ISAS - INTERNATIONAL SCHOOL
FOR ADVANCED STUDIES

Dynamical Charges at Surfaces and Interfaces: Their Role in the Schottky Barrier Problem

Thesis submitted for the degree of

"Doctor Philosophiæ"

CANDIDATE

Alice Ruini

SUPERVISORS

Prof. Raffaele Resta

Prof. Stefano Baroni

October 1997

SISSA - SCUOLA INTERNAZIONALE SUPERIORE DI STUDI AVANZATI
TRIESTE
Strada Costiera 11

TRIESTE

This thesis is available on the Web at the URL: http://www.sissa.it/ cm/thesis/1997/

SISSA ISAS

SCUOLA INTERNAZIONALE SUPERIORE DI STUDI AVANZATI
INTERNATIONAL SCHOOL FOR ADVANCED STUDIES

Dynamical Charges at Surfaces and Interfaces: Their Role in the Schottky Barrier Problem

Thesis submitted for the degree of

"Doctor Philosophiæ"

CANDIDATE

Alice Ruini

SUPERVISORS

Prof. Raffaele Resta

Prof. Stefano Baroni

October 1997

to my parents

Table of Contents

Table of Contents i

1 **Introduction** 1

2 **The dynamical effective charges** 5

2.1 The concept of atomic charge 5

2.2 The static charge: is it a well-defined quantity? 5

2.3 The dynamical charge from lattice dynamics 7

2.3.1 Equations of motion and dynamical matrix 8

2.3.2 Dispersion relations 11

2.3.3 The case of polar crystals: the dynamical charges 13

2.4 Some remarks about the dynamical charge concept 16

2.4.1 Acoustic sum rule 18

3 **Computational tools** 20

3.1 Density-Functional Theory 20

3.1.1 The plane-wave pseudopotential method 22

3.2 Density Functional Perturbation Theory 24

3.2.1 Linear response in crystals 25

3.2.2 Implementation of DFPT for a periodic solid 27

3.2.3 Calculation of the dynamical charges within DFPT 28

3.3 Direct method for the calculation of the dynamical charges 30

4 The dynamical-charge neutrality condition for polar surfaces 33

4.1 Formulation of the dynamical-charge neutrality condition at a crystal surface . 34

4.1.1 The (001) case 36

4.1.2 The (111) case 38

4.2 A case-study calculation for β-SiC surfaces 39

4.2.1 Bulk properties of β-SiC 40

4.2.2 The β-SiC(001) surface 41

4.2.3 The H-covered β-SiC(111) surface 44

5 Metal-semiconductor interfaces: Al/GaAs(001) Schottky barrier 48

5.1 The Schottky barrier problem: models and experiments 48

5.2 Method for the ab-initio calculation of the Schottky barrier height 53

5.3 Bulk properties of GaAs and Al 57

5.3.1 Gallium arsenide 57

5.3.2 Aluminum 58

5.4 Al/GaAs(001) junction 59

6 Schottky barrier interface morphology: role of dynamical charges 66

6.1 Thick layer of vacuum between metal and semiconductor 67

6.2 Bond elongations 68

6.3 Bulk strain of the metal overlayer 69

6.4 Interface study based on the dynamical charges 71

6.5 Dynamical charge neutrality at Schottky barriers 73

6.6 Discussion 74

7 Interface electronic states at Schottky barriers **76**

7.1 Calculation of interface states for Al/GaAs(001) 77

7.2 The (uncovered) GaAs(001) surface 80

7.3 The monolayer-coverage case 81

7.4 Al/ZnSe(001) interfaces 84

7.4.1 Bulk properties of ZnSe 86

7.4.2 Schottky barrier height calculation 87

7.4.3 Dynamical charges 88

7.4.4 MIGS-based analysis 88

8 Conclusions **91**

Acknowledgments **94**

A The macroscopic–average technique **97**

B The smearing technique in Brillouin-zone integration for metals **101**

Bibliography **104**

1 Introduction

During the past years, components based on metal-semiconductor junctions, also called Schottky barrier junctions, have been increasingly used in microelectronics [1], and research activity has continued with the aim of obtaining a full understanding of the physics of interfacial potential barrier formation (which is at the origin of the rectifying properties of the junction) and of current transport across the interface. In particular, the Schottky barrier height is the fundamental parameter for characterizing the electronic properties of related devices.

Experimentally it has been found that the barrier height depends on the specific semiconductor and is generally only weakly sensitive to the metal [2]: the Fermi level is therefore said to be *pinned* within the semiconductor gap for a number of metal-semiconductor interfaces. This property holds in particular for covalent semiconductors, while ionic compounds exhibit a relatively broad range of barrier heights for different contact metals. Moreover, the Fermi-level pinning is observed even at very low metal coverages: actually it has been found [3] that at more than a monolayer coverage the position of the Fermi energy converges rapidly to a unique value.

Despite several decades of extensive experimental and theoretical work [2], the key factors affecting the Fermi-level pinning at metal-semiconductor contacts have not yet been clearly assessed. The controversy concerns even the very basic issue as to whether the pinning is determined by *intrinsic* interface states which exist even at an abrupt ideal interface, or by *extrinsic*

electronic states arising from native defects. A large number of phenomenological models has been proposed, in order to get an explanation of the physical mechanisms responsible for the Fermi level pinning. Despite the assessment of the above facts, the experimental data are of little help in discriminating between the different theoretical pictures, given that the latter provide similar predictions from a quantitative point of view, and mostly in agreement with experimental findings.

Unfortunately, there is essentially no experimental access to the microscopic morphology of a given interface: were this known, the actual Schottky barrier height would be unambiguously determined by the laws of electrostatics and of quantum mechanics. The problem of Fermi-level pinning at semiconductor-metal contacts can be readdressed starting from *first-principles* calculations, that—though necessarily limited to rather idealized situations and affected by basic approximations necessary to cope with the complexity of the electronic many-body problem—allow, by their very nature, to have full control on the way the details at the atomic scale of a given system affect the various physical properties under investigation. In this sense, *ab initio* calculations are complementary to the experimental investigations and, in the specific case of metal-semiconductor contacts, they have in fact provided in recent years a great deal [4–10] of quantitative information that any successful model will have to account for. Because of this theoretical work, the following facts are now well established: the barrier height *does* depend on the nature of the metal, on the crystallographic direction [4], and furthermore for a given crystallographic direction of growth it even depends on the microscopic morphology of the interface [7]. Nevertheless, the electronic mechanisms actually governing the value of the Schottky barrier height—as well as their specific dependence on the microscopic morphology of the interface—have not been systematically investigated so far and are basically unknown.

In this thesis we provide a contribution in this direction, by studying the barrier-height variations induced in Al/GaAs(001) by several structural and morphological perturbations

which can be easily switched on and off in our computational framework. Our *ab initio* calculations, based on density functional theory, provide a microscopic probe for the nature of the interface—including its "effective" thickness—and for the electronic response phenomena responsible for the barrier height. In a microscopic description of insulating materials, the basic constants which couple electrostatic potentials to ionic displacements go under the name of dynamical effective charges. In the present thesis the Schottky barrier problem is investigated by means of the interface dynamical effective charges: we will elucidate their crucial role in determining the variations of the interface dipole and we will prove that they are the dominant factor affecting a distortion-induced variation of the barrier height.

For both molecules and periodic solids, the ionic dynamical charge tensors are known to obey a dynamical neutrality condition, which enforces their sum to vanish (over the whole finite system, or over the crystal cell, respectively). The application of the dynamical charge concept to the study of the Schottky barrier problem suggested us to extend the dynamical charge neutrality to systems where the periodicity is broken in one direction. For the sake of simplicity, we focus first our interest on the surface case and we show that a crystalline surface must actually be dynamically neutral in order to ensure that a rigid translation of the semiinfinite solid as a whole does not affect the work function. An explicit formulation of such neutrality requires in general the regularization of a nonconvergent sum, which results in a constraint for the dynamical charges of the surface ions; for a polar surface—such as (001) or (111) in the zincblende structure—the novel sum rule *forbids* surface ions of a given chemical species to have the same dynamical charge as in the bulk.

This novel formulation for the surface dynamical charge neutrality is found to hold also for metal-semiconductor interfaces. The detailed information about the interface morphology that we obtained by means of the dynamical charges is then conveniently coupled with an analysis of the electronic states at the interface, which allows one to go more inside to the physical

mechanisms at the origin of Fermi level pinning.

The second Chapter of this thesis is mainly devoted to a review of the dynamical charge concept in solid state physics. In Chapter 3 we briefly discuss the theoretical tools used in this work, namely density functional theory and density functional perturbation theory. Chapter 4 provides an extension of the dynamical charge neutrality condition to the case of a crystalline surface, which gives rise to a novel formulation for the acoustic sum rule. Our results concerning the Schottky barrier height calculation for the Al/GaAs(001) junction are presented in Chapter 5, while Chapter 6 is devoted to a detailed description of the morphology induced variations of the Schottky barrier, and a rationale for the results is given in terms of the dynamical charges. In Chapter 7 we present a study of the electronic states at the interface, of their decay length within the semiconductor, of their robustness under metal deposition and semiconductor ionicity. Finally, the last Chapter is devoted to our conclusions.

2 The dynamical effective charges

2.1 The concept of atomic charge

For a long time, there has been a continuing interest in the definition of atomic charges in solid state physics as well as in chemistry [11–13]. This interest lies essentially in the fact that this concept is helpful for a simple description of solids and molecules. The large variety of frameworks in which a concept of atomic charge naturally arises (definition of oxidation states, determination of electrostatic potential, theory of ionic conductivity of oxides, infrared spectrum analysis, XPS chemical shifts analysis...) underlines its central role. However, it reveals at the same time a problem: inspired by various models or by the description of various physical phenomena, many different definitions have been proposed that, unfortunately, are not equivalent [13]. However, it seems globally possible to separate the different concepts into static and dynamical charge.

2.2 The static charge: is it a well-defined quantity?

Intuitively, the atomic charge first appears as a static concept. The charge of an isolated atom is a well defined quantity; the purpose of defining atomic charges was therefore to extend this notion to molecules and solids, which is particularly nontrivial in the case of covalent bonds [14]. This challenge basically consists to replace the delocalized electron density by localized point

charges associated to each atom. This could *a priori* be performed from electronic density maps obtained experimentally or theoretically. A large variety of attempts have been made in order to define a unique way for partitioning the ground-state electronic density into contributes attributed to the different atoms, and we propose in this Section a brief overview.

A first group of procedures involves the basis functions that are used to represent the wave functions. The oldest of these methods is the Mulliken population analysis [15], which is unfortunately well known to be strongly dependent on the choice of the basis functions; an improvement of this technique, that eliminates most of its drawbacks, was proposed by Weinhold *et al.* [16], who introduced the concept of natural atomic orbitals.

Alternative approaches are based directly on the charge density distribution; these methods are usually preferred because they represent the first term of a multicenter multipole expansion and reproduce the dipole moments and the electrostatic potential in a satisfactory manner. In a first kind of definitions, like that of Hirshfeld [17], the charge is separated into *overlapping* contributions: the charge density at each point of space is separated between the constituent atoms in the same proportions as they contribute to the charge density to that point in a hypothetical compound, constructed by the superposition of the spherically symmetrized charge density distributions of the isolated atoms (various methods for fitting overlapping atomic densities have been proposed [13]). Another family of methods splits the electronic density between *non-overlapping* regions on the basis of topological arguments [18, 19], as firstly suggested by Bader [20, 21]. In this approach critical points are first identified that correspond to minimum charge density along the bond between the different atoms; from these points, paths for which the charge density decreases most rapidly are then developed in the direction perpendicular to the bond and the sets of these paths define zero-flux surfaces separating the atoms. This method is very elegant but it suffers from some problems. In particular, there is a possibility of empty portions of space that do not belong to any particular atom; furthermore,

since a high density on a particular atom repels the zero-flux surface, the Bader charge value usually exaggerates the atomic charges. In spite of these drawbacks, this charge is sometimes considered as the best choice.

All the previous methods are related to the electronic density and are probably the most commonly used. At variance with them, some atomic charges were also introduced in connection with other quantities, experimentally measured and related to an atomic charge via a simplified model (for example, consider the atomic charges deduced from the chemical shifts of core ionization energies in XPS or ESCA measurements [13]).

Finally, and without being exhaustive, it is important to mention that natural definitions also arise in the framework of the semi-empirical approaches; for example, in the bond orbital model of Harrison [22], the electronic interactions are modelized through a few parameters that monitor the charge transfer between the ions and allow to identify an effective static charge.

Although these procedures address in principle the same universal concept, each of them yields in practice a different quantitative result; it is therefore argued that the static charge is an ill-defined quantity that depends on the convention chosen [11, 12].

2.3 The dynamical charge from lattice dynamics

"*Whenever an ambiguity arises about the definition of a concept such as the atomic charge, it can be removed by discussing only quantities that can be experimentally determined at least in principle*" (from Ref. [22], p.125): the dynamical effective charges turn out to be well-defined quantities, because they are directly related to the change of polarization created by an atomic displacement, and are therefore experimentally measurable (at least in principle). The dynamical charge is a dynamical property in the sense that it concerns the response of the system to an atomic displacement. A complete description of the dynamical properties of the solids must be achieved by considering the presence of both positively charged nuclei

and negatively charged electrons (whose mutual interactions have a fundamental role), which requires the solution of the complete Schrödinger equation of the system. Both from analytical and computational points of view this is a very difficult task.

However, due to the large mass of the nuclei compared to that of the electrons, the nuclear motion can be decoupled from the electronic one for a great number of physical systems, in the spirit of the adiabatic approximation [23].

Following this approximation, the vibrational properties of the crystal can be derived by considering the ions as particles moving in a potential which integrates out the electronic degrees of freedom. This potential is the sum of a direct interaction and of an indirect one (also called ion-electron-ion interaction).

The assumption of small displacements from the equilibrium positions leads to the harmonic approximation [23]; the normal vibrational modes of a harmonic crystal are called *phonons*. In the following both the electron-phonon interaction (due to non-adiabatic effects) and the phonon-phonon interaction (due to anharmonicity) will be explicitly neglected.

2.3.1 Equations of motion and dynamical matrix

Let us consider an extended three-dimensional crystal made by N cells with n atoms in the unit cell. Then the position of the generic atom of cell L can be written, for the undistorted crystal, as:

$$\mathbf{R}_L^s = \mathbf{R}_L + \mathbf{d}_s \qquad s = 1, 2, ..., n \tag{2.1}$$

The lattice vector $\mathbf{R}_L$ can be expressed in terms of the basis vectors a_i as:

$$\mathbf{R}_L = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3, \tag{2.2}$$

where n_i are integer numbers, and the position of the s-th atom in the unit cell is given by:

$$\mathbf{d}_s = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 \qquad 0 < x_i < 1 \tag{2.3}$$

In the harmonic approximation one considers small displacements from the equilibrium position, so that the total potential energy per crystal cell can be expressed in terms of the displacements defined by:

$$\mathbf{R}_L^s \rightarrow \mathbf{R}_L^s + \mathbf{U}_L^s \tag{2.4}$$

as a Taylor expansion up to the second order:

$$E = E_0 + \frac{1}{2N} \sum_{LL'ss'} \mathbf{U}_L^s \mathbf{C}_{LL'}^{ss'} \mathbf{U}_{L'}^{s'}. \tag{2.5}$$

The coefficients $\mathbf{C}_{LL'}^{ss'}$ appearing in Eq. (2.5) are called *interatomic force constants* and are given by:

$$\mathbf{C}_{LL'}^{ss'} = \left. \frac{N\partial^2 E}{\partial \mathbf{U}_L^s \partial \mathbf{U}_{L'}^{s'}} \right)_0, \tag{2.6}$$

where the second derivatives are calculated at the equilibrium positions. Differentiation with respect to $\mathbf{U}_L^s$ of Eq. (2.5) allows us to write the force acting on an atom (s, L):

$$\mathbf{F}_L^s = -N\frac{\partial E}{\partial \mathbf{U}_L^s} = -\sum_{L's'} \mathbf{C}_{LL'}^{ss'} \mathbf{U}_{L'}^{s'}. \tag{2.7}$$

The interatomic force constants defined in Eq. (2.5) are not independent quantities, but they are connected to each other, due to the invariance properties of the crystal. In particular the following relations hold:

a. because of the translational invariance of the crystal, the force constants do not depend on the absolute positions in the crystal so that:

$$\mathbf{C}_{LL'}^{ss'} = f^{ss'}(\mathbf{R}_L - \mathbf{R}_L') \tag{2.8}$$

b. the potential energy remains unchanged if all atoms are displaced of the same amount

$$\mathbf{U}_L^s \rightarrow \mathbf{U}_L^s + \mathbf{v}. \tag{2.9}$$

This must be true for any $\mathbf{v}$ and $\mathbf{U}_L^s$, hence:

$$\sum_{L's'} \mathbf{C}_{LL'}^{ss'} = 0 \tag{2.10}$$

for any (s, L).

In the harmonic approximation the equations for the ionic motion are given by:

$$M_s \ddot{\mathbf{U}}^s_L = -\sum_{L's'} \mathbf{C}^{ss'}_{LL'} \mathbf{U}^{s'}_{L'}, \tag{2.11}$$

which constitute an infinite set of coupled equations. They may be obtained either classically from Hamilton's equations:

$$\begin{aligned} \dot{\mathbf{U}}^s_L &= \frac{\partial \mathcal{H}}{\partial \mathbf{P}^s_L} = \frac{\mathbf{P}^s_L}{M_s} \\ \dot{\mathbf{P}}^s_L &= \frac{\partial \mathcal{H}}{\partial \mathbf{U}^s_L} = -\sum_{L's'} \mathbf{C}^{ss'}_{LL'} \mathbf{U}^{s'}_{L'} \end{aligned} \tag{2.12}$$

where the Hamiltonian is:

$$\mathcal{H} = \sum_{Ls} \frac{\mathbf{P}^s_L}{2M_s} + NE, \tag{2.13}$$

or via the quantum mechanical equations of motion:

$$\begin{aligned} \dot{\mathbf{U}}^s_L &= \frac{i}{h}[H, \mathbf{U}^s_L] = \frac{\mathbf{P}^s_L}{M_s} \\ \dot{\mathbf{P}}^s_L &= -\frac{i}{h}[H, \mathbf{P}^s_L] = -\sum_{L's'} \mathbf{C}^{ss'}_{LL'} \mathbf{U}^{s'}_{L'} \end{aligned} \tag{2.14}$$

with H as the quantum counterpart of $\mathcal{H}$.

The solutions of Eqs. (2.11) can be expressed as:

$$\mathbf{U}^s_L = \frac{1}{\sqrt{M_s}} \mathbf{u}_s e^{i\mathbf{q}\mathbf{R}_L - i\omega t}. \tag{2.15}$$

This form is suggested by the fact that, because of translational invariance (2.8), the atomic displacements $\mathbf{U}^s_L$ must simultaneously be eigenfunctions of the force constants matrix and of the translation operator; the polarization vectors $\mathbf{u}_s$ and the frequency ω have to be properly determined in order to satisfy (2.11). The allowed values of $\mathbf{q}$ are chosen according to the Born-Von Karman periodic boundary conditions. Due to the periodicity in the reciprocal space of (2.15), one can limit his interest to $\mathbf{q}$ vectors belonging to the Brillouin zone of the crystal. By

substitution of (2.15) in (2.11) one obtains the equations:

$$\omega^2 \mathbf{u}_s = \sum_{s'} \mathbf{D}_{ss'}(\mathbf{q})\mathbf{u}'_s \tag{2.16}$$

where the *dynamical matrix* of the crystal $\mathbf{D}_{ss'}(\mathbf{q})$ has been introduced as a discrete Fourier transform:

$$\mathbf{D}_{ss'}(\mathbf{q}) = \frac{1}{\sqrt{M_s M_{s'}}} \sum_{l'} \mathbf{C}^{ss'}_{LL'} e^{i\mathbf{q}(\mathbf{R}_L - \mathbf{R}_{L'})} \tag{2.17}$$

Because of the crystal periodicity, Eq. (2.17) does not depend on index L. Thus, for a periodic crystal the infinite set of equations of motion (2.11) is reduced, for each $\mathbf{q}$, to a system of $3n$ equations in the $3n$ unknowns $\mathbf{u}_s$. Due to the symmetry properties of the force constants, $\mathbf{D}_{ss'}(\mathbf{q})$ is a $3n \times 3n$ hermitian matrix,

$$\mathbf{D}_{ss'}(\mathbf{q}) = (\mathbf{D}_{ss'}(\mathbf{q})^*)^T \tag{2.18}$$

which, according to (2.17), has the following property:

$$\mathbf{D}_{ss'}(-\mathbf{q}) = \mathbf{D}_{ss'}(\mathbf{q})^* \tag{2.19}$$

2.3.2 Dispersion relations

The solution of the eigenvalue problem (2.16) provides the frequency and the polarization vector as functions of the wavevector $\mathbf{q}$ in the Brillouin zone. The allowed values of the square of the frequency ω for a given value of $\mathbf{q}$ are the eigenvalues of the dynamical matrix, hence there are $3n$ solutions for ω^2, at each $\mathbf{q}$, which will be denoted by $\omega_j^2(\mathbf{q})$ (where $j = 1, 2, ..., n$) and can be interpreted as the branches of a multivalued function $\omega^2(\mathbf{q})$. Because of hermiticity of the dynamical matrix, the eigenvalues $\omega_j^2(\mathbf{q})$ are real; furthermore the stability of the crystal requires $\omega_j^2(\mathbf{q})$ to be positive (or zero). The relations expressed by the equations $\omega = \omega_j(\mathbf{q})$ are known as *dispersion relations.* Due to the hermiticity of $\mathbf{D}_{ss'}(\mathbf{q})$ and to the fact that the polarization vectors are defined by an eigenvalue problem, Eq. (2.16), we can choose them in

such a way to satisfy the orthonormality and the closure relations. It is easy to show that, close to the zone center ($\mathbf{q} \to 0$), Eq. (2.16) has three distinct solutions with zero frequency. Setting $\mathbf{q} = 0$ in (2.16) and in (2.17) we have that:

$$\omega_j^2(0)\mathbf{u}_s^j(0) = \sum_{L's'} \frac{1}{\sqrt{M_s M_{s'}}} \mathbf{C}_{LL'}^{ss'} \mathbf{u}_{s'} j(0) \tag{2.20}$$

If the ratio between the polarization vector $\mathbf{u}_s^j(0)$ and the square root of the mass $\sqrt{M_s}$ does not depend on s, that is:

$$\frac{\mathbf{u}_s^j(0)}{\sqrt{M_s}} = \mathbf{w}^j \qquad s = 1, 2, ..., n \tag{2.21}$$

then the eigenvalue must be zero, due to relationship (2.10). Since in a three-dimensional space we can chose three independent vectors satisfying (2.21) and because of the orthonormality relation among the polarization vectors, we have three branches with zero frequency at $\mathbf{q} = 0$. The corresponding displacements, that is the solutions of the equations of motion (2.11), are written as:

$$\mathbf{U}_L^j = \mathbf{w}^j e^{-i\omega t} \qquad j = 1, 2, 3 \tag{2.22}$$

and are obviously independent of s. Equations (2.22) say that all n particles in the unit cell move in phase and with equal amplitude; since this is characteristic of the displacements in an elastic continuum upon which a sound wave is impressed, such modes are called *acoustic modes*; moreover, the acoustic branches can be proved to have a linear dispersion near the zone center. In the limit $\mathbf{q} \to 0$ the eigenfrequencies of the remaining $(3n - 3)$ branches of a crystal have a finite limit: they are called *optical modes* because in a ionic crystal such a motion gives rise to a macroscopic fluctuating dipole moment, through which the crystal can interact with the (infrared) electromagnetic radiation. For such modes, the atoms in different sublattices vibrate rigidly against each other.

2.3.3 The case of polar crystals: the dynamical charges

In the case of a polar crystal the ionic displacements are accompanied by the onset of electric dipole moments: the related macroscopic field affects the motion of the ions by acting as an external force and, furthermore, it can polarize the ions and enhance their dipole moments. Electric interactions play a peculiar role in the limit $\mathbf{q} \to 0$ because of their long range nature and the dynamical matrix has in general a non-analytic behavior: in this case it is convenient to derive the equations of motion in a phenomenological approach [23], where, beside the ionic displacements $\mathbf{U}_s$, the macroscopic electric field in the crystal, $\mathbf{E}$, is treated as an independent dynamical variable. However, by using the Maxwell equations the latter variable can be expressed by the former, so that the final equations of motion, which determine the eigenmodes, again contain the displacement variables only. The key-quantity to consider in this case is the thermodynamic potential $\tilde{F}$ [24] (the electric enthalpy), which coincides with the total energy studied so far when the electric field is zero. For a generic displacement

$$\mathbf{U}_L^s = \mathbf{U}_s(\mathbf{q})e^{i\mathbf{q}\mathbf{R}_L} \tag{2.23}$$

and an electric field

$$\mathbf{E} = \mathbf{E}(\mathbf{q})e^{i\mathbf{q}\mathbf{r}}, \tag{2.24}$$

the electric enthalpy may be written as:

$$\tilde{F} = E_0^{\mathrm{tot}} + \frac{1}{2}\sum_{ss'} \mathbf{U}_s(\mathbf{q})\mathbf{C}^0_{ss'}(\mathbf{q})\mathbf{U}_{s'}(\mathbf{q}) - \frac{\Omega}{8\pi}\mathbf{E}(\mathbf{q})\epsilon_\infty\mathbf{E}(\mathbf{q}) - \sum_s e\mathbf{U}_s(\mathbf{q})\mathbf{Z}^*_s(\mathbf{q})\mathbf{E}(\mathbf{q}), \tag{2.25}$$

where Ω is the volume of the unit cell, while $\mathbf{C}^0_{ss'}(\mathbf{q})$ represents the dynamical matrix in absence of a macroscopic electric field and therefore describes the short-range forces including that connected with the Lorentz field (associated with the local field corrections [25]). The second order tensors appearing in the other terms of Eq. (2.25) are related to the dielectric properties of the crystal. In fact, ϵ_∞ is the high-frequency dielectric tensor, which describes the polarization

of the electrons only, whereas the ions are considered fixed in their equilibrium positions; from the experimental point of view its diagonal components are measured by considering oscillating electric fields with frequencies which are large compared to the lattice frequencies but which are smaller than the frequency of the lowest electronic transition. Finally, Z_s^* is the effective charges tensor for s-th atom, which is connected to the mixed second order derivative of the total energy with respect to atomic displacements and macroscopic electric field.

