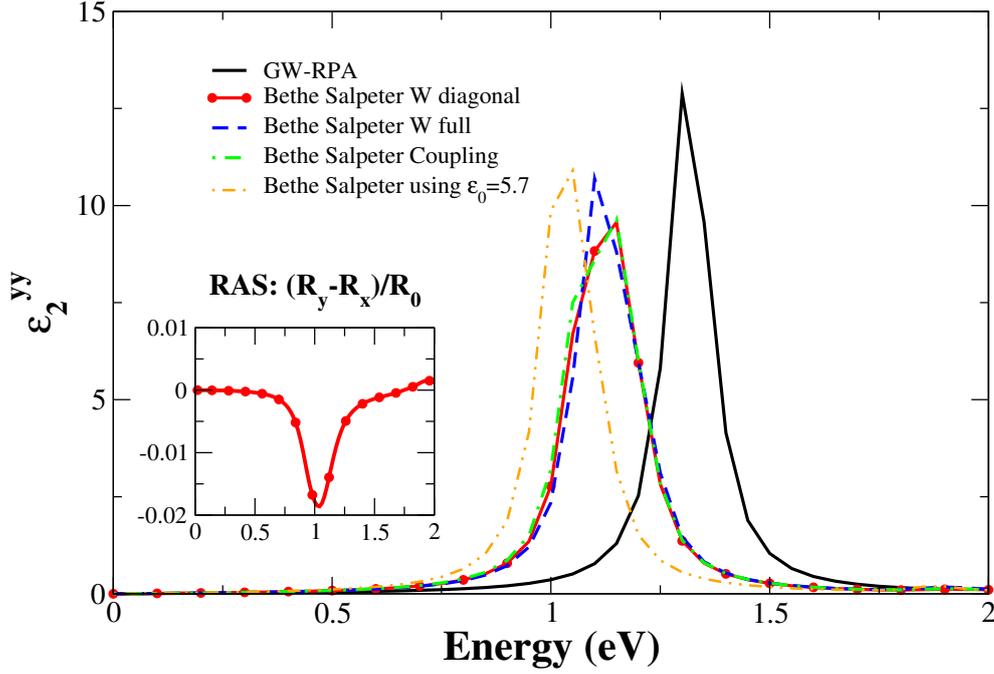
Figure 4.6: The low energy peak calculated including the off diagonal terms of the screened Coulomb interaction, the coupling between positive and negative transition term, and using a constant screening $1/\epsilon_0$. The effects of both the coupling and off diagonal terms is negligible. In the inset the RAS spectra corresponding to the Bethe Salpeter W diagonal calculation is shown. This calculations, being referred only to the low energy peak, were performed using 2 conductions and two valence bands only, the remaining parameters are kept fixed to the previous values. (Surface chains are along the y direction, all calculations are carried out using a 2×16 k point mesh in the IBZ).

In order to compute the final spectra, a new RPA screening was computed. This screening calculation used the quasiparticle updated energies instead of the DFT ones. In fact we have to remember that the final band structure that we have obtained, was the result of an iterative procedure in which the self energy was $\Sigma_{n+1} = G_n W_n$. Following ref. [58], in this case the functional derivative of the self energy with respect to the Green's function present in the

kernel of the Bethe Salpeter equation 4.39, should be $\frac{\partial \sigma_{n+1}}{\partial G_n} = W_n$. This RPA updated screening was calculated using 200 empty bands and a 1003×1003 matrix inversion. The results for the low energy peak are shown in fig. 4.7. As we can see, the effects of the updated screening enhances the excitonic effects, this is not surprising since the RPA-DFT screening is the screening of a semimetal and thus should be quite stronger than the updated screening. The excitonic binding energy is, in this case, 0.4 eV. The binding energy of the exciton is in this case quite smaller than in the (100) surface [118] (in that case the binding energy was 0.9 eV) the explanation for this might be found in the fact that the dispersion of the band gap is in this case very large yielding a much smaller effective mass and thus a smaller binding energy; we have also to take into account, however that the low energy peak is still not well converged at this level. This problem will be analysed in more detail in the next section. The higher energy side of the spectra is already well converged with the 2×16 k point mesh, as we can see from fig. 4.2. The convergency test on the number of bands to be used, look at fig. 4.4, show that including 15 valence and conductions band, already yields well converged spectra 8 eV. In fig. 4.8, we can see that also in the for the higher energy structures the excitonic effects shift and modify the spectra; in particular the small structures at ~ 2, ~ 3, and ~ 4 eV are shifted at lower energies, while in the structures around 6 eV most of the spectral weight is moved at lower energy in a single peak at ~ 5.2 eV.

4.7.3 Excitonic effects (2)

We can see from fig. 4.2 that the low energy peak is converged using the 4×64 k-point mesh, at least concerning its position. However such a k-point grid requires the knowledge of the inverse dielectric constant for 325 inequivalent q-vectors of the irreducible Brillouin zone, which would require months of computer time. Assuming that the dielectric constant is not so strongly varying, we can think to get its value for a broader mesh of q-points and interpolate it for the q point that are actually needed. In what comes next we will see how this interpolation scheme is actually implemented and tested in the case of bulk silicon and diamond. We will see that the results on bulk are encouraging: the scheme is hence subsequently applied to the case of diamond (111).

