

Maximally-localized Wannier Functions in III-V Semiconductors

By

Hazem Omar Abdollah Abu-Farsakh

Supervisor

Prof. Dr. Abdallah Qteish

December 23, 2003

Maximally-localized Wannier Functions in III-V Semiconductors

By

Hazem Omar Abdollah Abu-Farsakh

B.Sc. Applied Physics, Jordan University of Science and Technology, 2000

Thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science in Physics Department at Yarmouk University, Irbid, Jordan

Approved By:

Prof. Dr. Abdallah Qteish.....Chairman

Dr. Abdullah Al-Sharif.....Member

Dr. Abdul Raouf Al-Dairy.....Member

Prof. Dr. Jamil Khalifeh.....Member

December 23, 2003

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

إلى

إلى والصدى الحبيبين....

إلى أخواتي وإخواني الأعزاء

إلى كل من طلب العلم والمعرفة

ACKNOWLEDGMENTS

I would like to express my deep thanks to my supervisor Prof. Abdallah Qteish, to whom I am most grateful for his guidance, support, encouragement, and fruitful discussions throughout this work.

Many thanks go to the committee members (Dr. Abdullah Al-Sharif, Dr. Abdul Raouf Al-Dairy, Prof. Dr. Jamil Khalifeh) for their critical reading of the thesis and for their comments.

I am indebted to Prof. Nabil Laham, Mr. Yacoub Afeef, and the Center of theoretical and Applied Physics (CTAPS) at Yarmouk University for providing all the available facilities required to accomplish this work.

I wish also to thank Prof. Jörg Neugebauer and Matthias Wahn for cooperation, and Prof. Matthias Scheffler and Sixten Boek for hospitality, and others among the members of the Fritz-Haber Institute der Max-Planck Gesellschaft, Berlin, Germany.

I would like also to thank Dr. Ilan Schnell for many fruitful discussions.

Finally, I wish to express my deepest gratitude to my family for love, encouragements and spiritual support, and my friends (Omar Al-Obeidat, Hussain Abu-Kharma, Tayseer Abu-Alrub, Basil Sa'd, Qatad Samara, Burhan Ershaid, Othman Tarboush, and many others) for their encouragement and support.

Thank you all very much.

Irbid, Dec. 2003

Hazem Abu-Farsakh

ملخص

دالات فانيير الأكثر تمركزاً لمركبات III-V شبه الموصلة

رسالة ماجستير

مقدمه من

حازم عمر أبو فرسخ

بإشراف

أ.د. عبد الله قطيش

تم حساب دالات فانيير الأكثر تمركزاً لأربعة من مركبات III-V شبه الموصلة وهي GaAs و AlAs و GaN و AlN، وذلك باستخدام طريقة مرزاري وفاندر بلت. لتحقيق ذلك، قمنا بتطوير برنامجي كمبيوتر بلغتي فورتران 77 و ++C. ومن أجل المقارنة تم أيضاً حساب دالات فانيير الأكثر تمركزاً لأشباه الموصلات التالية: C و Si و Ge و SiC. في كل المواد أعلاه، وجدنا أن دالات فانيير الأكثر تمركزاً محصورة في اللبنة التركيبية البدائية، مما يظهر أهميتها للاستخدام في الكثير من التطبيقات النظرية. كما وجدنا أن هنالك علاقة بين أيونية الرابطة الكيميائية و بعد مركز دالة فانيير عن مركزها. بناءً على هذه العلاقة فقد تم وضع مقياس جديد لأيونية الرابطة، ووجدنا أن هذا المقياس يتفق مع مقياس الأيونية الناتج من استخدام الكثافة الإلكترونية الناتجة من الحسابات المتوافقة ذاتياً أكثر من اتفاهه مع مقياس فيليبس.

CONTENTS

CONTENTS.....	VII
ABSTRACT.....	1
CHAPTER 1.....	2
INTRODUCTION	2
CHAPTER 2.....	5
DENSITY FUNCTIONAL THEORY.....	5
2.1 INTRODUCTION	5
2.2 THE HOHENBERG-KOHN THEOREMS.....	8
2.3 THE KOHN-SHAM FORMALISM	10
2.4 THE LOCAL DENSITY APPROXIMATION.....	12
CHAPTER 3.....	14
THE PLANE-WAVE PSEUDOPOTENTIAL METHOD.....	14
3.1 INTRODUCTION	14
3.2 JUSTIFICATION OF THE PSEUDOPOTENTIAL APPROACH.....	16
3.3 NORM-CONSERVING PSEUDOPOTENTIALS.....	17
3.3.1 KERKER SCHEME.....	19
3.3.2 TROULLIER AND MARTINS SCHEME.....	19
3.4 KLEINMAN-BYLANDER FORM OF IONIC PSEUDOPOTENTIAL	20
3.5 MOMENTUM SPACE FORMALISM	20
CHAPTER 4.....	24
WANNIER FUNCTIONS.....	24
4.2 DEFINITIONS	24
4.3 ARBITRARINESS IN DEFINITION OF WF'S	25
4.4 MARZARI-VANDERBILT METHOD.....	26
4.4.1 SPREAD FUNCTIONAL IN REAL SPACE	26
4.4.2 SPREAD FUNCTIONAL IN MOMENTUM SPACE	28
4.4.3 GRADIENT OF SPREAD FUNCTIONAL	31
4.4.4 STEEPEST-DESCENT MINIMIZATION	32

CHAPTER 5	34
COMPUTATIONAL DETAILS	34
5.1 INTRODUCTION	34
5.2 LATTICE	34
5.3 STRUCTURES	35
5.4 RECIPROCAL LATTICE	36
5.5 SELF-CONSISTENT CALCULATIONS	37
5.6 CONSTRUCTION OF THE MLWF'S	39
5.6.1 RECIPROCAL SPACE GRIDS	39
5.6.2 WAVEFUNCTIONS INITIALIZATION	42
5.6.3 PLOTTING THE MLWF'S	46
CHAPTER 6	47
RESULTS AND DISCUSSION	47
6.1 INTRODUCTION	47
6.2 MAXIMALLY LOCALIZED WANNIER FUNCTIONS	48
6.2.1 FOR C, SI AND GE	48
6.2.2 FOR SIC	54
6.2.3 FOR GAAS, ALAS, GAN AND ALN	57
6.3 BOND IONICITY	64
6.4 CONCLUSIONS	68
APPENDIX A	69
SPECIAL POINTS	69
REFERENCES	71

ABSTRACT

Maximally-localized Wannier Functions in III-V semiconductors

Master of Science Thesis

By

Hazem Omar Abu-Farsakh

Department of Physics, Yarmouk University, Irbid, Jordan

Dec. 2003

Supervisor

Prof. Dr. Abdallah Qteish

The maximally-localized Wannier functions (MLWF's) of four III-V semiconductors (GaAs, AlAs, GaN and AlN) have been constructed, using the scheme of Marzari and Vanderbilt. To do that, we have developed FORTRAN77 and C++ codes. For comparison, we have constructed also the MLWF's of C, Si, Ge and SiC. In the eight studied systems, the MLWF's are confined in the primitive unit cell, which demonstrates their high potential application as accurate and minimal basis for various theoretical approaches. The deviation of the Wannier center from the center of the bond is found to be proportional to the bond ionicity. Based on this deviation, a new bond ionicity scale has been introduced. The obtained ionicities of the studied systems are in much better agreement with those obtained self-consistently (from the calculated charge density) than with the empirical Phillips ionicity scale.

Key words: maximally-localized Wannier functions, III-V semiconductors, bond ionicity, density functional theory, pseudopotential plane-wave.

CHAPTER 1

INTRODUCTION

Modern computational methods applied to condensed matter systems are mainly based on Density Functional Theory (DFT) of Hohenberg and Kohn [1]. The Kohn-Sham (KS) formalism [2] of DFT provides a way to *exactly* transform the many-body problem into a single-body one. However, the Exchange and Correlation (XC) potential have to be approximated. The most widely used approximation for the XC potential is the Local Density Approximation (LDA) [2,3]. The LDA is found to give surprisingly good results for many physical and chemical properties of a wide range of materials. The most serious shortcoming of LDA is the so-called *band-gap problem*: the band gaps of insulators and semiconductors are underestimated. Another approximation is the so-called Generalized Gradient Approximation (GGA) [4-6]. The GGA is found to improve the cohesive energies of solids and molecules. Recently, the Exact Exchange (EXX) [7,8] scheme has been introduced. In this scheme, the exchange effects are treated exactly, and only the correlation potential is approximated using LDA or GGA. The EXX formalism is found to give band-gaps, effective masses, and cohesive properties in good agreement with experiment. In its present implementation (using Plane-Wave (PW) basis set), the EXX calculations are very expensive in terms of both memory and CPU time.

Within the single-particle approximation, the electronic states of a periodic crystalline solid are usually described in terms of extended Bloch orbitals [9,10]. An alternative representation can be derived in terms of Wannier functions (WF's) [11]. WF's are formally defined via a unitary transformation of the Bloch orbitals. The most serious drawback of WF's, which so far has limited their applications compared to Bloch representation, is their non-uniqueness. This is a consequence of the phase indeterminacy that Bloch orbitals have at any wave vector \mathbf{k} . The non-uniqueness of WF's can be

utilized to construct the *maximally-localized* set of WF's (MLWF's). Recently, Marzari and Vanderbilt [12] developed a very practical method for generating MLWF's starting from the knowledge of the occupied Bloch states. The new technique has been successfully applied to crystal systems and small molecules [12]. Furthermore, a disentanglement procedure was introduced [13] to extend the original algorithm to the case of systems without gaps and to remove the limitation to the isolated group of bands that are separated by gaps from higher and lower bands.