The equations of motion for the displacement variables give the force acting on the generic atom:

$$\mathbf{F}_s(\mathbf{q}) = -\sum_{s'} \mathbf{C}^0_{ss'}(\mathbf{q})\mathbf{U}_{s'}(\mathbf{q}) + \mathbf{Z}^*_s(\mathbf{q})\mathbf{E}(\mathbf{q}), \tag{2.26}$$

whereas for the electric field variable we obtain:

$$\frac{\partial \tilde{F}}{\partial \mathbf{E}} = -\sum_s e\mathbf{Z}^*_s(\mathbf{q})\mathbf{U}_s(\mathbf{q}) - \frac{\Omega}{4\pi}\epsilon_\infty \mathbf{E}(\mathbf{q}) \tag{2.27}$$

It is now convenient to introduce the dielectric displacement $\mathbf{D}$ defined from the electrostatic law:

$$\nabla \cdot \mathbf{D} = 4\pi\rho \tag{2.28}$$

where ρ is the free charge density. The field $\mathbf{D}$ is connected to the electric field $\mathbf{E}$ by the relationship:

$$\mathbf{D} = -\frac{4\pi}{\Omega}\frac{\partial \tilde{F}}{\partial \mathbf{E}} \tag{2.29}$$

Thus, Eq. (2.27) can be written as:

$$\mathbf{D}(\mathbf{q}) = \frac{4\pi}{\Omega}\sum_s e\mathbf{Z}^*_s(\mathbf{q})\mathbf{U}_s(\mathbf{q}) + \epsilon_\infty \mathbf{E}(\mathbf{q}) \tag{2.30}$$

Let us now consider the Maxwell equations in the case of vanishing free charge density for a generic phonon of wavevector $\mathbf{q}$:

$$\begin{aligned} &\mathbf{q} \cdot \mathbf{D} = 0 \\ &\mathbf{q} \times \mathbf{E} = 0 \qquad \Rightarrow \mathbf{E} = \hat{\mathbf{q}}(\hat{\mathbf{q}} \cdot \mathbf{E}). \end{aligned} \tag{2.31}$$

If these relationships are inserted in the above expression for $\mathbf{D}(\mathbf{q})$, we find

$$\mathbf{E}(\mathbf{q}) = -\frac{4\pi e}{\Omega} \sum_s \frac{\hat{\mathbf{q}}(\hat{\mathbf{q}} \cdot \mathbf{Z}^*_s(\mathbf{q}) \cdot \mathbf{U}_s(\mathbf{q}))}{(\hat{\mathbf{q}}\epsilon_\infty \hat{\mathbf{q}})} \tag{2.32}$$

Substituting the latter equation in the expression for $\mathbf{F}_s(\mathbf{q})$, we obtain:

$$\mathbf{C}_{ss'}(\hat{\mathbf{q}}) = \mathbf{C}^0_{ss'}(\hat{\mathbf{q}}) + \frac{4\pi e^2}{\Omega} \frac{\mathbf{Z}^*_s(\mathbf{q}) \cdot \hat{\mathbf{q}} \;\; \hat{\mathbf{q}} \cdot \mathbf{Z}^*_{s'}(\mathbf{q})}{\hat{\mathbf{q}}\epsilon_\infty \hat{\mathbf{q}}} \tag{2.33}$$

So, the interatomic force-constants (and hence the dynamical matrix) have been evaluated accounting for the long-range Coulomb field within a treatment that is valid when the electric field is slowly varying in space ($\mathbf{q} \sim 0$); the last term in right hand side of Eq. (2.33) corresponds to the non-analytical contribution to the dynamical matrix and is therefore responsible for the dependence on the direction of the $\mathbf{q} \to 0$ limit of the dispersion modes.

It is hence evident that Z^* is a fundamental quantity in lattice dynamics: it governs, together with the optical dielectric constant ϵ_∞, the strength of the long-range Coulomb interaction responsible for the splitting between longitudinal and transverse optic modes.

As the polarization is a first derivative of the energy with respect to a ionic displacement, we can infer from Eq. (2.25) a definition for the dynamical charge tensor $Z^*_{s,\alpha\beta}$, as the coefficient of proportionality relating, to linear order and under the condition of zero electric field, the macroscopic polarization per unit cell created along the direction α, and the displacement along the direction β of the atoms belonging to the sublattice s; $Z^*_{s,\alpha\beta}$ equivalently describes the linear relationship between the force $F_{s,\beta}$ induced on atom s and the macroscopic electric field E_α:

$$Z^*_{s,\alpha\beta} = -\frac{\partial^2 \tilde{F}}{\partial E_\alpha \partial u_{s\beta}(\mathbf{q}=0)} = \Omega \frac{\partial P_\alpha}{\partial u_{s\beta}(\mathbf{q}=0)} = \frac{\partial F_{s,\beta}}{\partial E_\alpha}. \tag{2.34}$$

The previous findings of the phenomenological theory are supported by the rigorous and exact microscopic theory developed by Pick, Cohen, and Martin [26], where the crystal is treated as a dielectric medium in which the dynamical properties are described by the dielectric response of the electronic charge density to the ionic displacements in the lattice: the dynamical

effective charges, introduced in this Section as phenomenological quantities, receive in this way their formal legitimation and their first-principles definition.

Also for the case of molecules, a quantity similar to Z^* (the so-called *atomic polar tensor*) was introduced [27] in order to interpret the infrared intensities, which is defined as the change of the total dipole moment of the molecule with respect to an atomic displacement.

2.4 Some remarks about the dynamical charge concept

Essentially, the dynamical charge measures the current flowing along the bonds when the bond-lengths are varied; for this reason, it can in general differ from zero even if there is no ionicity: the deformation dipole moment can give rise to nonzero effective charges even on elemental crystals. This effect must be symmetry allowed, and an important theorem due to Zallen [28] shows that this is the case if the crystal structure has at least three atoms in the primitive unit cell. One of the simplest systems where this effect is observable is selenium, whose elemental crystal exhibits infrared active modes, and the effective charges are rather large [29].

By considering longitudinal and transverse optic waves in the absence of an external field, one can accordingly define longitudinal (Callen) and transverse (Born) dynamical charges; the relationship between the two is [23]:

$$Z_s^{*(L)} = \frac{Z_s^{*(T)}}{\varepsilon_\infty}. \tag{2.35}$$

The difference between $Z_s^{*(T)}$ and $Z_s^{*(L)}$ can be easily expressed also in terms of the boundary conditions assumed in the solution of Poisson equation: in a crystalline solid one actually refers to the rigid displacement of a whole sublattice and the induced charge is a periodic function. The transverse charge $Z_s^{*(T)}$ is then defined via the macroscopic polarization $\Delta\mathbf{P}$ linearly induced in the solid by a rigid displacement of the s sublattice by an amount $\mathbf{u}_s$ while the field is kept vanishing; the longitudinal charge $Z_s^{*(L)}$ is analogously defined, but the sublattice

displacement is performed in a depolarizing field $\Delta \mathbf{E} = -4\pi\Delta \mathbf{P}$.

It is worth noting that $Z_s^{*(T)}$ and $Z_s^{*(L)}$ represent the more convenient and natural choice to express the concept of dynamical charge and to discuss the infrared absorption activity of the crystal, because they are defined in terms of the crystal response to the macroscopic electric field, which is the experimentally accessible physical quantity of the problem.

Actually, it is possible to build up many alternative definitions of dynamical charge, by parameterizing the theory in terms of model-dependent, not directly measurable quantities. As an example, we recall here the Szigeti charge e_s^*, introduced by Szigeti [30] in order to investigate the effects of the electronic-charge deformation on the lattice vibrations. In this approach it is in particular assumed that in addition to the elastic restoring force an effective electric field $\mathbf{E}_{\text{eff}}$ acts on each ion, exerting a force on it equal to $e_s^*\mathbf{E}_{\text{eff}}$. However, the problem of evaluating the effective field acting on a ion is a complicated one, which depends on the details of the geometrical distribution of the ionic dipoles in the crystal and on the extent to which the electrons are localized about their nuclei; e_s^* turns out to express a model-dependent quantity which relates to Z^* through an assumption on the form of the Lorentz effective field associated to the transverse optic modes. In the particular case of a local spherical symmetry one gets:

$$Z_s^{*(L)} = [(\epsilon_\infty + 2)/3\epsilon_\infty] e_s^* \tag{2.36}$$

The Szigeti charge describes therefore the response of the system to an effective field, which cannot be obtained directly from experiments, and is hence a quantity defined with reference to a specific model.

On the other hand, in the case of an isolated molecule the solution of the Poisson equation is unique and it is not necessary to specify the boundary conditions; for this reason the definition of Z^* is straightforward and unique [31] in the molecular case (where the charge perturbation induced by a ionic displacement is localized).

The relation between $Z_s^{*(\mathrm{T})}$ (or $Z_s^{*(\mathrm{L})}$) and the dynamical properties of the crystal can be easily realized in the simple case of a cubic polar crystal: the longitudinal optical phonon frequency is higher than the transverse one due to the longitudinal electric field present in that case and the relationship between the two is given by:

$$\omega_{LO} = \sqrt{\omega_{TO}^2 + \frac{4\pi e^2}{\Omega} \frac{Z_s^{*(\mathrm{T})^2}}{\mu \varepsilon_\infty}}, \tag{2.37}$$

where μ is the reduced mass. Eq. (2.37) explicitly shows the connection between the dynamical charges and the splitting of longitudinal and transverse optic modes at the Γ point, and can be recast [32] in the Lyddane-Sachs-Teller form:

$$\frac{\varepsilon_0}{\varepsilon_\infty} = \frac{\omega_{LO}^2}{\omega_{TO}^2} \tag{2.38}$$

where ε_0 and ε_∞ are, respectively, the static and optical dielectric constants of the crystal.

2.4.1 Acoustic sum rule

In a neutral molecule the sum over all the atoms of the dynamical charges must vanish [33], since a rigid translation of the molecule as a whole induces no dipole:

$$\sum_s Z_s^* = 0, \tag{2.39}$$

where the index s runs over the atoms within the molecule. The analogous relationship holds in the case of a crystalline dielectric and is spelled out in an equally simple manner: the dynamical charges sum to zero over the crystal cell. In the latter case, the relationship expressed by Eq. (2.39) goes under the name of *acoustic sum rule* (ASR); it was actually derived [26] from the condition of vanishing frequency for an acoustic mode in the limit $\mathbf{q} \to 0$, which is due to the translational invariance of the crystal (as discussed in Section 2.3.2). The ASR imposes a requirement on a specific *collective* displacement of the different atoms.

The sum rule is usually stated for the transverse charges [26], but equivalently holds for the longitudinal charges as well. As will be explained in detail in Chapter 4, we aim at dealing

with molecules, surfaces, and bulk solids all on the same footing; the *longitudinal* charge is then the quantity of choice in the crystalline case. Suppose in fact we displace by an amount $\mathbf{u}_s$ only a single ion of species s in an infinite—and otherwise unperturbed—crystal: the induced charge is localized [34], and it is easy to prove that its dipole $\mathbf{d}$ is (to linear order):

$$d_\alpha = \sum_\beta Z^{*(\mathrm{L})}_{s,\alpha\beta} u_{s,\beta}. \tag{2.40}$$

This can be regarded as an equivalent definition of the longitudinal charge. An identical definition as Eq. (2.40) holds for the (unique) dynamical charge in the molecular case. We therefore drop the subscript (L) in the following of this thesis, and we assume Eq. (2.40) as a uniform definition of the dynamical charge for a molecule, a bulk solid, and a surface.

We will discuss in subsequent chapters the dynamical neutrality at surfaces and interfaces, always making reference to the longitudinal dynamical charge $Z^{*(\mathrm{L})}$.

3 Computational tools

3.1 Density-Functional Theory

In order to describe the actual electron distribution at the interface, one has to perform calculations in which the electrons are allowed to readjust to the specific environment determined by the different ionic species which are present.

In the spirit of the Born Oppenheimer (or *adiabatic*) approximation [23], the electronic degrees of freedom can be decoupled from the ionic ones. Ionic cores behave as classical particles and, given a certain ionic configuration, their potential energy is provided by the ground state energy of the electrons, treated as a quantum many-body system. This approach relies on the large mass difference between ions and electrons, which implies that electrons remain very close to their instantaneous ground state configuration while the ions move. For this reason, the knowledge of the electronic ground state allows a complete description of the properties of the system.

Density-Functional Theory (DFT) provides a fundamental theoretical framework [35] to calculate from *first principles* the ground state properties of a quantum many-electron system without solving the many-body Schrödinger equation, which would be an impossible task due to the large number of the degrees of freedom involved in the calculation. A central result [36] of DFT is that the ground state energy of a system of interacting electrons can be obtained,

at least in principle, by minimizing the functional:

$$E[n] = F[n] + \int V(\mathbf{r})n(\mathbf{r})d\mathbf{r}, \tag{3.1}$$

where $F[n]$ is a universal function of the electronic density (*i.e.* it is independent on $V(\mathbf{r})$), and $V(\mathbf{r})$ is the external potential acting on the electrons, that in crystals corresponds to the potential generated by the ionic cores. The form of $F[n]$ is in general unknown; in order to apply this theory to actual calculations, Kohn and Sham proposed to recast this functional by separating out of it a term, $T_0[n]$, defined as the kinetic energy of a non interacting electron system of density $n(\mathbf{r})$, and the Hartree term, which is the classical electrostatic interaction between the electrons:

$$F[n(\mathbf{r})] = T_0[n(\mathbf{r})] + \frac{1}{2}\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' + E_{xc}[n(\mathbf{r})]. \tag{3.2}$$

All our ignorance is now confined to the exchange-correlation energy, $E_{xc}[n(\mathbf{r})]$, which is *defined* by Eq. (3.2) as the difference between the unknown functional $F[n]$ and the remaining terms in its right hand side. Minimizing the total energy functional $E[n]$ with the constraint $\int n(\mathbf{r})d\mathbf{r} = N$, we obtain a set of self-consistent equations [37]:

$$\underbrace{\left[-\frac{\nabla^2}{2} + V_{SCF}(\mathbf{r})\right]}_{H_{SCF}} \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}) \tag{3.3}$$

$$V_{SCF} = V(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + v_{xc}(\mathbf{r}), \tag{3.4}$$

$$n(\mathbf{r}) = \sum_i |\psi_i(\mathbf{r})|^2 \, \theta(\epsilon_i - \epsilon_F). \tag{3.5}$$

These are the well-known Kohn-Sham (KS) equations (expressed in a. u.), where the Fermi energy ϵ_F is defined by the constraint on the number of electrons in the system and $v_{xc}(\mathbf{r}) = \delta E_{xc}[n]/\delta n(\mathbf{r})$ is the exchange-correlation potential. Due to this unknown term, the KS equations are of no practical use, unless an approximation for v_{xc} is specified. The most used

approximation in practical calculations is the so called *local density approximation* (LDA), in which the exchange-correlation energy is taken as a local function of the density itself:

$$E_{xc}[n] = \int n(\mathbf{r})\epsilon_{xc}(n(\mathbf{r}))d\mathbf{r}, \tag{3.6}$$

where $\epsilon_{xc}(n(\mathbf{r}))$ is the exchange-correlation energy per particle of the homogeneous electron gas at a density equal to the local density $n(\mathbf{r})$, which has been computed once for all with quantum MonteCarlo techniques [38]. In this way the potential v_{xc} is given by:

$$v_{xc}(\mathbf{r}) = \mu_{xc}(n(\mathbf{r})) = \frac{d}{dn}[n\epsilon_{xc}(n)]. \tag{3.7}$$

From the iterative solution of the KS equations one can get the electronic charge density distribution, while the ground state energy of the system reads:

$$\begin{aligned} E^{tot} &= -\frac{1}{2}\sum_i \theta(\epsilon_i - \epsilon_F)\int \psi_i^*(\mathbf{r})\nabla^2\psi_i(\mathbf{r})d\mathbf{r} + \int V(\mathbf{r})n(\mathbf{r})d\mathbf{r} \\ &+ \frac{1}{2}\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' + \int n(\mathbf{r})\epsilon_{xc}(n(\mathbf{r}))d\mathbf{r} \\ &+ \sum_{\mathbf{R},s,s'}{}' \frac{Z_s Z_s'}{|\mathbf{R}+\tau_s-\tau_{s'}|}\,. \end{aligned} \tag{3.8}$$

3.1.1 The plane-wave pseudopotential method

In the study of infinite solids it is necessary to exploit the crystal symmetry to circumvent the problem of the infinite number of degrees of freedom. The Bloch theorem states that in a periodic crystal the electronic wavefunctions are given by the product of a function having the same periodicity times a phase factor. This property allows to map the problem onto a single cell, repeated in space through the so-called periodic boundary conditions, having a small number of atoms and thus, of electrons.

In order to solve in practice the KS equations, one can write the KS orbitals in terms of a suitable finite basis set. A standard choice is that of using plane waves (PW), which have the

great advantage of being translationally invariant:

$$\psi_i(\mathbf{r}) = \psi_{n,\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n,\mathbf{k}+\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}}, \tag{3.9}$$

where k belongs to the first Brillouin Zone (BZ) of the crystal, $\mathbf{G}$ is a reciprocal–lattice vector, and n is a band index. The use of Bloch functions would then require the knowledge of the electronic wavefunctions at every point in the first BZ of the reciprocal space. Nevertheless, efficient k-sampling methods have been proposed by different authors [39–41], where very restricted sets of "special points" are taken as representative of the entire zone.

The electronic density is then calculated self-consistently through Eq. (3.5). For insulating systems at zero temperature the occupation of the Bloch states can be fixed to two or zero electrons per spin-degenerate state. In metals a fractional occupation of states at the sampling k-points close to the Fermi energy is commonly introduced to mimic the properties of the Fermi surface. The calculations reported in this thesis employ the smearing technique [42] to deal with BZ integration in the presence of a Fermi surface, as described in Appendix B.

The dimension of the PW basis set is determined by fixing the kinetic energy cutoff, E_{cut}, through the condition:

$$|\mathbf{k} + \mathbf{G}|^2 \leq E_{cut}. \tag{3.10}$$

The choice of a PW basis has the advantage that the matrix elements of the Hamiltonian in Eq. (3.3) are particularly simple and that the accuracy of the expansion can be easily checked and systematically improved by increasing the value of E_{cut}; furthermore, PW's are independent of the structure of the crystal. In order to obtain an accurate description of the system with a reasonably small number of PW's the core electrons are frozen in the atomic configuration around the nuclei, and only the chemically active valence electrons are treated explicitly. To this end, a smooth angular-momentum-dependent *pseudopotential* is then introduced to describe the interaction between valence electrons and ionic cores (nuclei + core electrons). There are many different schemes to generate ionic pseudopotentials from

first principles (see for example Refs. [43, 44]). Basically, all of them satisfy the following requirements: (*i*) the lowest pseudo-energy levels are equal to the valence all-electron energies; (*ii*) each pseudo-wave function coincides with the corresponding all-electron one outside a properly chosen core radius; (*iii*) as a consequence of (*ii*), the real and pseudo charge inside the core radius agree for each valence state. This last condition is called *norm conservation*, and ensures the transferability of the pseudopotential to different chemical environments [45]. The accuracy of the results obtained with *norm-conserving* pseudopotentials is comparable with those from all-electron calculations [46, 47]. The price to be payed for norm conservation is nonlocality, which shows up in the explicit l-dependence of the radial ionic pseudopotential. A computationally convenient form for the pseudopotential has been introduced by Kleinman and Bylander [48], who pointed out that a significant reduction of the numerical effort can be achieved if the nonlocality of the potential is not restricted to the angular part, but if also the radial potential is replaced by a suitable non local separable operator. The Kleinman-Bylander form must be used with some caution, because in some cases it can lead to a wrong description of the chemical properties of the system, due to the appareance of unphysical states in the energy spectrum of the isolated atom. These spurious states, known in the literature as *ghosts* [49], may occur because—due to the non locality of the radial potential—it is no longer guaranteed that the radial wave-functions can be ordered in terms of the increasing number of their nodes. In particular, the nodeless radial (pseudo)functions, which are assumed to describe the relevant atomic states for different angular numbers, are not necessarily the ground states of the pseudo–Hamiltonian from which they are obtained.

3.2 Density Functional Perturbation Theory

The Density Functional Perturbation Theory (DFPT) provides a very efficient approach to the perturbation theory in the DFT framework: it basically consists [50, 51] in a self-consistent

scheme which directly provides the derivatives of the KS wave-functions appearing in Eqs. (3.3–3.5) with respect to an external parameter upon which the potential depends. In particular, DFPT provides a method to obtain the variation of the ground-state energy due to an external perturbation from the knowledge of the induced electronic density variation within the linear response theory. This allows e.g. to obtain dynamical matrices at arbitrary wavevectors with a computational effort comparable to that of a self-consistent calculation for the unperturbed system, as well as several response functions.

3.2.1 Linear response in crystals

Suppose that the *bare* external potential acting on the electrons becomes $V_{\boldsymbol{\lambda}}(\mathbf{r})$, that is a continuous function of some parameters $\boldsymbol{\lambda} \equiv \{\lambda_i\}$, so that $V_{\boldsymbol{\lambda}=0}$ corresponds to the unperturbed external potential of the reference system. The ground-state energy of the perturbed system is:

$$E_{\boldsymbol{\lambda}} = \langle \Psi_{\boldsymbol{\lambda}} | H_{\boldsymbol{\lambda}} | \Psi_{\boldsymbol{\lambda}} \rangle. \tag{3.11}$$

The Hellmann-Feynman theorem [52] states that the generalized force associated with the variation of the external parameters $\boldsymbol{\lambda}$ is given by the ground-state expectation value of the derivative of $V_{\boldsymbol{\lambda}}(\mathbf{r})$:

$$\frac{\partial E_{\boldsymbol{\lambda}}}{\partial \lambda_i} = \int n_{\boldsymbol{\lambda}}(\mathbf{r}) \frac{\partial V_{\boldsymbol{\lambda}}(\mathbf{r})}{\partial \lambda_i} d\mathbf{r} \tag{3.12}$$

where $E_{\boldsymbol{\lambda}}$ is the electron ground-state energy relative to given values of the $\boldsymbol{\lambda}$ parameters, and $n_{\boldsymbol{\lambda}}$ is the corresponding electron density distribution. Total energy variations are obtained from Eq. (3.12) by integration. In order to have energy variations correct up to second order in $\boldsymbol{\lambda}$, it is necessary that the right hand side of Eq. (3.12) is correct to linear order:

$$\begin{aligned} \frac{\partial E_{\boldsymbol{\lambda}}}{\partial \lambda_i} = \int \Bigg(n_0(\mathbf{r}) \frac{\partial V_{\boldsymbol{\lambda}}(\mathbf{r})}{\partial \lambda_i} + \sum_j \lambda_j \frac{\partial n_{\boldsymbol{\lambda}}(\mathbf{r})}{\partial \lambda_j} \frac{\partial V_{\boldsymbol{\lambda}}(\mathbf{r})}{\partial \lambda_i} + \\ + n_0(\mathbf{r}) \sum_j \lambda_j \frac{\partial^2 V_{\boldsymbol{\lambda}}(\mathbf{r})}{\partial \lambda_i \lambda_j} \Bigg) d\mathbf{r} + \mathcal{O}(\boldsymbol{\lambda}^2), \end{aligned} \tag{3.13}$$

all derivatives being calculated at $\boldsymbol{\lambda} = 0$. Integration of Eq. (3.13) gives:

$$E_{\boldsymbol{\lambda}} = E_0 + \sum_i \lambda_i \int n_0(\mathbf{r}) \frac{\partial V_{\boldsymbol{\lambda}}(\mathbf{r})}{\partial \lambda_i} d\mathbf{r} + \\ + \frac{1}{2} \sum_{ij} \lambda_i \lambda_j \int \left(\frac{\partial n_{\boldsymbol{\lambda}}(\mathbf{r})}{\partial \lambda_j} \frac{\partial V_{\boldsymbol{\lambda}}(\mathbf{r})}{\partial \lambda_i} + n_0(\mathbf{r}) \frac{\partial^2 V_{\boldsymbol{\lambda}}(\mathbf{r})}{\partial \lambda_i \partial \lambda_j} \right) d\mathbf{r} \, . \tag{3.14}$$

The previous expression indicates that the linear response of the system to an external perturbation is sufficient to evaluate the second order correction to the total energy (indeed even the third order one, due to the so-called "*2n + 1*" theorem [53]). So, once the nature of the perturbative parameters $\boldsymbol{\lambda}$ is specified and the first-order variation of the electronic density due to the perturbation is known, the calculation of the second derivatives of the energy E formally allows, according to Eq. (2.25), to access either the matrix of the force constants, or the dielectric tensor, or the dynamical effective charge, or any other response function. This linear response of the system is calculated by linearizing the set of Eqs. (3.3–3.5). The first order variations of the wave-functions are the solutions of the equation:

$$\left[-\frac{\nabla^2}{2} + V_{SCF}(\mathbf{r}) - \epsilon_i \right] \Delta\psi_i(\mathbf{r}) = - \left[\Delta V_{SCF}(\mathbf{r}) - \langle \psi_i | \Delta V_{SCF} | \psi_i \rangle \right] \psi_i(\mathbf{r}) \tag{3.15}$$

The self-consistent variation of the potential ΔV_{SCF}, in turn, must account for the variation of the external potential and for the effects due to the variation of the charge density; it is then written as:

$$\Delta V_{SCF}(\mathbf{r}) = \Delta V(\mathbf{r}) + \int \frac{\Delta n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{d\mu_{xc}(n(\mathbf{r}))}{dn} \Delta n(\mathbf{r}). \tag{3.16}$$

The variation of the self-consistent potential depends therefore upon the variation of the valence charge density which in turn depends on the variation of the Kohn-Sham orbitals:

$$\Delta n(\mathbf{r}) = 2 \sum_i \psi_i^*(\mathbf{r}) \Delta\psi_i(\mathbf{r}) \theta(\epsilon_i - \epsilon_F). \tag{3.17}$$

3.2.2 Implementation of DFPT for a periodic solid

The advantage of the above formulation appears when it is applied to a periodic solid. This direct approach allows the computation of the charge–density variation due to a perturbation of any wavelength with the same workload as required for a lattice–periodic perturbation. The variation of the external potential can be written in the form:

$$\Delta V(\mathbf{r}) = e^{i\mathbf{qr}} \Delta\tilde{V}(\mathbf{q},\mathbf{r}), \tag{3.18}$$

where $\Delta\tilde{V}(\mathbf{q},\mathbf{r})$ has the same periodicity of the direct lattice, and

$$\Delta\tilde{V}^*(-\mathbf{q},\mathbf{r}) = \Delta\tilde{V}(\mathbf{q},\mathbf{r}). \tag{3.19}$$

This is the typical form of the general perturbation due to a phonon of $\mathbf{q}$ wave-vector, but it also includes the limit $\mathbf{q} \to 0$, which provides a description both of the effects of a macroscopic perturbation on the system and of its dynamical properties at the Brillouin zone center. The perturbation couples states at $\mathbf{k}$ with states at $\mathbf{k}+\mathbf{q}$; to the first perturbative order the variation of the KS potential and of the charge density have the same periodicity of the external potential. Rewriting Eq. (3.15) on the basis of the unperturbed wave-functions, we find that the solution $\Delta\psi_n(\mathbf{k}+\mathbf{q},\mathbf{r})$ (which is the variation of the orbital $\psi_n(\mathbf{k},\mathbf{r})$) has components on the $\psi_n(\mathbf{k}+\mathbf{q},\mathbf{r})$ wave-functions, and where no degeneracy occurs one can formally write:

$$\Delta\psi_n(\mathbf{k}+\mathbf{q},\mathbf{r}) = \sum_{m\neq n} \frac{\int_V \psi_m^*(\mathbf{k}+\mathbf{q},\mathbf{r})\Delta V_{SCF}(\mathbf{r})\psi_n(\mathbf{k},\mathbf{r})d\mathbf{r}}{\epsilon_n(\mathbf{k}) - \epsilon_m(\mathbf{k}+\mathbf{q})} \psi_m(\mathbf{k}+\mathbf{q},\mathbf{r}). \tag{3.20}$$

In the computation of the total charge density we need only the projection of the first order wave-function on the conduction-band manifold. In the case of a nonmetallic solid, where there is a finite energy gap, this projection is well defined because all the denominators appearing in Eq. (3.20) are nonzero if n is a valence state. This fact can be demonstrated inserting Eq. (3.20) in Eq. (3.17), which in terms of Bloch wave-functions reads:

$$\Delta n(\mathbf{r}) = 2\sum_{\mathbf{k}}\sum_{n=1}^{N_b} \left[\psi_n^*(\mathbf{k},\mathbf{r})\Delta\psi_n(\mathbf{k}+\mathbf{q},\mathbf{r}) + \Delta\psi_n^*(\mathbf{k}-\mathbf{q},\mathbf{r})\psi_n(\mathbf{k},\mathbf{r})\right] . \tag{3.21}$$

We finally obtain the relationship:

$$\Delta n(\mathbf{r}) = 4\sum_{\mathbf{k}}\sum_{n=1}^{N_b} \psi_n^*(\mathbf{k},\mathbf{r}) P_c(\mathbf{k}+\mathbf{q}) G_0(\mathbf{k}+\mathbf{q},\epsilon_n(\mathbf{k})) \times \\ P_c(\mathbf{k}+\mathbf{q})\Delta V_{SCF}(\mathbf{r})\psi_n(\mathbf{k},\mathbf{r}), \tag{3.22}$$

where

$$G_0(\mathbf{k},\epsilon_n(\mathbf{k}_1)) = [\epsilon_n(\mathbf{k}_1) - H_{SCF}(\mathbf{k})]^{-1} \tag{3.23}$$

is the Green function of the unperturbed system, while $P_c(\mathbf{k})$ is the projector over the conduction-state manifold:

$$P_c(\mathbf{k}) = 1 - \sum_{i=n}^{N_b} |\psi_n(\mathbf{k})\rangle\langle\psi_i(\mathbf{k})|. \tag{3.24}$$

According to Eq. (3.22), Δn has nonvanishing Fourier components only in correspondence of wave-vectors $\mathbf{q}+\mathbf{G}$. The expansion in Fourier components yields:

$$\Delta\tilde{n}(\mathbf{q}+\mathbf{G}) = \frac{4}{V}\int_V \psi_n^*(\mathbf{k},\mathbf{r}) e^{i(\mathbf{q}+\mathbf{G})\mathbf{r}} P_c(\mathbf{k}+\mathbf{q})\Delta\psi_n(\mathbf{k}+\mathbf{q},\mathbf{r})d\mathbf{r}, \tag{3.25}$$

where $P_c(\mathbf{k}+\mathbf{q})\Delta\psi_n(\mathbf{k}+\mathbf{q},\mathbf{r})$ is the solution of Eq. (3.15), which is in practice expanded in a PW-representation.