Interpolation scheme for the screening

Within the interpolation scheme for the dielectric function, we compute the RPA screening for a certain grid of q-vectors; then this screening is used in a Bethe Salpeter calculation where a denser mesh of q points is required. The
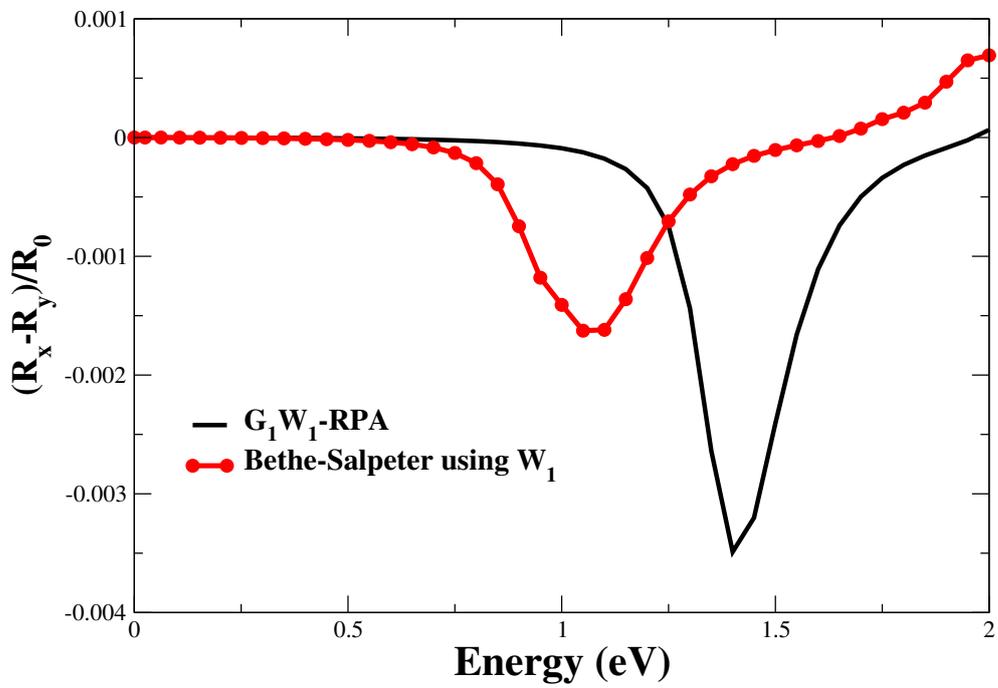
Figure 4.7: RAS spectra for the low energy part of the spectra including excitonic effects computed using an updated screening. The coupling and the off-diagonal terms of the excitonic hamiltonian were not taken into account; the IBZ is sampled using a 2×16 k point mesh. (Surface chains are along the y direction).
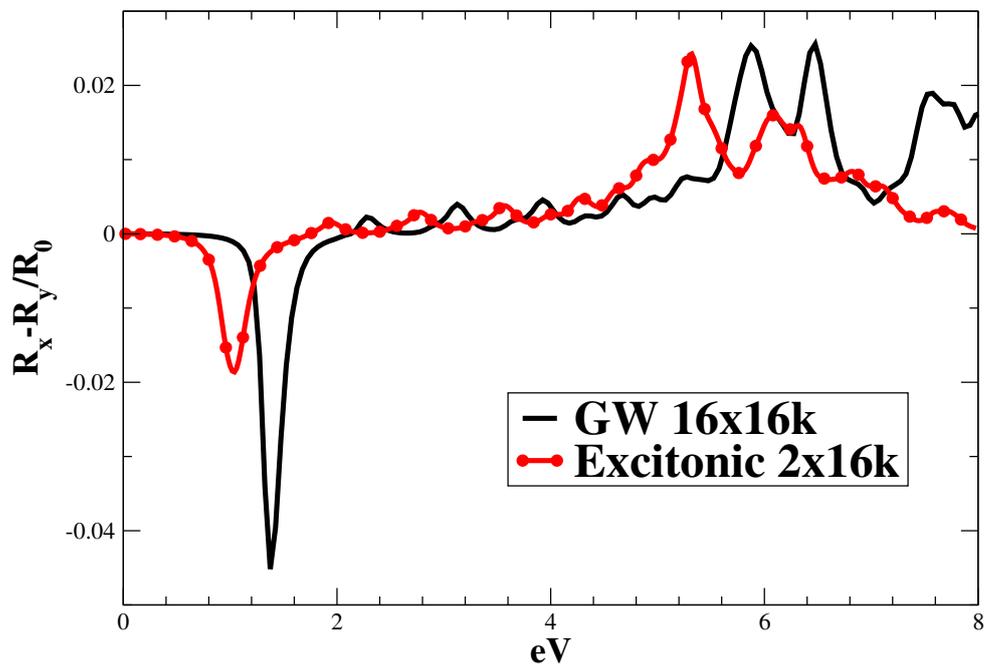
Figure 4.8: RAS spectra including excitonic effects. The calculation was performed with a 2×16 k-point sampling of the IBZ, and including 15 valence and conduction bands to get convergency up to 8 eV. The screening in this case was not updated, since the high energy structures are less sensitive to the variation of the quasiparticle energies in the screening due to the GW corrections. (Surface chains are along the y direction).

value of the dielectric matrix for each q point needed is replaced by the one of the closest q-point that has been actually calculated; care is put in order to choose q-vectors grids that are lying in the same irreducible Brillouin zone. In fig. 4.9 we see the absorption spectra of diamond and silicon bulk computed with this interpolation scheme. The diamond calculation is performed using a grid of 2048 shifted k-points in the first Brillouin zone. The calculation requires the knowledge of the screening for 85q points in the irreducible Brillouin zone, the curve in red is the result obtained using the 85 q-points grid, while the curve in black is the spectra obtained using a dielectric matrix computed for a 19 q-points grid (the q-vectors grid associated to the 10 Chadi Cohen k-point mesh); the two curves are very similar. Also for the case of silicon we see that in the case of the 256 shifted k-points mesh, using the proper dielectric matrix calculated for a 19 q-vector grid or the dielectric matrix computed on a 6 q vectors grid is quite the same. In the case of silicon it is also shown the result of the absorption spectra for the 32 shifted k-point mesh from which the 6 q-vectors mesh is obtained; we see that in this case the spectra differ strongly from the 256 k-points grid. In the case of bulk silicon and diamond, it seems like if the inverse microscopic dielectric matrix is smooth enough in order to allow the application of this interpolation scheme; this might not be so obvious at a surface and care should be put to check this assumption; the results on these tests systems are however encouraging.

Preliminary results on a dense k-points grid

In what follows, we have applied the interpolated dielectric matrix scheme for the case of the (111) surface of diamond using a 4×64 k-point mesh. The RPA static dielectric matrix has been computed for a 25 q vectors mesh obtained from a 4×4 k-point grid, using 200 empty bands, the inversion of a 1003×1003 G-vectors matrix, and updated energies. In fig. 4.10 the two grids of q points are shown. We can see that associated to each q vector of the broader grid there is an area of the irreducible Brillouin zone, the q vectors of the denser grid that are found within this area are associated to it. The Bethe-Salpeter calculation was then carried out with the convergency parameter of the previous calculation, i.e. using 503 G-vectors for the wavefunction and the exchange term, and 401 G-vectors for the screened Coulomb interaction term. Since now we are interested only in the low energy peak, only two valence and two conduction bands are taken into account. The reduction of the number of bands used has been carried out after checking that the single particle states contributing to the low energy peak in the 2×16 calculation are only the surface state within the gap.