Since their introduction in 1937, WF's have played an important role in the theoretical study of the properties of periodic solids. Moreover, the representation of electronic ground state of periodic systems in terms of localized Wannier orbitals has recently attracted considerable attention due to, mainly, two recent developments. First, the development of "order-N" or "linear scaling" electronic structure methods [14]. In these methods the computational time scales only as the first power of the system size [15], instead of the third power typical of conventional methods based on solving for Bloch states. Moreover, MLWF's are being used as very accurate minimal basis in a variety of other applications, such as the construction of the Hamiltonian in the form of a multi-band Hamiltonian in second quantization (a kind of extended, multi-band Hubbard model) such that all the standard many-body methods can be applied [16]. Second, the formulation of the modern theory of bulk polarization of crystalline solids [17,18], which directly relates the vector sum of the centers of the MLWF's to the macroscopic electronic polarization of a crystalline insulator. Furthermore, the piezoelectric tensor defined as polarization derivative with respect to strain can be easily calculated [18] within this approach.

The main driving force behind the present work is the formulation of the EXX potential, within the KS formalism, in terms of MLWF's. This is expected to reduce dramatically the computational efforts of the EXX calculations. The first step in this ambitious project is the construction of the MLWF's, which is done in this present work.. Thus, the main aim of this work is of two folds:

- I- The implementation of the Marzari-Vanderbilt scheme of constructing MLWF's. We have developed two codes: one written in FORTRAN77 and takes its input (mainly Bloch wavefunctions) from data obtained using the CASTEP code [19], and the other in C++ and takes its input from data

obtained using the SFHingX code [20]. The two codes are found to give identical results, and, thus, the reported results are obtained by using the C++ code.

- II- As a preliminary application, the developed code is used to construct the MLWF's of some III-V compounds. This is done for GaAs, AlAs, AlN, and GaN. For comparison, we have constructed also the MLWF's of SiC, Si, Ge and C. Moreover, the use of Wannier centers as a measure of the ionicity is investigated.

This thesis is organized as follows. In the following chapter, we review the DFT and the LDA for the XC potential. In chapter 3, we present the plane-wave pseudopotential method, together with a review of some modern *ab-initio* pseudopotential generation methods. Chapter 4 is devoted to introduce both the WF's and the Marzari-Vanderbilt scheme for constructing the MLWF's. In chapter 5, we show the used computational details. Finally, in chapter 6, we report, discuss our results and provide our conclusions.

CHAPTER 2

DENSITY FUNCTIONAL THEORY

2.1 Introduction

Density Functional Theory (DFT) of Hohenberg and Kohn [1] is one of the most popular and successful quantum mechanical approaches. It is an *exact* theory for the *ground state* of an interacting many-particle system. It is nowadays routinely applied for calculating, e.g., the binding energy of molecules in chemistry and the band structure of solids in physics. Some applications relevant for fields traditionally considered more distant from quantum mechanics, such as biology and mineralogy are beginning to appear. DFT owes this versatility to the generality of its fundamental concepts and the flexibility one has in implementing them.

To get a first idea of what density functional theory is about, it is useful to recall some elementary quantum mechanics. In quantum mechanics it is known that all information we can possibly have about a given system is contained in its many-body wavefunction Ψ . A system of N interacting electrons in an external potential is described by the many-electron Schrödinger equation which, in the time independent and adiabatic approximations¹ [21], is

$$\hat{H}\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = E\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N). \quad (2.1)$$

\hat{H} is the Hamiltonian operator:

$$\hat{H} = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 + \sum_{i=1}^N V_{ext}(\mathbf{r}_i) + \frac{1}{2} \sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \quad (2.2)$$

¹ In the adiabatic approximation the nuclei are treated as objects at fixed positions. It is based on the fact that typical electronic velocities are much greater than typical ionic velocities.

where \mathbf{r}_i is the position of the i -th electron and $V_{ext}(\mathbf{r}_i)$ is the external potential. If the Hamiltonian \hat{H} is spin independent, the spin part of the wavefunction cancels out, and, therefore, the spin variables are suppressed for simplicity. Hereafter, we will use the Hartree atomic units ($\hbar = e = m_e = 1$).

The usual quantum-mechanical approach to Schrödinger's equation (SE) can be summarized by the following sequence

$$V_{ext}(\mathbf{r}) \xrightarrow{SE} \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \xrightarrow{\langle \Psi | \dots | \Psi \rangle} \text{observables}, \quad (2.3)$$

i.e., one specifies the system by choosing $V_{ext}(\mathbf{r})$, plugs it into Schrödinger's equation, solves that equation for the wavefunction Ψ , and then calculates expectation values of observables with this wave function. For example, the electronic charge density is given by

$$n(\mathbf{r}) = 2N \int d\mathbf{r}_2 \int d\mathbf{r}_3 \dots \int d\mathbf{r}_N \Psi^*(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) \Psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N). \quad (2.4)$$

Powerful methods for solving Schrödinger's equation have been developed during decades of struggling with the many-body problem. The problem with these methods is the great demand they place on one's computational resources: it is simply *impossible* to apply them efficiently to large and complex systems.

It is here where DFT provides a practical alternative. The Hohenberg-Kohn theorems [1] have promoted the *charge density*, $n(\mathbf{r})$, from just one among many observables to the status of a basic variable, as Ψ . The DFT approach can be summarized by the sequence

$$n(\mathbf{r}) \longrightarrow \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \longrightarrow V_{ext}(\mathbf{r}). \quad (2.5)$$

That is, the knowledge of $n(\mathbf{r})$ implies knowledge of the wavefunction and the external potential, and hence of all other observables. The above equation describes only the conceptual structure of DFT.

The Kohn-Sham (KS) formalism of DFT [2] provided a way to *exactly* transform the many-body problem into a single-body one. Kohn and Sham proved that solving Schrödinger equation of an auxiliary system of non-interacting electrons moving in an effective potential,

$$\left[-\frac{1}{2} \nabla^2 + V_{eff}[n(\mathbf{r})] \right] \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}),$$

would give the same result of the real interacting system moving in an external potential $V_{ext}(\mathbf{r})$. Here $\{\psi_i\}$ are the single body wavefunctions. The effective potential $V_{eff}[n]$ is given by

$$V_{eff}[n] = V_{ext}(\mathbf{r}) + V_H[n] + V_{XC}[n],$$

where $V_H[n]$ is due to classical electron-electron interactions, and $V_{XC}[n]$ is due to the quantum mechanical exchange and correlation (XC).

However, The XC energy functional is unknown. In fact, it depends in a complicated way on the charge density distribution of the system, and thus, no simple exact expression for it is available. Therefore, in practice, an approximation for it is required. The most widely used one is the Local Density Approximation (LDA) [2,3] where the XC energy functional is considered to be locally equal to that of a uniform electron gas of the same local density. Another approach is the so-called Generalized Gradient Approximation (GGA) [4-6] where the XC energy functional depends locally on the density and on its gradient. Recently, an Exact Exchange (EXX) has been introduced [7,8], where the exchange part of the XC energy functional is treated exactly and only the correlation part needs to be approximated. In this thesis, we limit ourselves to the LDA.

This DFT approach forms the basis of the large majority of electronic structure calculations in physics and chemistry. Much of what we know about the electrical, magnetic, and structural properties of materials has been calculated using DFT, and the extent to which DFT has contributed to the science of molecules is reflected by the 1998 Nobel Prize in Chemistry, which was awarded to Walter Kohn [22], the founding father of DFT, and John Pople [23], who was instrumental in implementing DFT in computational chemistry.

The following sections are devoted to explain the DFT. First, in section 2, the Hohenberg-Kohn theorems will be introduced. Then, in section 3, the KS formalism of the DFT will be discussed. Finally, section 4 is devoted to the LDA.

2.2 The Hohenberg-Kohn Theorems

At the heart of DFT is the Hohenberg-Kohn (HK) theorems. HK theorems was introduced in 1964 by Hohenberg and Kohn [1]. The first theorem can be stated as follows. *The complete many-body wavefunction, Ψ , of an electronic system is a unique functional $\Psi[n(\mathbf{r})]$ of the electronic charge density $n(\mathbf{r})$.* As a consequence, the expectation value of any observable is also a functional of $n(\mathbf{r})$

$$O = O[n] = \langle \Psi[n] | \hat{O} | \Psi[n] \rangle. \quad (2.6)$$

Hohenberg and Kohn [1] gave a straightforward proof of this theorem, which was generalized to include systems with degenerate states by Levy in 1979 [24].

Hohenberg and Kohn defined a *universal* functional $F[n]$:

$$F[n] = \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle, \quad (2.7)$$

where \hat{T} and \hat{V}_{ee} are the kinetic and the electron-electron interaction energy operators of the many-body system respectively. The functional $F[n]$ is *universal* in the sense that it does not depend on the external potential $V_{ext}(\mathbf{r})$ which represents a particular system of interest. With the help of $F[n]$ Hohenberg and Kohn further defined, for a given external potential $V_{ext}(\mathbf{r})$, the total energy functional

$$E[n] = \int d\mathbf{r} V_{ext}(\mathbf{r}) n(\mathbf{r}) + F[n]. \quad (2.8)$$

The second theorem of Hohenberg and Kohn can be stated as follows. *The total energy functional of a system assumes its minimum value (the ground-state energy) at the true ground state density of that system (n_0), i.e. it obeys a variational property*

$$E_0 = E[n_0] \leq E[n], \quad (2.9)$$

where n is some other density. This is very similar to the usual variational principle for wavefunctions. If we calculate the expectation value of a Hamiltonian with a trial wavefunction Ψ' that is not its Ground State (GS) wavefunction Ψ_0 we can never obtain an energy below the true GS energy,

$$E_0 = E[\Psi_0] = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle \leq \langle \Psi' | \hat{H} | \Psi' \rangle = E[\Psi']. \quad (2.10)$$

Similarly, in DFT, if we calculate the GS energy of a Hamiltonian using a density that is not its GS density we can never find a result below the true GS energy. From Eq. (2.8) and Eq. (2.9) it is clear that

$$E_0 = \int d\mathbf{r} V_{ext}(\mathbf{r})n_0(\mathbf{r}) + F[n_0]. \quad (2.11)$$

After these abstract considerations let us now consider one way in which one can make practical use of DFT. Assume we have specified our system (i.e., $V_{ext}(\mathbf{r})$ is known). Assume further that we have a reliable approximation for $F[n]$. All one has to do then is to minimize $E[n]$ (Eq. (2.8)) with respect to $n(\mathbf{r})$ under the constraint that the total number of electrons in the system, N , is constant, that is

$$N = \int d\mathbf{r} n(\mathbf{r}). \quad (2.12)$$

The best function $n_0(\mathbf{r})$ is the system's GS charge density and the value $E_0 = E[n_0]$ is the GS energy. Assume now that $V_{ext}(\mathbf{r})$ depends on a parameter a . This can be, for example, the lattice constant in a solid or the angle between two atoms in a molecule. Calculation of E_0 for many values of a allows one to plot the $E_0(a)$ curve and to find the equilibrium value of a . In this way one can calculate quantities like molecular geometries and sizes, lattice constants, unit cell volumes, charge distributions, total energies, etc. By looking at the change of $E_0(a)$ with a one can, moreover, calculate compressibilities and bulk moduli (in solids) and vibrational frequencies (in molecules). By comparing the total energy of a composite system (e.g., a molecule) with that of its constituent systems (e.g., individual atoms) one obtains dissociation energies. By calculating the total energy for systems with one more (or less) electron one obtains the electron affinity (or ionization energy). All this follows from DFT without having to solve the many-body Schrödinger equation.