3.2.3 Calculation of the dynamical charges within DFPT

As already pointed out in Section 2.3.3, the dynamical charge tensor $Z^*_{s,\alpha\beta}$ is defined as the second derivative of the energy with respect to the displacement along the direction β of the atoms belonging to the sublattice s and to the macroscopic electric field E_α. As the polarization is a first derivative of the energy with respect to a ionic displacement, $Z^*_{s,\alpha\beta}$ can be calculated as the coefficient of proportionality relating, to linear order and under the condition of zero electric field, the macroscopic polarization per unit cell induced along the direction α, and the displacement along the direction β of the atoms belonging to the sublattice s [see also Eq. (2.34)]:

$$Z^*_{s,\alpha\beta} = \Omega\frac{\partial P_\alpha}{\partial u_{s,\beta}}. \tag{3.26}$$

The same reasoning used in the derivation of the variation of the charge density for a periodic solid [see Eq. (3.22)] can be applied in this case and we arrive to the result:

$$\frac{\partial P_\alpha}{\partial u_{s,\beta}} = \frac{4}{N\Omega}\sum_{\mathbf{k}}\sum_{i=1}^{N_b}\int \psi_i^*(\mathbf{k},\mathbf{r}) r_\alpha P_c(\mathbf{k}+\mathbf{q}) G_0(\mathbf{k}+\mathbf{q},\epsilon_i(\mathbf{k})) P_c(\mathbf{k}+\mathbf{q}) \Delta V_{SCF}(\mathbf{r})\psi_i(\mathbf{k},\mathbf{r}) d\mathbf{r}. \tag{3.27}$$

This formula shows that the derivative of the polarization is defined in terms of the projection on the conduction-state manifold of the variation of the KS wavefunctions, as determined by the selfconsistent iterative procedure previously described. It requires an expression for the variation of the KS potential due to an atomic displacement of s sublattice in the β direction.

The matrix elements of the $\mathbf{r}$ operator appearing in Eq. (3.27) are boundary dependent in any finite sample and therefore ill-defined in an infinite solid. However it happens that in the first order derivative of the polarization only off-diagonal elements of the $\mathbf{r}$ operator are involved, and they can be readily transformed into a boundary-insensitive form using the relationship [54, 55]

$$\langle\psi_n(\mathbf{k})|\mathbf{r}|\psi_m(\mathbf{k})\rangle = \frac{\langle\psi_n(\mathbf{k})|[H_{SCF},\mathbf{r}]|\psi_m(\mathbf{k})\rangle}{\epsilon_n(\mathbf{k})-\epsilon_m(\mathbf{k})}, \tag{3.28}$$

where:

$$[H_{SCF},\mathbf{r}] = \frac{-i\hbar\mathbf{p}}{m} + i[V_{ext},\mathbf{r}], \tag{3.29}$$

and i and j states are non degenerate eigenstates of the Hamiltonian. Strictly speaking, the left hand side of Eq. (3.28) is defined only for a finite system, and Eq. (3.28) should be considered as a definition of the off-diagonal matrix elements of $\mathbf{r}$ for an infinite system. In practice one defines the functions [51]:

$$\phi_{n\alpha}(\mathbf{k},\mathbf{r}) = P_c(\mathbf{k}) r_\alpha \psi_n(\mathbf{k},\mathbf{r}) = iP_c(\mathbf{k}) G_0(\mathbf{k},\epsilon_n(\mathbf{k})) P_c(\mathbf{k}) [H_{SCF}, r_\alpha]\psi_n(\mathbf{k},\mathbf{r})\,, \tag{3.30}$$

which are computed only once because they are independent of the variation of the charge density, and then we recast Eq. (3.26) in the form:

$$Z^*_{s,\alpha\beta} = \frac{4}{N}\sum_{\mathbf{k}}\sum_{n=1}^{N_b}\int_V \phi^*_{n\alpha}(\mathbf{k},\mathbf{r})\Delta\psi_{ns,\beta}(\mathbf{k},\mathbf{r}) d\mathbf{r}\,. \tag{3.31}$$

This equation gives the electronic contribution to the dynamical effective charge, because we considered so far the response of the system in terms of the electronic charge-density variation. The ionic contribution to the polarization (and hence to the effective charge) is trivially given by the ionic pseudo-charge Z_s, which has to be added to expression (3.31) in order to recover the total dynamical charge.

3.3 Direct method for the calculation of the dynamical charges

In the first 80's, Martin and Kunc [56–58] developed a practical method to obtain directly from self-consistent calculations the dynamical effective charges, as the first moment of the electronic charge density variation induced by the displacement of an atom. In particular, they found out that it is possible to deal with localized properties of the induced charge density and, moreover, to treat the electron-electron interaction in a self-consistent way.

Let us define the induced charge density (ionic plus electronic) caused by displacement of the atom at $\mathbf{R}_{ls}$ in the cell l by:

$$F_{s,\beta}(\mathbf{r} - \mathbf{R}_{ls}) = \frac{\partial n(\mathbf{r})}{\partial R_{ls,\beta}} \ . \tag{3.32}$$

The total induced charge density δn is the sum of the contributions caused by the displacement of each atom; therefore the $\mathbf{k}$ Fourier component of the longitudinal polarization is given by:

$$- i\mathbf{k} \cdot \mathbf{P}_L = \frac{1}{V} \sum_{l\beta} \int \exp\left(-i\mathbf{k} \cdot \mathbf{r}\right) F_{s,\beta}(\mathbf{r} - \mathbf{R}_{ls}) u_{ls,\beta} d\mathbf{r} \ , \tag{3.33}$$

where $V^{-1} \int d\mathbf{r}$ denotes the average over the entire crystal. For a long-wavelength optic mode, with $u_{ls,\beta} = u_{s,\beta} \exp(i\mathbf{k} \cdot \mathbf{R}_{ls})$, the integral in (3.33) is independent of cell l, and to lowest order in $\mathbf{k}$ may be written as

$$\mathbf{k} \cdot \mathbf{P}_L = \frac{1}{\Omega} \sum_{\beta} \int (\mathbf{k} \cdot \mathbf{r}) F_{s,\beta}(\mathbf{r}) u_{s,\beta} d\mathbf{r} \ , \tag{3.34}$$

where Ω is the volume of the unit cell. So, by recalling the definition of Z^* [see Eq. (2.34)], one gets:

$$Z^*_{s,\alpha\beta} = \int r_\alpha F_{s,\beta}(\mathbf{r}) d\mathbf{r} \, . \tag{3.35}$$

As a first issue, it is interesting to show that the integral which defines Z^* is convergent. To this aim, let us define the coordinate r_α to be z (i.e. α=3), and the integral of $F_{s,\beta}(\mathbf{r}) = \frac{\partial n(\mathbf{r})}{\partial \mathbf{u}_{s,\beta}}$ over the plane perpendicular to $\hat{z}$ to be $\overline{F}_{s,\beta}(z)$, which has dimensions of charge per unit area. At large z the functions F and $\overline{F}$ may be considered in the continuum limit of a dipole at the origin screened by the local macroscopic dielectric tensor $\epsilon_{\alpha\beta}$. The dipole field has the simple property that the integral $\overline{F}(z)$ vanishes for any $z \neq 0$ since it is the average over a plane that does not pass through the dipole. Thus, despite the fact that $F(\mathbf{r})$ is long ranged, $\overline{F}(z)$ is nonzero only for a small range of z, of the order of atomic dimensions, where the local continuum arguments do not apply. Therefore, the macroscopic effective charge can be calculated in terms of well-defined integrals over the function $F_{s,\beta}(\mathbf{r})$.

Finally, it is necessary to establish a convenient method for the calculation of $\overline{F}_{s,\beta}(z)$. Define $P(0)$ to be the plane perpendicular to $\hat{z}$ and passing through atoms of type s. If z is a high-symmetry direction, the atoms in the plane $P(0)$ form a two-dimensional periodic array which can be labeled by a cell index m. Now, $\overline{F}_{s,\beta}(z)$ is the integral of $F_{s,\beta}(\mathbf{r} - \mathbf{R}_{ms})$ over infinite planes $P(z)$ parallel to $P(0)$, which is independent of the cell index m. Let us define:

$$H_{s,\beta}(\mathbf{r}) = \sum_m F_{s,\beta}(\mathbf{r} - \mathbf{R}_{ms}) \, , \tag{3.36}$$

which is a function periodic in two dimensions and results from equal displacements of each atom (ms) in the plane $P(0)$. It can be represented by its Fourier components $H(\mathbf{G}, z)$, where $\mathbf{G}$'s are the reciprocal vectors of the two-dimensional lattice:

$$H(\mathbf{G}, z) = \int H(\mathbf{r}) \exp(i\mathbf{G} \cdot \mathbf{r}) dx dy \, . \tag{3.37}$$

It is easy to see that $\overline{F}_{s,\beta}(z)$ is exactly the same as the zero Fourier component of H, i.e. the

average $H(0,z)$. Eventually we get:

$$Z^*_{s,3\beta} = \int z H_{s,\beta}(0,z) dz. \tag{3.38}$$

One can determine $H(\mathbf{G},z)$ by numerically calculating the charge density of the crystal with a plane of atoms displaced. The most practical and successful method for calculating the charge distribution of a two-dimensional periodic system is the supercell method, which corresponds to build up a lattice of displaced planes with periodicity in the z direction large enough so that the potential and charge density converge to their bulk limits in the regions between the disturbances. The calculation of the charge density variation, and consequently of $H_{s,\beta}(\mathbf{G},z)$ is exactly the same as the calculation of the self-consistent charge density at an interface or at a surface; the determination of the first moment of $H_{s,\beta}(0,z)$ is exactly equivalent to the calculation of the dipole contribution to the work function at an interface or a surface. So, from elementary electrostatics the change in the average potentials from one side of a plane of displaced atoms to the other is given by:

$$\Delta V = 4\pi\sigma = (4\pi/A_0)Z^* u \ , \tag{3.39}$$

where σ is the dipole moment per unit area, A_0 is the area of the plane per atom, and $Z^* u$ is the dipole moment per atom in the linear approximation.

4 The dynamical-charge neutrality condition for polar surfaces

As reviewed in the second Chapter, the dynamical charge Z^*_s is a Cartesian tensor with the point symmetry of the ionic site, whose components (α, β) measure the dipole linearly induced (in the α direction) by a unit displacement of the ion s (in the β direction) at zero electric field or, equivalently, the force (in the β direction) linearly induced on the given ion s by a unit electric field (in the α direction) at zero ionic displacement [see Eq. (2.34)].

We already pointed out in Section 2.4.1 that the Z^*_s's obey a *dynamical neutrality* condition for both molecules and periodic solids, which enforces their sum to vanish (over the whole finite system, or over the crystal cell, respectively). The two cases of molecules [31, 59] and of bulk solids [26, 60] have received previous attention in the literature, both as a matter of principle and as a subject of practical calculations. The case of a crystal surface has never been considered [61], so our aim is to extend this sum rule to the non trivial case of the surface of a semiinfinite solid.

4.1 Formulation of the dynamical-charge neutrality condition at a crystal surface

In a semiinfinite crystalline insulator it is in general possible to distinguish a surface region and a bulk region; the condition of vanishing macroscopic field both in the bulk of the solid and outside in the vacuum region amounts to require that no static charge is present at the surface. In these hypotheses, the average of the electrostatic potential in the bulk of the solid with respect to the vacuum level is a well defined quantity, which actually determines the work function of the given surface. We impose the physical requirement that the surface work function is not affected by a rigid translation of the semiinfinite solid as a whole: a necessary condition for this to occur is a novel sum rule for the surface dynamical charges, which we are going to formulate.

To this aim, let us consider a rigid displacement $\mathbf{u}_s$ of an ionic layer: this induces a dipole per unit area [see Eq. (2.40)], hence a potential drop across the layer whose value is given by:

$$\Delta\phi = \frac{4\pi}{A} \sum_{\beta} Z^*_{s,3\beta} u_{s,\beta}, \tag{4.1}$$

where the surface is normal to the z axis; the system has then a two-dimensional periodicity normally to z and A is the area of the corresponding unit cell.

A rigid translation of the semiinfinite crystal by an amount $\mathbf{u}$ induces a total potential drop which is, by linearity, the sum over s of the expressions in Eq. (4.1). As anticipated above, we explicitly require this quantity to vanish, hence a naïf expression for the constraint appears to be:

$$\sum_{s} Z^*_{s,3\beta} = 0 \quad (\text{any } \beta). \tag{4.2}$$

This formal expression for the dynamical charge neutrality of the surface cannot be used as such, given that the infinite sum in general does not converge: in the bulk region it oscillates periodically, owing to bulk dynamical-charge neutrality [see Eq. (4.6)].

Therefore the problem is to regularize the indeterminate sum in Eq. (4.2) by using the appropriate physical criterion. The panacea for this class of problems at large is the macroscopic average (see Appendix A), which yields, by construction, the correct electrostatic-potential average in the bulk region. Application of this criterion to the present case is straightforward and can be basically performed in three steps.

One first maps the problem into a simple electrostatic one by assigning a point charge of magnitude $Z^*_{s,3\beta}$ to each ion. Secondly, one builds up the planar average of this charge distribution, which takes the general form:

$$\overline{\rho^*}(z) = \frac{1}{A} \sum_s Z^*_{s,3\beta} \, \delta(z - z_s), \tag{4.3}$$

where z_s are the positions of the ionic planes along the z axis. Finally, one filters $\overline{\rho^*}(z)$ through the convolution [see also Eq. (A.5)]:

$$\overline{\overline{\rho^*}}(z) = \frac{1}{b} \int_{z-b/2}^{z+b/2} dz' \, \overline{\rho^*}(z'), \tag{4.4}$$

where b is the one-dimensional periodicity in the bulk region. The result is a piecewise constant function, vanishing both in the vacuum and in the bulk region, and whose integral in the surface region is trivial.

We are now able to state the novel sum rule by requiring the integral of the macroscopically averaged dynamical-charge distribution to vanish:

$$\int_{-\infty}^{+\infty} \overline{\overline{\rho^*}}(z) dz = 0. \tag{4.5}$$

It is important to realize that—at variance with the naïf expression of Eq. (4.2)—the surface sum rule (4.5) involves *both* the values of the dynamical charges *and* the coordinates of the planes z_s.

The nontrivial content of Eq. (4.5) will be made clear with two examples: two different kinds of simple polar surfaces of a cubic binary crystal will be taken into account, whose bulk

dynamical charges have the following form:

$$Z^{*\mathrm{bulk}}_{s,\alpha\beta} = (-1)^s |Z^*|\, \delta_{\alpha\beta}. \tag{4.6}$$

The $\beta \neq 3$ components of the sum rule expressed by Eq. (4.2) are correct, because of periodicity, in the same sense and for the same reason why this equation is correct in the bulk. We focus then on the $\beta = 3$ component, and we further limit ourselves to the cases where the ionic planes parallel to z contain either cations only, or anions only, alternately. This is the case for the (001) and (111) surfaces in the zincblende structure, which are going to be illustrated separately in the following.

4.1.1 The (001) case

We start with the (001) surface, where b is one half of the cubic lattice constant and the ionic planes are equally spaced by an amount $b/2$. We assume that the crystal lies in the positive-z half-space, so that the coordinate of the s-th plane are $(s-1)b/2$.

Equation (4.4) *forbids* ions of a given chemical species to have the same dynamical charge at the surface and in the bulk; actually, if it were so, the integral of the macroscopically averaged dynamical charge distribution would then differ from zero, as schematically shown in Fig. (4.1a), for the case where the surface is cation-terminated. As a first issue, one can consider the special case where only the outermost cation has a dynamical charge $Z^*_{1,33}$ different from the bulk value Z^*_{bulk}; this means that the surface region, monitored through the effective charge distribution, includes one atomic layer only. Figure (4.1b) illustrates that in such case $Z^*_{1,33}$ *must* be one half of Z^*_{bulk} in order to fulfill Eq. (4.4).

For the most general (001) surface, where several Z^*_s in the surface region may deviate from their bulk value, the piecewise constant function $\overline{\overline{\rho^*}}(z)$ vanishes from $-\infty$ to $-b/2$; it assumes the value $Z^*_{1,33}$ (apart for a constant factor) up to $z = 0$; then $\overline{\overline{\rho^*}}(z) = Z^*_{1,33} + Z^*_{2,33}$ up to $b/2$, and in general $Z^*_{s,33} + Z^*_{s+1,33}$ in the following intervals, until it vanishes again in the bulk

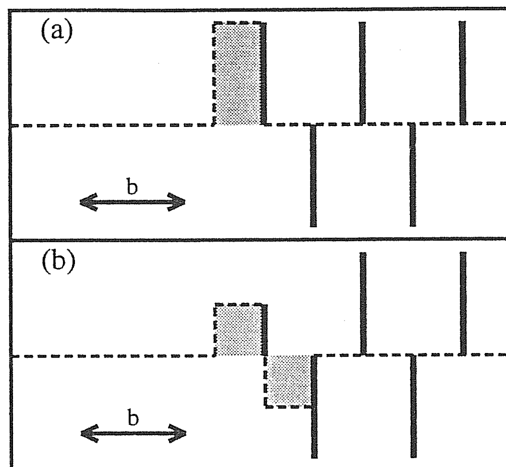

Figure 4.1: Case of (001) cation-terminated polar surface. Examples for the δ-like function $\overline{\rho^*}(z)$ (spikes, whose height measures the δ strength) and for the corresponding piecewise constant function $\overline{\overline{\rho^*}}(z)$ (dashed line), in arbitrary units. The shaded region corresponds to the integral of the macroscopic charge distribution, which is required to vanish. Cases (a) is then forbidden, while case (b) is not.

region, say for $s > N$. Since the the intervals are of equal length, the sum rule for the integral of $\overline{\overline{\rho^*}}(z)$ requires:

$$2\sum_{s=1}^{N} Z^*_{s,33} + Z^*_{N+1,33} = 0. \tag{4.7}$$

The outstanding information driven by this last equation is that the dynamical charges of the ions in the surface region sum up to *one half* of the bulk dynamical charge (with the appropriate sign). The neutrality relationship is precisely the same as by replacing the dynamical charges with static point-like charges and requiring the surface to be neutral.

In the previous derivation we assumed for the sake of simplicity an ideal truncated-bulk geometry. Nevertheless, it is easy to demonstrate that the new sum rule is not affected by any surface relaxation; if one actually computes the macroscopic average of the dynamical charges distribution making the hypothesis of some kind of relaxation of the outermost layers, the resulting form of the constraint given by Eq. (4.7) is unchanged.

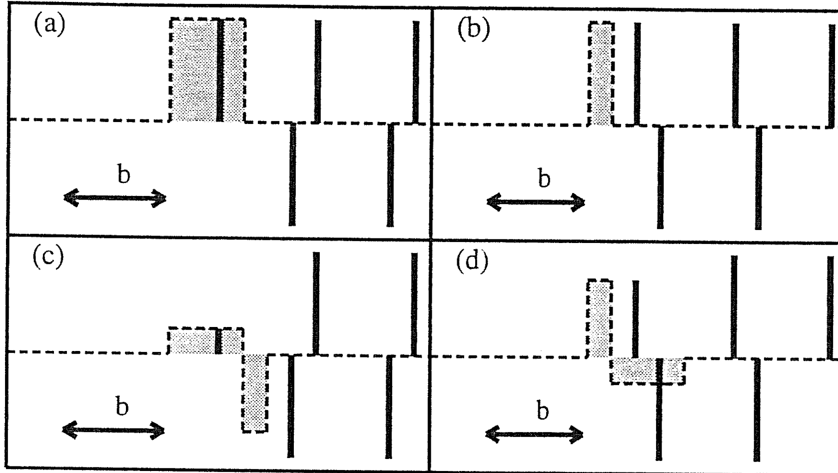

Figure 4.2: Case of (111) cation-terminated polar surface. Examples for the δ-like function $\overline{\rho^*}(z)$, (spikes, whose height measures the δ strength) and for the corresponding piecewise constant function $\overline{\overline{\rho^*}}(z)$ (dashed line), in arbitrary units. The shaded region corresponds to the integral of the macroscopic charge distribution, which is required to vanish. Cases (a) and (b) are then forbidden, while cases (c) and (d) are not.

4.1.2 The (111) case

The (111) surface represents a more complex situation: the ionic planes are in fact no longer equally spaced in this case. By defining the function mod$(i, 2)$ as 0 when i is an even number and 1 when i is odd, then the coordinate of the s-th plane are $[(s-1)b/2 + b/4\text{mod}(s-1, 2)]$ (or $[(s-1)b/2 - b/4\text{mod}(s-1, 2)]$, depending on the surface termination), where the linear period b is $1/\sqrt{3}$ times the lattice constant and the crystal is again assumed to lie in the positive-z half-space.

For both non-equivalent cation-terminated surfaces we show in Figs.(4.2a) and (4.2b) that the surface sum rule is violated if the surface ions have the same dynamical charge as in the bulk; on the contrary it is fulfilled if, for example, the first-layer ionic dynamical charge is $\frac{1}{4}$ (Fig.4.2c) or $\frac{3}{4}$ (Fig.4.2d) of its bulk value.

Following the same procedure as in the (001) case, it is easy to show that the piecewise

constant function $\overline{\overline{\rho^*}}(z)$ vanishes from $-\infty$ to $-b/2$; it assumes the value $Z^*_{1,33}$ (apart for a constant factor) up to $z = b/4$ (or $z = -b/4$); then $\overline{\overline{\rho^*}}(z) = Z^*_{1,33} + Z^*_{2,33}$ up to $b/2$, and in general $Z^*_{s,33} + Z^*_{s+1,33}$ in the following intervals, until it vanishes again in the bulk region, say for $s > N$. The intervals separating the ionic layers are now of different length, namely they alternatively amount to $3b/4$ and $b/4$ (or $b/4$ and $3b/4$); therefore the sum rule for the integral of $\overline{\overline{\rho^*}}(z)$ requires:

$$
\begin{aligned}
&4\sum_{s=1}^{N} Z^*_{s,33} + 3Z^*_{N+1,33} = 0 \qquad \text{if } N \text{ is even} \\
&4\sum_{s=1}^{N} Z^*_{s,33} + Z^*_{N+1,33} = 0 \qquad \text{if } N \text{ is odd}
\end{aligned}
\tag{4.8}
$$

(or viceversa, always depending on the surface termination). Then, even for the more complex (111) surface, the sum of the dynamical charges in the surface region is a well-defined fraction of the bulk dynamical charge, which actually amounts to $\pm 1/4$ or $\pm 3/4$, depending both on surface termination and on parity in the number of layers included in the sum.

4.2 A case-study calculation for β-SiC surfaces

For any different polar surface a strong constraint has been proved therefore to hold, which forbids the surface ions to have the same dynamical charges as in the bulk and gives rise accordingly to a nontrivial relationship among the surface dynamical charges.

We are able to confirm our predictions about the surface dynamical-charge neutrality [expressed by Eqs. (4.7) and (4.8)] through first-principles calculations for some paradigmatic test cases. To keep matter simple we want to deal with insulating surfaces; for this reason we chose the cubic polytype β-SiC, whose (001) Si-terminated surface is insulating. We have also studied the β-SiC(111) surface, where the insulating character is ensured by saturating the dangling bonds with H atoms.

Table 4.1: Structural, optic and vibrational parameters of bulk β-SiC.

	a_0 [a.u.]	$B_0^{th}(Mbar)$	ϵ_∞	$Z^{*(T)}$ (Si)	$Z^{*(T)}$ (C)	$Z^{*(L)}$
Theory	8.24	2.14	7.40	2.58	−2.57	.34
Expt.	8.25	2.24	6.52	2.1–2.7	2.1–2.7	–

4.2.1 Bulk properties of β-SiC

As a preliminary step, we considered some bulk properties of β-SiC (in the zincblende structure). Although a group-IV semiconductor, SiC exhibits a relatively large ionicity of $g = 0.475$ on the Garcia-Cohen scale [62]. The heteropolarity of SiC bond stems from the very different strengths of the C and Si potentials giving rise to very different covalent radii ($r_c^{\rm C} = 0.77$ Å and $r_c^{\rm Si} = 1.17$ Å, respectively). Accordingly, the electronegativity of C ($e_{\rm C} = 2.5$) is considerably larger than that of Si ($e_{\rm Si} = 1.7$); the stronger C potential, as compared to that of Si, leads to a charge transfer $\delta\rho_{\rm Si\rightarrow C}$ from Si to C, so that the electronic charge density distribution around the midpoint of the Si-C bond is greatly asymmetric [63]. As a consequence, Si atoms act as cations while C atoms act as anions in β-SiC crystals.

All our calculations are performed using DFT within LDA, as described in Chapter 3. Plane waves up to a kinetic energy cutoff of 18 Ry were used; this basis set is too small if one aims at a precise prediction of the physical properties of any carbon-based materials, but it is enough to demonstrate the points of principle addressed here, because the ASR in the bulk is well satisfied (largely within 1%), as shown in Table 4.1. We use a set of 28 irreducible special points (a uniform Monkhorst-Pack (12,12,12) grid [40]) in the bulk calculations, and a consistent set in the supercell ones.

The theoretical lattice constant of 8.24 a.u. has been obtained with the same procedure as will be described in more detail in Section 5.3; density-functional perturbation theory [50, 51],

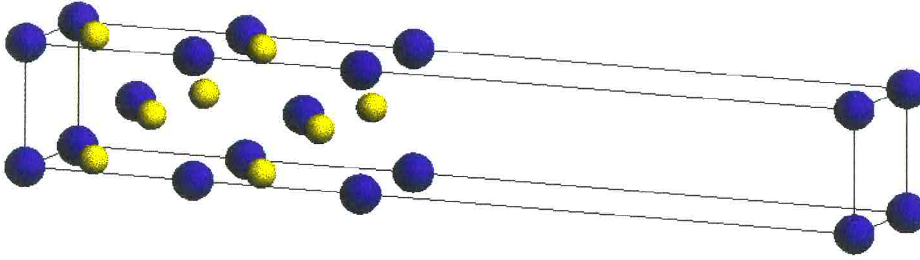

Figure 4.3 (color): Supercell used for modeling the β-SiC(001) surface. Yellow balls identify Si atoms (cations), while blue balls identify C atoms (anions).

implemented as shown in Section 3.2.3, provides for the bulk solid $Z^{*(T)}_{\rm Si} = +2.58$, $Z^{*(T)}_{\rm C} = -2.57$, and $\varepsilon_\infty = 7.40$. The calculated longitudinal dynamical charge of β-SiC is therefore $|Z^*_{\rm bulk}| = 0.34$. Our results are collected in Table 4.1, where the corresponding experimental data [64, 65] are also reported for comparison.

It should be noticed that our computational algorithm does not enforce dynamical neutrality: the error actually monitors the numerical accuracy of the calculation.

4.2.2 The β-SiC(001) surface

The (001) Si-terminated surface of β- SiC has been the subject of recent theoretical [63, 66] and experimental [67] work. The actual structure is 2×1 reconstructed, but even the ideal (truncated bulk) one is insulating (see Fig. 4.4), and fits well our purpose of dealing with a test case as simple as possible, without loss of generality (as previously pointed out). We are going to show that the sum rule takes precisely the form of Eq. (4.7) in this test-case.