The resulting RAS spectra is shown in fig. 4.11. We can see that now the
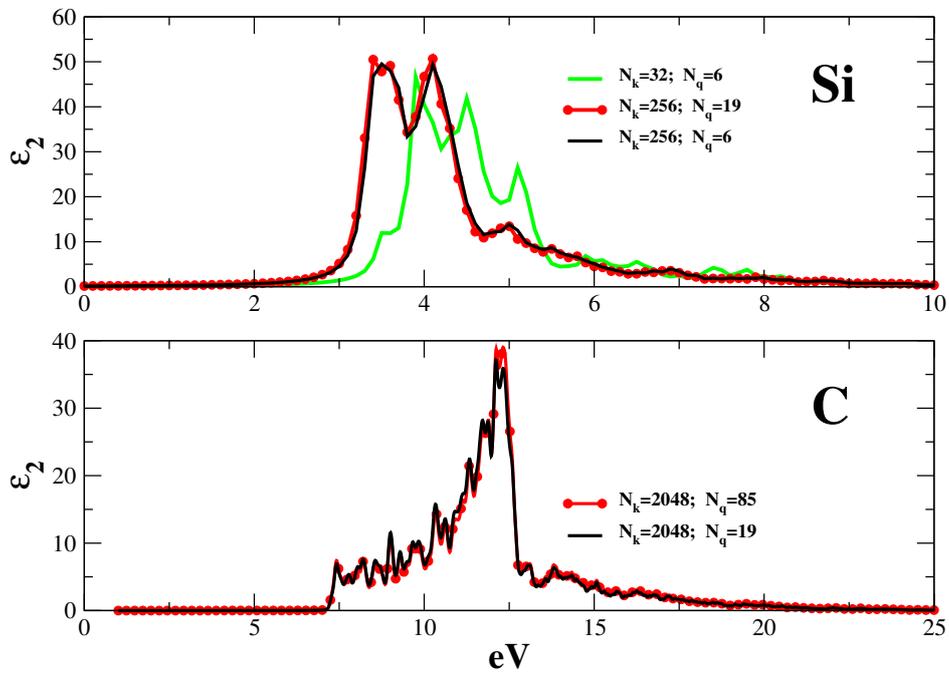
Figure 4.9: Absorption spectra for diamond and silicon bulk. The spectra obtained with the interpolated dielectric matrix (full black line) are compared to the full calculations (red circles).
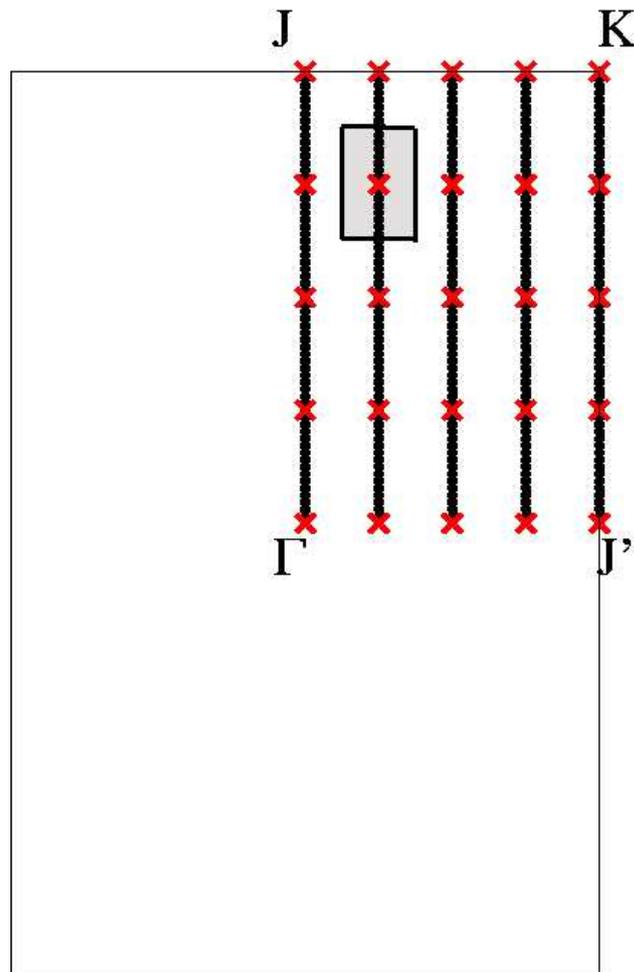
Figure 4.10: The mesh of q-vectors generated by the 4×4 and by the 4×64 k-points grids are compared. The 4×4 k-points grids produces a 25 q-point mesh (red crosses), while the 4×64 k-points grid produces a 325 q-point (black circles). To each of the 25 q points is associated an area, all the q points of the 325 set that are found within this area are associated to it.

binding energy, 0.5 eV, is quite larger than in the case of the 2×16 mesh, even if it is still less than the one of the (100) surface. The most important thing, however, is that now that the lineshape of the spectra is changed. In the 2×16 k point spectra, concerning the low energy peak, the effects of the introduction of the electron-hole interaction was an almost rigid shift of 0.2 eV toward lower energies. In this case, instead, the spectral weight is redistributed in a different way, the main weight is at energies close to the GW-RPA peak, but a series of peaks appear at lower and higher energies.
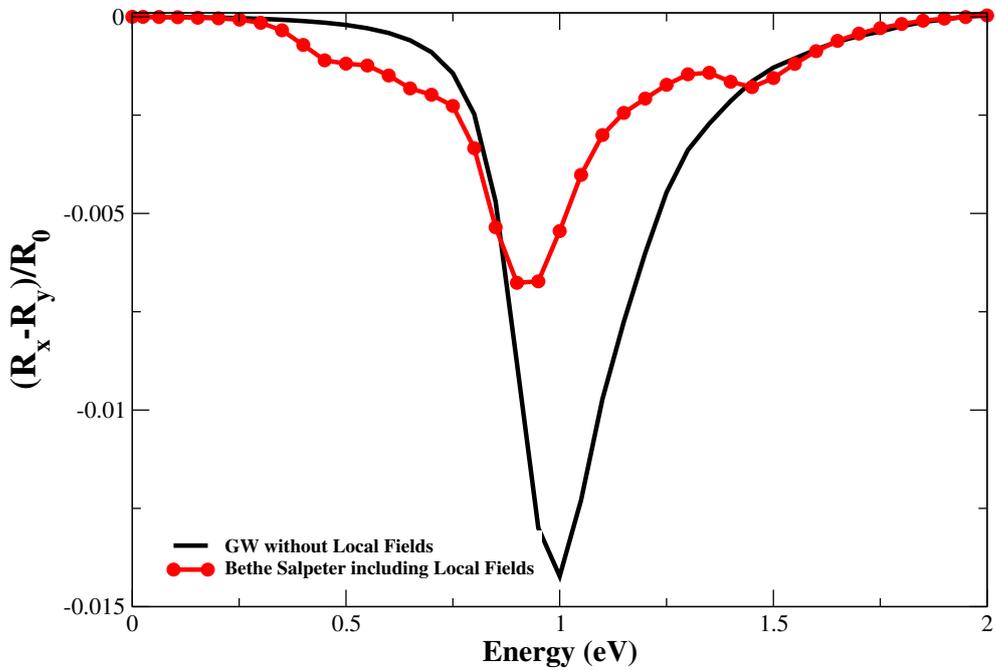
Figure 4.11: RAS spectra of the (111) surface of diamond using a 4×64 k-point grid. The spectral weight is now redistributed giving rise to a series of peaks. The binding energy of the exciton is 0.5 eV. (Surface chains are along the y direction)

4.8   Conclusions

In this chapter we have first reviewed the theory behind the calculation of optical spectra and then we have seen its application to the RAS spectra of

diamond (111). There is the need of going beyond the single particle picture to be able to describe the optical properties of solids, and, in fact, we have seen how the solution of the Bethe Salpeter equation can be cast into a two body Hamiltonian problem, and how the macroscopic dielectric matrix can be expressed in terms of its eigenvalues and eigenvectors.