In principle it should be possible to calculate all observables, since the HK theorem guarantees that they are all functionals of $n_0(\mathbf{r})$. However, in practice, one does not know how to do this explicitly. Another problem is that the minimization of $E[n]$ is, in general, a tough numerical problem of its own. Moreover, one needs a reliable approximation for $F[n]$ to begin with. In the next section, on the KS formalism, we will see one widely used method for solving these problems.

2.3 The Kohn-Sham Formalism

In 1965, Kohn and Sham [2] have introduced the following separation of the functional $F[n]$,

$$F[n] = T_s[n] + E_H[n] + E_{XC}[n], \quad (2.13)$$

where $T_s[n]$ is the kinetic energy of *non-interacting* electrons with density $n(\mathbf{r})$

$$T_s = -\frac{1}{2} \sum_{i=1}^N \int d\mathbf{r} y_i^*(\mathbf{r}) \nabla^2 y_i(\mathbf{r}) \quad (2.14)$$

(here y_i is the single particle wave functions of the non-interacting system), $E_H[n]$ is the classical electron-electron interaction energy (or Hartree energy)

$$E_H[n] = \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \quad (2.15)$$

The term $E_{XC}[n]$ contains the remaining electron-electron interaction energies, and the difference between the kinetic energies of the interacting and non-interacting systems, $T - T_s$. This functional is, by definition, the quantum mechanical *exchange and correlation* energy of the interacting system, which is unknown and needs to be approximated. Nonetheless, note that Eq. (2.13) is formally exact.

Using equations (2.8) and (2.13), the total energy functional can be written as

$$E[n] = T_s[n] + E_H[n] + E_{ext}[n] + E_{XC}[n] \quad (2.16)$$

where

$$E_{ext}[n] = \int d\mathbf{r} V_{ext}(\mathbf{r})n(\mathbf{r}). \quad (2.17)$$

The variational principle applied to Eq. (2.16) yields

$$\begin{aligned} \frac{dE[n]}{dn(\mathbf{r})} &= \frac{dT_s}{dn(\mathbf{r})} + \frac{dE_{ext}}{dn(\mathbf{r})} + \frac{dE_H}{dn(\mathbf{r})} + \frac{dE_{XC}}{dn(\mathbf{r})} \\ &= \frac{dT_s}{dn(\mathbf{r})} + V_{ext}(\mathbf{r}) + V_H[n] + V_{XC}[n] \\ &= m, \end{aligned} \quad (2.18)$$

where m is the Lagrange multiplier associated with the requirement of constant number of particles, $V_H[n]$ and $V_{XC}[n]$ are the Hartree and the XC potentials, respectively.

Consider now a *non-interacting* system of particles moving in external potential $V_{eff}(\mathbf{r})$ having the same $n(\mathbf{r})$ as the above interacting system. For this system the minimization condition is

$$\frac{dT_s}{dn_s(\mathbf{r})} + V_{eff}[n(\mathbf{r})] = m, \quad (2.19)$$

since there are no Hartree and XC terms in the absence of interactions. Comparing this equation with the previous one (Eq. (2.18)) we conclude that they are mathematically identical if $V_{eff}[n]$ is chosen to be

$$V_{eff}[n] = V_{ext}(\mathbf{r}) + V_H[n] + V_{XC}[n]. \quad (2.20)$$

Consequently, one can calculate the density of the interacting (many-body) system, in the potential $V_{ext}(\mathbf{r})$, by solving the equations of a non-interacting (single-body) system, in the potential $V_{eff}[n]$. In particular, the Schrödinger equation of this auxiliary system is

$$\left[-\frac{1}{2}\nabla^2 + V_{eff}[n] \right] y_i(\mathbf{r}) = e_i y_i(\mathbf{r}), \quad (2.21)$$

yields wave functions that reproduce the density $n(\mathbf{r})$ of the original system (these are the same wave functions employed in Eq. (2.14)),

$$n(\mathbf{r}) = 2 \sum_{i=1}^N |y_i(\mathbf{r})|^2. \quad (2.22)$$

Equations (2.20), (2.21), and (2.22) are known as KS equations, and $V_{eff}[n]$ is called the KS potential. The problem of minimizing $E[n]$ was replaced by that of solving a set of single-particle Schrödinger equations. These equations have to be solved self-consistently. In this case, one starts with an initial guess for $n(\mathbf{r})$, calculates the corresponding $V_{eff}[n(\mathbf{r})]$, and then solves the differential equation (2.21) for the y_i 's. From these y_i 's one calculates a new density, using (2.22), and starts again. The process is repeated until a reasonable convergence is reached.

Note, from the derivation of Eq. (2.21), that the e_i 's are the eigenvalues of the auxiliary single-body Schrödinger equation whose eigenfunctions yield the correct density. They are usually called the KS eigenvalues. DFT assigns no formal interpretation to KS orbitals and KS eigenvalues, except that of the highest occupied state which is

equal to the ionization energy. It is only the ground state charge density that has strict physical meaning in the KS equations. Nevertheless, most band-structure calculations in solid-state physics are actually calculated from the KS eigenvalues, and it comes as a great (and welcome) surprise that in many situations the KS eigenvalues, e_i , do provide a quite good approximation to the actual energy levels of extended systems [25]. Moreover, it was found recently [26] that the energy gap between the highest occupied KS eigenvalue and the lowest unoccupied one agrees quite well with experimental values for the true energy gap, even for difficult cases, when an extremely accurate V_{eff} is used (within the EXX approach).

2.4 The Local Density Approximation

The LDA [2,3] is the most common approach to approximate the unknown XC functional $E_{XC}[n]$. The idea behind LDA is to replace the XC energy $E_{XC}[n]$ of the non-uniform $n(\mathbf{r})$ with $E_{XC}[n]$ computed as if *locally* the interacting electron gas has the same XC energy of a *uniform* interacting electron gas with the same local density. Now, let $e_{XC}^{\text{hom}}(n)$ be the XC energy per electron for a homogeneous (uniform) electron gas with a density n . Then, the total XC energy of this system, $E_{XC}[n]$, can be written as

$$E_{XC}[n] = \int d\mathbf{r} n e_{XC}^{\text{hom}}(n). \quad (2.23)$$

In LDA one imagines the real inhomogeneous system (with density $n(\mathbf{r})$ in a potential $V_{ext}(\mathbf{r})$) is decomposed in small cells in each of which $n(\mathbf{r})$ and $V_{ext}(\mathbf{r})$ are approximately constant. In each cell (i.e., locally) one can then use the expression of a homogeneous system to approximate the contribution of the cell to the real inhomogeneous one. Making the cells infinitesimally small and summing over all of them yields

$$E_{XC}^{LDA}[n] = \int d\mathbf{r} n(\mathbf{r}) e_{XC}^{\text{hom}}(n(\mathbf{r})), \quad (2.24)$$

and the corresponding exchange-correlation potential is given by

$$V_{XC}^{LDA}(\mathbf{r}) = \frac{dE_{XC}^{LDA}[n(\mathbf{r})]}{dn(\mathbf{r})} = \frac{d[n(\mathbf{r})e_{XC}^{\text{hom}}(n(\mathbf{r}))]}{dn(\mathbf{r})}. \quad (2.25)$$

Usually, e_{XC} is splitted into an exchange, e_X , and a correlation, e_C , part

$$e_{XC}^{LDA}[n(\mathbf{r})] = e_X^{\text{hom}}[n(\mathbf{r})] + e_C^{\text{hom}}[n(\mathbf{r})]. \quad (2.26)$$

The exchange energy per electron for a homogeneous system is given by

$$e_X^{\text{hom}}(n) = -\frac{3}{4} \left(\frac{3}{p} n \right)^{\frac{1}{3}}. \quad (2.27)$$

Usually e_X^{hom} and e_C^{hom} are expressed in terms of the parameter r_s , which is related to n through the relation

$$n = \left(\frac{4p}{3} r_s^3 \right)^{-1}. \quad (2.28)$$

The correlation energy per electron has been calculated by Quantum Monte Carlo (QMC) calculations performed by Cerperly and Alder [27] and got parameterized by Perdew and Zunger [28]:

$$e_C = \begin{cases} \frac{-0.1423}{1 + 1.0529\sqrt{r_s} + 0.3334r_s} & \text{for } r_s \geq 1 \\ -0.048 + (0.0311 + 0.002r_s) \ln(r_s) - 0.0116r_s & \text{for } r_s < 1 \end{cases}. \quad (2.29)$$

LDA is exact for a homogeneous system and is expected to be valid for systems with slowly varying charge density. Nevertheless, LDA is found to be a very good approximation even in the case when the charge density is rapidly varying with respect to \mathbf{r} [3,29,30,31]. LDA also has some drawbacks. For example, lattice constants and band gaps are underestimated (*band gap problem*). Several approaches appeared to improve the LDA. Examples are the GGA [4-6] and the EXX [7,8] methods.