The surface calculations have been performed in the supercell geometry shown in Fig. 4.3, where the SiC slabs, that contain 9 atomic layer and have a double Si termination, are separated by a $\simeq 20$ a.u. thick vacuum region (approximately equivalent to 11 layers), in order to decouple the surfaces. The planar and the macroscopic average of the pseudopotential valence-electron density are plotted in Fig. 4.5, which clearly shows that the two surfaces are well separated

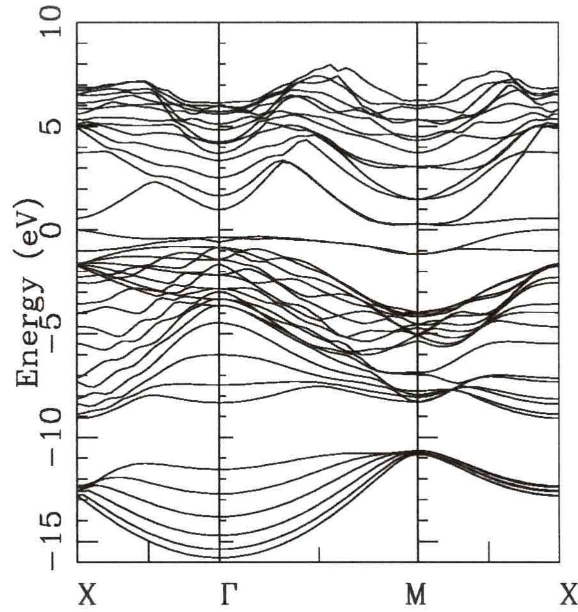

Figure 4.4: Surface band structure of β-SiC(001) surface from a 9-atoms slab calculation, along the $\overline{\Gamma X}$, $\overline{\Gamma M}$, $\overline{XM}$ lines in the SBZ. The zero of the energy corresponds to the valence band top. The insulating character of the surface is evident because of the presence of a band gap ($\approx$ 0.3 eV).

and that the central region of the slab is bulklike, with eight electrons per cell in average. We have also accurately checked that each of the two Si-terminated surfaces is statically neutral.

Starting from the reference equilibrium ground state of Fig. 4.5, we calculate the dynamical charges of all the ions in the supercell using again DFPT for our composite structure of periodically repeated slabs. Obviously, the calculated macroscopic dielectric tensor is a rather artificial quantity, only indirectly related to genuine material properties; equally artificial are the transverse charge tensors $Z^*_{s,\alpha\beta}$ when one considers the $\beta = 3$ component. As already stressed, the longitudinal dynamical charges are the relevant physical quantities in this problem: they are anisotropic tensors in the surface region, and they must converge to their (isotropic) value $\pm 0.34\,\delta_{\alpha\beta}$ in the bulk region.

The convergence proves to be rather slow: insofar as the dynamical charges are concerned, the surface region is much larger than the equilibrium charge density of Fig. 4.5 would suggest. Despite this fact, our slab is large enough to recover the bulk Z^*_{33} value for the center Si

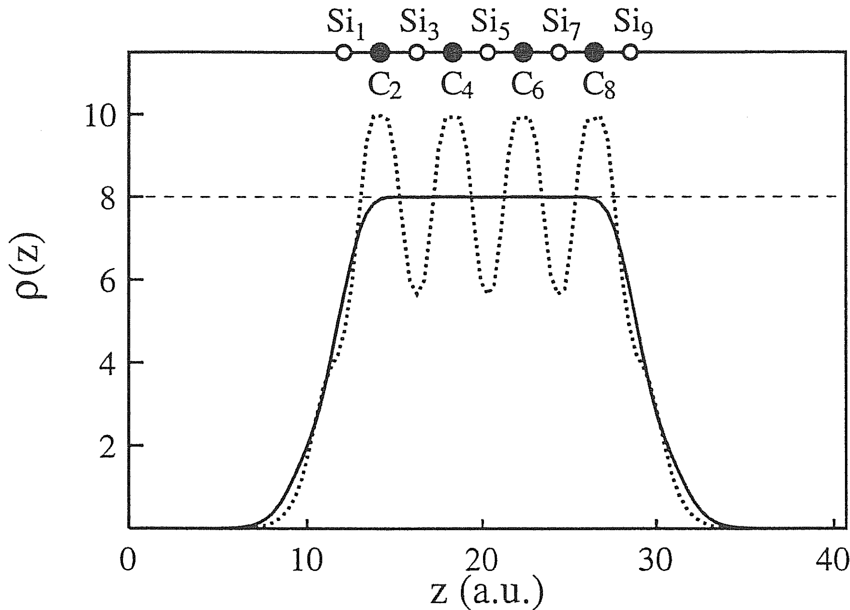

Figure 4.5: Self-consistent valence electron density for the β-SiC(001) surface (from a 9-atom supercell calculation), in units of electrons per bulk cell. The dotted line shows the average over planes parallel to the surface, as a function of the normal coordinate z. The solid line shows the same function after the macroscopic–average filtering.

Table 4.2: Calculated dynamical effective charges for the 9-atom SiC(001) supercell.

Atom_s	Si_1	C_2	Si_3	C_4	Si_5	C_6	Si_7	C_8	Si_9
Z^*	+0.13	−0.27	+0.29	−0.31	+0.33	−0.31	+0.29	−0.27	+0.13

ion (0.33 vs. 0.34 from the bulk calculation). The calculated relevant dynamical charges are collected in Table 4.2. The robustness of these figures has been then checked in two different ways. First we have performed similar calculations on a fully relaxed (though unreconstructed) structure: the outermost SiC bond length increases by 4%. The picture is unchanged, but the bulklike limit is recovered faster, with $Z^*_{33} = 0.34$ for the center Si ion. Secondly the adequacy of our 9-atom supercell has been checked against a few test calculations performed with a slab of 13 atoms. Using the data reported above, the sum of the dynamical charges $Z^*_{s,33}$ over five layers in the surface region, up to the central Si, is 0.17, thus demonstrating our main finding

of Eq. (4.7).

The present (001) geometry is a particularly simple example, where the meaning of the surface sum rule can be made clear without any formal derivations. One first observes that the usual ASR [26] requires the sum of all Z^* in the supercell to vanish: actually, our calculations comply with ASR within a few times 0.01. The sum rule can be interpreted as a "dynamical neutrality" of the supercell as a whole: since our supercell contains two equivalent interfaces, the ASR obviously implies the dynamical neutrality of each of them separately. We may assume each of the interface regions to be one half of the supercell, and clearly the sum of the Z^* vanish in each of them. The key point is that our semiconductor slab has n cations and $n+1$ anions ($n = 4$ in the actual calculation), and therefore the central anion must be reckoned with weight *one half* in summing the dynamical charges of each interface. One arrives therefore, with a simple and alternative argument, at the conclusion that the sum of the dynamical charges Z^* in the interface region equals *one half* the bulk dynamical charge of the semiconductor (with the appropriate sign). Therefore, in this particular example the surface sum rule looks like equivalent to the previously known (molecular) sum rule for the finite slab. But one has to bear in mind that our novel sum rule is indeed an important property of each surface individually: as such, it applies to a semiinfinite solid with only one surface, and it applies to less simple supercells as well.

4.2.3 The H-covered β-SiC(111) surface

The last issue of previous section is is confirmed when a (111) surface is considered, where four different truncations are possible, and the two surfaces are necessarily inequivalent. We then focus on the β-SiC(111); the supercell that describes the surface is made by 8 layers of SiC, and the dangling bonds (on each surface) are saturated with hydrogen atoms in order to deal with a semiconducting system (the computational supercell is shown in Fig. 4.6). It

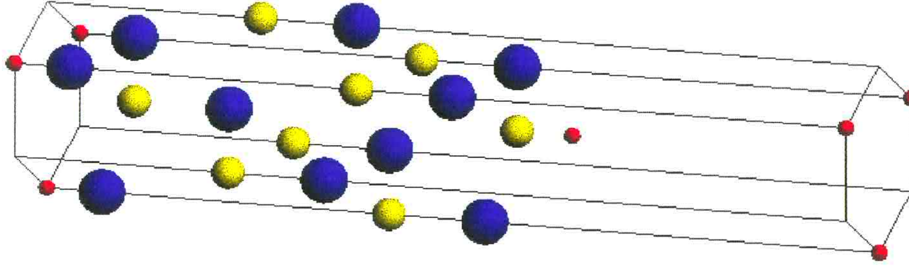

Figure 4.6 (color): Supercell used for modeling the H covered β-SiC(111) surface. Yellow balls identify Si atoms (cations), blue balls identify C atoms (anions), while red balls correspond to H atoms (needed to saturate the dangling bonds).

Table 4.3: Calculated dynamical effective charges for the 10-atom H-SiC(111) supercell.

Atom_s	H_1	Si_2	C_3	Si_4	C_5	Si_6	C_7	Si_8	C_9	H_{10}
Z^*	-0.06	$+0.25$	-0.33	$+0.35$	-0.34	$+0.34$	-0.33	$+0.36$	-0.18	-0.04

is straightforward to show that the presence of the H atoms does not affect the form of the surface sum rule, as stated in Eq. (4.8), provided that the dynamical charge of H is reckoned as well.

The calculated Z^* are shown in Table 4.3: the dynamical charges recover their bulk Z^*_{33} value for the center atoms (C_5 and Si_6), thus confirming that the 10-atom supercell is large enough.

However, some technical complications occur. The (111) surfaces of binary compounds are seldom studied in first-principles calculations, unless one is interested in spontaneous polarization effects [68]. Essentially, the difficulty stems from the fact that two non-equivalent surfaces are there, with different values for the work-functions: a supercell-based description of the system (where periodic boundary conditions are used) implies then that fictitious electric fields arise in order to periodically match the work function difference. These fields are equivalent to

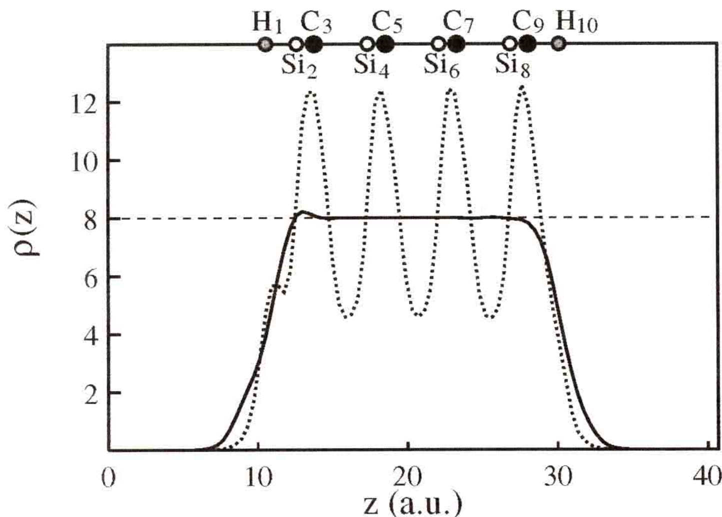

Figure 4.7: Self-consistent valence electron density for the H-covered β-SiC(111) surface (from a 10-atom supercell calculation), in units of electrons per bulk cell. The dotted line shows the planar average, as a function of the normal coordinate z, while the solid line shows the macroscopic average.

static uncompensated charge at the two surfaces, and are an artifact of the periodic supercell geometry: the fields and the surface charges would disappear in the limit of an infinitely thick vacuum layer.

This monopole charge accumulation is evident from the macroscopic average of the electronic charge (plotted in Fig.4.7 together with the planar average), and even more from the slope of the average electrostatic potential (shown in Fig. 4.8). In other words, owing to the work-function difference, the surfaces of a finite slab appear us charged due to the supercell periodic boundary conditions.

It is easy to evaluate the static charge monopole Z_{surf}, which is given in fact by the non-compensation between the ionic and the electronic charge on the surface region:

$$Z_{\text{surf}} = \int_{\text{surf.reg.}} [\, \bar{\bar{\rho}}_{\text{el}}(z) - \bar{\bar{\rho}}_{\text{ion}}(z)\,]\, dz = 0.03 \tag{4.9}$$

At this point, the dynamical neutrality condition can be generalized by simply including Z_{surf} in Eq. (4.8) as an additional charge contribution (which is equal and opposite on the two

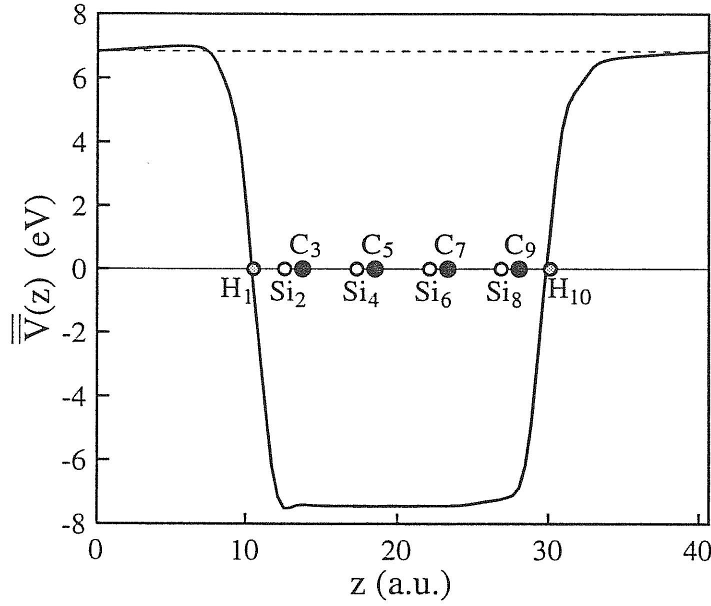

Figure 4.8: Macroscopic average of the total electrostatic potential for for the H-covered β-SiC(111) surface (from a 10-atom supercell calculation).

surfaces due to the neutrality of the whole supercell). For the present H-covered β-SiC(111) the surface sum rule therefore read:

$$4Z_{\text{surf}} + 4\sum_{s=1}^{N} Z^*_{s,33} + 3Z^*_{N+1,33} = 0 \qquad \text{if } N \text{ is even} \tag{4.10}$$

$$4Z_{\text{surf}} + 4\sum_{s=1}^{N} Z^*_{s,33} + Z^*_{N+1,33} = 0 \qquad \text{if } N \text{ is odd}$$

The values of the surface dynamical charges (collected in Table 4.3) can now be inserted in the previous expression, by taking into account that Z_{surf} is positive on the H-Si bond terminated surface, while it is negative on the H-C bond terminated one: the equality (4.10) holds within our numerical accuracy (few times 0.01).

We then proved that it is possible to formulate the surface dynamical neutrality condition for the (111) surface through Eq. (4.10), which provides a generalization of Eq (4.8) accounting for the presence of some surface static charge.

5 Metal-semiconductor interfaces: Al/GaAs(001) Schottky barrier

As already stressed in the Introduction, the detailed mechanisms involved in Schottky barrier (SB) formation at a metal-semiconductor interface are difficult to be identified, so that this yet remains one of the most active areas of solid state physics.

The complexity of the physical problem arises from the messy features of semiconductor interfaces, such as relaxation, lattice mismatch and disordering, chemical reaction, interdiffusion, exchange of atoms, presence of defects and metal-induced gap states.

5.1 The Schottky barrier problem: models and experiments

In 1938 Schottky explained the rectifying property of metal-semiconductor junctions through a space-charge layer on the semiconductor side of the contact which is depleted of mobile majority carriers. The current of electrons (holes) through the junction is caused by the energy difference between the Fermi level and the bottom (top) of the conduction (valence) band, when the semiconductor is n-doped (p-doped), as showed in Fig. 5.1.

In their model [69], Mott and Schottky identified the barrier height with the difference between the metal work function Φ_m and the semiconductor electronic affinity χ_s, so the

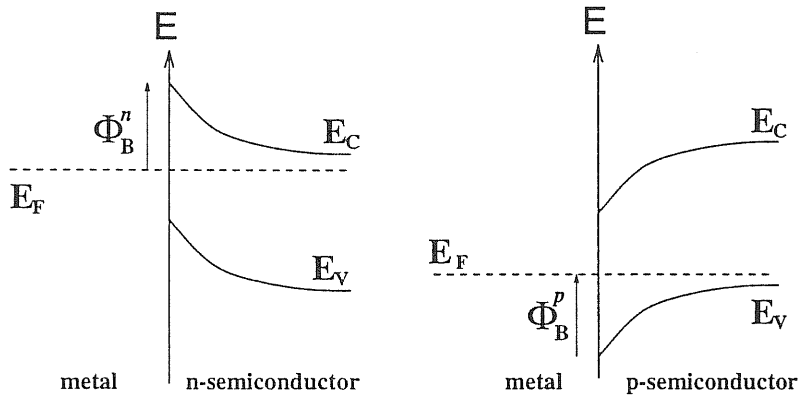

Figure 5.1: Usual schematic representation of n- and p-type barriers. The band-bending due to semiconductor doping occurs typically at distances of the order of a few hundreds of Å and can be therefore neglected in our atomic-scale study of the region very near to the interface.

Schottky-Mott equations read:

$$\Phi_B^n = \Phi_m - \chi_s \tag{5.1}$$

$$\Phi_B^p = E_g - \Phi_m + \chi_s = E_g - \Phi_B^n$$

The sum of the p-barrier and the n-barrier equals the semiconductor gap E_g.

Experience shows that these equations cannot be verified because, in general, the SB height shows typical variations from one metal to the other of the order of 0.2 eV, whereas the range of variation of the work function over different metals is of some eV. Moreover, the assumption that the work functions of the metal and semiconductor remain unchanged after the interface is formed seems not plausible when one considers that the interfacial dipoles are likely to be formed as a result of the localized charge transfer accompanying the formation of bonds between the metal and the semiconductor.

To explain this weak dependence of the barrier height on Φ_m Bardeen [70] attributed the Schottky barrier mechanism to the presence of localized states in the semiconductor, which are

able to accommodate the extra-charge coming from the metal contact and thus to effectively pin the Fermi level.

After that, virtually all other subsequent workers in the SB field agree that the Fermi-level pinning is caused by interface electronic states in the band gap, and the only controversial issue is the nature and the origin of these states. We can distinguish two different classes of models.

The first one is based on intrinsic properties of the semiconductor and suggests that it is possible for the wave functions of those electrons in the metal with energies corresponding to the forbidden gap in the semiconductor to penetrate into the semiconductor in the form of exponentially damped evanescent waves [71]: they thus become Bloch states of the semiconductor with complex wave-vector components normal to the interface. This gives rise to a transfer of charge from the metal to the semiconductor and the states so arising induced by the presence of the metal (*metal induced gap states*, or MIGS) might dominate the mechanism of Fermi level pinning.

The second one is instead based on extrinsic interface properties and attributes the pinning to some gap states provided by the formation of interfacial defects. Wieder [72] and Spicer [73], for example, suggest that point defects might occur as a result of the energy released during interface formation. Freeouf and Woodall [74] propose to replace the metal work function with an effective one: they suppose the existence of a microscopic anion pack at the interface created during the interface formation and modifying the semiconductor electron affinity.

At present time, it has not been established in a definitive way which picture is the right one. On one hand the approach underlying the intrinsic model has been questioned because of the simplicistic treatment of surface structure [75]. On the other hand it has been claimed [76] that the density of defects needed for Fermi level pinning may be much higher than the one

known to exist at interfaces. Moreover, if SB heights are determined by Fermi level pinning associated with defects, then they should be unrelated to heterojunction lineups (actually, at device quality heterojunctions there are very few defects); but the fact that band lineups and barrier heights are so closely connected experimentally argues strongly for a theoretical approach which deals with both of these on a unified footing and gives the relationship between SB and band lineup as a natural result.

Among the models belonging to the first class, it is important to mention the Tersoff's one [77]. The basic idea underlying this approach is to introduce a neutrality level for the semiconductor, playing a role similar to the one of the Fermi level of the metal. According to Tersoff, this neutrality level must fall at the energy E_B (the *branching point* of the complex band structure), where MIGS wave-functions cross over from being mostly valence-band derived to being conduction-band derived. The MIGS create the dipole needed to align the neutrality level and the Fermi level. Tersoff extended this model also to heterojunctions between semiconductors, concluding that the energies E_B of both semiconductors have to be nearly aligned, because the screening in the semiconductor should be similar to the one in the metal, owing to the big dielectric constants of the semiconductors ($\epsilon > 10$).

Following Tersoff's model, the mechanism governing the band alignment for a semiconductor-semiconductor heterojunction is the same as for a metal-semiconductor junction, since it forces the reference energy levels of both materials to be aligned and this alignment condition is transitive. Moreover, if the same mechanism is responsible for the band line-up in both types of junctions, then the difference between n- or p-Schottky barriers between a metal M and two semiconductors A and B should equal the valence band discontinuity (VBO) or the conduction band discontinuity (CBO) for the A/B heterojunction:

$$\Phi_B^n(M/A) - \Phi_B^n(M/B) = CBO(A/B) \tag{5.2}$$
$$\Phi_B^p(M/A) - \Phi_B^p(M/B) = VBO(A/B)$$

Besides, Φ_B should be metal-independent; equation (5.2) is often referred to as transitivity rule. The Tersoff's model is appealing because of its simplicity, but it lacks of a formal physical basis. Particularly in the case of semiconductor heterojunctions, where our theoretical understanding is much better, it has been proved [78] that the Tersoff's model is only a rough guideline, accurate no better than 0.3 eV.

A wide range of surface-science techniques has been applied to study the microscopic interactions and Schottky barrier formation for metals deposited on clean III-V semiconductors. Extensive studies of Schottky barrier heights for thick metal layers on III-V semiconductors have been carried out using I/V, C/V, and photoresponse (internal photoemission, IPE) methods, the samples being mostly obtained growing the metal on a semiconductor substrate with *molecular beam epitaxy* (MBE) technique.

Nevertheless, also from an experimental point of view the scenario is complicated, because the application of modern experimental techniques has clearly shown that interfaces formed between metals and semiconductors are complex regions whose physical properties are highly dependent on the preparation procedure. In order to understand the fundamental physics of contacts to semiconductors the most important studies are those where metals have been deposited onto atomically clean and ordered surfaces and only in a few cases (e.g. aluminum on gallium arsenide) it has been possible to prepare epitaxial single crystal contacts with virtually perfect interfaces to the semiconductor. However, even in such a case of epitaxial growth, Al atoms have been demonstrated to be highly mobile both on GaAs(110) and on GaAs(001) surfaces and chemical reactions have been detected [79].

The Schottky barrier problem is thus relatively messy, because of uncertainties in the microscopic structure and in the chemistry of the interface, as well as because of the difficulty of performing reliable spectroscopic measurements of the semiconductor in the presence of a metal.

Moreover, it turns out to be difficult to choose one class of models simply by comparing theoretically predicted barrier heights with experiments because both of them, even if based upon very different physical ideas, give similar predictions from a quantitative point of view, and mostly in agreement with experimental findings.

5.2 Method for the ab-initio calculation of the Schottky barrier height

We determine SB heights within the present calculations by focusing on the p-type barrier (relevant for hole carriers and hence indicated as Φ_B^p), which is the difference between the metal Fermi level and the semiconductor valence band top (see Figure 5.1):

$$\Phi_B^p = E_F - E_v. \tag{5.3}$$

This energy difference has to be evaluated at a distance across the interface which is small compared to any band-bending length (typically hundreds of lattice spacings) but larger than the 4-5 monolayer distance on either side of the interface where the constituents have a charge density (electronic and nuclear) that is not identical to that in their respective bulk materials. Following the same approach as the well-known one used for computing the valence band offset for semiconductor heterojunctions [80], where the barrier height was also partitioned into an electrostatic *potential lineup* contribution and a *band-structure term* contribution, Eq. (5.3) can be expressed as:

$$\begin{aligned}\Phi_B^p &= [E_F - \langle \mathrm{V} \rangle_m] - [E_v - \langle \mathrm{V} \rangle_s] \\ &+[\langle \mathrm{V} \rangle_m - \langle \mathrm{V} \rangle_s]_{IF} \\ &= \Delta E_F - \Delta E_v + \Delta \mathrm{V}.\end{aligned} \tag{5.4}$$

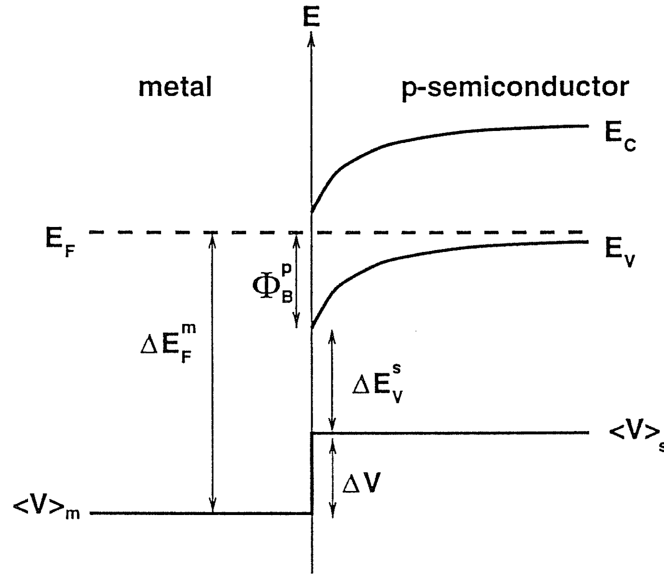

Figure 5.2: Contributions to the Schottky barrier (schematic). ΔE_F^m is the Fermi level of bulk metal, with respect to the corresponding average electrostatic potential; ΔE_v is the valence band top of the semiconductor, with respect to the bulk average electrostatic potential; Δ V represents the potential generated by the interface dipole.

as made explicit in figure 5.2. The first term on the right-hand side is the Fermi energy in the bulk metal with respect to the average electrostatic potential $\langle \mathrm{V} \rangle_m = \mathrm{V}_m(\mathrm{G} = 0)$ of that crystal. The second term is the energy of the valence-band edge in the bulk semiconductor with respect to its average potential $\langle \mathrm{V} \rangle_s = \mathrm{V}_s(\mathrm{G} = 0)$. Both of these band-structure terms are simply properties of the bulk materials and they have nothing to do with the metal-semiconductor interface, except for the fact that the strain configurations of the constituent materials in the pure bulk must be identical to the strain configuration on either side of the interface.

The average potentials $\langle \mathrm{V} \rangle$ in Eq. (5.4) are ill-defined quantities, but their difference $\Delta \mathrm{V}$ is well-defined and can be obtained from a supercell calculation. Actually, $\Delta \mathrm{V}$ represents the interfacial dipole, i.e. the difference in the average electrostatic potential on either side of the metal-semiconductor interface (at the intermediate length scale defined above). It is simple to

argue that ΔV is the sum of two contributions: the first is the Hartree potential discontinuity, due to the electron charge distribution, while the second is provided by the Coulomb potential discontinuity, due to ionic point charges. Their calculation requires, respectively, the knowledge of electronic charge density $\rho_{\rm el}(\mathbf{r})$ and of ionic charge density $\rho_{\rm ion}(\mathbf{r})$.

The plane-wave pseudopotential method developed within DFT (described in Section 3.1) provides the Fourier coefficients $\rho_{\rm el}(\mathbf{G})$ of the electronic charge density, so that:

$$\rho_{\rm el}(\mathbf{r}) = \sum_{\mathbf{G}} \rho_{\rm el}(\mathbf{G}) \exp(i\mathbf{G}\cdot\mathbf{r}). \tag{5.5}$$

The ionic density can be described as the superposition of Gaussian functions centered on the atomic positions τ_s:

$$\rho_{\rm ion}(\mathbf{r}) = \sum_{s=1}^{N} Z_s \left(\frac{\alpha}{\pi}\right)^{3/2} exp(-\alpha \mid \mathbf{r} - \tau_s \mid^2), \tag{5.6}$$

where the sum runs over all the atoms s within the supercell Ω, τ_s and Z_s represent, respectively, the positions and the charges of the N ions in the supercell and α is such to ensure that the ionic charge density is relevant only in the region nearly surrounding the atoms and goes rapidly to zero elsewhere.

The electrostatic potential V($\mathbf{r}$) is related to the charge density by the Poisson's equation

$$\nabla^2 \mathrm{V}(\mathbf{r}) = 4\pi\rho(\mathbf{r}), \tag{5.7}$$

where the charge density ρ is made of an electronic contribution, due to the valence electrons, and an ionic contribution, due to the ionic cores of positive charge:

$$\rho(\mathbf{r}) = (\rho_{\rm el}(\mathbf{r}) - \rho_{\rm ion}(\mathbf{r})). \tag{5.8}$$

The integral of the valence electron density $\rho_{\rm el}$ and of the ionic cores density $\rho_{\rm ion}$ over the supercell volume Ω equals the number of valence electrons $\mathrm{N_{el}}$ within the supercell:

$$\int_\Omega \rho_{\rm el}(\mathbf{r}) d\mathbf{r} = \int_\Omega \rho_{\rm ion}(\mathbf{r}) d\mathbf{r} = N_{\rm el}, \tag{5.9}$$

and the system is globally neutral.