In the case of the (111) surface of diamond we see that the excitonic effects are strong in both the low and high energy part of the spectra. The binding energy of the exciton ranges from 0.2 to 0.5 eV, and the spectral weight is redistributed. Care should be put in analysing this results, especially for the low energy size of the spectra, since convergency with respect to the k-point grid is in this case very delicate and an interpolation scheme for the dielectric matrix has been used; however we can conclude that the evidence of the presence of an electronic gap could be given by optical spectra, and that a strong anystropy signal at low energy would be the signature of the transition within surface states close to the gap.

Conclusions

Despite the fact that the (111) surface of diamond represents a prototype, and quite simple system, its electronic band structure and underlying geometry were up to now not fully determined. As a matter of fact no *ab-initio* calculation was able to reproduce its insulating character. The deficiency of the theory in this case could not be easily addressed to a definite reason like a wrong equilibrium geometry nor to the unfeasibility of the usual G_0W_0 scheme for this specific band structure.

Confirming previous ground state results, we found that this surface reconstructs with π-bonded chains without any significant dimerisation, buckling, or antiferromagnetic ordering. The resulting band structure is semimetallic in contrast with experimental data that predict a gap of at least 0.5 eV (see fig. 2.2).

Experiments rule out the dimerised geometry for symmetry reasons [22]; and, regarding buckling, there is no definite indication: medium energy ion scattering data [23], X-ray diffraction structure analysis [27], and infrared-visible sum frequency generation measurements [20] find best agreement for an atomic arrangement featuring tilted chains, but can not completely rule out an unbuckled undimerized geometry; at the same time, core level binding-energy measurements [28], and a recent LEED study [26], see no evidence for buckling. We have seen, however, that even in the presence of a reasonable buckling no gap opens at the DFT level. The mechanism of the formation of the electronic must hence be found elsewhere.

DFT (within LDA or GGA schemes) is known to underestimate the gap of semiconductors. DFT itself, in fact, is an exact theory when applied to ground state properties but, in principle, there is no formal justification to the use of this theory for excited state properties. The absence of the gap between surface states at the C(111)(2×1) surface may be due to this reason. A formalism in which the single particle excitations energy can be more straightforwardly

accessed is given by the Green's function theory. Within this approach the single particle excitation energies are computed calculating the quasiparticle correction to the DFT band structure using a first order perturbation scheme, and the self energy operator is approximated as the product of the independent particle Green's function G_0 and the RPA screened Coulomb interaction W_0. As we have seen in chapter 3, the G_0W_0 corrections to the band structure do not change the situation; this is due to the coincidence of two facts specific of this surface:

1. At the DFT level, along the JK direction of the Brillouin zone (see fig. 2.2) the surface bands are crossing the Fermi level. Both surface states at J (K) are so occupied (unoccupied).

2. Along the JK direction the higher and lower surface bands' wavefunction are equivalent, exhibiting in both case a dangling bond character.

The combination of these two facts implies that the energy of both states at J (K) are corrected downward (upward) by approximately the same amount. As a result the dispersion along the JK line is enhanced and the surface stays semimetallic.

After checking that a non perturbative approach including the off diagonal terms of the self-energy operator could not improve the results, we have carried out GW calculations within an iterative scheme. The starting configuration corresponded to the Kohn-Sham band structure (i.e. using the Kohn and Sham eigenvalues and eigenfunctions) using an artificial population in which the states with the correct symmetry were occupied. The GW calculation were then repeated, updating the quasiparticle energies while building the self-energy, till self consistency was reached. As a result a direct gap of about 1 eV opens between the surface states [13]. The insulating character of this surface thus is not a consequence of asymmetric geometrical changes, rather of many body effects.

In the absence of inverse photoemission experiments to directly compare with our band gap calculations we have performed also the calculations of the optical spectra [14] including the effect of the electron hole interaction. We have found that strong anisotropy signals at low energy are connected to the transitions between surface states, the excitonic binding energy ranging from

---

[13]We have also checked that this value was independent from the starting configuration, being it either an artificial *ad hoc* occupied configuration, or a configuration in which the conduction band was shifted rigidly by a small or big amount. The indirect gap is ~ 0.5 eV.

[14]RAS experiments are currently being carried out in Rome and Berlin [81].

0.3 to 0.5 eV. If this signal could be detected it would give important indications on the presence and magnitude of the electronic gap.

Appendix A

The haydoch recursive method for optical spectra

As we have seen in chapter 4, in order to obtain an optical spectra including the electron-hole attraction, one has to diagonalise the excitonic Hamiltonian. The macroscopic dielectric function is then written in terms of the excitonic eigenfunctions and eigenvalues. However, this procedure is not only very time consuming, but it requires the storage in the main memory of at least half of the excitonic Hamiltonian whose number of elements scales as the square of the number of transitions $N_t^2 \sim N_k^2 N_v^2 N_c^2$, being N_k, N_c, and N_v the number of k points in the first Brillouin zone, of conduction and valence bands respectively. This fact can be understand intuitively thinking that in order to diagonalise a matrix, all its elements must be known at one time.

In the case of large systems, as surfaces, the memory requirement for the diagonalisation of the excitonic Hamiltonian can become prohibitive. In this appendix we want to present a method that allows to obtain the macroscopic dielectric function within an iterative scheme in a continued-fraction representation. During this procedure the excitonic Hamiltonian is repetitively applied to an initial state. In other words the only operation that has to be performed is a matrix-vector product, which requires the knowledge of only a matrix line at the time (scaling now as the number of transitions). This procedure is a generalisation of the Lanczos method [119] [1] to the computation of any dynamical properties of a system that can be cast into a form given by:

$$I(\omega) = Im[\langle \Psi_0 | \hat{O}^\dagger \frac{1}{\omega - \hat{H} + i\eta} \hat{O} | \Psi_0 \rangle] = Im[\langle P | \frac{1}{\omega - \hat{H} + i\eta} | P \rangle]. \qquad (A.1)$$

---

[1]The Lanczos method allows to iteratively find a basis in which the Hamiltonian has a tridiagonal representation

Appendix A

This method has been introduced by R. Haydock [111], and in the same paper it has been illustratively applied to the computation of the local density of states of a system of a given Hamiltonian. Then it has been applied to the computation of the density-density correlation function of a model Hamiltonian by E.R. Gagliano and coworkers [120], and finally it has been used to compute the macroscopic dielectric function by L.X. Benedict and E.L. Shirley in ref. [121]. We have implemented this method in the excitonic code *EXC* [109] [2].