CHAPTER 3

THE PLANE-WAVE PSEUDOPOTENTIAL METHOD

3.1 Introduction

In practice, numerical solution of the KS differential equation (2.23) proceeds by expanding the KS orbitals in a suitable set of basis functions. One among various existing options is the Plane-Wave (PW) basis set. Unfortunately, the PW basis set is usually very poorly suited to expanding the electronic wavefunctions because an enormous basis size is required to accurately describe the rapidly oscillating wavefunctions of the valence electrons in the core region. On the other hand, the energies associated with core wavefunctions are orders of magnitude higher than the energies associated with the valence ones. Moreover, it is well known that most physical properties of solids depend on the valence electrons to a much greater degree than that of the tightly bound core electrons. It is for these reasons that the pseudopotential (PP) approximation is introduced. This approximation uses these facts to remove the core electrons and the strong nuclear potential and replace them with a weaker pseudopotential which acts on a set of pseudo-wavefunctions rather than the true valence wavefunctions. In some cases there are some states which can be considered as semicore: they are highly localized, but they contribute to the chemical bonds. These states should be treated as valence.

The main advantage of such a formulation is that one can in principle replace the all-electron problem with an effective Hamiltonian acting on smooth wavefunctions representing the valence electrons. This is shown schematically in Fig. 3.1. Thus, the computational load is reduced in two respects: firstly, by reducing the number of the self-consistent wavefunctions to be determined; secondly, by avoiding to represent the rapid oscillations of the all-electron valence orbitals close to the nucleus, while keeping the long tails that are mainly responsible for the formation of the chemical bonds. A justification of the pseudopotential approach has been provided by the orthogonalized PW (OPW) method.

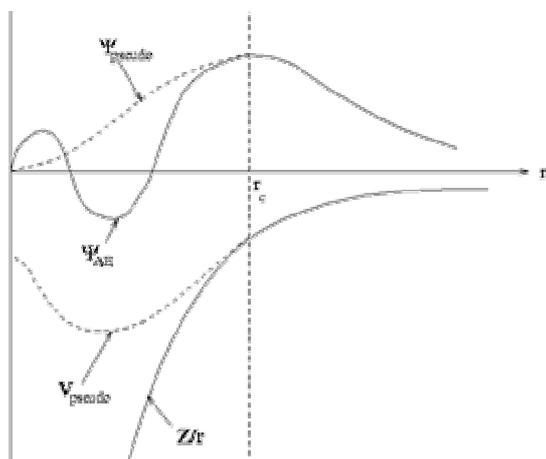


Fig. 3.1: Schematic illustration of all-electron (solid lines) and pseudo (dashed line) potentials, and their corresponding wavefunctions.

Several types of pseudopotentials have been applied to the study of electronic structure and ground state properties of solids. Examples are: (i) empirical pseudopotentials [32], and (ii) *ab initio* (first principles) pseudopotentials [33,34,35]. Empirical pseudopotentials were widely used in the 60s and 70s. The main disadvantage of the empirical pseudopotentials is that they are non-transferable, that is, they are not able to accurately describe the valence electrons in different atomic, molecular, and solid state environments. Moreover, they are not able to lead to accurate valence charge density, which is the basic quantity in DFT. Empirical pseudopotentials were fashionable before the introduction of the *ab initio* pseudopotentials.

In 1979, the norm-conserving pseudopotentials have been introduced by Hamann, Schlüter, and Chiang [33]. These pseudopotentials need to be constructed only once using the all-electron valence wavefunctions and eigenvalues obtained from self-consistent atomic calculations. The pseudo wavefunctions match, by construction, the all-electron

ones beyond a certain core radius (r_c) (see Fig. (3.1)). This ensures good charge density distribution outside the core region. Moreover, the norm-conserving leads to good transferability of these pseudopotentials. These potentials are known as semilocal, since they depend on both \mathbf{r} and angular momentum quantum number l . Kleinman and Bylander [36] have shown that the computational cost of the matrix elements of these potentials between PW's, can be greatly reduced by transferring the semilocal form to a truly nonlocal one. Finally, the expansion of the valence wavefunctions of highly localized states, such as the p-state of the first row elements, in terms of PW's requires a very large number of these basis. Several methods have been devised [37,38,39] to optimize the pseudopotential, in terms of the required number of PW's. Here, we will describe the one used in this work, namely, that of Troullier and Martins [40].

This chapter is organized as follows. In the next section, we provide a justification of the pseudopotential approach. Then, in section 3, we introduce the construction of norm-conserving pseudopotentials, and we describe the Kerker and Troullier-Martins schemes. Next, in section 4, we present Kleinman-Bylander form of pseudopotentials. Finally, in section 4 we focus on the momentum space formalism.

3.2 Justification of The Pseudopotential Approach

To see this, consider a solid as a collection of ion cores and valence electrons. The core wavefunctions are well localized while the valence ones are extended and hopefully can be expanded by a reasonable number of PW's. The one electron Schrödinger equation reads

$$Hy = ey \quad (3.1)$$

where the Hamiltonian is the sum of the kinetic energy, T , and an effective potential operators, V_A . Let us expand the true valence electronic wavefunction, y , as

$$y = f + \sum_c b_c f_c \quad (3.2)$$

where f is a smooth wavefunction, f_c are the core wavefunctions corresponding to the bound states in the ion core, and b_c is determined from the condition that y and f_c are orthogonal to each other;

$$\langle y | f_c \rangle = 0 \quad (3.3)$$

which leads to

$$b_c = -\langle f_c | f \rangle. \quad (3.4)$$

Using the above three equations, and the fact that y and f_c are solutions of the Schrödinger equation with eigenvalues e and e_c respectively, Eq. (3.1) can be manipulated to take the form

$$\left[H + \sum_c (e - e_c) |f_c\rangle\langle f_c| \right] f = e f, \quad (3.5)$$

that is, the smooth valence wavefunction f is the solution of a modified Schrödinger equation with the same eigenvalue as the all-electron valence wavefunction. Eq. (3.5) can be expressed as

$$(H + V_R) f = e f \quad (3.6)$$

where V_R is a repulsive potential operator. The *pseudopotential* is defined to be the sum of the actual potential, V_A , and V_R

$$V_{ps} = V_A + V_R \quad (3.7)$$

which represents a weakly attractive potential as a consequence of the cancellation between V_A and V_R . Thus, Eq. (3.6) can be written as

$$(T + V_{ps}) f = e f \quad (3.8)$$

where f is regarded as a pseudo-wavefunction.

3.3 Norm-Conserving Pseudopotentials

The norm conservation means that the total amount of charge inside the core region ($r < r_{cl}$) is correctly given by the pseudo-wavefunction. This condition is related to the transferability of the pseudopotential, see below. Such a condition is very important, because it ensures that the electrostatic energy associated with valence electrons is well approximated and that the long-range tail of the electrostatic potential from the nucleus plus the core electrons is correct. In the following, we give the basic requirements of the construction of norm-conserving pseudopotentials [33,41], which have been very widely employed in computational materials science during the last two decades.

A common starting point for generating modern *ab initio* pseudopotentials is self-consistent all-electron (AE) calculation for the atom in a reference configuration. This is

done by solving the radial Schrödinger equation for the valence wavefunctions and their corresponding eigenvalues;

$$\left[-\frac{1}{2} \frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} + V[n] \right] rR_{nl}(r) = e_{nl} rR_{nl}(r) \quad (3.9)$$

where

$$V[n] = -\frac{Z}{r} + V_H[n] + V_{xc}[n] . \quad (3.10)$$

This yields the eigenvalues e_l^{AE} and radial wavefunctions $R_l^{AE}(r)$ (hereafter, the quantum number n will be suppressed). A semilocal screened pseudopotential is then constructed such that:

- (i) The self-consistent calculation for the pseudo atom in the same valence configuration yield the same valence eigenvalues as the real atom,

$$e_l^{ps} = e_l^{AE} . \quad (3.11)$$

- (ii) The normalized radial pseudo wavefunctions are nodeless and equal to the normalized radial all-electron wavefunction at and beyond a chosen core radius r_{cl} , that is

$$R_l^{ps}(r) = R_l^{AE}(r) \quad \text{for } r \geq r_{cl} . \quad (3.12)$$

This condition implies that the charge enclosed in a sphere of radius r_{cl} in the pseudo atom must be identical to that of the real atom. This is called norm conservation

$$\int_0^{r_{cl}} |rR_l^{ps}(r)|^2 dr = \int_0^{r_{cl}} |rR_l^{AE}(r)|^2 dr . \quad (3.13)$$

- (iii) The first and second derivatives of the pseudo wavefunction are equal to the corresponding derivatives of the real wavefunction at $r = r_{cl}$.

The total charge in the sphere of radius r_{cl} is related to the first energy derivative of the logarithmic derivative of R at the sphere boundary [42]

$$-\left[r^2 R^2(r) \frac{d}{de} \frac{d}{dr} \ln R \right]_{r=r_{cl}} = \int r^2 R^2(r) dr . \quad (3.14)$$

Condition (ii), which was first introduced by Hamann *et al* [33], thus implies that the first energy derivative of the pseudo logarithmic derivative is identical to the exact result. As a consequence, the scattering properties of the real ion core are transferred to the pseudo core, which ensures the transferability of the norm-conserving pseudopotentials.

The screened pseudopotential, $V_{scr,l}^{ps}(r)$, is then obtained by inverting the radial Schrödinger equation

$$V_{scr,l}^{ps}(r) = e_l - \frac{l(l+1)}{2r^2} + \frac{1}{2rR_l^{ps}(r)} \frac{d^2}{dr^2} [rR_l^{ps}(r)] . \quad (3.15)$$

Finally, the ionic pseudopotential $V_{ion,l}^{ps}(r)$ is obtained by unscreening: subtracting the Hartree plus exchange and correlation potential calculated from the pseudo charge density,

$$V_{ion,l}^{ps}(r) = V_{scr,l}^{ps}(r) - V_H[n^v(r)] - V_{xc}[n^v(r)] . \quad (3.16)$$

3.3.1 Kerker Scheme

The simplest scheme for generating norm-conserving pseudopotentials, which satisfy the above conditions, have been introduced by Kerker [34]. He has defined a radial pseudo wavefunction inside the core region of the form

$$R_l^{ps}(r) = r^l e^{p(r)} , \quad (3.17)$$

where $p(r)$ is a 4th degree polynomial;

$$p(r) = ar^4 + br^3 + gr^2 + d . \quad (3.18)$$

Note that the linear term in r is absent to avoid a singularity in the pseudopotential at $r = 0$. The four coefficients are determined easily by applying the above conditions. For more details see Ref. [34].