The Poisson's equation is easily solved in reciprocal space

$$V(\mathbf{G}) = 4\pi e \frac{\rho(\mathbf{G})}{|G|^2}, \tag{5.10}$$

where $\mathbf{G}$ is a non vanishing reciprocal-lattice vector. The $|\mathbf{G}| = 0$ term is not included because all the energies are calculated with respect to the average electrostatic energy $V(\mathbf{G} = 0)$, that is set to zero.

Both the electrostatic potential $V(\mathbf{r})$ and the density $\rho(\mathbf{r})$ are periodic functions and they oscillate rapidly on the scale of the interatomic bulk distance. Since we are interested in the macroscopic properties of these quantities, we have to average out their microscopic oscillations. This is done through the macroscopic average concept (as explained in Appendix A).

For Al/GaAs junctions the periodicity on the semiconductor side of the interface is $1/\sqrt{2}$ times the periodicity on the metal side; in this case the planar and the macroscopic average of the electrostatic potential can be obtained through Eqs. (A.6), (A.9) and (A.10) and taking advantage of the simple expression of the Poisson equation (5.10) in reciprocal space:

$$V(\mathbf{G}_{||} = 0, G_z) = \frac{4\pi e}{G_z^2} \rho(\mathbf{G}_{||} = 0, G_z), \tag{5.11}$$

where, from Eq. (5.6), the Fourier transform of the ionic charge density ρ_{ion} is:

$$\rho_{\text{ion}}(\mathbf{G}_{||} = 0, G_z) = \sum_s \frac{Z_s}{\Omega} e^{-iG_z \tau_{s,z}} e^{-G_z^2/4\alpha}, \tag{5.12}$$

The discontinuity ΔV [the third term in Eq. (5.4)] is finally obtained by evaluating the difference:

$$\Delta V = \overline{\overline{V}}(z_1) - \overline{\overline{V}}(z_2), \tag{5.13}$$

where z_1 identifies a position within the semiconductor, far from the interface, while z_2 identifies a position within the metal, far from the interface.

5.3 Bulk properties of GaAs and Al

As a preliminary step towards the study of Al/GaAs junctions, we have analyzed the structural properties both of bulk aluminum and of bulk gallium arsenide. The properties of the unperturbed crystals are studied within the formalism of the DFT as thoroughly reported in Chapter 3.

The equilibrium structure is determined by minimizing the total energy of the crystal with respect to the lattice parameter a. For this purpose, the values of the total energy, calculated with a fixed kinetic-energy cutoff at different lattice parameters, have been fitted to a Murnaghan's equation of state [81]:

$$E(\Omega) = \frac{\Omega_0 B_0}{B_0'} \left[\frac{1}{B_0' - 1} \left(\frac{\Omega_0}{\Omega} \right)^{B_0' - 1} + \frac{\Omega}{\Omega_0} \right] + \text{const}, \tag{5.14}$$

where B_0 is the bulk modulus, B_0' its derivative with respect to the pressure, and $\Omega_0 = a_0^3/4$ the equilibrium volume of the unit cell.

5.3.1 Gallium arsenide

GaAs is a binary compound semiconductor, belonging to the III-V group and displaying a zincblende structure. To describe the interaction between valence electrons and ionic cores, we use both for Ga and for As the norm-conserving pseudopotential of Bachelet, Hamann, and Schlüter [43] in Kleinman-Bylander [48] form (in order to reduce the computational effort) and made ghost-free by Gonze [49]. According with a convergence study for the total energy of bulk GaAs with respect to the energy cutoff and to special point sampling at the experimental lattice constant (a_0 = 10.68 a.u.), the PW basis set used has a kinetic energy cutoff of 20 Ry and the BZ-integration is performed using a uniform Monkhorst-Pack [40] (10,10,10) grid (i.e. 19 k-points in the irreducible wedge of the BZ). It has also been checked that these parameters are sufficient to obtain accurate results (converged within $\sim 1\%$) for the structural parameters,

achieved by interpolating Eq. (5.14); the outcomes are displayed in Table 5.1 in comparison with the experimental values. The discrepancy with the experimental values is 1.9% for the lattice parameter and 1.8% for the bulk modulus. This agreement is good and comparable with the typical accuracy of LDA calculations for semiconductors. The top of the valence band for GaAs, which is necessary for the calculation of the Schottky barrier height [see Eq. (5.4)], is $\Delta E_v = 5.17$ eV.

5.3.2 Aluminum

Crystalline Al is a group IIIA metal and displays a FCC structure. The pseudopotential used to describe aluminum has been obtained from the method originally proposed by von Barth and Car [82]; for this pseudopotential we checked the absence of ghost-states. As for all other metals, the main difficulty is in relation with the description of its Fermi surface and the special point technique must be used with care in this case. The BZ integration is then performed with the smearing technique described in Appendix B, using the Hermite-Gauss smearing function of order $N = 1$ and a smearing width $\sigma = 10$ mRy. As a first step, a kinetic-energy cutoff of 18 Ry has been chosen, by imposing a total-energy convergence of the order of mRy and verifying that it was enough to ensure the structural parameters to be converged within 1% . Convergence with respect BZ sampling and smearing width σ has been further controlled, keeping in mind that the smaller is the value of the smearing width σ, the finer is the **k**-points mesh needed to achieve convergence. For $\sigma = 10$ mRy a mesh of 182 points (corresponding to a (24,24,24) grid) in the in the irreducible wedge of the BZ is enough to ensure a satisfactory convergence of the calculated quantities (i.e. the total energy, the Fermi energy and the structural parameters). The values so obtained are shown in Table 5.1, compared with the corresponding experimental ones. Following the recipe described by Eq. (B.6), the value of the Fermi level with respect to the average electrostatic potential in the cell has been self-consistently calculated: $\Delta E_F =$

Table 5.1: Lattice parameters and bulk modulus for GaAs and Al.

	$a_0^{th}(a.u.)$	$a_0^{exp}(a.u.)$	diff.	$B_0^{th}(GPa)$	$B_0^{exp}(Gpa)$	diff.
GaAs	10.48	10.68	1.9%	76.8	75.5	1.8%
Al	7.48	7.64	2.1%	82.5	79.3	4.0%

8.65 eV.

5.4 Al/GaAs(001) junction

As pointed out in Section 5.2, the barrier height is partitioned into two contributions: the electrostatic *potential lineup* across the interface Δ V, and the *band-structure term*, which is the difference between bulk properties of the two constituents, and hence it is obtained from independent calculations for crystalline GaAs and Al, as described respectively in Sections 5.3.1 and 5.3.2.

The potential lineup, on the other hand, is an interface-specific property: Al/GaAs interfaces are then described with the supercell method: they are modeled through periodically repeated and alternating slabs of metal and semiconductor. This approach allows a reciprocal space formulation of the interface problem and the use of PW's, otherwise not possible because of the loss of translational symmetry along the growth direction. The isolated interface configuration is well represented, provided that the two adjacent interfaces are sufficiently spatially separated in order to not interact. This is verified *a posteriori*, by examining the charge density and the Hartree potential in the regions far from the interface and by showing that they are identical to those of the bulk metal and of the bulk semiconductor.

Extensive experience with semiconductor-semiconductor heterojunctions has shown that a 6+6 superlattice (i.e. $(GaAs)_3(AlAs)_3$) is more than sufficient to extract accurate band offset

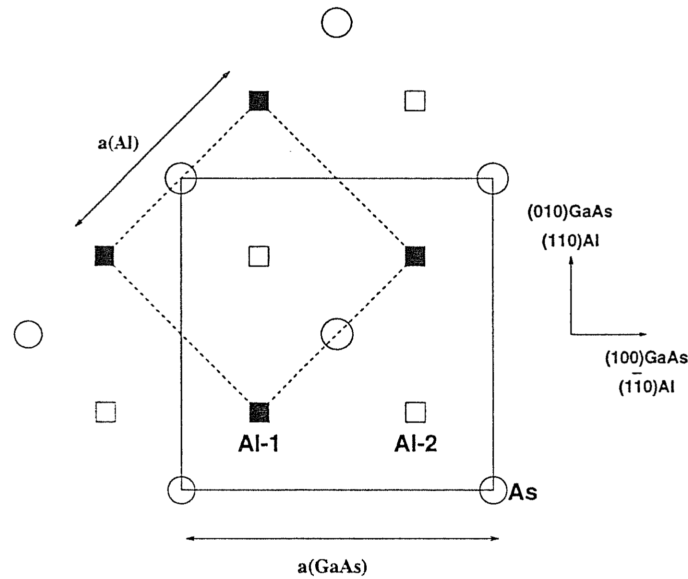

Figure 5.3: Top-view for the (001)Al/GaAs supercell.

values [83]. At variance with this, we expect that larger supercells are needed for accurate SB calculations, because of the presence of the MIGS, whose charge decays evanescently in the semiconductor region.

Actually we found out that certainly a suitable supercell is a 13+9 superlattice, where the semiconductor slab is made of 13 atomic layers (6 Ga layers and 7 As layers), and is therefore As-terminated on both sides, while the metallic slab is made of 9 layers of Al (each contributing two atoms per supercell). The supercell contains therefore 31 atoms altogether, and has in it two equivalent interfaces with Al-As bonds (as shown in Fig. 5.4).

Al/GaAs junctions have been extensively studied from both a theoretical and an experimental point of view, as already pointed out in Section 4.1. In the case where they are grown along the (001) direction, the (001) axes of Al and GaAs are obviously parallel and, since the Al and GaAs lattice constants approximately satisfy $a_{\mathrm{GaAs}} \approx \sqrt{2}a_{\mathrm{Al}}$. It is possible for Al to be grown on a GaAs(001) substrate provided it is rotated 45° about its (001) axis, so that the Al (110) direction points along the GaAs (100) one (see Fig. 5.3).

The actual Al/GaAs(001) interface is sp-bonded and almost perfectly lattice-matched (the

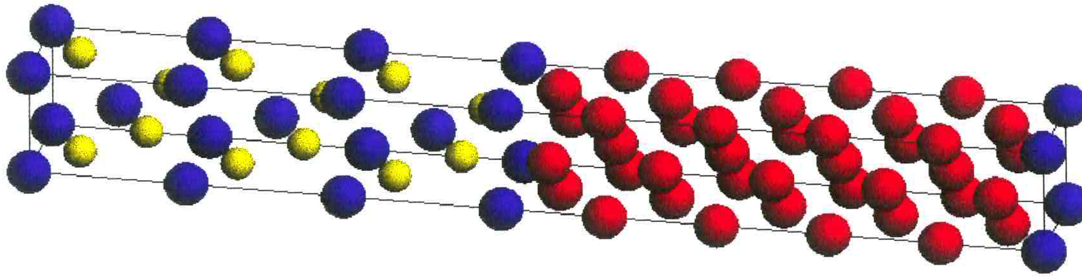

Figure 5.4 (color): Supercell used for modeling the Al/GaAs(001) interface. Yellow balls identify Ga atoms, blue balls identify As atoms, while red balls correspond to Al atoms.

1% experimental mismatch is well reproduced by our theoretical calculations). At variance with previous first-principles work, we do not aim at a detailed modeling of the the real interface; instead, we concentrate on a reference system as simple as possible. The reason for this kind of approach is that our challenge is to evidentiate the leading effects induced by controlled variations of the interface morphology, which will be shown in next Chapter. *Ab-initio* calculations allow to study ideal interfaces (hardly obtainable in practice) and therefore to "isolate" the most relevant contributions to the *pinning* effect.

We assume therefore a defect-free epitaxial geometry as a working hypothesis and, on the same ground, we study here an ideal, unstrained interface where the metal is a "fake" Al, perfectly lattice-matched to GaAs, and hence retaining its cubic structure in the epitaxial overlayer. Strain effects, although quantitatively sizeable, are considered spurious in the present analysis (see however some considerations below). Furthermore, we choose to take as distance between the last As-plane and the first Al-plane the average plane spacing in Al region and in GaAs region: this recipe minimizes the force acting on interfacial atoms, and hence the relaxations effects.

The calculations have been performed with the same theoretical tools as for the bulk Al- and GaAs-calculations, using the same pseudopotentials and plane waves up to 20 Ry; reciprocal space integrations are performed on a Monkhorst-Pack special-point (24,24,2) grid, still using

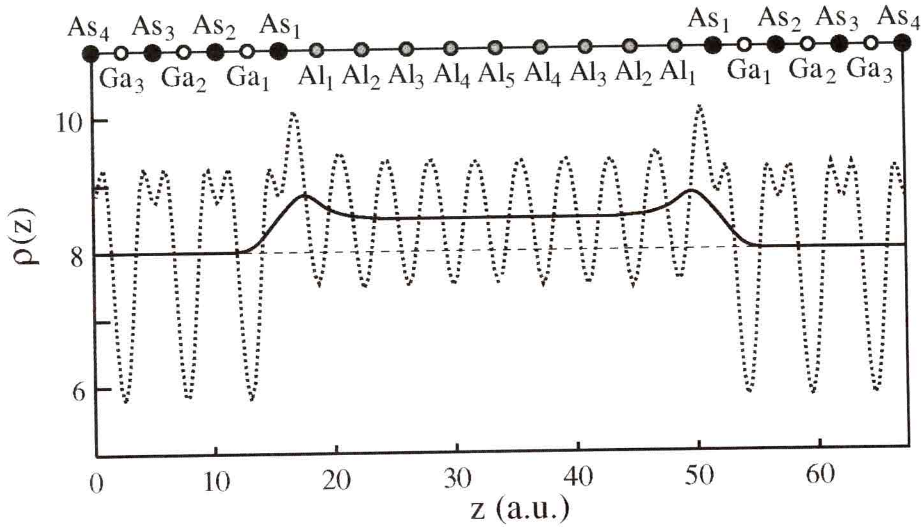

Figure 5.5: Planar (dotted line) and macroscopic (solid line) average of the electronic charge density for the (001)Al/GaAs supercell.

the smearing technique (see Appendix B) with a smearing parameter $\sigma = 0.01$ Ry. The relevant results extracted from a typical 31-atom supercell are shown in Figs. ?? and ?? in terms of planar and macroscopic average of the electron density, in units of (valence) electrons per semiconductor cell. In these units the charge density in the semiconductor region is 8, whereas the one in the metal region reaches the value of 8.485, which actually equals 6 (number of electrons in one periodicity of Al) times $\sqrt{2}$ (ratio between the periodicity in the GaAs region and the one in the Al region). It is easily realized that the actual density reaches its bulk unperturbed value very close to the junction, thus showing that the supercell is large enough to model the isolated (and neutral) interface. Moreover, one is allowed to infer the dimension of the interfacial region, which in fact extends over less than about 5 Å.

The discontinuity of the average potential is the sum of two compensating terms (see Fig. 5.6), i.e. the Hartree potential and the ionic potential, obtained as solution of the Poisson equation for the electronic and the ionic charge, respectively. The fact that their sum is nearly constant far from interface (as is evident from Fig. 5.7) ensures *a posteriori* that the dimension

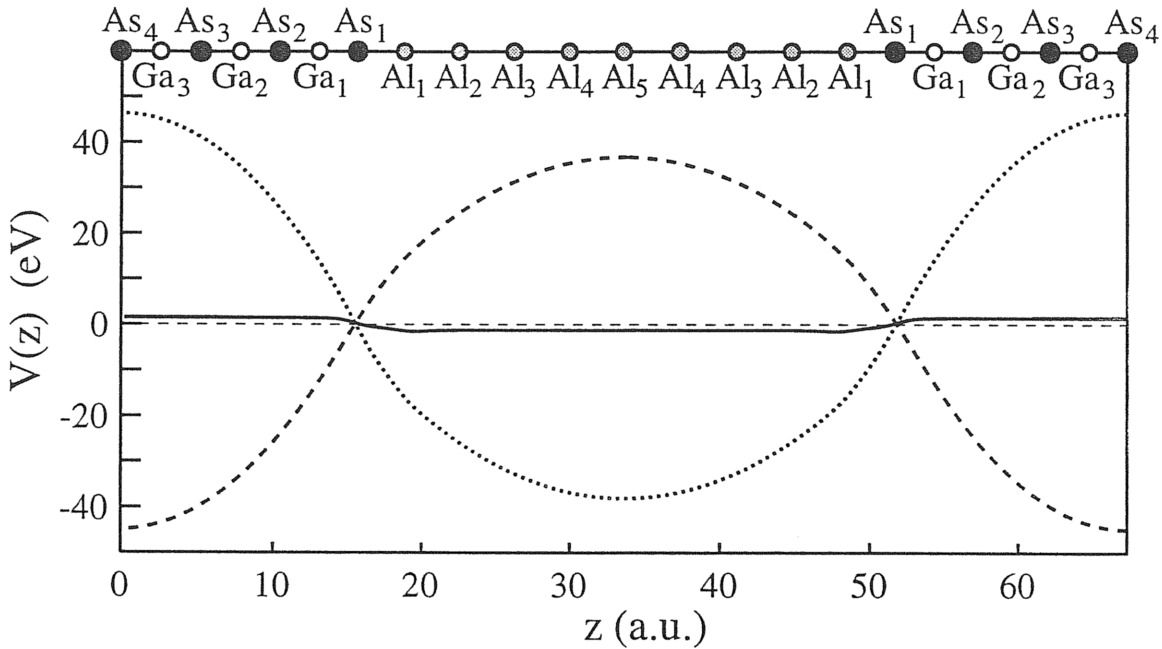

Figure 5.6: Macroscopic average of the electronic (dotted line), ionic (dashed line) and total (solid line) average electrostatic potential for the (001)Al/GaAs supercell. The first two quantities are perfectly compensating far from the interface.

of our supercell is sufficient. The macroscopic average of the total electrostatic potential is shown in Fig. 5.7: the lineup between the plateaus in the two bulks coincides with the ΔV discussed above: its value for this calculation is 2.76 eV. Putting this figure together with the ones obtained by the two independent self-consistent calculations for the individual bulks (see Sections 5.3.1 and 5.3.2), we get the value Φ_B^p=0.73 eV for the Schottky barrier at our ideal junction between GaAs and fake Al. When we compare different (001) calculations amongst themselves, as extensively done in next Chapters, our estimated numerical accuracy for Φ_B^p is 0.01 eV. We stress that this is a *relative* accuracy, for a given set of technical ingredients. Variation of the latter, as for instance by adopting different pseudopotentials, would affect the results by much more.

As already discussed above, our SB analysis is intentionally focused on an ideal, unstrained interface, in order to identify the genuine effects of controlled variations of the interface mor-

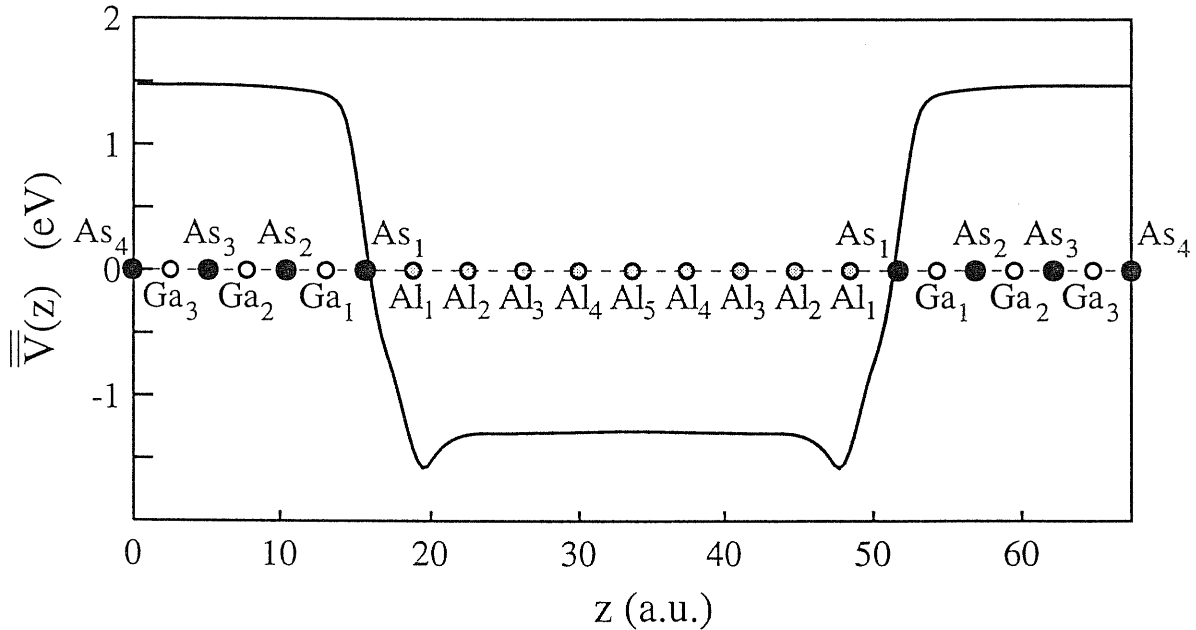

Figure 5.7: Macroscopic average of the total electrostatic potential for the (001)Al/GaAs supercell. The difference between the two plateaus corresponds to ΔV.

phology on the barrier height. Nevertheless, for the sake of completeness, we also briefly summarize the main features of the actual interface: they have to be considered if one would aim at a comparison with the experiments (which requires a realistic description of Al/GaAs(001) junction).

First of all, in order to gradually recover in our theoretical calculation the actual interfacial geometry, it might be important to consider that Al needs to be slightly strained to accommodate the lattice mismatch, according to the elasticity theory for dislocation-free, coherent epitaxy. We will come back to this point in Section 6.3, where all computational details are also reported. We anticipate here that the SB height is found to be completely insensitive (within our numerical accuracy) to the tetragonal strain of the metal overlayer.

An important feature of the actual interfacial geometry, that we did not account for, is the presence of two inequivalent Al atoms per layer (as it is explicitly displayed in Fig. 5.3); namely, half of the Al atoms in the layer immediately adjacent to the top GaAs layer sit nearby

the sites of continuation of the zincblende GaAs (we call these Al-1 atoms), while the other half (Al-2 atoms) sit in interstitial sites (see Fig. 5.3); the Al-1 atoms are thus bonded covalently to the GaAs substrate, while the Al-2 atoms are not. Hence, the two independent Al sublattices are expected to not lie in the same plane. We verified this behavior by controlling the different intensities of the forces acting on both types of Al when they are forced to sit in the same plane (as in our supercell). Nevertheless, recently it has been found that the relaxation effect is to change the p-barrier of only about 2 meV [84].

Moreover, our calculation within DFT-LDA framework is in principle affected by few kinds of imprecisions, but they tend to compensate each other. First of all, the spin-orbit interaction removes the valence-band top degeneration at Γ point (causing a band splitting $\Delta = 0.33$ eV) for GaAs. This effect decreases the p-barrier of ~ 0.11 eV. The many-body effects, beyond LDA, also modify the band terms and increase the p-barrier of about 0.1 eV. The relaxation of 3d electrons of GaAs plays a minor role. Needs *et al.* [85] showed that these three corrections compensate nearly perfectly (within about 0.01 eV). As a consequence, our calculated SB height ($\Phi_B^p = 0.73$ eV) is in fair agreement with the experimental value ($\Phi_B^p = 0.66$ eV [86]).

6 Schottky barrier interface morphology: role of dynamical charges

The electronic mechanisms governing the value of the Schottky barrier—as well as their dependence on the microscopic morphology of the interface—have not been systematically investigated so far and are basically unknown. Here we provide a contribution in this direction, by studying the barrier-height variations induced in Al/GaAs(001) by several structural and morphological perturbations which are switched on and off in our computational framework.

In this Chapter we are going then to present a first-principles study of the effect (in terms of Schottky barrier height variation) induced in ideal Al/GaAs(001) junctions by several structural and morphological perturbations which are easily switched on and off in our computational framework. We now investigate how our calculated value of SB height, Φ_p^B, depends on different perturbations which alter the interface morphology.

This analysis will bring to the conclusion that the dynamical effective charges Z^* (introduced in Chapter 2) in the interface region are the key quantity for rationalizing morphology-induced variations of the Schottky barrier: a detailed study of these charges actually allows one to predict which distortions affect (or do not affect) the barrier height.

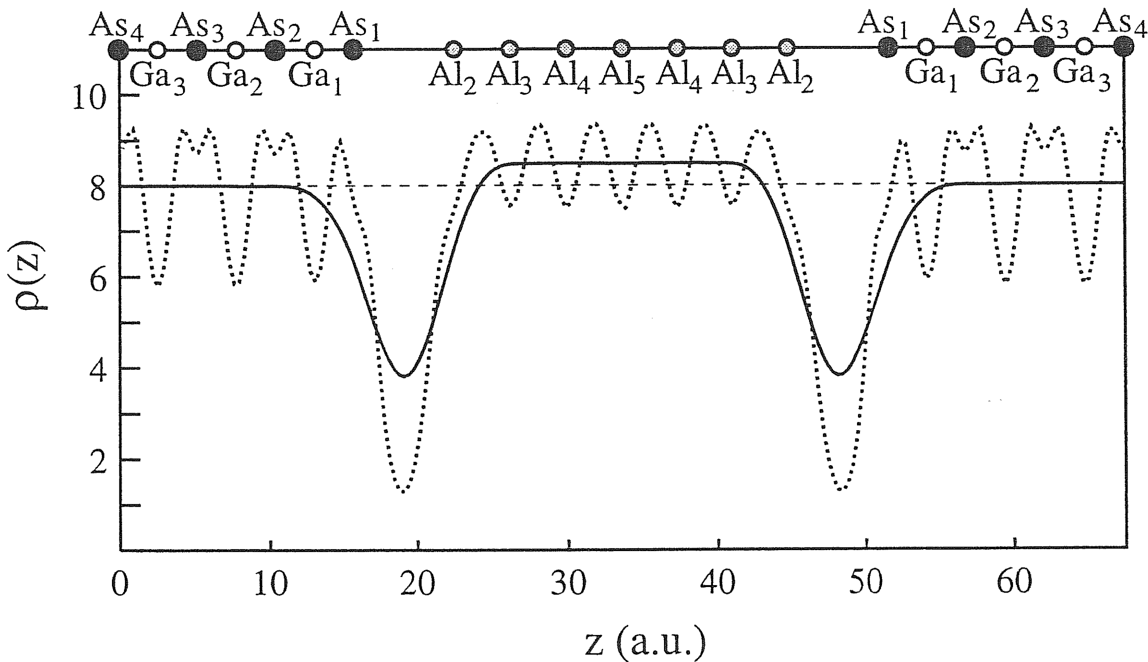

Figure 6.1: Planar (dotted line) and macroscopic (solid line) average of the electronic charge density for the (001)Al/GaAs supercell where two layers of Al have been removed.

6.1 Thick layer of vacuum between metal and semiconductor

Using the same supercell geometry as in the previous Chapter and maintaining all the technical ingredients described therein unchanged, we apply as a first step a very strong morphology perturbation by breaking the Al-As interface bond, in order to investigate to which extent the Al-As distance affects the SB height. Technically, we perform the calculation removing the Al layer nearest to the interface, so that the layer of vacuum inserted amounts to 7.0 Å.

The planar and macroscopic average of the electronic charge are displayed respectively in Fig. 6.1. We find in this way a barrier of -0.24 eV, very much different from the previously calculated value of $\Phi_B^p = 0.73$ eV: in this case the barrier is even not present, because the Fermi level is not pinned within the semiconductor gap. This result provides further evidence (if any was needed) that the early Mott-Schottky model—where the identity of the two quantities was postulated—is invalid; the barrier formation is then strongly dependent on the interfacial bonding and cannot be regarded as a property of the two separate bulk solids.