In this appendix we will first write the macroscopic dielectric function as eq. A.1, then we will see how it can be computed as a continued fraction within the Haydock recursive scheme.

In the Coulomb gauge and neglecting non linear terms, the part of the Hamiltonian concerning the interaction of the electromagnetic field with particles can be in general written as [122]:

$$H_{int} = -\frac{1}{c} \sum_i \mathbf{v}_i \cdot \mathbf{A} \tag{A.2}$$

where $\boldsymbol{v_i}$ are the velocity operator of each electron and $\boldsymbol{A}$ is the vector potential of the electromagnetic field. The electron are thus coupled to the electromagnetic field through their current, and it is very interesting to notice that in the presence of a local non perturbed Hamiltonian the usual $\boldsymbol{A} \cdot \boldsymbol{P}$ form for the interaction term is recovered.

If the interaction between light and matter is given by eq. A.2, for a monochromatic perturbation of frequency ω, and amplitude A_0 the Fermi golden rule gives a transition rate from the initial state $|i\rangle$ to the final state $|i\rangle$ of the form:

$$P_{i \to f} = 2\pi \frac{A_0^2}{c^2} |\langle f|\mathbf{v} \cdot \hat{e}|i\rangle|^2 \delta(E_f - E_i - \omega), \tag{A.3}$$

were $\hat{e}$ is the polarisation vector. Following the derivation of ref. [123] p. 35, the macroscopic dielectric function can be connected to the transition rates through the absorption coefficients giving:

$$\epsilon_2 = \frac{4\pi^2}{\omega^2} \sum_f |\langle f|\mathbf{v} \cdot \hat{e}|0\rangle|^2 \delta(E_f - E_i - \omega) \tag{A.4}$$

where $|0\rangle$ is the many body ground state. The velocity operator can be written in a more convenient way [124] as:

$$v_\alpha = \lim_{q \to 0} \frac{[H, e^{iqr_\alpha}]}{q}, \tag{A.5}$$

---

[2]This work has been performed in collaboration with F. Sottile during a one month stage in L. Reining's group at the *Ecole Polytechnique* in Palaiseau

and inserting this expression in eq. A.4 we get:

$$\epsilon_2 = 4\pi^2 \lim_{q \to 0} \frac{1}{q^2} \hat{e} \cdot \sum_f |\langle f|[H, e^{i q \mathbf{r}}]|0\rangle|^2 \frac{1}{E_f - E_0} \delta(E_f - E_0 - \omega) =$$

$$= 4\pi^2 \lim_{q \to 0} \sum_f \hat{e} \cdot \langle f|e^{i q \mathbf{r}}|0\rangle|^2 \delta(E_f - E_0 - \omega). \qquad (A.6)$$

Exploiting the properties of the principal part of an integral, namely

$$\frac{1}{x + i\epsilon} = P\left(\frac{1}{x}\right) - i\pi\delta(x)$$

and unfolding the expression of the square modulus, we can rewrite eq. A.6 as:

$$-4\pi \lim_{q \to 0} Im\langle 0|\hat{e} \cdot \frac{e^{-iqr}}{q} \frac{1}{\omega - \hat{H} + i\eta} \frac{e^{iqr}}{q} \cdot \hat{e}|0\rangle. \qquad (A.7)$$

We have finally recovered an expression for the macroscopic dielectric function that is equivalent to eq. A.1, where $|P\rangle$ is defined as: $|P\rangle = \lim_{q \to 0} \hat{e} \cdot \frac{e^{iqr}}{q}|0\rangle$. Now we want to show how eq. A.1 can be evaluated in terms of a continued fraction by an iterative procedure.

We start from the state $|P\rangle = |1\rangle$ and we build an orthonormal basis following this scheme:

$$a_1 = \langle 1|H|1\rangle; \qquad b_1 = |H|1\rangle - a_1|1\rangle|$$

$$|2\rangle = \frac{H|1\rangle - a_1|1\rangle}{|H|1\rangle - a_1|1\rangle|} = \frac{H|1\rangle - a_1|1\rangle}{b_1}$$

it is straightforward to see that $\langle 1|2\rangle = 0$, and that $\langle 1|H|2\rangle = b_1$. We call $a_2 = \langle 2|H|2\rangle$. Next $|3\rangle$ is defined as:

$$|3\rangle = \frac{H|2\rangle - a_2|2\rangle - b_1|1\rangle}{|H|2\rangle - a_2|2\rangle - b_1|1\rangle|} = \frac{H|2\rangle - a_2|2\rangle - b_1|1\rangle}{b_2}$$

where we have implicitely defined b_2.
From this three initial states it is possible give an iterative definition for the rest of the basis:

$$a_i = \langle i|H|i\rangle; \qquad b_i = |H|i\rangle - a_i|i\rangle - b_{i-1}|i - 1\rangle|$$

$$|i + 1\rangle = \frac{H|i\rangle - a_i|i\rangle - b_{i-1}|i - 1\rangle}{b_i}; \qquad (A.8)$$

it is easy to show that this is an orthonormal set and that the Hamiltonian is written in a tridiagonal form in terms of the coefficients $\{a_i, b_i\}$ as:

$$H = \begin{pmatrix} a_1 & b_1 & 0 & \ldots & & & & \\ b_1 & a_2 & b_2 & 0 & \ldots & & & \\ 0 & b_2 & \ddots & & & & & \\ \vdots & \vdots & & & & & & \\ & & & a_i & b_i & 0 & \ldots & \\ & & & b_i & a_{i+1} & & & \\ & & & 0 & b_{i+1} & \ddots & & \\ & & & & \vdots & \vdots & & \end{pmatrix} \qquad (A.9)$$

Let's now define the operator

$$(z - \hat{H}) = (\omega + i\eta - \hat{H}). \qquad (A.10)$$

In the Hilbert space connected by the Hamiltonian to the state $|1\rangle$ the set of states $|i\rangle$ that we have just generated is complete so that we can write:

$$\sum_n \langle m|(z - \hat{H})|n\rangle \langle n|(z - \hat{H})^{-1}|p\rangle = \delta_{mp} \qquad (A.11)$$

in particular, choosing $p = 1$ we get:

$$\sum_n \langle m|(z - \hat{H})|n\rangle \langle n|(z - \hat{H})^{-1}|1\rangle = \sum_n (z - \hat{H})_{mn}(z - \hat{H})^{-1}_{n1}\delta_{m1}. \qquad (A.12)$$