3.3.2 Troullier and Martins Scheme

The Troullier and Martins [40] have modified the Kerker scheme, to highly improve the convergence of the valence pseudo wavefunction in terms of PW's. This is done by imposing further conditions, namely:

1. The third and fourth derivatives of the pseudo wavefunctions at the core radius, r_{c1} , are equal to the corresponding ones of the all-electron wavefunctions.
2. The pseudopotential is analytic at the origin, $V_{l,scr}'(0) = 0$.

These extra conditions lead to pseudopotentials that decay rapidly ($\sim q^{-4}$) in Fourier space. Whereas, the Kerker scheme yields a decay as q^{-2} . Note that the above additional

conditions will increase the number of adjustable parameters that are used in Kerker scheme.

3.4 Kleinman-Bylander Form of Ionic Pseudopotential

The ionic pseudopotential is usually decomposed into a purely local (*loc*) and a semilocal (*sl*) parts;

$$V_{ion}^{ps}(\mathbf{r}) = V_{ion,loc}^{ps}(\mathbf{r}) + \sum_l V_{sl,l}^{ps}(\mathbf{r}) \hat{P}_l. \quad (3.19)$$

Here, \hat{P}_l an angular momentum projection operator for the component of the wavefunction and is given in terms of spherical harmonics, Y_{lm} , as

$$\hat{P}_l = \sum_{m=-l}^l |Y_{lm}\rangle \langle Y_{lm}|. \quad (3.20)$$

$V_{ion,loc}^{ps}(\mathbf{r})$ is a local potential (usually chosen to be one of the components of V_{ion}^{ps}), and $V_{sl,l}^{ps}(\mathbf{r})$ is the semilocal potential for the angular momentum component l ;

$$V_{sl,l}^{ps}(\mathbf{r}) = V_{ion,l}^{ps}(\mathbf{r}) - V_{ion,loc}^{ps}(\mathbf{r}) \quad r > r_{cl}, \quad (3.21)$$

where r_{cl} is a core radius. Kleinman and Bylander (KB) [36] have shown that the semilocal part of the potential can be transformed into a truly nonlocal or (separable in reciprocal space), which highly reduce the computational efforts to calculate the matrix elements of the pseudopotentials between PW's. Their separable form is

$$V_{non-loc,l}^{KB}(\mathbf{r}, \mathbf{r}') = \frac{|V_{sl,l}(\mathbf{r}) \mathbf{y}_l^{ps}(\mathbf{r})\rangle \langle \mathbf{y}_l^{ps}(\mathbf{r}') V_{sl,l}(\mathbf{r}')|}{\langle \mathbf{y}_l^{ps}(\mathbf{r}) | V_{sl,l}(\mathbf{r}) | \mathbf{y}_l^{ps}(\mathbf{r}) \rangle}. \quad (3.22)$$

Here $\mathbf{y}_l^{ps}(\mathbf{r})$ is the atomic pseudo wavefunction of angular momentum l .

Note that KB potential and the semilocal potential give identical results when applied to the pseudo atom. However, they don't produce identical results (but equally valid) when applied in another environment.

3.5 Momentum Space Formalism

In the momentum space formalism the solution of the second order differential single particle Schrödinger equation is transformed to that of an eigenvalue matrix problem. This is done by expanding both the periodic KS potential, $V_{eff}(\mathbf{r})$, and the Bloch

wavefunctions, $y_{nk}(\mathbf{r})$, in terms of PW's. The PW basis set is the simplest and natural choice to expand the pseudo-wavefunction of the solid. Moreover, it provides a high numerical efficiency and accuracy.

The KS potential is periodic with the same periodicity as the Bravais lattice [9,10] of the solid:

$$V_{eff}(\mathbf{r} + \mathbf{R}) = V_{eff}(\mathbf{r}), \quad (3.23)$$

where \mathbf{R} is a lattice translation vector. For this reason, we can apply the Bloch's theorem [9,10], which says that the solutions of Eq. (2.21) have the form

$$y_{nk}(\mathbf{r}) = u_{nk}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}}, \quad (3.24)$$

where \mathbf{k} is a vector inside the first Brillouin Zone (BZ) of the reciprocal lattice [9], n denotes the band index, and $u_{nk}(\mathbf{r})$ is a periodic function with the same periodicity as $V_{eff}(\mathbf{r})$:

$$u_{nk}(\mathbf{r} + \mathbf{R}) = u_{nk}(\mathbf{r}). \quad (3.25)$$

Thus, $u_{nk}(\mathbf{r})$ can be written as

$$u_{nk}(\mathbf{r}) = \sum_{\mathbf{G}} C_{nk}(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}, \quad (3.26)$$

where \mathbf{G} is a reciprocal lattice vector, and $C_{nk}(\mathbf{G})$ are the Fourier expansion coefficients.

The single particle Schrödinger equation can be transformed, after some mathematical manipulations, to the eigenvalue matrix problem:

$$\sum_{\mathbf{G}'} \left[\frac{1}{2} |\mathbf{G}' + \mathbf{k}|^2 d_{\mathbf{G}\mathbf{G}'} + V_{eff}(\mathbf{G} + \mathbf{k}, \mathbf{G}' + \mathbf{k}) \right] C_{nk}(\mathbf{G}') = e_n(\mathbf{k}) C_{nk}(\mathbf{G}). \quad (3.27)$$

Here,

$$V_{eff}(\mathbf{G} + \mathbf{k}, \mathbf{G}' + \mathbf{k}) = \frac{1}{V} \int d\mathbf{r} e^{-i(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}} V_{eff}(\mathbf{r}) e^{i(\mathbf{G}'+\mathbf{k})\cdot\mathbf{r}} \quad (3.28)$$

where V is the volume of the solid. $V_{eff}(\mathbf{G} + \mathbf{k}, \mathbf{G}' + \mathbf{k})$ can be separated into

$$V_{eff}(\mathbf{G} + \mathbf{k}, \mathbf{G}' + \mathbf{k}) = V_{ion}^{ps}(\mathbf{G} + \mathbf{k}, \mathbf{G}' + \mathbf{k}) + V_{XC}(\mathbf{G} - \mathbf{G}') + V_H(\mathbf{G} - \mathbf{G}'), \quad (3.29)$$

where

$$V_{ion}^{ps}(\mathbf{G} + \mathbf{k}, \mathbf{G}' + \mathbf{k}) = \frac{1}{V} \int d\mathbf{r} e^{-i(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}} V_{ion}^{ps}(\mathbf{r}) e^{i(\mathbf{G}'+\mathbf{k})\cdot\mathbf{r}}, \quad (3.30)$$

$$V_{XC}(\mathbf{G} - \mathbf{G}') = \frac{1}{V} \int d\mathbf{r} V_{XC}(\mathbf{r}) e^{i(\mathbf{G} - \mathbf{G}')\mathbf{r}}, \quad (3.31)$$

and

$$V_H(\mathbf{G} - \mathbf{G}') = \frac{1}{V} \int d\mathbf{r} V_H(\mathbf{r}) e^{i(\mathbf{G} - \mathbf{G}')\mathbf{r}}. \quad (3.32)$$

Note that both $V_H(\mathbf{r})$ and the LDA form of $V_{XC}(\mathbf{r})$ are local potentials, and hence, their Fourier transforms depends only on $(\mathbf{G} - \mathbf{G}')$. However, $V_{ion}^{ps}(\mathbf{G} + \mathbf{k}, \mathbf{G}' + \mathbf{k})$ can be written as

$$V_{ion}^{ps}(\mathbf{G} + \mathbf{k}, \mathbf{G}' + \mathbf{k}) = V_{loc}^{ps}(\mathbf{G} - \mathbf{G}') + \sum_l V_{KB,l}^{ps}(\mathbf{G} + \mathbf{k}, \mathbf{G}' + \mathbf{k}), \quad (3.33)$$

where the local part is written as

$$V_{loc}^{ps}(\mathbf{G} - \mathbf{G}') = \frac{1}{V} \int d\mathbf{r} V_{loc}^{ps}(\mathbf{r}) e^{i(\mathbf{G} - \mathbf{G}')\mathbf{r}}, \quad (3.34)$$

and the KB separable form of the pseudopotential leads to

$$\begin{aligned} V_{KB,l}^{ps}(\mathbf{G} + \mathbf{k}, \mathbf{G}' + \mathbf{k}) &= \frac{4p(2l+1)P_l(\cos g)}{\langle \mathbf{y}_l^{ps}(\mathbf{r}) | V_{sl,l}^{ps}(r) | \mathbf{y}_l^{ps}(\mathbf{r}) \rangle} \\ &\times \left[\int_0^\infty \mathbf{y}_l^{ps}(\mathbf{r}) V_{sl,l}^{ps}(r) j_l(|\mathbf{G} + \mathbf{k}|r) r^2 dr \right] \\ &\times \left[\int_0^\infty \mathbf{y}_l^{ps}(\mathbf{r}) V_{sl,l}^{ps}(r) j_l(|\mathbf{G}' + \mathbf{k}|r) r^2 dr \right]. \end{aligned} \quad (3.35)$$

Here, j_l are the order l spherical Bessel functions, P_l are the Legendre polynomials, and

$$\cos g = \frac{(\mathbf{G} + \mathbf{k}) \cdot (\mathbf{G}' + \mathbf{k})}{|\mathbf{G} + \mathbf{k}| \cdot |\mathbf{G}' + \mathbf{k}|}.$$

Note that the introduction of the KB form results in a factorization of the original integral into a product of integrals each involve only \mathbf{G} or \mathbf{G}' .