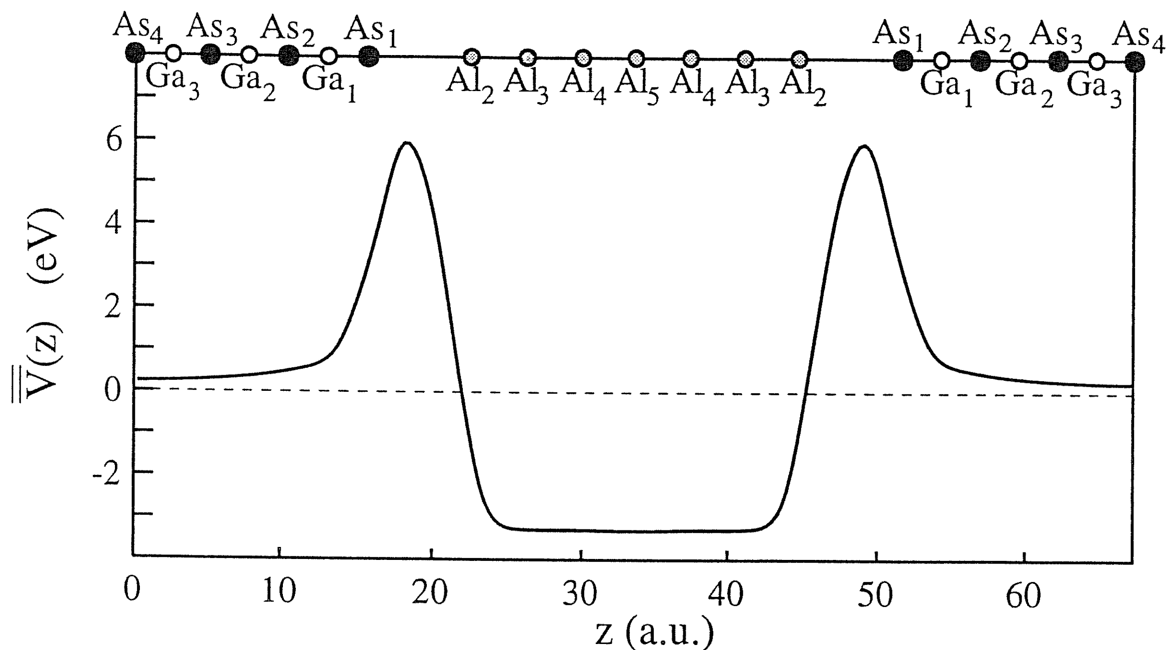

Figure 6.2: Macroscopic average of the total potential for the (001)Al/GaAs supercell where two layers of Al have been removed.

6.2 Bond elongations

We consider then a very *thin* layer of vacuum inserted between the metal and the semiconductor: instead of breaking the Al–As bond, we thus gently elongate it while keeping the rest of the structure rigid (the length of the supercell is elongated accordingly). Such a displacement is commonly referred to as *interfacial strain.* The Schottky barrier is found to depend very weakly upon interfacial strain: it takes in fact a strain as large as 3% in order to vary Φ_B^p by 0.01 eV, our estimated numerical accuracy. With the (enormous) value of 10%, Φ_B^p varies by about 0.04 eV.

Next we perform an analogous 10% elongation, but on the Ga–As bond nearest to the interface: we get in this case the much larger variation of 0.09 eV. The more this (10%) bond elongation involves adjacent layers placed towards the center of the semiconductor slab, the more the barrier height comes out to be affected. On the other hand, the application of the same kind of perturbation within the metal slab leaves the barrier unchanged (within our

numerical accuracy).

We give below a simple rationale for such different dependence of Φ_B^p on different local strains: we will see that the dynamical charges of interface ions play a major role.

6.3 Bulk strain of the metal overlayer

The next step is to consider the effect of *bulk* strain on the metal side. By the way, this kind of morphological perturbation goes in the direction to recover in our theoretical calculation the actual interfacial geometry (as already pointed out in previous Chapter, actual Al needs to be slightly strained to accommodate the lattice mismatch, according to the elasticity theory for dislocation-free, coherent epitaxy).

In our case, Al is biaxially compressed in the interface plane, and therefore, due to the epitaxial geometry, it is tetragonally elongated along the z direction (perpendicularly to the interface); this elongation can be obtained using the Hooke law of elasticity theory, by considering Al as a linear, homogeneous medium. Its cubic symmetry reduces the number of independent elastic constants to three, namely C_{11}, C_{12}, and C_{44}. The deformation in the interface plane ϵ_{xx} and ϵ_{yy} are identical and imposed by the substrate, while the value of ϵ_{zz} to minimize the elastic energy at zero pressure is:

$$\epsilon_{zz} = \mu\epsilon_{xx}, \tag{6.1}$$

where $\mu = 2C_{12}/C_{11}$ in the case of (001) orientation.

Using the experimental [87] elastic constants of Al (i.e. $C_{11} = 108$, $C_{12} = 62.0$ and $C_{44} = 28.3$) and the fact that $\epsilon_{xx} = 0.92\%$, we obtain $\mu = 1.15$ and therefore $\epsilon_{zz} = 1.06\%$.

Following the same lines as in Section 5.4, we set the distance between metal and semiconductor as the average of the metal-metal (strained) distance and the semiconductor-semiconductor distance: this choice minimizes the forces acting on interfacial atoms, and hence the relaxations

effects on the SB height, as also shown in Ref. [84].

The calculated Φ_B^p is completely insensitive to the tetragonal strain ϵ: we actually found a Φ_B^p variation of 0.01 eV. Compared to previous theoretical work demonstrating that the Schottky barrier for a given semiconductor seems to vary with the nature of the metal [4], this finding is rather unexpected. We find in fact that the barrier is unchanged in the special case considered, namely two metals having the same chemical composition but different lattice parameters, hence different electronic densities.

We elaborate a little bit more about these findings, which give insight into the robustness of Φ_B^p and shed some light on the very important—although disturbingly vague—concept that the barrier is formed extremely close to the semiconductor [2]. Imagine an ideal double interface, where the semiconductor is joined to a first metal, and then the first metal is joined to a second metal. The barrier forms at the semiconductor-metal interface, and then—if the middle slab is thick enough—remains constant through the second interface, since the Fermi level is aligned across any metal-metal contact. This *transitivity rule* is not expected to hold when the thickness of the middle slab is reduced. Instead, in our case study a macroscopic slab is not needed—not even a microscopic one—in order for the barrier to be robustly established.

As a double check of our transitivity finding, we scrutinize separately the two contributions [see Eq. (5.4)] to the barrier height ΔE_p (which correspond to the whole band structure term, $\Delta E_p = \Delta E_F - \Delta E_v$) and ΔV. While their sum turns out to be ϵ-independent, their individual variation is sizeable. With the above value of $\epsilon = 0.01$, the calculated ΔE_p varies by -0.10 eV: we wish to compare this to the ΔV value at an ideal strained/unstrained metal homojunction. To this aim, a supercell calculation is unnecessary: ΔV is a pure volume effect, and we get it by calculating the deformation potential [88] of bulk Al, *i.e.* the linear variation of the Fermi energy, measured with respect to the average of the electrostatic potential. We find in this way $\Delta V = -0.11$ eV, in very good agreement with the above value.

6.4 Interface study based on the dynamical charges

The next probe we are going to use in order to test the robustness of the barrier height, are displacements of individual atoms, while the rest of the structure is kept fixed. The basic quantities measuring the response of the electronic system to such perturbations are the effective charges for lattice dynamics. As already pointed out [see Eq. (2.40)], a displacement of an ionic plane in the bulk semiconductor by an amount u creates a dipole per unit area; this induces a potential lineup, which is proportional to the longitudinal dynamical charge of the given ionic species [as shown in Eq. (3.39)].

Bulk Z^* calculation have been performed using planar displacements in pure Al (or pure GaAs) supercells, and ingredients analogous to those used throughout in this thesis. For Al we get $|Z^*|$ smaller than 0.005 (in any bulk metal Z^* vanishes due to perfect screening); for GaAs we get instead $Z^* = \pm 0.18$, which agrees with the experimental value within a few percent (remind that we are dealing here with *longitudinal* charges).

The calculation of the dynamical effective charges of the different ions across the metal-semiconductor junction gives a way to monitor the transition between the two bulk materials and provides a very meaningful measure of the interface thickness. Actually, a structural distortion may affect (to linear order) the electrostatic lineup—and hence the barrier Φ_B^p—only if it displaces ions whose Z^* is nonvanishing.

Our computational supercell, considered as a whole, is a metallic system. Although nothing prevents in principle the use of DFPT for a metal, our present implementation is limited to insulators and does not allow the use of smearing techniques. Furthermore, the quantities which are directly computed by the present code are Z_T^* and ϵ_∞, where ϵ_∞ refers to the dielectric constant of the supercell and therefore diverges. The most appropriate way to calculate the Z_L^*'s for such systems is then the direct method described in Section 3.3. The dynamical charges Z_L^* are in fact directly provided through Eq. (3.39), with a typical u value of 0.03 a.u.;

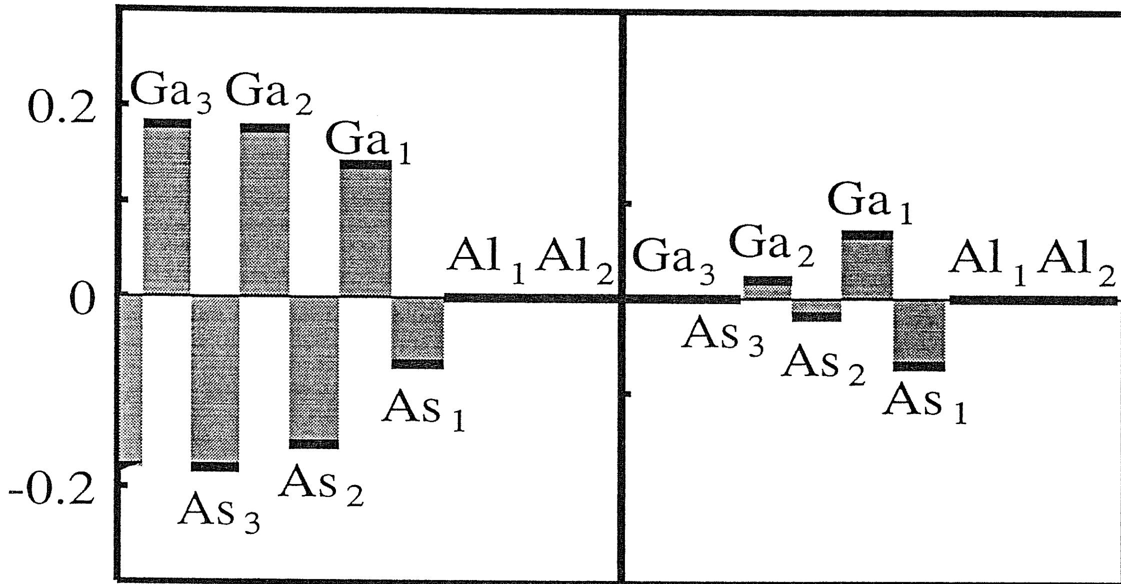

Figure 6.3: The left panel shows the calculated dynamical charges for the Al/GaAs(001) interface, in the form of an hystogram; the thickness of black lines indicate our numerical accuracy in the calculation. The right panel is the macroscopic average of the left one: it shows the averages of the Z*'s over a segment, centered at a running point, and whose length equals the periodicity of the bulk semiconductor region. The plot illustrates the dynamical neutrality of the interface, and also shows that the interface region is more extended on the semiconductor side than on the metal one.

a conservative estimate of the numerical accuracy of our Z*'s is 0.01. When approaching the interface from the semiconductor side, our calculated Z^* values are: +0.18 (Ga_2), −0.15 (As_2), +0.14 (Ga_1), −0.07 (As_1). Entering into the metal, the calculated Z^* drop rapidly to their (vanishing) bulk value. Since there are two nonequivalent Al atoms per plane, we displace each of them at a time. We get ±0.01 for Al_1, and −0.01, +0.02 for the Al_2 atoms. The figures obtained for Al atoms (also shown in Fig. 6.3) have been rounded to 0.01: their apparent differences being of the order of our numerical accuracy.

One important message emerging from our calculated Z*'s is that—as far as the dynamical charges are concerned—the interface is very sharp on the metal side, while instead a semicon-

ductor ion "feels" the presence of the metal up to a depth of a few layers: the closest cation (Ga_1) is already strongly "nonmetallic", though not yet bulklike. Although Z^* is a *linear* property of lattice distortions, our calculations indicate that a structural defect on the metal side—even very close to the junction—would have a little effect on Φ_B^p; while on the contrary a defect on the semiconductor side is likely to have a sizeable effect. Of particular importance to the barrier height are therefore the detailed arrangements of the semiconductor atoms closest to the metal (given that noncentrosymmetric structural defects deep in the semiconductor can be ruled out). This sensitivity of the barrier height to the morphology of the first few semiconductor layers is in qualitative agreement with the findings of other authors [4–9], who have considered chemical defects in an otherwise undistorted structure.

6.5 Dynamical charge neutrality at Schottky barriers

In this section we recall and generalize to the case of an interface between a pair of semiinfinite crystals (like the case of Schottky barriers) our findings concerning the dynamical charge neutrality at the surface of a semi-infinite crystal, which were the main topic of Chapter 4.

We can focus on the semiconductor dynamical charges, because the metal region of the junction plays a role similar to the vacuum region in the surface case analyzed in Chapter 4, due to the fact that the metal dynamical charges vanish. Our computational findings strongly support the predictions of the novel surface sum rule, as stated in Eq. (4.5) [or, equivalently, in Eq. (4.7), valid in the case of (001) growth direction]. If the sum of the GaAs dynamical charges is performed over four monolayers (up to to the Ga_2 plane in Fig. 6.3), we get in fact:

$$\sum_{s=1}^{4} Z_s^* = 0.09 = \frac{Z_{\mathrm{GaAs}}^*}{2} \tag{6.2}$$

The (001) growth-direction case, where two equivalent polar interfaces are possible, is a very special case, as pointed out in Section 4.2.2. In fact Eq. (4.7) can be inferred in this case even by

simple and intuitive arguments. Actually, the interface sum rule can be directly derived from the previously known (molecular) sum rule for the finite slab, which can be interpreted as a "dynamical neutrality" of the supercell as a whole; since our supercell contains two equivalent interfaces, the ASR obviously implies the dynamical neutrality of each of them separately. Each of the interface regions can be considered as one half of the supercell, and clearly the sum of the Z^* must vanish in each of them. Because the semiconductor slab has n cations and $n+1$ anions ($n = 6$ in the actual calculation), the central anion must be reckoned with weight *one half* in summing the dynamical charges of each interface. From this reasoning, Eq. (4.7) follows. Nevertheless, it is worth stressing here again that the dynamical neutrality condition expressed by Eq. (4.7) is indeed a property of each interface individually.

6.6 Discussion

The dynamical charges are very closely related to the lineup induced (to linear order) by interface strain, as first shown in Ref. [89] for the similar case of a semiconductor-semiconductor heterojunction. In the present case we have independently calculated the effect of interface strain (see above) and found that it is very small. More precisely, we find zero Φ_B^p variation (within our computational tolerance) when the bond-length elongation is comparable to the one used in calculating the Z^*'s.

We are now able to provide a rationale for such findings. Actually, the explanation lies in the fact that all the dynamical charges on the metal side are extremely small. Let us think of an isolated junction between two semiinfinite bulks: the interface strain amounts to a rigid relative translation. Suppose first that the semiconductor is kept fixed, and that the metal is displaced: by linearity, the lineup induced by the displacement of the semiinfinite metal is the sum of the lineups induced by the displacement of individual metal planes, and this sum is close to zero using our calculated Z^* values. We wish to recover an identical result when

we keep the metal fixed, and we displace the semiconductor instead: this looks less trivial, since the dynamical charges oscillate indefinitely in the semiconductor bulk. We have shown in Chapter 4 how to regularize such an indeterminate sum using the appropriate physical criterion: the dynamical-neutrality condition for the interface, which is fulfilled for the present Al/GaAs(001) calculation, is then the crucial property ensuring the correct result.

7 Interface electronic states at Schottky barriers

In the previous Chapter we performed a study of the interface morphology by means of the dynamical charges and we found out that they monitor the thickness of the interface region in a very perspicuous way. The main message is that the interface is very sharp on the metal side, while instead the semiconductor "feels" the presence of the metal up to a depth of few layers. The reason for this is the presence of evanescent states in the forbidden gap of the semiconductor, which are filled whenever their energy is lower than the Fermi energy of the metal. In this Chapter we investigate the basic features of these electronic states, their decay length within the semiconductor, and their robustness under metal deposition. Moreover, we will investigate whether and how these properties change for a SB junction whose Fermi level is not strongly pinned in the semiconductor gap. With strong (weak) pinning we mean that the SB height variations are small (large) when the metal overlayer is varied. The case of GaAs is one of strong pinning (variations of less than about 1 eV), whereas ZnSe is one of weak pinning (variations of some tenths of eV [90]).

7.1 Calculation of interface states for Al/GaAs(001)

As a first step, the local density of states has been calculated from the KS orbitals of the supercell:

$$N(\epsilon, \mathbf{r}) = \sum_{n,\mathbf{k}} \psi^*_{n,\mathbf{k}}(\mathbf{r})\psi_{n,\mathbf{k}}(\mathbf{r})\delta(\epsilon - \epsilon_{n,\mathbf{k}}) \tag{7.1}$$

[see also Eq. (B.4)]. The δ-function is smeared according to the prescriptions reported in Appendix B, and we have chosen a width $\sigma = 0.14$ eV. Then we make use once more of the macroscopic–average technique (Appendix A) in order to filter out irrelevant oscillations having the periodicity of the two lattices (metal and semiconductor): we indicate the macroscopic average of the local density of states as $\overline{\overline{N}}(\epsilon, z)$. The analysis of this quantity allows to unambiguously demonstrate the exponential character of the evanescent states and hence to rigorously measure their decay length in the semiconductor region. Moreover, it is possible to monitor the whole band gap by checking the exponential character of interfacial states while ϵ is swept across the gap, from the valence band top to the conduction band bottom.

The macroscopic average of the local density of states for the Al/GaAs(001) 31-atom supercell is shown in Fig. 7.1, top panel. The function has been computed at the Fermi energy ϵ_F which is selfconsistently provided by the supercell ground–state calculation, Eq. B.6. In order to display the exponential tail of the evanescent states, the same function $\overline{\overline{N}}(\epsilon_F, z)$ is plotted in a semilogarithmic scale in the bottom panel. It is easily realized that the fit with a straight line is almost perfect in the region of the first semiconductor layers: the fit indicates a decay length of $\simeq$ 4.9 a. u. We are thus able to detect the metal induced gap states (MIGS) (see Section 5.1) through the computation of $\overline{\overline{N}}(\epsilon, z)$. The exponential character of these states is preserved also when ϵ is swept below the Fermi energy (for an energy interval of $\simeq$ 0.8 eV, hence up to the valence band edge of the semiconductor) and above it (for an energy interval of $\simeq$ 0.7 eV, hence approximately up to the conduction band edge); we remind that our DFT–LDA gap is 1.45 eV.

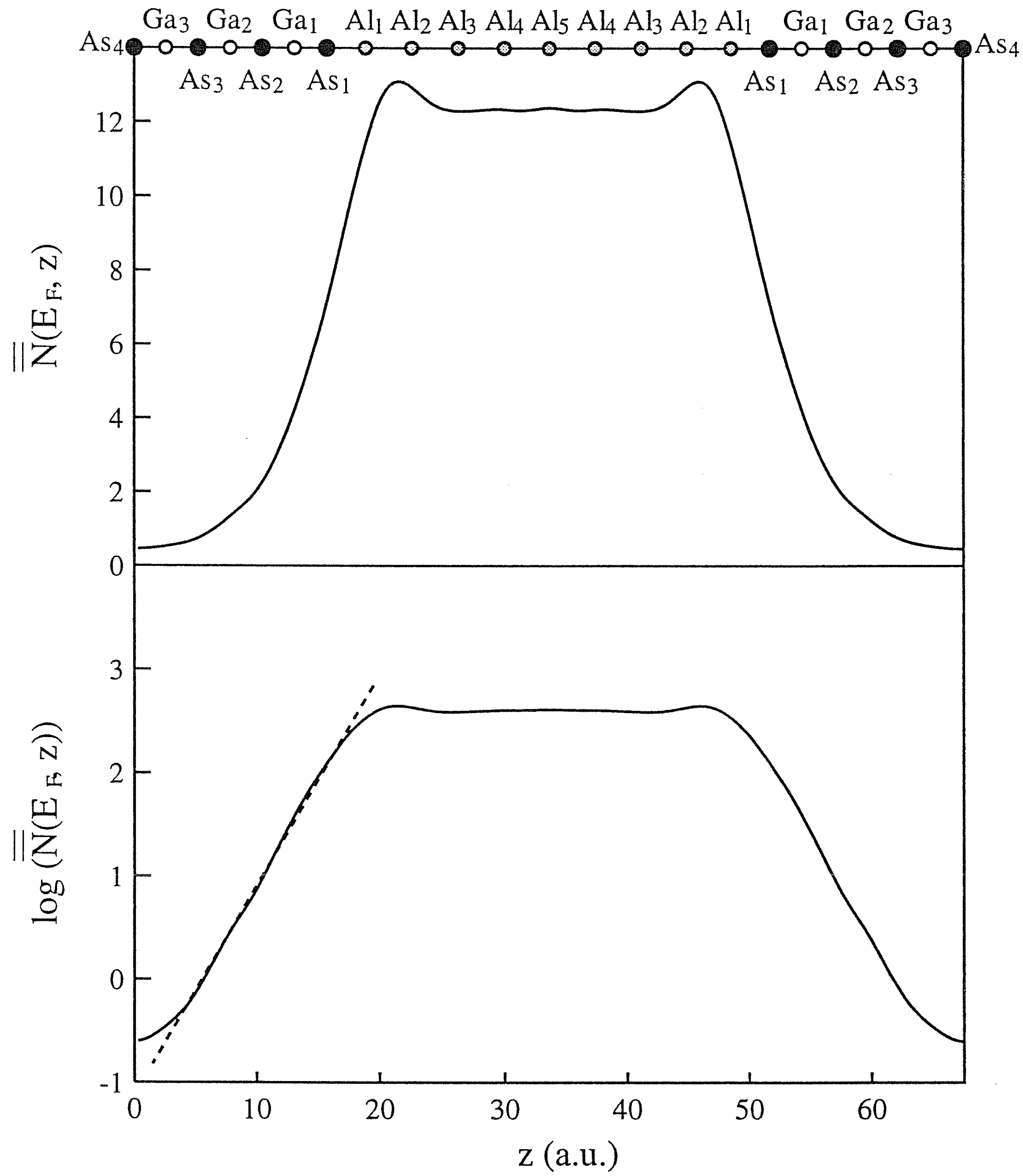

Figure 7.1: Plot of the macroscopic average of the local density of states $\overline{\overline{N}}(\epsilon_F, z)$ (top panel)), and semilogarithmic plot of the same function (bottom panel). In the semiconductor region adjacent to the interface the function is accurately linear.

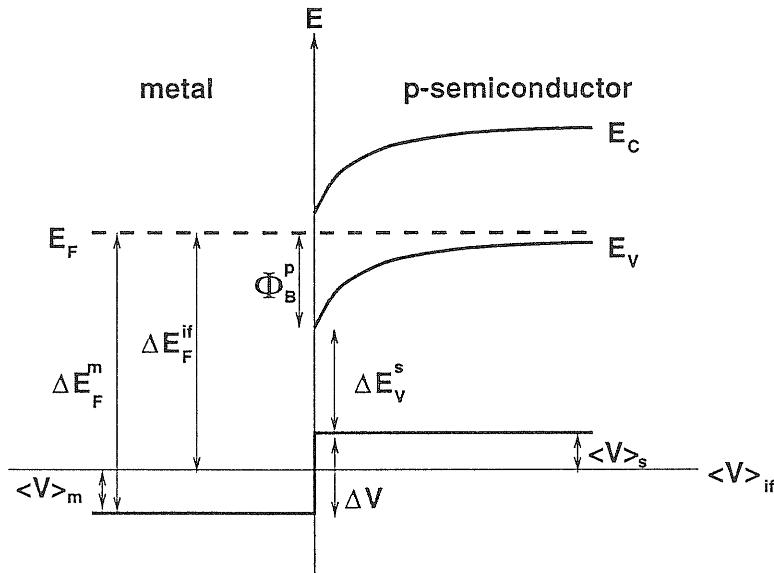

Figure 7.2: The Schottky barrier height (schematic). The specific contribution of the supercell Fermi energy, calculated with respect to the average electrostatic potential of the interface, is identified (ΔE_F^{if}).

The accuracy of the interface–states analysis could be affected by some uncertainties in the evaluation of the Fermi level position ϵ_F in the supercell calculation. In our previous determination of the SB, as explained the previous Chapters, we have *not* used this ϵ_F value, and we have instead performed an independent calculation for the bulk metal in order to achieve greater numerical accuracy. In the present Chapter it is instead expedient to use homogeneous data, all originated by the same selfconsistent calculation: we check however that the two approaches provide similar results (in fact within 0.01 eV) for the SB value. In order to make this point better clear, we have redrawn Fig. 7.2 in a slightly different way as Fig. 5.2. Using the same symbols as in the figure, the raw Fermi energy ΔE_F^{if} in the computer output is measured with respect to the average of the electrostatic potential in the supercell $\langle V \rangle_{\text{if}}$, a quantity which has no physical meaning. The SB height can equivalently be expressed as:

$$\Phi_B^p = \Delta E_F^{\text{if}} - \Delta E_v + \langle V \rangle_s, \tag{7.2}$$

where $\langle V \rangle_s$ is also referred to the average electrostatic potential of the interface and is computed from the value of $\overline{\overline{V}}(z)$, Fig. 5.7, at a z value well within the semiconductor, far from the interface. Our supercell calculation provides ΔE_F^{if}=7.39 eV and $\langle V \rangle_s = 1.47$ eV; by recalling the value for ΔE_v (=5.17eV, as reported in Section 5.3.1), we finally get Φ_B^p=0.74. This value for the SB height agrees very well with the one obtained in Chapter 5 through Eq. 5.4 (Φ_B^p=0.73), and this fact unequivocally demonstrates the reliability of the calculated interface Fermi energy used in the present analysis.

7.2 The (uncovered) GaAs(001) surface

The As-terminated GaAs(001) surface is already metallic even with no coverage, and has therefore evanescent states in the gap. It is very interesting to investigate to which extent these states bear some resemblance (if any) to the ones which exist in the case of metallic coverage, displayed in the previous Section. The results reported in Section 6.1 already demonstrated that the barrier formation is dramatically dependent on the presence of the metal-semiconductor bond at the interface: once this is broken, the Fermi level is no more pinned within the semiconductor gap. Therefore we do not expect much resemblance.

In order to investigate this issue, we use the same supercell geometry as in the previous Section, where all Al atoms are removed. As a preliminary step, we compute the planar average of the electronic charge density, and the corresponding macroscopic average: the charge profile of GaAs/vacuum interface is shown in Fig. 7.3. The figure makes evident that that the GaAs bulk-like behavior is recovered up to near the interface. Another important point is worth mentioning. Since our supercell is metallic, we calculate a selfconsistent Fermi energy even when no bulk metal is present. This defines in fact a surface Fermi energy: its calculated position is 0.01 eV higher than the valence band top. This is much different from the SB heigth (0.73 eV), which by definition is the analogous quantity when the bulk metal is present.

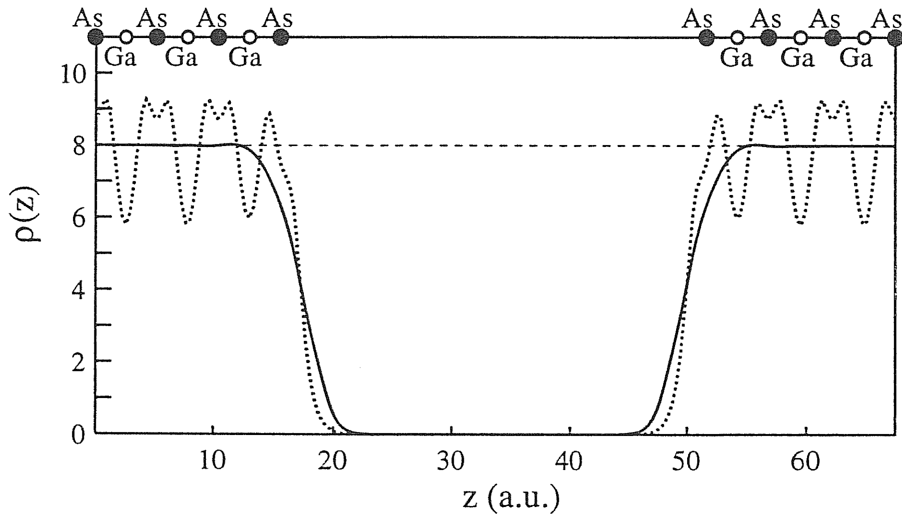

Figure 7.3: Planar (dotted line) and macroscopic (solid line) average of the electronic charge density for the GaAs(001) surface.