We rewrite everything more conveniently defining:

$$X_n = (z - \hat{H})^{-1}_{n1} = \langle n|(z - \hat{H})^{-1}|1\rangle$$

and noting that $Im(X_1)$ is exactly $I(\omega)$ we were looking for. If we write eq. A.12 in terms of $\{X_i\}$ we have:

$$\sum_n (z - H)_{mn}X_n = \delta_{m1} \qquad (A.13)$$

that is a linear system of equation for $\{X_i\}$. So applying basic linear algebra theorems we have that:

$$X_1 = \frac{\det B_0}{\det(z - H)} \qquad (A.14)$$

120

where B_0 is a matrix identical to $(z - H)$ except for its first column that has

been substituted with the vector $\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}$. Namely the two matrices are:

$$z - H = \begin{pmatrix} z - a_1 & -b_1 & 0 & \cdots \\ -b_1 & z - a_2 & -b_2 & 0 & & \cdots \\ 0 & -b_2 & \ddots \\ \vdots & \vdots \\ & & & z - a_i & -b_i & 0 & \cdots \\ & & & -b_i & z - a_{i+1} \\ & & & 0 & -b_{i+1} & \ddots \\ & & & & \vdots & \vdots \end{pmatrix} \qquad (A.15)$$

and

$$z - H = \begin{pmatrix} 1 & -b_1 & 0 & \cdots \\ 0 & z - a_2 & -b_2 & 0 & & \cdots \\ 0 & -b_2 & \ddots \\ \vdots & \vdots \\ & & & z - a_i & -b_i & 0 & \cdots \\ & & & -b_i & z - a_{i+1} \\ & & & 0 & -b_{i+1} & \ddots \\ & & & & \vdots & \vdots \end{pmatrix} \qquad (A.16)$$

If now we call D_i the matrix obtained from $(z - H)$ removing the first i rows and columns, we can see that:

$$\det(z - \hat{H}) = (z - a_1) \det D_1 - b_1^2 \det D_2 \qquad (A.17)$$

and that

$$\det B_0 = \det D_1. \qquad (A.18)$$

Inserting this results in the expression for X_0 given by eq. A.14 we get that:

$$I(\omega) = Im\left(\frac{\det D_1}{(z - a_1)\det D_1 - b_1^2 \det D_2}\right) =$$

$$= Im\left(\frac{1}{(z - a_1) - b_1^2 \frac{\det D_2}{\det D_1}}\right) \qquad (A.19)$$

121

now in the same way:

$$\det D_1 = (z - a_2) \det D_2 - b_2^2 \det D_3 \qquad (A.20)$$

so that

$$I(\omega) = Im \frac{1}{(z - a_1) - \frac{b_1^2}{z - a_2 - \frac{\det D_3}{\det D_2}}} \qquad (A.21)$$

in general it can be seen that:

$$\det D_i = (z - a_{i+1}) \det D_{i+1} - b_{i+1}^2 \det D_{i+2} \qquad (A.22)$$

so that finally we can express $I(\omega)$ as a continued fraction:

$$I(\omega) = Im \left(\frac{1}{z - a_1 - \frac{b_1^2}{z - a_2 - \frac{b_2^2}{z - a_3 - \cdots}}} \right). \qquad (A.23)$$

To summarise, in order to obtain the spectra one has:

1. Define the starting state $|1\rangle = |P\rangle = \lim_{q \to 0} \hat{e} \cdot \frac{e^{iqr}}{q} |0\rangle$

2. Iteratively apply the Hamiltonian to the state generating the orthonormal set $\{|i\rangle\}$ and storing the coefficients $\{a_i, b_i\}$ defined by the relations:

$$a_i = \langle i|H|i \rangle; \qquad b_i = |H|i\rangle - a_i|i\rangle - b_{i-1}|i-1\rangle|$$

$$|i + 1\rangle = \frac{H|i\rangle - a_i|i\rangle - b_{i-1}|i-1\rangle}{b_i} \qquad (A.24)$$

3. Build the continued fraction A.23, and compute its imaginary part.

4. Check the convergency of the spectra with the number of iteration used to build the continued fraction.

Bibliography

[1] F. P. Bundy, H. T. Hall, H. M. Strong, and R. H. Wentorf, Nature **176**, 51 (1995).

[2] J. C. Angus, H. A. Will, and W. S. Stanko, J. Appl. Phys. **39**, 2915 (1968).

[3] L. K. Bigelow and M. P. D' Evelyn, Surf. Sci. **500**, 986 (2002).

[4] G. A. J. Amaratunga, Science **297**, 1657 (2002).

[5] E. Kohn, M. Adamschik, P. Schmid, A. Denisenko, A. Aleksov, and W. Ebert, J. Phys. D: Appl. Phys. **34**, R77 (2001).

[6] F. Spaziani, M. C. Rossi, S. Salvatori, G. Conte, and P. Ascarelli, Appl. Phys. Lett. **82**, 3785 (2003).

[7] H. Watanabe, K. Hayashi, D. Takeuchi, S. Yamanaka, H. Okushi, K. Kajimura, and T. Sekiguchi, Appl. Phys. Lett. **73**, 981 (1998).

[8] S. Koizumi, K. Watanabe, M. Hasegawa, and H. Kanda, Science **292**, 1899 (2001).

[9] M. Pagels, C. E. Hall, N. S. Lawrence, A. Meredith, T. G. J. Jones, H. P. Godfried, C. S. J. Pickles, J. Wilman, C. E. Banks, R. G. Compton, and L. Jiang, Anal. Chem. **77**, 3705 (2005).

[10] J. Isberg, J. Hammersberg, E. Johansson, T. Wilkström, D. Twitchen, A. J. Whitehead, S. Coe, and A. Scarsbrook, Science **297**, 1670 (2002).

[11] F. Maier, M. Riedel, B. Mantel, J. Ristein, and L. Ley, Phys. Rev. Lett. **85**, 3472 (2000).

[12] P. Strobel, M. Riedel, J. Ristein, and L. Ley, Nature **430**, 439 (2004).

[13] K. Tsugawa, K. Kitatani, H. Noda, A. Hokazono, K. Hirose, M. Tajima, and H. Kawarada, Diamond Relat. Mater. **8**, 927 (1999).

[14] *Artificial Retina gets Diamond Coating*, 2005, NewsNature.

[15] A. Härtl, E. Schmich, J. A. Garrid, J. Hernando, S. C. R. Catharino, S. Walter, P. Feulner, A. Kromka, D. Steinmüller, and M. Stutzmann, Nature Materials **3**, 736 (2004).