Eq. (3.27) is the momentum space representation of the Schrödinger equation. Non-trivial solutions are obtained by solving the determinantal equation

$$\left| \left(\frac{1}{2} |\mathbf{G} + \mathbf{k}|^2 - e_n \right) d_{\mathbf{G}\mathbf{G}'} + V_{eff}(\mathbf{G} + \mathbf{k}, \mathbf{G}' + \mathbf{k}) \right| = 0. \quad (3.36)$$

The practical solution of Eq. (3.27) is achieved by using a finite basis set. To provide such a set, the expansion in Eq. (3.26) is truncated: one can select the \mathbf{G} -basis such that

$$\frac{1}{2}|\mathbf{G} + \mathbf{k}|^2 \leq E_{cut}, \quad (3.37)$$

E_{cut} is a certain energy cutoff, which depends on the ionic potentials considered. Finally, Eq. (3.27) is iteratively solved, for the required eigenstates, using the conjugate gradient method [43]. Then, the self-consistent charge density, potential, and the total energy are calculated. The band structure calculations are performed by solving Eq. (3.27), with fixed V_{eff} (the self-consistent one), for the considered \mathbf{k} -points, using also a conjugate gradient method.

CHAPTER 4

WANNIER FUNCTIONS

4.1 Introduction

This chapter is devoted first to introduce the Wannier functions (WF's), their definition, and properties. Then, to introduce the Marzari-Vanderbilt method for determining the maximally-localized set of Wannier functions.

4.2 Definitions

It is possible to define a function, denoted by $w_n(\mathbf{r} - \mathbf{R})$ or $|\mathbf{R}n\rangle$, such that the Bloch functions $y_{n\mathbf{k}}(\mathbf{r})$ in the n -th band are given by

$$y_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} w_n(\mathbf{r} - \mathbf{R}). \quad (4.1)$$

$w_n(\mathbf{r} - \mathbf{R})$ is the WF in cell defined by \mathbf{R} associated with the band n [11]. This can be justified by noting that any Bloch function, $y_{n\mathbf{k}}(\mathbf{r})$, considered as a function of \mathbf{k} for a fixed \mathbf{r} , is periodic in the reciprocal space, therefore, it has a Fourier series expansion in the real space. The WF's play the role of the localized atomic orbitals in the tight-binding approximation [9]. However, their construction does not depend on the localization of the corresponding atomic wavefunctions, but, in the case of delocalized crystal wavefunctions, they don't resemble these atomic wavefunctions. Eq. (4.1) can be inverted easily by multiplying it by $e^{-i\mathbf{k}\cdot\mathbf{R}'}$ and integrating over all values of \mathbf{k} in the first Brillouin Zone (BZ) [9];

$$\begin{aligned} \int_{BZ} d\mathbf{k} y_{n\mathbf{k}}(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{R}'} &= \sum_{\mathbf{R}} \int_{BZ} d\mathbf{k} e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{R}')} w_n(\mathbf{r}-\mathbf{R}) \\ &= \sum_{\mathbf{R}} \frac{(2p)^3}{v_c} d_{\mathbf{R},\mathbf{R}'} w_n(\mathbf{r}-\mathbf{R}) \end{aligned}$$

that is

$$w_n(\mathbf{r}-\mathbf{R}) = \frac{v_c}{(2p)^3} \int_{BZ} d\mathbf{k} e^{-i\mathbf{k}\cdot\mathbf{R}} y_{n\mathbf{k}}(\mathbf{r}). \quad (4.2)$$

Here v_c is the real space primitive cell volume. Note that we have used the relation

$$\int_{BZ} d\mathbf{k} e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{R}')} = \frac{(2p)^3}{v_c} d_{\mathbf{R},\mathbf{R}'}. \quad (4.3)$$

The transformation in Eq. (4.2) is unitary, which implies that WF's form a *complete orthonormal set*. To see this

$$\begin{aligned} \int d\mathbf{r} w_n^*(\mathbf{r}-\mathbf{R}) w_{n'}(\mathbf{r}-\mathbf{R}') &= \langle \mathbf{R}n | \mathbf{R}'n' \rangle \\ &= \left(\frac{v_c}{(2p)^3} \right)^2 \int_{BZ} d\mathbf{k} \int_{BZ} d\mathbf{k}' e^{i\mathbf{k}\cdot\mathbf{R}} e^{-i\mathbf{k}'\cdot\mathbf{R}'} \langle y_{n\mathbf{k}} | y_{n'\mathbf{k}'} \rangle \\ &= \left(\frac{v_c}{(2p)^3} \right)^2 \int_{BZ} d\mathbf{k} \int_{BZ} d\mathbf{k}' e^{i\mathbf{k}\cdot\mathbf{R}} e^{-i\mathbf{k}'\cdot\mathbf{R}'} d(\mathbf{k}'-\mathbf{k}) d_{n,n'} \\ &= d_{n,n'} d_{\mathbf{R},\mathbf{R}'}, \end{aligned} \quad (4.4)$$

Note also that $w_n(\mathbf{r}-\mathbf{R})$ and $w_n(\mathbf{r}-\mathbf{R}')$ transform into each other with a translation of a lattice vector $\mathbf{R}-\mathbf{R}'$.

4.3 Arbitrariness in Definition of WF's

The WF's are not unique. This is a consequence of the phase indeterminacy, $e^{if_n(\mathbf{k})}$, of the Bloch orbitals, $y_{n\mathbf{k}}(\mathbf{r})$, at every wave vector \mathbf{k} ;

$$y_{n\mathbf{k}}(\mathbf{r}) = e^{if_n(\mathbf{k})} u_{n\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}}, \quad (4.5)$$

where $f_n(\mathbf{k})$ is a real periodic function in reciprocal space. However, the phase (or ‘‘gauge’’) transformation does not result into a simple change of the overall phases of the WF's. Their shape and spatial extent will in general be affected. For a single isolated band, the gauge transformation

$$u_{n\mathbf{k}} \rightarrow e^{if_n(\mathbf{k})} u_{n\mathbf{k}} \quad (4.6)$$

is the only allowed type of gauge transformation for changing the set of WF's associated with that band. For a composite group of J bands (i.e. bands that are connected among themselves by degeneracies, but are isolated from all lower or higher bands), the allowed gauge transformation has a more general form:

$$u_{nk} \rightarrow \sum_{m=1}^J U_{mn}^{(\mathbf{k})} u_{mk}, \quad (4.7)$$

where $U^{(\mathbf{k})}$ is a unitary matrix that mixes the bands at wave vector \mathbf{k} . Note that Eq. (4.6) can be regarded as a special case of Eq. (4.7). Thus, the most general operation that transforms the Bloch orbitals into WF's is given by

$$w_n(\mathbf{r} - \mathbf{R}) = \frac{v_c}{(2p)^3} \int_{BZ} d\mathbf{k} \sum_m U_{mn}^{(\mathbf{k})} y_{m\mathbf{k}}(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{R}}. \quad (4.8)$$

This arbitrariness of WF's can be utilized to construct the *maximally-localized* set of WF's. In the next section we introduce the Marzari and Vanderbilt [12] scheme to do just that.

4.4 Marzari-Vanderbilt Method

Marzari and Vanderbilt [12] have introduced a spread functional (Ω) as a measure of the total delocalization or spread of the WF's. Given a set of Bloch orbitals, $y_{m\mathbf{k}}(\mathbf{r})$, the choice of the unitary matrix, $U^{(\mathbf{k})}$, in Eq. (4.7), that minimizes the spread functional can be found. Then, the determination of the maximally-localized WF's (MLWF's) can be established.

4.4.1 Spread Functional in Real Space

The spread functional, Ω , is defined as

$$\Omega = \sum_n \left[\langle r^2 \rangle_n - \bar{\mathbf{r}}_n^2 \right]. \quad (4.9)$$

Here, $\langle r^2 \rangle_n$ and $\bar{\mathbf{r}}_n$ are the expectation values $\langle \mathbf{0}n | r^2 | \mathbf{0}n \rangle$ and $\langle \mathbf{0}n | \mathbf{r} | \mathbf{0}n \rangle$ respectively.

The choice of this form of spread functional is justified in Ref. [12]. This functional is to be minimized with respect to the unitary transformation $U^{(\mathbf{k})}$. It is convenient to decompose Ω into two terms,

$$\begin{aligned}\Omega &= \sum_n \left[\langle r^2 \rangle_n - \left(\sum_{\mathbf{R}m} |\langle \mathbf{R}m | \mathbf{r} | \mathbf{0}n \rangle|^2 - \sum_{\mathbf{R}m \neq \mathbf{0}n} |\langle \mathbf{R}m | \mathbf{r} | \mathbf{0}n \rangle|^2 \right) \right] \\ &= \Omega_I + \tilde{\Omega},\end{aligned}\quad (4.10)$$

where

$$\Omega_I = \sum_n \left[\langle r^2 \rangle_n - \sum_{\mathbf{R}m} |\langle \mathbf{R}m | \mathbf{r} | \mathbf{0}n \rangle|^2 \right] \quad (4.11)$$

and

$$\tilde{\Omega} = \sum_n \sum_{\mathbf{R}m \neq \mathbf{0}n} |\langle \mathbf{R}m | \mathbf{r} | \mathbf{0}n \rangle|^2. \quad (4.12)$$

It is clear that $\tilde{\Omega}$ is positive definite. It can be shown [12] that Ω_I is also positive definite, and, moreover, *gauge-invariant*: invariant under any arbitrary unitary transformation (4.8) of the Bloch orbitals. To see this, from (4.11) we can write

$$\begin{aligned}\Omega_I &= \sum_n \left[\sum_a \langle \mathbf{0}n | r_a^2 | \mathbf{0}n \rangle - \sum_a \sum_{\mathbf{R}m} \langle \mathbf{0}n | r_a | \mathbf{R}m \rangle \langle \mathbf{R}m | r_a | \mathbf{0}n \rangle \right] \\ &= \sum_{na} \left[\langle \mathbf{0}n | r_a^2 | \mathbf{0}n \rangle - \langle \mathbf{0}n | r_a P r_a | \mathbf{0}n \rangle \right] \\ &= \sum_{na} \left[\langle \mathbf{0}n | r_a (1 - P) r_a | \mathbf{0}n \rangle \right]\end{aligned}$$

that is

$$\Omega_I = \sum_{na} \langle \mathbf{0}n | r_a Q r_a | \mathbf{0}n \rangle. \quad (4.13)$$

Here, P is the projection operator onto the group of bands under consideration and Q is the projection operator onto all other bands, namely

$$P = \sum_{\mathbf{R}n} |\mathbf{R}n\rangle \langle \mathbf{R}n| = \sum_{nk} |y_{nk}\rangle \langle y_{nk}| \quad (4.14)$$

and

$$Q = 1 - P. \quad (4.15)$$

Then, from (4.13) we get

$$\begin{aligned}\Omega_I &= \sum_a \text{tr}[P r_a Q r_a] \\ &= \|PxQ\|^2 + \|PyQ\|^2 + \|PzQ\|^2.\end{aligned}\quad (4.16)$$

Here $\|A\|^2 = \text{tr}[A^\dagger A]$. It is clear from the last expression that Ω_1 is positive definite. It is also gauge invariant as mentioned previously. Accordingly, the minimization of Ω corresponds to the minimization of only $\tilde{\Omega}$.