We accordingly expect that a study based on the inspection of $\overline{\overline{N}}(\epsilon, z)$ for the GaAs(001) surface should evidentiate large differences with respect to the Al/GaAs(001) interface. Actually, this comparison (shown in Fig. 7.4) demonstrates that the As-terminated GaAs(001) surface, although metallic, displays electronic states within the semiconductor gap which are completely different with respect to the case where the metal is there.

7.3 The monolayer-coverage case

The low-coverage regime plays an important role in the SB problem, because it is experimentally well-known that Fermi-level pinning occurs even at very low metal coverages; in particular it has been found [3] that at more than a monolayer-coverage the position of the Fermi energy converges rapidly to its bulk value.

We investigate in particular the monolayer-coverage case. To this aim, we use the same supercell described in Section 5.4, where we remove all metallic layers but the one nearest to the semiconductor substrate; in this way we maintain unchanged all the technical details as

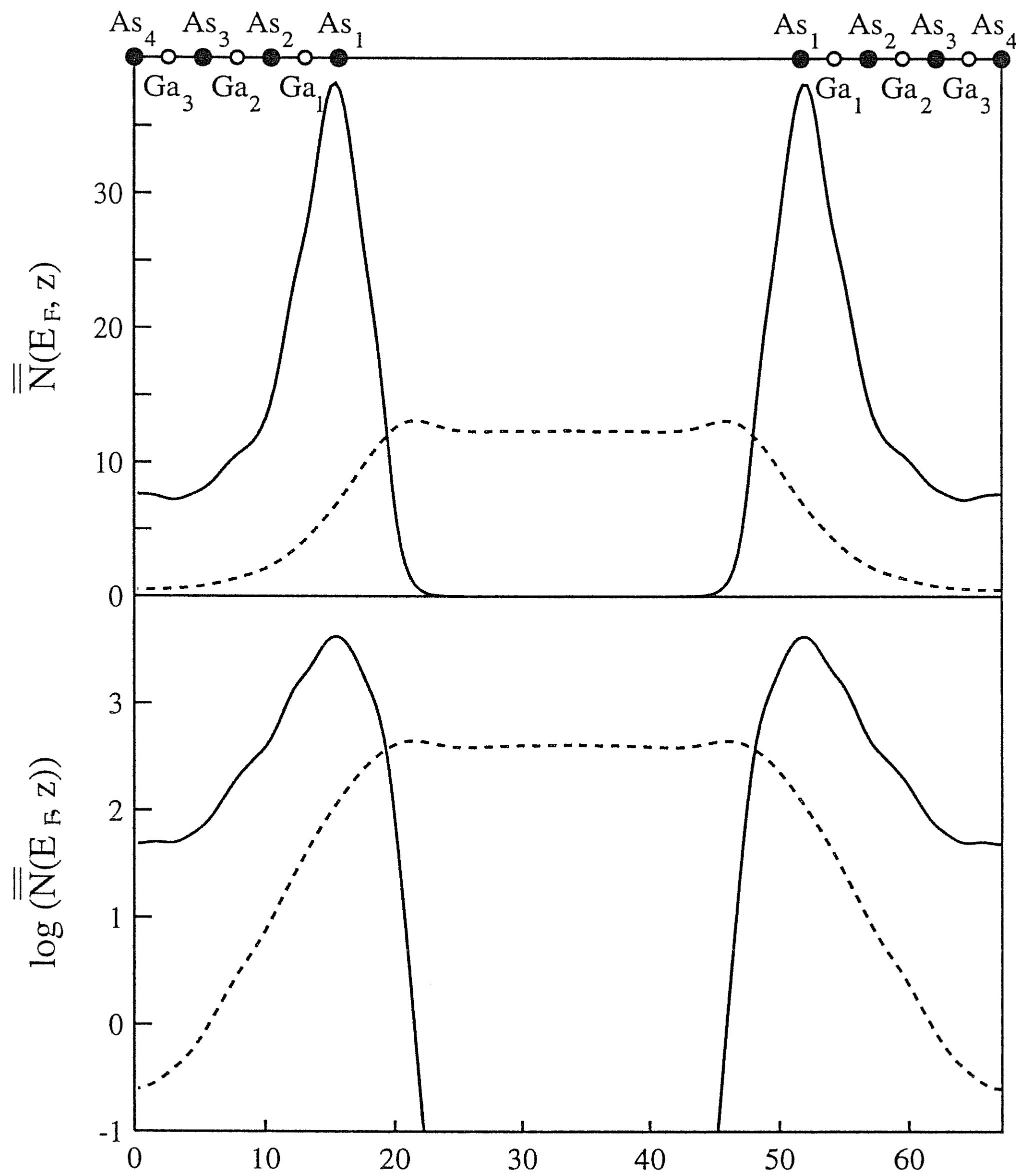

Figure 7.4: In the top panel, $\overline{\overline{N}}(\epsilon_F, z)$ for the GaAs(001) surface is reported (solid line). As a comparison, $\overline{\overline{N}}(\epsilon_F, z)$ for the full-coverage case is also shown (dashed line). The bottom panel displays the same quantities on a semilogarithmic scale.

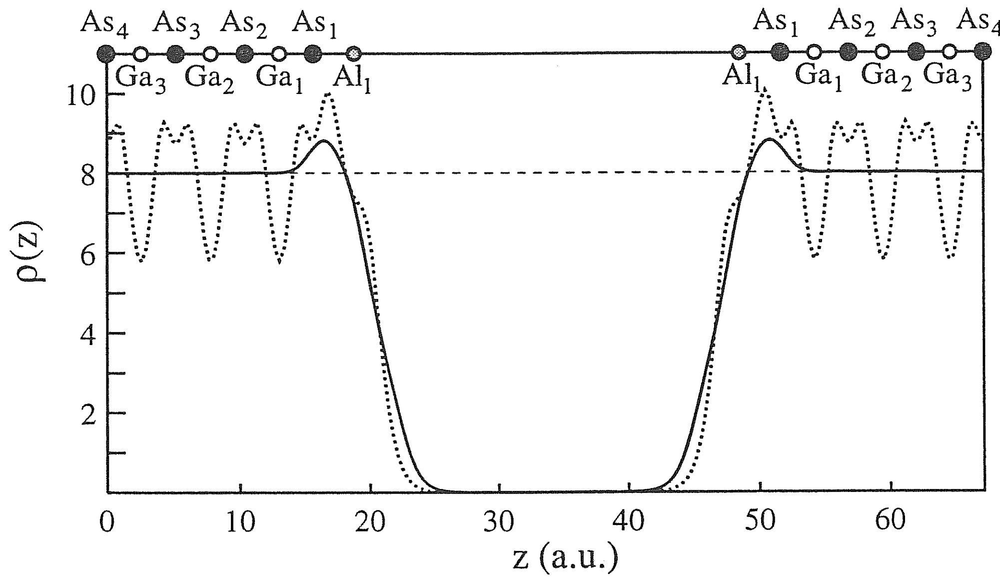

Figure 7.5: Planar (dotted line) and macroscopic (solid line) average of the electronic charge density for the Al/GaAs(001) junction, in the monolayer-coverage regime.

for the full-coverage case.

In Fig. 7.5 the calculated planar and macroscopic averages of the electronic charge density are shown, from which, through the Poisson equation (along the lines expounded in Section 5.4), it is easy to derive the total macroscopically averaged electrostatic potential (see Fig. 7.6).

Also in this case of metal low-coverage, the Fermi level position (and therefore the SB height) cannot be determined through Eq. 5.4, because the bulk properties of the metal monolayer are ill-defined quantities. The SB height can then be provided again by Eq. 7.2: with the values of $\Delta E_F^{\mathrm{if}} = 2.26$ eV and $\langle V \rangle_s = 1.61$ eV (and recalling that $\Delta E_v = 5.17$ eV), the SB value is $\Phi_B^p = 0.65$ eV. The SB is therefore well-established (actually in the full-coverage regime we get $\Phi_B^p = 0.73$ eV). In other words, this result demonstrates that the Fermi level is already pinned (within less than 0.1 eV) in the semiconductor gap even at low-coverage regimes, in semiquantitative agreement with the experimental findings. This feature receives further evidence from the calculation of the MIGS for the present case: the macroscopic average of the LDOS, $\overline{\overline{N}}(\epsilon, z)$, shows in fact exactly the same behavior as for the full-coverage

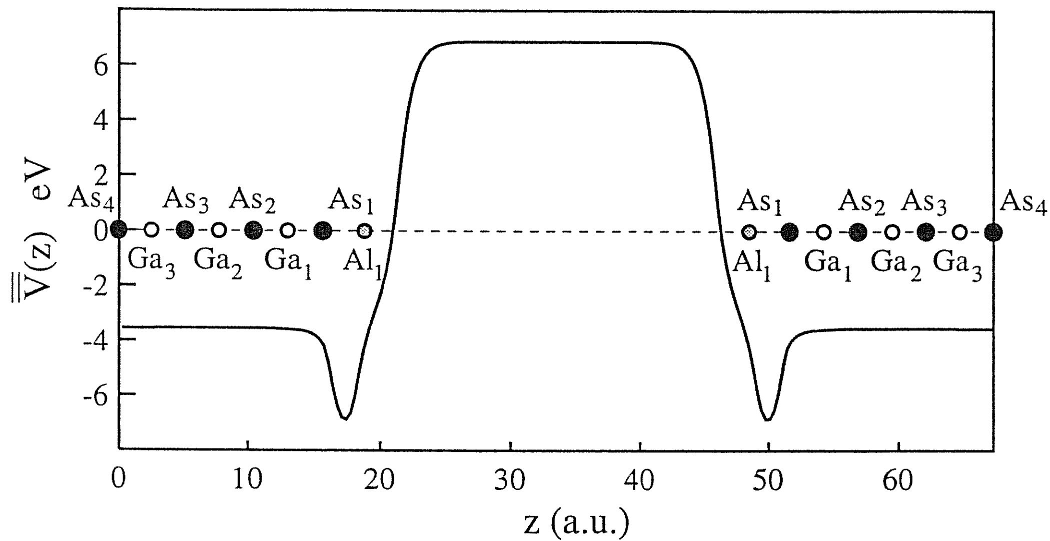

Figure 7.6: Macroscopic average of the total electrostatic potential for the Al/GaAs(001) junction, in the monolayer-coverage regime.

regime. As displayed in Fig. 7.7, the MIGS at the Fermi energy are in fact practically identical (notwithstanding the fact that the Fermi energy ϵ_F is slightly different in the two cases). In the semilogarithmic plot (bottom panel) one notices almost the same slope (reflecting the same exponential decay length), and even almost the same values (reflecting the same prefactor).

7.4 Al/ZnSe(001) interfaces

The pinning of the Fermi-level is typically encountered when the semiconductor involved in the junction has a covalent character, as is the case for GaAs which exhibits the relatively small ionicity of $g = 0.317$ on the Garcia-Cohen scale [62]. It could be interesting to analyze how the properties above reported for Al/GaAs(001) change when GaAs is replaced by a more ionic semiconductor, like ZnSe; actually, it is experimentally assessed [1] that ZnSe-based Schottky barrier junctions present a nonnegligible dependence on the metal involved, so that the Fermi level pinning is not well-established. Moreover ZnSe, which exhibits a large ionicity $g = 0.597$ [62], is lattice-matched with GaAs, allowing then a meaningful comparison between ideal,

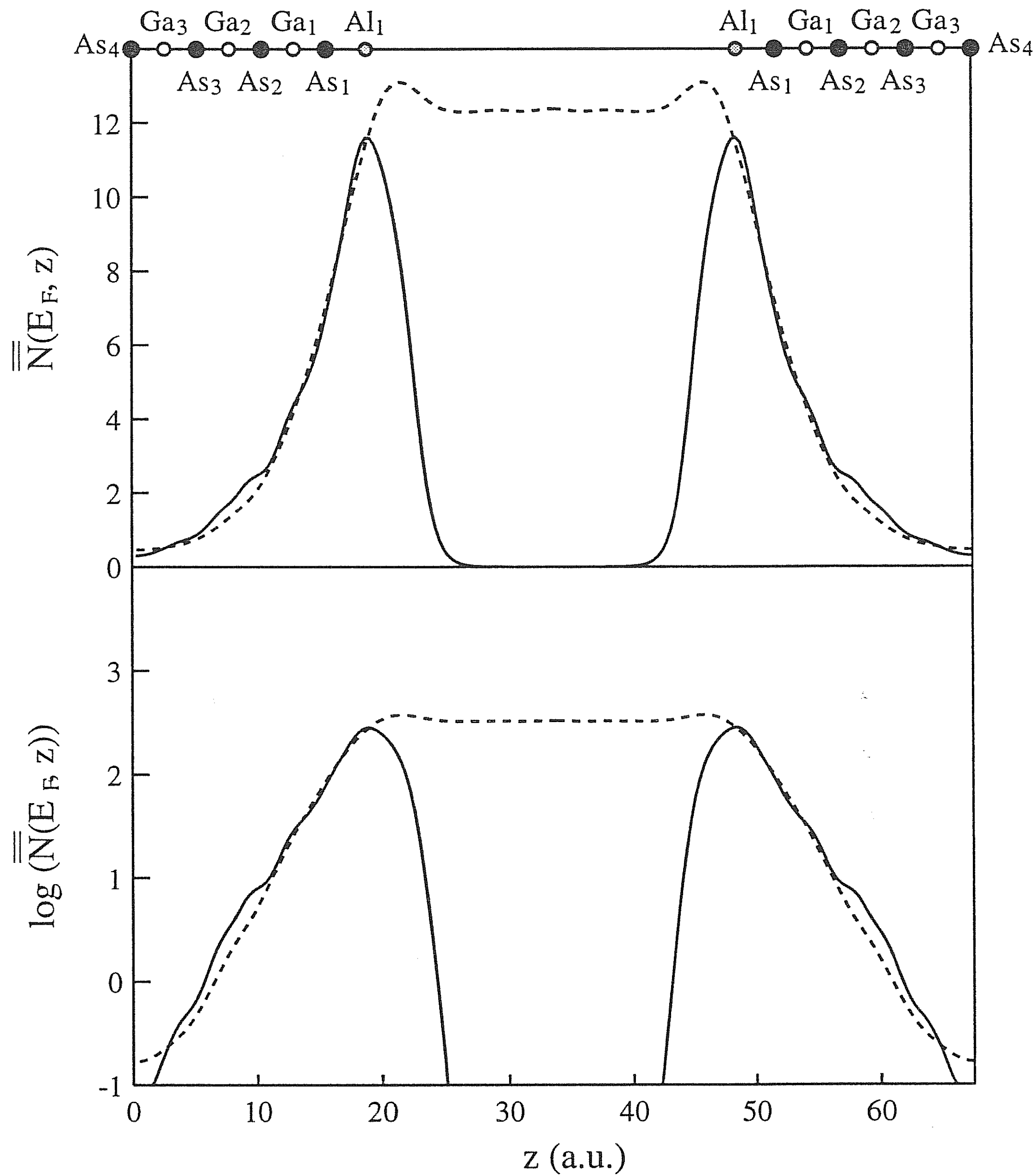

Figure 7.7: In the upper panel $\overline{\overline{N}}(\epsilon_F, z)$ for the monolayer-coverage case is reported (solid line). As a comparison, $\overline{\overline{N}}(\epsilon_F, z)$ for the full-coverage case is shown (dashed line). The lower panel displays the same quantities on a semilogarithmic scale.

Table 7.1: Bulk properties of ZnSe.

	$a_0(a.u.)$	$B_0(GPa)$	ϵ_∞	$Z^{*(T)}$	$Z^{*(L)}$
Theory	10.47	76.4	6.3	2.01	0.31
Expt.	10.71	66.7	6.3	2.03	-

defect-free interfaces.

7.4.1 Bulk properties of ZnSe

In order to describe Zn and Se atoms, we use pseudopotentials generated with the von Barth-Car (VBC) recipe [82]. The case of Zn pseudopotential deserves a special digression to explain how $3d$ electrons in zinc atoms can be conveniently treated in a pseudopotential scheme. Actually, the *frozen-core* approximation, which is at the basis of "standard" pseudopotential theory, holds when the energies of the core electrons are much lower than the valence electrons energies; in such a case, the overlap between the core and the valence charge is negligible and it is possible to separate the non-linear exchange-correlation term as $\epsilon_{xc}(n_c + n_v) \approx \epsilon_{xc}(n_c) + \epsilon_{xc}(n_v)$, including the core part in the pseudopotential. The case of zinc atoms is however more difficult to deal with, being the closed-shell of $3d$ orbitals quite close in energy to the $4s$ and $4p$ electrons energy; moreover, there is a large overlap between the core and the valence charges. In principle one may consider as valence electrons also d electrons, but this would require the use of a very large kinetic-energy cutoff, and hence, of high computational workload, within the standard norm-conserving pseudopotential technique or of the complex *ultra-soft* pseudopotential scheme [91]. A simpler solution of this problem is to adopt the non-linear core-correction (NLCC) scheme [92, 93], which consists in including the *total* charge rather than the *valence* only in the density-dependence of the exchange-correlation term, so that

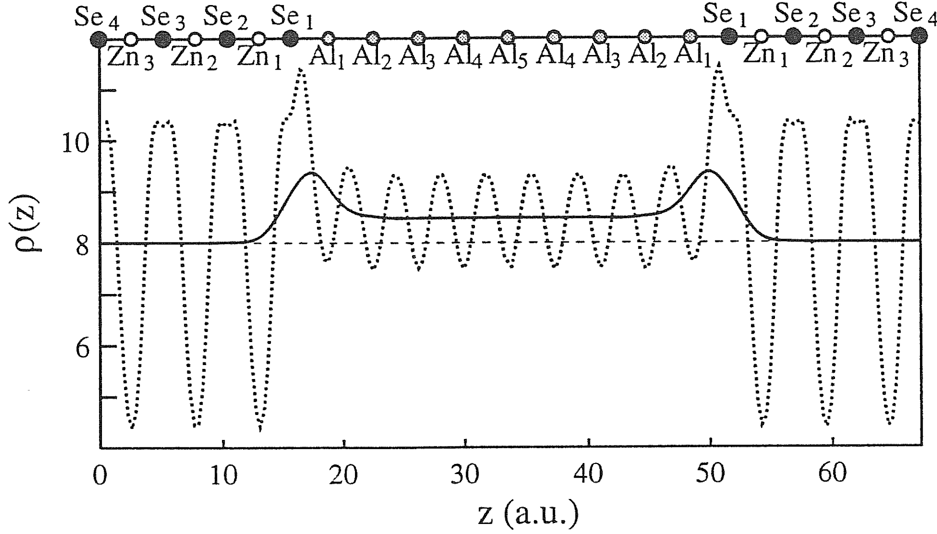

Figure 7.8: Planar (dotted) and macroscopic (solid) average of the electronic charge density for the Al/ZnSe(001) junction.

$E_{xc}[n] = \int d\mathbf{r}\epsilon_{xc}(n_c + n_v)(n_c + n_v)$.

The results are summarized in Table 7.1; the pseudopotentials used are able to reproduce the almost perfect lattice-matching condition with GaAs, and hence with Al (see also Table 5.1). The valence band edge for bulk ZnSe is $\Delta E_v = 3.47$ eV.

7.4.2 Schottky barrier height calculation

The Al/ZnSe interface has been described by using a 31-atom supercell (with 7 Se atoms, 6 Zn atoms, and 18 Al atoms), where the semiconductor slab has a double Se (anion) termination. Following the procedure reported in Chapter 5 for the Al/GaAs junction, we calculate the electronic charge density, the electrostatic potential drop across the junction, and hence the SB height for Al/ZnSe. We get $\Delta V = 3.12$ eV, and hence $\Phi_B^p = 2.06$ eV (through Eq. 5.4). Some recent accurate measurements [94] of the barrier height for Al/ZnSe(001) junctions gave for the p-type SB the result of 2.1±0.1 eV.

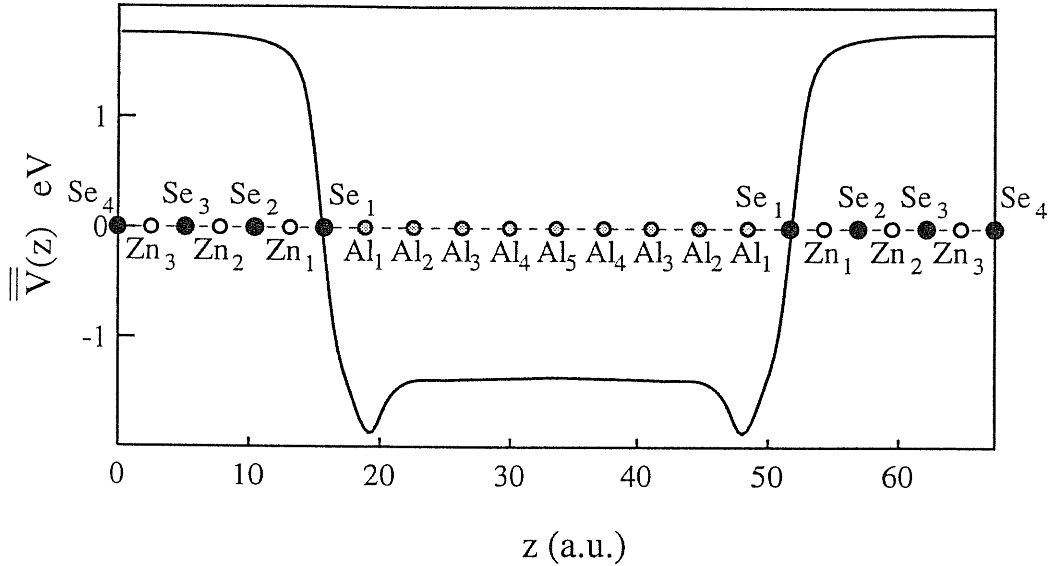

Figure 7.9: Macroscopic average of the total electrostatic potential for the Al/ZnSe(001) junction.

7.4.3 Dynamical charges

An analysis based on the interface dynamical charges (as reported in Section 6.4 for the Al/GaAs(001) case), is now applied to the Al/ZnSe(001) interface, in order to explore its morphological features. Using the recipe given by Eq. (3.39), the calculated Z^* values are: +0.30 (Zn_2), −0.27 (Se_2), +0.24 (Zn_1), −0.12 (Se_1), where Se_1 corresponds to the anionic plane closest to the interface with Al. Entering into the metal, the calculated Z^* vanish (and so recover their bulk value). The trend that we obtain is similar to the case of the Al/GaAs(001), and the dynamical charge neutrality condition is fulfilled [in the form given by Eq. (4.7)], but in this case the dynamical charges seem to converge to their bulk value (Z^*_{bulk}=0.31) in a slightly faster way. The interface thickness, when monitored through the dynamical charges distribution, is therefore slightly less extended in the ZnSe region than in the GaAs region.

7.4.4 MIGS-based analysis

The study of the metal induced gap states for the Al/ZnSe junction gives a quantitative, unambiguous support to the idea that weak pinning is related to small penetration of the

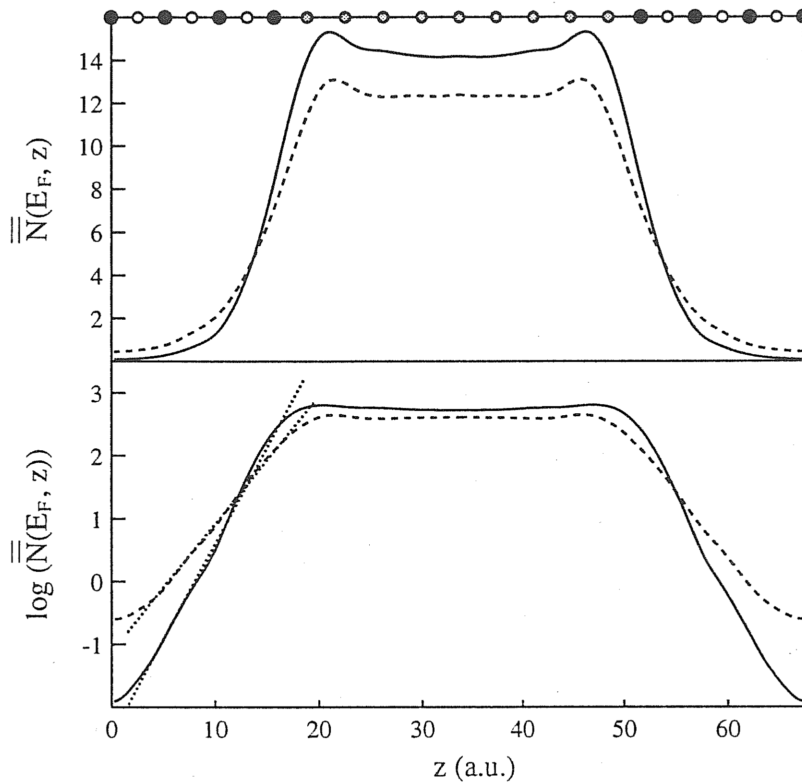

Figure 7.10: In the upper panel, $\overline{\overline{N}}(\epsilon_F, z)$ is plotted for the Al/ZnSe(001) interface (solid line) and for the Al/GaAs(001) case (dotted line). In the lower panel the same quantities are reported in a semilogarithmic scale. Black dots correspond to anion (Se or As), white dots to cation (Zn and Ga), while grey dots correspond to metal atoms (Al).

evanescent states in the gap. The exponential tails are shown in Fig. 7.10, where the analogous plots for the Al/GaAs case are shown for the sake of comparison. The decay length is about 3.3 a. u.. This is in qualitative agreement with previous theoretical calculations [95], but we are now able to unambiguously extract a value for the decay length, because of the power of our macroscopic–average technique.

The Fermi level pinning (i. e. the insensitivity of the SB height with respect to the metal) is related to the penetration depth of the MIGS in the semiconductor region. This can be

explained by considering that the less the MIGS penetrate the semiconductor, the less they will be involved in charge transfer between atoms of the semiconductor; in other words the less effective they will be at screening the interior of the semiconductor from local contributions to the interface dipole, which are dependent on the specific interaction between metal and semiconductor at the interface [77].

8 Conclusions

In this thesis we have investigated some basic issues concerning the Schottky barriers at metal-semiconductor heterojunctions. We have introduced in this study several innovative tools and concepts. The most important of them is the ionic dynamical charge tensor, which we have imported into the theory of the Schottky barriers, and we have shown to provide basic insight into the morphology–induced variations of the Schottky barrier height. As a by–product of our investigations, we have also found a result of more general scope, which applies to interfaces and surfaces on the same ground: we have called our theorem "dynamical neutrality".

Quite generally, the dynamical charges are the very fundamental constants measuring the coupling of electrical effects to ionic displacements in a given material. In particular, for crystalline solids, the dynamical charges are a dominating ingredient in the microscopic analysis of piezoelectricity, ferroelectricity, zone–center lattice dynamics, and more.

For both periodic solids and neutral molecules the dynamical charges are well-known to obey a dynamical neutrality condition, which enforces their sum to vanish (over the crystal cell, or over the whole finite system, respectively), while the case of a crystal surface (in general of a system where the translational symmetry is broken in one direction, like metal–semiconductor interfaces) was never considered before.

We have extended the dynamical-charge sum rule to the non trivial case of a crystalline surface, by showing that it *must* be dynamically neutral in order to ensure that a rigid trans-

lation of the semiinfinite solid as a whole does not affect the work function. In the interface case, our dynamical neutrality condition ensures that a rigid translation of the interface as a whole does not affect the Schottky barrier height.

Explicit formulation of our surface (or interface) neutrality requires in general the regularization of a nonconvergent sum. To this aim, it is necessary to take advantage of the appropriate physical criterion. We did this using another tool which is a recurrent leitmotiv all along the present thesis: the macroscopic average.