[16] W. Yang, O. Auciello, J. E. Butler, W. Cai, J. A. Carlisle, J. E. Gerbi, D. M. Gruen, T. Knickerbocker, T. L. Lasseter, J. N. Russell, L. M. Smith, and R. J. Hamers, Nature Materials **1**, 253 (2002).

[17] F. Bechstedt, *Principles of Surface Physics* (Springer, Berlin Heidelberg, 2003).

[18] F. J. Himpsel, D. E. Eastman, P. Heimann, and J. F. van der Veen, Phys. Rev. B **24**, 7270 (1981).

[19] B. B. Pate, P. M. Stefan, C. Binnis, P. J. Jupiter, I. Lindau, and W. E. Spicer, J. Vac. Sci. Technol. **19**, 349 (1981).

[20] R. P. Chin, J. Y. Huang, Y. R. Shen, T. J. Chuang, and H. Seki, Phys. Rev. B **52**, 5985 (1995).

[21] K. Pandey, Phys. Rev. B **25**, R4338 (1982).

[22] R. Graupner, M. Hollering, A. Ziegler, J. Ristein, L. Ley, and A. Stampfl, Phys. Rev. B **55**, 10841 (1997).

[23] W. J. Huisman, J. F. Peters, and J. F. van der Veen, Surf. Sci. **396**, 253 (1998).

[24] T. E. Derry, L. Smit, and J. F. Van Der Veen, Surf. Sci. **167**, 502 (1986).

[25] E. C. Sowa, G. D. Kubiak, R. H. Stulen, and M. A. VanHove, J. Vac. Sci. Technol. A **6**, 832 (1988).

[26] S. Walter, J. Bernhardt, U. Starke, K. Heinze, F. Maier, J. Ristein, and L. Ley, J. Phys. Condens. Matter **14**, 3085 (2002).

[27] W. J. Huisman, M. Lohmeier, H. A. van der Vegt, J. F. Peters, S. A. de Vries, E. Vlieg, V. H. Etgens, T. E. Derry, and J. F. van der Veen, Surf. Sci. **396**, 241 (1998).

[28] J. F. Morar, F. J. Himpsel, G. Hollinger, J. L. Jordon, G. Hughes, and F. McFeely, Phys. Rev. B **33**, 1340 (1985).

[29] A. A. Stekolnikov, J. Furthmüller, and F. Bechstedt, Phys. Rev. B **65**, 115318 (2002).

[30] G. D. Kubiak and K. W. Kolasinski, Phys. Rev. B **39**, 1381 (1989).

[31] J. F. Morar, F. J. Himpsel, G. Hollinger, J. Jordon, G. Hughes, and F. R. McFeely, Phys. Rev. B **33**, 1346 (1985).

[32] R. E. Schlier and H. E. Farnsworth, J. Chem. Phys. **30**, 917 (1959).

[33] P. W. Palmberg and W. T. Peria, Surf. Sci. **6**, 57 (1967).

[34] M. Born and J. M. Oppenheimer, Ann. Physik **84**, 457 (1927).

[35] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[36] R. M. Dreizler and E. K. U. Gross, *Density Functional Theory* (Springer Verlag, Berlin Heidelberg, 1990).

[37] E. H. Lieb, A Brief Review of Thomas-Fermi Theory, 2000, www.physics.nyu.edu/LarrySpruch/Lieb.pdf.

[38] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[39] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[40] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[41] J. P. Perdew and S. Kurth, in *A Primer in Density Functional Theory*, Vol. 620 of *Lecture Notes in Physics*, edited by C. Fiolhais, F. Nogueira, and M. Marques (Springer, Berlin, 2003).

[42] J. P. Perdew, K. Burke, and Y. Wang, Phys. Rev. B **54**, 6533 (1986).

[43] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992).

[44] G. B. Bachelet, *Density Functional Theory and First-Principles Pseudopotentials* (1988), dipartimento di Fisica dell' Università di Trento.

[45] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Saunders College Publishing, Orlando, 1976).

[46] D. J. Chadi and M. L. Cohen, Phys. Rev. B **8**, 5747 (1973).

[47] M. Bockstedte, A. Kley, J. Neugebauer, and M. Scheffler, Comp. Phys. Comm. **107**, 187 (1997).

[48] M. Rohlfing, M. Palummo, G. Onida, and R. Del Sole, Phys. Rev. Lett. **85**, 5440 (2000).

[49] A. Scholze, W. G. Schmidt, and F. Bechstedt, Phys. Rev. B **53**, 13725 (1996).

[50] G. Kern, J. Hafner, J. Furthmüller, and G. Kresse, Surf. Sci. **357-358**, 422 (1995).

[51] S. Iaorli, G. Galli, F. Gygi, M. Parrinello, and E. Tosatti, Phys. Rev. Lett. **69**, 2947 (1992).

[52] F. Bechstedt, A. A. Stekolnikov, J. Furthmüller, and P. Käckell, Phys. Rev. Lett. **87**, 16103 (2001).

[53] J. P. Perdew, R. G. Parr, M. Levy, and J. L. B. Jr, Phys. Rev. Lett. **49**, 1691 (1982).

[54] J. F. Janak, Phys. Rev. B **18**, 7165 (1978).

[55] J. P. Perdew and M. Levy, Phys. Rev. Lett. **51**, 1884 (1983).

[56] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. Lett. **56**, 2415 (1986).

[57] M. Grüning, A. Marini, and A. Rubio, 2005, submitted.

[58] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).

[59] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, San Francisco, 1971).

[60] L. D. Landau and E. M. Lifŝits, *Fisica Statistica, teoria dello stato condensato* (Editori Riuniti, Roma, 1981).

[61] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Dover, New York, 1975).

[62] R. W. Godby and P. Garcï¿$\frac{1}{2}$-Gonzàlez, in *A Primer in Density Functional Theory*, Vol. 620 of *Lecture Notes in Physics*, edited by C. Fiolhais, F. Nogueira, and M. Marques (Springer, Berlin, 2003).

[63] B. Farid, in *Electron Correlation in the Solid State*, edited by N. H. March (Imperial College Press, London, 1999), pp. 103–262.

[64] L. Hedin, Phys. Rev. **139**, 796 (1965).

[65] L. Hedin and S. Lundqvist, Solid State Physics **23**, 1 (1969).

[66] R. Del Sole, L. Reining, and R. W. Godby, Phys. Rev. B **49**, 8024 (1994).

[67] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).

[68] R. W. Godby and R. J. Needs, Phys. Rev. Lett. **62**, 1169 (1989).

[69] The GW code that has been used is the original Godby-Reining code, with the contribution of Dott. O. Pulci for the extension to the calculations of GW corrections on surfaces.