Note that $\tilde{\Omega}$ can, in turn, be decomposed into two parts

$$\tilde{\Omega} = \Omega_{\text{OD}} + \Omega_{\text{D}}, \quad (4.17)$$

where

$$\Omega_{\text{OD}} = \sum_{m \neq n} \sum_{\mathbf{R}} \left| \langle \mathbf{R}m | \mathbf{r} | \mathbf{0}n \rangle \right|^2 \quad (4.18)$$

and

$$\Omega_{\text{D}} = \sum_n \sum_{\mathbf{R} \neq \mathbf{0}} \left| \langle \mathbf{R}n | \mathbf{r} | \mathbf{0}n \rangle \right|^2. \quad (4.19)$$

Ω_{OD} is the band-off-diagonal part and Ω_{D} is the band-diagonal part.

4.4.2 Spread Functional in Momentum Space

It is shown in Ref. [44] that the matrix elements of the position operator between WF's take the form

$$\langle \mathbf{R}n | \mathbf{r} | \mathbf{0}m \rangle = i \frac{v_c}{(2p)^3} \int d\mathbf{k} e^{i\mathbf{k} \cdot \mathbf{R}} \langle u_{n\mathbf{k}} | \nabla_{\mathbf{k}} | u_{m\mathbf{k}} \rangle, \quad (4.20)$$

and

$$\langle \mathbf{R}n | r^2 | \mathbf{0}m \rangle = - \frac{v_c}{(2p)^3} \int d\mathbf{k} e^{i\mathbf{k} \cdot \mathbf{R}} \langle u_{n\mathbf{k}} | \nabla_{\mathbf{k}}^2 | u_{m\mathbf{k}} \rangle. \quad (4.21)$$

The above two equations can be inverted by multiplying by $e^{-i\mathbf{k}' \cdot \mathbf{R}}$ then summing over \mathbf{R} , and using

$$\sum_{\mathbf{R}} e^{i(\mathbf{k}-\mathbf{k}') \cdot \mathbf{R}} = \frac{(2p)^3}{v_c} d_{\mathbf{k}, \mathbf{k}'} \quad (4.22)$$

we get,

$$\langle u_{n\mathbf{k}} | \nabla_{\mathbf{k}} | u_{m\mathbf{k}} \rangle = -i \sum_{\mathbf{R}} e^{-i\mathbf{k} \cdot \mathbf{R}} \langle \mathbf{R}n | \mathbf{r} | \mathbf{0}m \rangle, \quad (4.23)$$

$$\langle u_{n\mathbf{k}} | \nabla_{\mathbf{k}}^2 | u_{m\mathbf{k}} \rangle = - \sum_{\mathbf{R}} e^{-i\mathbf{k} \cdot \mathbf{R}} \langle \mathbf{R}n | r^2 | \mathbf{0}m \rangle. \quad (4.24)$$

Thus, $\bar{\mathbf{r}}_n$ and $\langle r^2 \rangle_n$ can be written as

$$\bar{\mathbf{r}}_n = i \frac{v_c}{(2p)^3} \int d\mathbf{k} \langle u_{n\mathbf{k}} | \nabla_{\mathbf{k}} | u_{n\mathbf{k}} \rangle, \quad (4.25)$$

$$\langle r^2 \rangle_n = - \frac{v_c}{(2p)^3} \int d\mathbf{k} \langle u_{n\mathbf{k}} | \nabla_{\mathbf{k}}^2 | u_{n\mathbf{k}} \rangle$$

Now, integrating the last equation by parts gives

$$\langle r^2 \rangle_n = \frac{v_c}{(2p)^3} \int d\mathbf{k} \left| \nabla_{\mathbf{k}} u_{n\mathbf{k}} \right|^2. \quad (4.26)$$

We shall now derive expressions for Ω and its different parts in terms of a discretized \mathbf{k} -space mesh. Accordingly, we make the substitution

$$\frac{v_c}{(2p)^3} \int d\mathbf{k} \rightarrow \frac{1}{N} \sum_{\mathbf{k}}, \quad (4.27)$$

where N is the number of \mathbf{k} -points in the first BZ, or equivalently, the number of real-space cells in the system. If the Brillouin Zone has been discretized into a uniform Monkhorst-Pack mesh [45] (see Appendix A) containing Z \mathbf{k} -points, then the simplest finite difference expressions for, the gradient, $\nabla_{\mathbf{k}}$ (i.e. involving the smallest possible shell) of a smooth function of \mathbf{k} , $f(\mathbf{k})$, is given by [12]

$$\nabla f(\mathbf{k}) = \sum_{\mathbf{b}} w_b \mathbf{b} [f(\mathbf{k} + \mathbf{b}) - f(\mathbf{k})], \quad (4.28)$$

and

$$|\nabla f(\mathbf{k})|^2 = \sum_{\mathbf{b}} w_b [f(\mathbf{k} + \mathbf{b}) - f(\mathbf{k})]^2. \quad (4.29)$$

Here, \mathbf{b} is a vector connecting each \mathbf{k} -point to one of its nearest neighbors and w_b is the weight associated with each shell $|\mathbf{b}| = b$. For any given \mathbf{k} -point in a cubic mesh $w_b = 3/Zb^2$, with $Z = 6, 8$, or 12 for simple cubic, body centered cubic and face centered cubic \mathbf{k} -space meshes respectively [12].

Therefore, starting from (4.25) and (4.28) we now have

$$\begin{aligned} \bar{\mathbf{r}}_n &= \frac{i}{N} \sum_{\mathbf{k}} \left\langle u_{n\mathbf{k}} \left| \sum_{\mathbf{b}} w_b \mathbf{b} (u_{n,\mathbf{k}+\mathbf{b}} - u_{n\mathbf{k}}) \right. \right\rangle \\ &= \frac{i}{N} \sum_{\mathbf{k}, \mathbf{b}} w_b \mathbf{b} \left[\langle u_{n\mathbf{k}} | u_{n,\mathbf{k}+\mathbf{b}} \rangle - \langle u_{n\mathbf{k}} | u_{n\mathbf{k}} \rangle \right] \end{aligned}$$

$$= \frac{i}{N} \sum_{\mathbf{k}, \mathbf{b}} w_b \mathbf{b} \left[\langle u_{n\mathbf{k}} | u_{n, \mathbf{k}+\mathbf{b}} \rangle - 1 \right], \quad (4.30)$$

and from (4.26) and (4.28) we get

$$\begin{aligned} \langle r^2 \rangle_n &= \frac{1}{N} \sum_{\mathbf{k}, \mathbf{b}} w_b \left[| \langle u_{n, \mathbf{k}+\mathbf{b}} \rangle - | u_{n\mathbf{k}} \rangle | \right]^2 \\ &= \frac{1}{N} \sum_{\mathbf{k}, \mathbf{b}} w_b \left[2 - \langle u_{n\mathbf{k}} | u_{n, \mathbf{k}+\mathbf{b}} \rangle^* - \langle u_{n\mathbf{k}} | u_{n, \mathbf{k}+\mathbf{b}} \rangle \right] \\ &= \frac{1}{N} \sum_{\mathbf{k}, \mathbf{b}} w_b \left[2 - 2 \operatorname{Re} \langle u_{n\mathbf{k}} | u_{n, \mathbf{k}+\mathbf{b}} \rangle \right]. \end{aligned} \quad (4.31)$$

To find the final expressions for Ω and its different parts, a desirable property should be taken into account. Namely, if the WF's are shifted by a lattice vector, \mathbf{R} , that is when $u_{n\mathbf{k}} \rightarrow u_{n\mathbf{k}} e^{-i\mathbf{k} \cdot \mathbf{R}}$, one should find

$$\begin{aligned} \bar{\mathbf{r}}_n &\rightarrow \bar{\mathbf{r}}_n + \mathbf{R} \\ \langle r^2 \rangle_n &\rightarrow \langle r^2 \rangle_n + 2 \bar{\mathbf{r}}_n \cdot \mathbf{R} + R^2 \end{aligned} \quad (4.32)$$

so that Ω remains unchanged. Expressions (4.30) and (4.31) can be modified to obey the above requirements. Thus, the expressions for the expectation values of \mathbf{r} and r^2 are given by [12]

$$\bar{\mathbf{r}}_n = -\frac{1}{N} \sum_{\mathbf{k}, \mathbf{b}} w_b \mathbf{b} \operatorname{Im} \ln M_{mn}^{(\mathbf{k}, \mathbf{b})} \quad (4.33)$$

$$\langle r^2 \rangle_n = \frac{1}{N} \sum_{\mathbf{k}, \mathbf{b}} w_b \left\{ \left[1 - |M_{mn}^{(\mathbf{k}, \mathbf{b})}|^2 \right] + \left[\operatorname{Im} \ln M_{mn}^{(\mathbf{k}, \mathbf{b})} \right]^2 \right\}, \quad (4.34)$$

where

$$M_{mn}^{(\mathbf{k}, \mathbf{b})} = \langle u_{m\mathbf{k}} | u_{n, \mathbf{k}+\mathbf{b}} \rangle. \quad (4.35)$$

Hence, one can find that equations (4.11), (4.18), and (4.19) become

$$\begin{aligned} \Omega_{\mathbf{I}} &= \frac{1}{N} \sum_{\mathbf{k}, \mathbf{b}} w_b \left(J - \sum_{mn} |M_{mn}^{(\mathbf{k}, \mathbf{b})}|^2 \right) \\ &= \frac{1}{N} \sum_{\mathbf{k}, \mathbf{b}} w_b \operatorname{tr} [P^{(\mathbf{k})} Q^{(\mathbf{k}+\mathbf{b})}], \end{aligned} \quad (4.36)$$

$$\Omega_{\text{OD}} = \frac{1}{N} \sum_{\mathbf{k}, \mathbf{b}} w_b \sum_{m \neq n} |M_{mn}^{(\mathbf{k}, \mathbf{b})}|^2, \quad (4.37)$$

and

$$\Omega_D = \frac{1}{N} \sum_{\mathbf{k}, \mathbf{b}} w_b \sum_n \left(-\text{Im} \ln M_{mn}^{(\mathbf{k}, \mathbf{b})} - \mathbf{b} \cdot \bar{\mathbf{r}}_n \right)^2. \quad (4.38)$$

Note that it is now clear from the expression (4.36) that Ω_I is gauge invariant.