In this way we have discovered a constraint for the dynamical charges of the surface ions for any polar surface, such as (001) and (111) in the zincblende structure. The specific formulation of the sum rule depends on the particular orientation, but in general it *forbids* ions of a given chemical species to have the same dynamical charge in the surface region and and in the bulk region. The surface dynamical neutrality condition has been then demonstrated through first-principles calculations for (001)- and (111)-oriented β-SiC surfaces.

We have then focussed our interest on the Schottky barrier problem. Since real interfaces are "complex and dirty", first–principle calculations performed on idealized geometries provide a deal of unique "experimental" information. In this work, we have thoroughly investigated variations of the Schottky barrier heights induced by variations of the microscopic morphology, and found systematic trends which are explained by the concept of dynamical charge, and by the actual values of the ionic dynamical charges in the neighborhhood of the junction. Because of our sum rule, their value *must* be different from the bulk one.

All the calculations have been performed within the DFT–LDA framework, and using first–principle ingredients. Our main test case has been Al/GaAs, but we have performed some complementary calculations for Al/ZnSe, and other (not reported in this thesis) for Al/AlAs. Among the most significant results, we give quantitative evidence that the Schottky barrier height is very little affected by any structural distortions on the metal side—including elon-

gations of the metal-semiconductor bond (i.e. interface strain)—whereas it strongly depends on the interface structure on the semiconductor side. The rationale for this is found in the dynamical–charge analysis.

In the final part of this thesis—whose results are still in preliminary form—we have investigated the nature of the evanescent states in the semiconductor gap, and their role in pinning the Fermi energy of the metal. Our innovative tool here is the macroscopic average of the local density of states, which allows a very perspicuous analysis of the evanescent states, and a precise determination of their decay length. The most important result of this final part is summarized as follows. The unreconstructed, uncovered, semiconductor surface is metallic, but its evanescent states at the Fermi energy bear little resemblance to those of the metal-semiconductor junction. Instead, it is enough to cover the semiconductor surface with one monolayer of metal in order to get essentially the same evanescent states as for the thick–coverage case. In agreement with both common wisdom and experimental hints, we find that the formation of the metal–semiconductor bond is crucial to the pinning: we confirm that the barrier is essentially established at monolayer coverage.

Acknowledgments

I am so much indebted to so many people that it is impossible to give here an explicit acknowledgment to all of them; it is also difficult to express in few (english!) words what I feel. However, I will try...

First, I would like to thank my supervisors, for introducing me into this field of research and for being a so important guide in my formation. Raffaele Resta has continuously and very patiently supported this work, through enlightening suggestions and ideas, and his contagious enthusiasm. Stefano Baroni has been a source of invaluable–not only scientific–advice, that I will never forget.

I am very grateful to Stefano de Gironcoli, who always responded with the utmost patience to all my questions, even the most elementary ones...he is an (unreachable) example!

I am strongly indebted to Nadia Binngeli, Julien Bardi, Christof Berthod, and Alfonso Baldereschi, for very fruitful suggestions, and to Johannes Pollmann for precious discussions. I have also benefited by the interaction with Maria Peressi (always so nice and patient), and Andrea Dal Corso.

Thanks also to my office-mate Sandro Scandolo, for his friendly help.

In these three years, I really enjoyed the stimulating and friendly "climate" within the CM Sector; in particular I want to thank Erio Tosatti.

A special thank to Sabrina (and Pietro too!), for her smiling friendship.

Acknowledgments are also devoted to the computer staff (Marina, Luisa, Roberto, Davide, Dario, and Fabio), and to the administrative staff (Andrea & Riccardo [very funny, indeed!], Fabia, Claudia, Patrizia and Alex), for their kindness and efficiency.

Moreover, I do not want to forget a lot of people, with whom I shared a so good time here in Trieste:

Marco Saitta (with whom I really shared a lot of stuff: physics, (lot of) laughing, games (!!!), troubles, and, above all, a precious friendship), Barbara Montanari (for teaching me that true friendship does not decay with distance, for being so close to me, and ... for being so), Matteo Calandra (M.D.C., as spelled in english), Daniele Passerone (a special friend: past year was somewhat worse without him), Lorenzo and Antonella (a wonderful page-couple), the "littol" Cecilia Clementi and the cool Giovanni Fossati, my home-mates (Elena, Cristina, Sophie and the Lo), Stefano Martinelli (who in particular knows why!), Marco Fornari (always so nice), Carlo Cavazzoni (my favourite driver), Claudia Bungaro, Claudio Tebaldi, Stefano Giovanazzi, Franz Di Tolla, Orion Ciftja, Catia Lavalle, Gianluca Giallone Israel, Valentina Dodorico. Silvia Zane, Dario Alfe', Alessandro Laio, Michele Lazzeri, Antonio Trovato, Federico Becca, Stefano Serra, Franck Celestini, and many many many more, for their help and especially for their true friendship.

Finally, I thank all the people who have been always close to me with their encouragements and their love: my parents, my brothers, and all my friends.

The final words are for Andrea, for being always with me.

Appendix A

The macroscopic–average technique

We fix the following notation for the Fourier transform of a function f periodic on the supercell Ω:

$$f(\mathbf{G}) = \frac{1}{\Omega} \int_{\Omega} f(\mathbf{r}) e^{-i\mathbf{G}\mathbf{r}} d\mathbf{r} \tag{A.1}$$

$$f(\mathbf{r}) = \sum_{\mathbf{G}} f(\mathbf{G}) e^{i\mathbf{G}\mathbf{r}}. \tag{A.2}$$

The macroscopic average concept, which was already anticipated and applied in Section (3.1) for a particularly simple case, is a basic one in classical electromagnetism. Any macroscopic quantity $f^{macro}(\mathbf{r})$ is related to its microscopic counterpart $f^{micro}(\mathbf{r})$ through a convolution integral

$$f^{macro}(\mathbf{r}) = \int d\mathbf{r}' w(\mathbf{r} - \mathbf{r}') f^{micro}(\mathbf{r}') \tag{A.3}$$

where $w(\mathbf{r})$ is a real function, which differs from zero in some neighborhood of $\mathbf{r} = 0$ and is normalized to unity over all the space [96]. It is a trivial matter to show that such averaging commutes with the space and time differentiation occurring in Maxwell equations. The choice of the filter function $w(\mathbf{r})$ is largely arbitrary; however each macroscopic problem has its own appropriate lower limit of relevant lengths and this sets the size of the w function to be used [97]. When studying electrostatics in crystalline materials, periodicity allows to take length scale as small as $\Omega^{1/3}$, where Ω is the unit cell volume.

The application of this concept to the interface problem [80, 98] is enlightening: macroscopically averaged quantities show no microscopic oscillations on either side of the interface and recover the macroscopic limit of the two bulks, so that deviations from the bulk macroscopic values unambiguously define the interface region. In particular, this approach allows to get rid of a long-standing confusion about what is the interface dipole: the genuine interface features (and in particular the *true* interface dipole) are in fact blown up using the macroscopically averaged interface charge as the only ingredient, with no subtraction of arbitrary bulk-contribution and no choice of boundary.

In the interface case, where the periodicity is broken in one direction (e.g. in the z direction), the first obvious simplification for calculating the macroscopic average of a given physical quantity $f(\mathbf{r})$ (the density $\rho(\mathbf{r})$ or the electrostatic potential $V(\mathbf{r})$) is to consider the *planar average* $\overline{f}(z)$:

$$\overline{f}(z) = \frac{1}{\Omega_{||}} \int_{||} f(x, y, z) dx dy, \tag{A.4}$$

where $\Omega_{||}$ is the area of the supercell in the plane (xy) parallel to the surface. The *macroscopic average* $\overline{\overline{f}}(z)$ is then obtained by filtering $\overline{f}(z)$ through the one-dimensional convolution:

$$\overline{\overline{f}}(z) = \frac{1}{b} \int_{z-\frac{b}{2}}^{z+\frac{b}{2}} \overline{f}(z') dz' = \frac{1}{b} \int \Theta(\frac{b}{2} - |z - z'|) \overline{f}(z') dz', \tag{A.5}$$

where Θ is the step function and b is the interplanar distance along the z direction. This procedure is equivalent to average $f(\mathbf{r})$ over a slab-adapted bulk unit cell centered at the point $\mathbf{r}$ and therefore corresponds to the usual definition of macroscopic quantities in electrostatics (Eq. A.3).

From Eqs. (A.2) and (A.4) follows that in a reciprocal space formulation of the problem the planar average $\overline{f}(z)$ can be conveniently calculated as:

$$\overline{f}(z) = \sum_{G_z \neq 0} e^{iG_z z} f(\mathbf{G}_{||} = 0, G_z), \tag{A.6}$$

while, from Eq. A.5, the Fourier transform of the macroscopic average $\overline{\overline{f}}(z)$ takes the form:

$$\overline{\overline{f}}(G_z) = \overline{f}(G_z)\frac{sin(G_z b/2)}{G_z b/2}. \tag{A.7}$$

Let's consider the more complicated case where the two materials A and B which constitute the interface have two different periodicities. A filter function w able to recover macroscopic electrostatics in the bulk region of both materials can be obtained by filtering twice in succession, using the function w_A and w_B appropriate to each material:

$$w(\mathbf{r}) = \int d\mathbf{r}' w_A(\mathbf{r} - \mathbf{r}') w_B(\mathbf{r}') \tag{A.8}$$

which explicitly shows commutativity. The macroscopic average of $f(z)$ can be written as:

$$\begin{aligned} \overline{\overline{f}}(z) &= \frac{1}{b_A b_B} \int\int \overline{f}(z)\Theta(\frac{b_A}{2} - |z - z'|)\Theta(\frac{b_B}{2} - |z - z''|)dz'dz'' \\ &= \sum_{G_z \neq 0} \overline{\overline{f}}(G_z)e^{iG_z z}, \end{aligned} \tag{A.9}$$

where $\overline{\overline{f}}(G_z)$ can be easily calculated in reciprocal space as:

$$\overline{\overline{f}}(G_z) = \overline{f}(G_z)\frac{sin(G_z b_A/2)}{G_z b_A/2}\frac{sin(G_z b_B/2)}{G_z b_B/2}. \tag{A.10}$$

Appendix B

The smearing technique in Brillouin-zone integration for metals

Calculation of properties of the electronic ground-state (such as charge density, forces, interatomic force constants,...) requires integration of periodic functions over k-vectors in the BZ. In the case of metals the functions to be integrated are discontinuous at the Fermi level, due to the partial filling of the energy bands. This fact leads to a very slow convergence with respect to the k points density, when the integral is evaluated on a uniform mesh in the BZ. On the other hand, the use of a very fine k mesh is prohibitive, in practical implementations, for time- and memory-limit reasons. To overcome this problem, Methfessel and Paxton proposed an high-precision scheme for BZ integration in metals [42], which consists in a modification of the simpler Gaussian smearing technique of Fu and Ho [99].

The integrals in the BZ that we wish to evaluate have the following general form:

$$I = \int_{BZ} f(\mathbf{k})\theta(\epsilon_F - \epsilon(\mathbf{k}))d\mathbf{k}, \tag{B.1}$$

where $\epsilon(\mathbf{k})$ represents an energy band as a function of the wave vector, ϵ_F is the Fermi energy, and θ is the step function for the Fermi cutoff:

$$\theta(\epsilon_F - \epsilon(\mathbf{k})) = \int_{\infty}^{\epsilon_F} \delta(\epsilon - \epsilon(\mathbf{k}))d\epsilon \tag{B.2}$$

The k-convergence can be easily improved by broadening the δ function in Eq. B.2 into Gaussian, Lorentzian or similar smooth functions with characteristic linewidth σ. In fact, provided that the average energy separation between neighboring computed eigenvalues is small with respect to σ, the discontinuity arising from the step function is smeared out, and the integral can be computed accurately on a discrete **k** grid of points in the BZ.

However, the only justification for this *ad hoc* procedure is that the limit $\sigma \rightarrow 0$ one recovers the absolutely converged result at the expense of using a prohibitively fine mesh. Thus, for each choice of σ, the **k** sum converges to a different value, and convergence with respect to the broadening width σ must be further checked.

Methfessel and Paxton suggested an efficient way to achieve absolute convergence, based on a more sophisticated choice of the broadening function. They expand the delta function as:

$$\delta(x) = \sum_{n=0}^{\infty} A_n H_{2n}(x) e^{-x^2} \tag{B.3}$$

where $x = \frac{\epsilon - \epsilon(\mathbf{k})}{\sigma}$, H_{2n} are Hermite polynomials, A_n are the expansion coefficients, and σ is an arbitrary "linewidth".

By truncating the sum in Eq. B.3 to a finite order N, one obtains a broadened $\tilde{\delta}$-function and a corresponding smooth approximation $\tilde{\theta}$ for the step function B.2. The order $N = 0$ corresponds to the simple Gaussian broadening [99]. By construction, truncation to the finite order N leads to a negligible error in the evaluation of the integral B.1, if the function is representable as a polynomial of degree $2N$ or less in an interval of $\approx 5\sigma$ around the Fermi energy. The k-converged result can therefore be made to approach the true value either by increasing N or by reducing σ. Unless f is a constant near the Fermi level, the simple Gaussian smearing gives results considerably far from the desired *zero-width* limit, and a good convergence is obtained only for a very small linewidth. In many instances, the first-order approximation ($N = 1$) in the smearing function is enough to have a significant enhancement in precision with a minimal extra effort in the BZ zone sampling [100].

Within this approach, the local density of states is convoluted with the smearing function $f(\epsilon) = (1/\sigma)\tilde{\delta}(\epsilon/\sigma)$ (an approximation to the Dirac's δ-function that becomes exact for vanishing linewidth σ). The local density of states can now be computed accurately on a discrete grid of **k** points, provided the average energy separation between neighboring computed eigenvalues is small with respect to σ:

$$n(\mathbf{r},\epsilon) = \sum_{n,\mathbf{k}} \frac{1}{\sigma}\tilde{\delta}\left(\frac{\epsilon - \epsilon_{n,\mathbf{k}}}{\sigma}\right)\psi^*_{n,\mathbf{k}}(\mathbf{r})\psi_{n,\mathbf{k}}(\mathbf{r}) \tag{B.4}$$

From this basic quantity the electronic charge density follows:

$$n(\mathbf{r}) = \int_{-\infty}^{\epsilon_F} n(\mathbf{r},\epsilon)d\epsilon = \sum_{n,\mathbf{k}} \tilde{\theta}\left(\frac{\epsilon_F - \epsilon_{n,\mathbf{k}}}{\sigma}\right)\psi^*_{n,\mathbf{k}}(\mathbf{r})\psi_{n,\mathbf{k}}(\mathbf{r}). \tag{B.5}$$

The Fermi energy is finally determined from the normalization to the total number of particles:

$$N_{\text{el}} = \int_{-\infty}^{\epsilon_F} n(\epsilon)d\epsilon = \sum_{n,\mathbf{k}} \tilde{\theta}\left(\frac{\epsilon_F - \epsilon_{n,\mathbf{k}}}{\sigma}\right). \tag{B.6}$$

Bibliography

[1] S. M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1969).

[2] See for instance, W. Mönch, *Electronic Structure of Metal-Semiconductor Contacts* (Kluver, Dordrecht, 1990).

[3] R. Cao, K. Miyano, T. Kendelewicz, I. Lindau, and W. E. Spicer, J. Vac. Sci. Technol. B, **5** (1987).

[4] M. van Schilfgaarde and N. Neuman, Phys. Rev. Lett. **65** 2728 (1990); J. Vac. Sci. Technol. B **9**, 2140 (1991).

[5] R.G. Dandrea and C.B. Duke, J. Vac. Sci. Technol. A **11**, 848 (1993); *ibid.* B **11**, 1553 (1993).

[6] M. van Schilfgaarde, E.R. Weber, and N. Neuman, Phys. Rev. Lett. **73**, 581 (1994).

[7] R.J. Needs, J.P.A. Charlesworth, and R.W. Godby, Europhys. Lett. **25**, 81 (1994).

[8] J. Bardi, N. Binggeli, and A. Baldereschi, Phys. Rev. B **54**, R11102 (1996).

[9] C. Berthod, J. Bardi, N. Binggeli, and A. Baldereschi, J. Vac. Sci. Technol. B **14(4)**, 3000 (1996).

[10] M. Peressi, N. Binggeli, and A. Baldereschi, J. Phys. D: Appl. Phys., in press.

[11] W. Cochran, Nature **191**, 60 (1961).

[12] R. S. Mulliken, J. Chem. Phys. **3**, 573 (1935).

[13] J. Meister, and W. H. E. Schwartz, J. Chem. Phys. **98**, 8245 (1994).

[14] M. Posternak, R. Resta, and A. Baldereschi, Phys. Rev. B **50**, 8911 (1994); R. Resta, S. Massidda, M. Posternak, and A. Baldereschi, Mat. Res. Soc. Symp. Proc. **409**, 9 (1996).

[15] R. S. Mulliken, J. Chem. Phys. **36**, 3428 (1962).

[16] A. E. Reed, R. B. Weinstock, and F. A. Weinhold, J. Chem. Phys. **83**, 735 (1985).

[17] F. L. Hirshfeld, Theor. Chim. Acta **44**, 129 (1977).

[18] D. L. Cooper, Nature **371**, 651 (1994).

[19] B. Silvi, and A. Savin, Nature **371**, 683 (1994).

[20] R. F. W. Bader, T. T. Nguyen-Dang, and Y. Tal, Rep. Prog. Phys. **44**, 893 (1981).

[21] R. F. W. Bader, *Atoms in molecules. A quantum theory*, Clarendon Press, Oxford, UK (1990). Rep. Prog. Phys. **44**, 893 (1981).

[22] W. A. Harrison, *Electronic structure and the properties of solids*, edited by W. H. Freeman and Co, San Francisco, (1980).

[23] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices*, Clarendon Press, Oxford (1954) — A. A. Maradudin, E. W. Montroll, G. H. Weiss, and I. P. Ipatova, *Solid State Physics, Supplement*, Vol. 3, Academic Press, New York (1971).

[24] L. D. Landau, and E. M. Lifshits *Electrodynamics of continuous Media*, Pergamon Press (1960).

[25] R. Bonneville, Phys. Rev. B **21**, 368 (1980).

[26] R. Pick, M.H. Cohen, and R.M. Martin, Phys. Rev. B **1**, 910 (1970).

[27] J. F. Biarge, J. Herranz, and J. Morcillo, An. R. Soc. Esp. Quim. A, **57**, 81 (1961); J. Morcillo, L. J. Zamorano, and J. M. V. Heredia, Spectrochim. Acta **22**, 1969 (1966); J. Morcillo, J. F. Biarge, J. M. V. Heredia, and A. Medina, J. Mol. Struct. **3**, 77 (1969).

[28] R. Zallen, Phys. Rev. **173**, 833 (1968).

[29] H. Wendel, *Lattice dynamics of trigonal selenium and tellurium–State of the art* in *the physics of Selenium and Tellurium* E. Gerlach and P. Grosse eds., Springer Series in Solid-State Sciences 13 (Springer, Berlin, 1979).

[30] B. Szigeti, *Trans. Faraday Soc.* **45**, 155 (1949); B. Szigeti, *Proc. Roy. Soc. (London)* **A204**, 51 (1950).

[31] R.D. Amos, in: *Ab-Initio Methods in Quantum Chemistry - I*, edited by K.P. Lawley (Wiley, New York, 1987), p. 99.

[32] R. M. Lyddane, R. G. Sachs, and E. Teller, Phys. Rev. **59**, 673 (1941).

[33] P. Lazzeretti and R. Zanasi, Chem. Phys. Lett. **112**, 103 (1984).

[34] The induced charge is only weakly localized (*e.g.* not exponentially). At large distances, it goes to zero as r^{-3} times a function which oscillates around zero with the lattice periodicity. This is a typical "local-field effect" in the microscopic dielectric response of a lattice-periodical medium.

[35] *Theory of the Inhomogeneous Electron Gas*, edited by S. Lundqvist and N.H. March (Plenum, New York, 1983).

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

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

[38] D.M. Ceperley and B.J. Alder, Phys. Rev. Lett. **45**, 566 (1980); J. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[39] A. Baldereschi, Phys. Rev. B **7**, 5212 (1973).

[40] H. J. Monkhorst and J.D. Pack, Phys. Rev. B **13**, 5188 (1976).

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

[42] M. Methfessel and A.T. Paxton, Phys. Rev. B **40**, 3616 (1989).

[43] G.B. Bachelet, D.R. Hamann and M. Schlüter, Phys. Rev. B **26**, 4199 (1982).

[44] N. Troullier and J. L. Martin, Phys. Rev. B **43**, 1993 (1991).

[45] D. R. Hamann, M. Schlüter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).

[46] G.B. Bachelet and N.C. Christensen, Phys. Rev. B **8**, 5747 (1973).

[47] W.E. Pickett, Comp. Phys. Rep. **9**, 115 (1989).

[48] L. Kleinman, and D.M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).

[49] X. Gonze, P. Käckell, and M. Scheffler Phys. Rev. B **41**, 12264 (1990).

[50] S. Baroni, P. Giannozzi, and A. Testa, Phys. Rev. Lett. **58**, 1861 (1987).

[51] P. Giannozzi, S. de Gironcoli, P. Pavone, and S. Baroni, Phys. Rev. B **43**, 7231 (1991).

[52] H. Hellmann, *Einführung in die Quantenchemie* (Deuticke, Leipzig, 1937); R. P. Feynman, Phys. Rev. **56**, 340 (1939).

[53] X. Gonze and J. P. Vigneron, Phys. Rev. B **39**, 13120 (1989); A. Debernardi and S. Baroni, Solid State Commun. **91**, 813 (1994); A. Dal Corso and F. Mauri, Phys. Rev. B **50**, 5756 (1994).

[54] S. Baroni, and R. Resta, Phys. Rev. B, **33**, 7017 (1986).

[55] M. S. Hybertsten, and S. G. Louie, Phys. Rev. B, **35**, 5585 (1987).

[56] R. M. Martin, and K. Kunc, Phys. Rev. B **24**, 2081 (1981).

[57] K. Kunc, and R. M. Martin, Phys. Rev. Lett. **48**, 406 (1982).

[58] K. Kunc, in *Electronic Structure, Dynamics and Quantum Structural Properties of Condensed Matter*, edited by J.T. Devreese and P. Van Camp (Plenum, New York, 1985), p. 227.

[59] P. Lazzeretti, M. Defranceschi, and G. Berthier, Adv. Quantum Chem. **26**, 1 (1995).

[60] R. Resta, Rev. Mod. Phys. **66**, 899 (1994).

[61] We are aware of only one calculation, with no discussion of the general problem: F. Ancilotto et al., Phys. Rev. B **43**, 8930 (1991).

[62] A. Garcia and M. L. Cohen, Phys. Rev. B **47**, 4215 (1993); **47**, 4221 (1993).

[63] M. Sabisch, P. Krüger, A. Mazur, M. Rohling, and J. Pollmann, Phys. Rev. B 53, 13121 (1996); M. Sabisch, P. Krüger, and J. Pollmann, Phys. Rev. B 51, 13367 (1995).

[64] D. Olego, M. Cardona, and P. Vogl, Phys. Rev. B **25**, 3878 (1982).

[65] Landolt & Börnstein, it Physics of Group IV Elements and III-V Compounds, ed. O. Madelung, Springer Verlag, Berlin (1982). Madelung, New Series, Group III, Vol. 17, Pt. A, (1982).

[66] A. Catellani, G. Galli, and F. Gygi, Phys. Rev. Lett. **77**, 5090 (1996).

[67] R. Kaplan, Surf. Sci. **215**, 111 (1989); V. Bermudez and J. Long, Appl. Phys. Lett. **66**, 475 (1995); M. Shek, Surf. Sci. **349**, 317 (1996).

[68] M. Posternak, A. Baldereschi, A. Catellani, and R. Resta, Phys. Rev. Lett., **64**, 1777 (1990).

[69] N. F. Mott, Proc. Cambridge Philos. Soc., **34**, 568 (1938).

[70] J. Bardeen, Phys. Rev. **71**, 717 (1947).

[71] V. Heine, Phys. Rev. A **138**, 1689 (1965).

[72] H. H. Wieder, J. Vac. Sci. Technol. **15**, 1498 (1978).

[73] W. E. Spicer, P. W. Chye, P. R. Skeath & I. Lindau, J. Vac. Sci. Technol. **16**, 1422 (1979).

[74] J. L. Freeouf & J. M. Woodall, Appl. Phys. Lett. **39**, 727 (1986).

[75] C. B. Duke and C. Mailhot, J. Vac. Sci. Tech. B **3**, 1170 (1985).

[76] W. A. Harrison, J. Vac. Sci. Tech. B **3**, 1231 (1985).

[77] J. Tersoff, Phys. Rev. Lett. **52**, 465 (1984); J. Tersoff, Phys. Rev. B **30**, 8 (1984).

[78] G. Ortiz, R. Resta, and A. Baldereschi, J. Phys.: Condens. Matter **2**, 10217 (1990).

[79] R. Ludeke, and G. Landgren, J. Vac. Sci. Technol. **19**, 667 (1981).

[80] S. Baroni, M. Peressi, R. Resta, and A. Baldereschi, in *Proceedings of the 21th International Conference on the Physics of Semiconductors*, edited by Ping Jiang and Hou-Zhi Zheng (World Scientific, Singapore, 1993), p.689.

[81] F. D. Murnaghan, *Deformation of an elastic solid*, chap.4, John Wiley, NY (1951).

[82] U. von Barth and R. Car (unpublished).

[83] S. Baroni, R. Resta, A. Baldereschi, and M. Peressi, in *Spectroscopy of Semiconductor Microstructures*, edited by G. Fasol, A. Fasolino, and P. Lugli, NATO ASI Ser. B, Vol. 206 (Plenum, NY, 1989), p.251.

[84] R. G. Dandrea and C.B. Duke, J. Vac. Sci. Technol. B **11**, 848 (1993); ibid. 1553.

[85] R. J. Needs, J. P. A. Charlesworth and R. W. Godby, Europhys. Letters **25**, 31 (1994).

[86] P. Revva, J. M. Langer, M. Missous, and A. R. Peaker, J. Appl. Phys. **74**, 416 (1993).

[87] Landolt & Börnstein, it Numerical Data and Functional Relationships in Science and Technology, ed. O. Madelung, Springer Verlag, Berlin.

[88] R. Resta, Phys. Rev. B **44**, 11035 (1991).

[89] M. Peressi, S. Baroni, A. Baldereschi, and R. Resta, Phys. Rev. B **41**, 12106 (1990).

[90] *Metallization and metal-Semiconductor Interfaces*, edited by P. Batra, NATO ASI Series B, vol. 195 (Plenum Publishing, New York, 1989).

[91] D. H. Vanderbilt, Phys. Rev. B **41**, 7892 (1990).

[92] S. G. Louie, S. Froyen, and M. L. Cohen, Phys. Rev. B **26**, 1738 (1982); G. E. Engel and R. J. Needs, Phys. Rev. B **41**, 7876 (1990); A. Qteish and R. J. Needs, Phys. Rev. B **43**, 4229 (1991).

[93] A. Dal Corso, S. Baroni, R. Resta, and S. de Gironcoli, Phys. Rev. B **47**, 3588 (1993).

[94] W. Chen, A. Kahn, P. Soukiassan, P.S. Mangat, J. Gaines, C. Ponzoni, and D. Olego, Phys. Rev. B, **51**, 14265 (1995).

[95] S. G. Louie, J. R. Chelikowsky, and M. L. Cohen, Phys. Rev. **15**, 2154 (1977).

[96] J. D. Jackson, *Classical Electrodynamics*, Wiley, New York (1975).

[97] F. N. H. Robinson, *Macroscopic Electromagnetism*, Pergamon, Oxford (1973).

[98] A. Baldereschi, S. Baroni, and R. Resta, Phys. Rev. Lett. **61**, 734 (1988). A thorough account is in: S. Baroni, R. Resta, A. Baldereschi, and M. Peressi, in: *Spectroscopy of semiconductor microstructures*, edited by G. Fasol, A. Fasolino and P. Lugli, NATO ASI Series B, vol 206 (Plenum Publishing, New York, 1989), p 251.

[99] C. L. Fu, and K. M. Ho, Phys. Rev. B **28**, 5480 (1983).

[100] S. de Gironcoli, Phys. Rev. B **51**, 6773 (1995).