[70] G. Onida, L. Reining, R. W. Godby, R. Del Sole, and W. Andreoni, Phys. Rev. Lett. **75**, 818 (1995).

[71] B. Arnaud and M. Alouani, Phys. Rev. B **62**, 4464 (2000).

[72] R. Shaltaf, *Are band offset calculations plasmon pole model dependent?*, Nanoquanta Bulletin, 2005.

[73] F. Aryasetiawan and O. Gunnarsson, Rep. Prog. Phys. **61**, 237 (1998).

[74] A. Marini, R. Del Sole, A. Rubio, and G. Onida, Phys. Rev. B **66**, 161104(R) (2002).

[75] L. Hedin, Int. J. Quantum Chem. **56**, 445 (1995).

[76] O. Pulci, F. Bechstedt, G. Onida, R. Del Sole, and L. Reining, Phys. Rev. B **60**, 16758 (1999).

[77] M. Rohlfing and S. G. Louie, Phys. Rev. B **62**, 4927 (2000).

[78] F. Bruneval, private communication.

[79] R. Del Sole, private communication.

[80] O. Pulci, M. Marsili, E. Luppi, C. Hogan, V. Garbuio, F. Sottile, R. Magri, and R. Del Sole, Phys. Stat. Sol. (b), **242**, 2737 (2005).

[81] W. Richter and P. Chiaradia, private communication.

[82] W. D. Schöne and A. G. Eguiluz, Phys. Rev. Lett. **81**, 1662 (1998).

[83] W. Ku and A. G. Eguiluz, Phys. Rev. Lett. **89**, 126401 (2002).

[84] K. Delaney, P. García-Gonzáles, A. Rubio, P. Rinke, and R. W. Godby, Phys. Rev. Lett. **93**, 249701 (2004).

[85] W. Ku and A. G. Eguiluz, Phys. Rev. Lett. **93**, 249702 (2004).

[86] M. L. Tiago, S. Ismail-Beigi, and S. G. Louie, Phys. Rev. B **69**, 125212 (2004).

[87] S. V. Faleev, M. van Schilfgaarde, and T. Kotani, Phys. Rev. Lett. **93**, 126406 (2004).

[88] A. Marini, R. Del Sole, A. Rubio, and G. Onida, Phys. Rev. B **66**, 161104(R) (2002).

[89] E. K. Chang, X. Blase, and S. G. Louie, Phys. Rev. B **64**, 155108 (2003).

[90] J. A. Alford, M. Y. Chou, E. K. Chang, and S. G. Louie, Phys. Rev. B **67**, 125110 (2003).

[91] A. Fleszar and W. Hanke, Phys. Rev. B **71**, 45207 (2005).

[92] W. Luo, S. Ismail-Beigi, M. L. Cohen, and S. G. Louie, Phys. Rev. B **66**, 195215 (2002).

[93] F. Bruneval, F. Sottile, V. Olevano, R. Del Sole, and L. Reining, Phys. Rev. Lett. **94**, 186402 (2005).

[94] D. Cahen and A. Kahn, Adv. Mater **15**, 271 (2003).

[95] L. Kronik and Y. Shapira, Surf. Sci. Rep. **37**, 1 (1999).

[96] S. Baroni, R. Resta, A. Baldereschi, and M. Persessi, in *Spectroscopy of Semiconductor Microstructures*, edited by G. Fasol, A. Fasolino, and P. Lugli (Plenum Publishing Corporation, New York, 1989), p. 251.

[97] C. Sgiarovello, N. Binggeli, and A. Baldereschi, Phys. Rev. B **64**, 195305 (2001).

[98] J. B. Cui, J. Reistein, and L. Ley, Phys. Rev. Lett. **81**, 429 (1998).

[99] F. Maier, J. Reistein, and L. Ley, Phys. Rev. B **64**, 165411 (2001).

[100] G. Grosso and G. Pastori Parravicini, *Solid State Physics* (Academic Press, London, 2000).

[101] M. A. L. Marques and E. K. U. Gross, Ann. Rev. Phys. Chem. **55**, 427 (2004).

[102] F. Sottile, Ph.D. thesis, Ecole Polytechnique, Palaiseau, 2003.

[103] G. Adragna, R. Del Sole, and A. Marini, Phys. Rev. B **68**, 165108 (2003).

[104] L. D. Landau, E. M. Lifšits, and L. P. Pitaevskij, *Elettrodinamica dei mezzi continui* (Editori Riuniti, Roma, 1986).

[105] J. D. Jackson, *Classical Electrodynamics* (John Wiley & Sons; 3rd edition, New York, 1999).

[106] Del Sole, in *Photonic Probes of Surfaces*, Vol. 2 of *Electromagnetic Waves: Recent Developments in Research*, edited by P. Halevi (Elsevier, Amsterdam, 1995), p. 133.

[107] S. L. Adler, Phys. Rev. **126**, 413 (1962).

[108] V. I. Gavrilenko and F. Bechstedt, Phys. Rev. B **55**, 4343 (1997).

[109] EXC code, http://theory.polytechnique.fr/codes/exc/.

[110] W. Hanke, in *Advances in Solid State Physics*, edited by J. Treusch (Vieweg, Braunschweig, 1979), p. 43.

[111] R. Haydock, Comput. Phys. Comm. **20**, 11 (1980).

[112] A. D. Papadopoulos and E. Anastassakis, Phys. Rev. B **43**, 5090 (1991).

[113] P. Weightman, D. S. Martin, R. J. Cole, and T. Farrell, Rep. Prog. Phys. **68**, 1251 (2005).

[114] A. Bagchi, R. G. Barrera, and A. K. Rajagopal, Phys. Rev. B **20**, 4824 (1979).

[115] R. Del Sole, Sol. State Comm. **37**, 537 (1981).

[116] F. Manghi, R. Del Sole, A. Selloni, and E. Molinari, Phys. Rev. B **41**, 9935 (1990).

[117] C. Noguez and S. E. Ulloa, Phys. Rev. B **53**, 13138 (1996).

[118] M. Palummo, O.Pulci, R. Del Sole, A. Marini, M. Schwitters, S. R. Haines, K. H. Williams, D. S. Martin, P. Weightman, and J. E. Butler, Phys. Rev. Lett. **94**, 087404 (2005).

[119] E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994).

[120] E. R. Gagliano and C. A. Balseiro, Phys. Rev. Lett. **59**, 2999 (1987).

[121] L. X. Bendict and E. L. Shirley, Phys. Rev. B **59**, 5441 (1999).

129

[122] A. F. Starace, Phys. Rev. A **3**, 1242 (1971).

[123] S. Albrecht, Ph.D. thesis, Ecole Polytechnique, Palaiseau, 1999.

[124] R. Del Sole and R. Girlanda, Phys. Rev. B **48**, 11789 (1993).