4.4.3 Gradient of Spread Functional

Consider the first-order change of the spread functional Ω arising from an infinitesimal gauge transformation (4.7), given by

$$U_{mn}^{(\mathbf{k})} = d_{mn} + dW_{mn}^{(\mathbf{k})} \quad (4.39)$$

where dW is an infinitesimal anti-hermitian matrix, $dW^\dagger = -dW$. Inserting (4.39) in (4.7) yields

$$u_{nk} \rightarrow u_{nk} + \sum_m dW_{mn}^\dagger u_{mk}. \quad (4.40)$$

Marzari and Vanderbilt [12] have utilized this change in $U^{(\mathbf{k})}$ to derive a very elegant expression for the gradient of Ω , given as

$$G^{(\mathbf{k})} = \frac{d\Omega}{dW^{(\mathbf{k})}} = 4 \sum_{\mathbf{b}} w_b \left(A[R^{(\mathbf{k}, \mathbf{b})}] - S[T^{(\mathbf{k}, \mathbf{b})}] \right). \quad (4.41)$$

Here, A and B are the operators

$$A[B] = \frac{B - B^\dagger}{2}, \quad \text{and} \quad S[B] = \frac{B + B^\dagger}{2i},$$

with $R^{(\mathbf{k}, \mathbf{b})}$ and $T^{(\mathbf{k}, \mathbf{b})}$ are defined as

$$R_{mn}^{(\mathbf{k}, \mathbf{b})} = M_{mn}^{(\mathbf{k}, \mathbf{b})} M_{mn}^{(\mathbf{k}, \mathbf{b})*},$$

$$T_{mn}^{(\mathbf{k}, \mathbf{b})} = \tilde{R}_{mn}^{(\mathbf{k}, \mathbf{b})} q_n^{(\mathbf{k}, \mathbf{b})},$$

and

$$\tilde{R}_{mn}^{(\mathbf{k}, \mathbf{b})} = \frac{M_{mn}^{(\mathbf{k}, \mathbf{b})}}{M_{nn}^{(\mathbf{k}, \mathbf{b})}}, \quad \text{and} \quad q_n^{(\mathbf{k}, \mathbf{b})} = \text{Im} \ln M_{nn}^{(\mathbf{k}, \mathbf{b})} + \mathbf{b} \cdot \bar{\mathbf{r}}_n.$$

The condition required for having found a minimum is that the gradient, $G^{(\mathbf{k})}$, given by Eq. (4.41), should vanish. The most important thing to note here is that the gradient depends only on the overlap matrices, $M_{mn}^{(\mathbf{k}, \mathbf{b})}$.

4.4.4 Steepest-Descent Minimization

The simplest (but not necessarily the most efficient) method to minimize Ω is the steepest-descent. Using small updates to the unitary matrices, as in Eq. (4.39), with the choice

$$dW^{(k)} = e G^{(k)} \quad (4.42)$$

where e is a positive infinitesimal, leads to

$$\begin{aligned} d\Omega &= \sum_{\mathbf{k}} \text{tr}[G^{(k)} dW^{(k)}] \\ &= -e \sum_{\mathbf{k}} \|G^{(k)}\|^2. \end{aligned}$$

Here $\|A\|^2 = \sum_{mn} |A_{mn}|^2$ and $G^\dagger = -G$. Thus, the choice (4.42) guarantees $d\Omega < 0$, which always leads to a reduction in Ω . We take $e = a/4w$, where $w = \sum_{\mathbf{b}} w_{\mathbf{b}}$, so that

$$\Delta W^{(k)} = \frac{a}{w} \sum_{\mathbf{b}} w_{\mathbf{b}} \left(A[R^{(k,\mathbf{b})}] - S[T^{(k,\mathbf{b})}] \right), \quad (4.43)$$

and, then, the wavefunctions are updated according to the matrix $\exp[\Delta W^{(k)}]$ which is unitary because ΔW is anti-hermitian. Following the choice of Marzari and Vanderbilt [12], The pre-factor a was set to 0.4.

The algorithm is as follows. First, we start with a reference set of Bloch orbitals and evaluate the overlap matrices

$$M_{mn}^{(0)(\mathbf{k},\mathbf{b})} = \left\langle u_{m\mathbf{k}}^{(0)} \left| u_{n,\mathbf{k}+\mathbf{b}}^{(0)} \right. \right\rangle, \quad (4.44)$$

and we choose our starting $U_{mn}^{(0)(\mathbf{k})}$ to be d_{mn} . Then we evaluate $\Delta W^{(0)(\mathbf{k})}$ given by Eq. (4.43) and update the unitary matrices, according to

$$U^{(N)(\mathbf{k})} = U^{(N-1)(\mathbf{k})} \exp[\Delta W^{(N-1)(\mathbf{k})}], \quad (4.45)$$

where N is used to denote the cycle number. After that, we evaluate a new set of the $M^{(\mathbf{k},\mathbf{b})}$ matrices according to

$$M^{(N)(\mathbf{k},\mathbf{b})} = U^{(N)(\mathbf{k})\dagger} M^{(0)(\mathbf{k},\mathbf{b})} U^{(N)(\mathbf{k}+\mathbf{b})}. \quad (4.46)$$

This cycle is repeated until convergence is obtained. It is clear that this evolution towards the minimum requires only relatively inexpensive updating of the unitary matrices. The overlap matrices, $M^{(\mathbf{k},\mathbf{b})}$, have to be calculated only once.

To highly improve the efficiency of the minimization procedure, the reference Bloch orbitals are prepared in order to make the starting Wannier functions somewhat localized. The other advantage of this step is to help to avoid getting trapped in a local minima [12]. We use the same preparation method used by Marzari and Vanderbilt [12]. In this method, the set of reference Bloch orbitals are prepared by projecting them onto Gaussian functions, $g_n(\mathbf{r})$, centered at mid-bond positions. This is done as follows, first

$$|\mathbf{f}_{nk}\rangle = \sum_m |\mathbf{y}_{mk}\rangle \langle \mathbf{y}_{mk} | g_n \rangle, \quad (4.47)$$

then, the resulting orbitals \mathbf{f}_{nk} are orthonormalized via Löwdin orthogonalization scheme

$$|\tilde{\mathbf{f}}_{nk}\rangle = \sum_m (S^{-1/2})_{mn} |\mathbf{f}_{mk}\rangle, \quad (4.48)$$

where $S_{mn} = \langle \mathbf{f}_{mk} | \mathbf{f}_{nk} \rangle$, and finally reconverted to cell-periodic functions with

$$u_{nk}^{(0)} = e^{-i\mathbf{k}\cdot\mathbf{r}} \tilde{\mathbf{f}}_{nk}(\mathbf{r}). \quad (4.49)$$

Another preparation method is used in Ref. [46]. This method involves a gauge transformation for each band given by

$$\mathbf{y}_{nk}(\mathbf{r}) \rightarrow \mathbf{y}_{nk}(\mathbf{r}) \exp[-i \operatorname{Im} \ln \mathbf{y}_{nk}(\mathbf{r}_n)].$$

To make this method work well, \mathbf{r}_n should be chosen where the Wannier functions are expected to be reasonably large [46].

Note that the exponential $\exp[\Delta W^{(k)}]$ is a matrix operation, which is performed as follows. First, we use the representation $\Delta W^{(k)} = -iH$, where H is a Hermitian matrix ($H = i\Delta W$). Next, we diagonalize H , i.e. $H = CDC^{-1}$ where D is diagonal. Then we have

$$\begin{aligned} \exp(-iH) &= \exp(-iCDC^{-1}) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} (-iCDC^{-1})^n \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} C(-iD)^n C^{-1} \\ &= C \exp(-iD) C^{-1} \end{aligned}$$

$\exp(-iD)$ is calculated by just taking the exponential of each element of $-iD$.

CHAPTER 5

COMPUTATIONAL DETAILS

5.1 Introduction

Our calculations are performed within the local density approximation (LDA) for the exchange-correlation potential, using a pseudopotential plane-wave (PP-PW) approach. In this chapter we introduce the used computational details. First we will describe the lattice, structure, and reciprocal lattice for the studied systems. Next, we will state some details used in the self-consistent calculations. Finally, we give some details related to the construction of the MLWF's.

5.2 Lattice

The underlying Bravais lattice of all the studied crystals is the face-centered cubic (FCC). In the FCC lattice, the lattice points are located at the corners and at the centers of the faces of a periodically repeated cub, as shown in Fig. 5.1. Thus each lattice point has 12 nearest neighbors. The conventional cell contains four lattice points. Any two lattice points are connected by a lattice translation vector

$$\mathbf{R} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3, \quad (5.1)$$

where $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ are the primitive translation vectors, and n_1, n_2, n_3 are integers. In the case of FCC Bravais lattice, we have

$$\mathbf{a}_1 = \frac{1}{2} a(\hat{\mathbf{x}} + \hat{\mathbf{y}}); \quad \mathbf{a}_2 = \frac{1}{2} a(\hat{\mathbf{y}} + \hat{\mathbf{z}}); \quad \mathbf{a}_3 = \frac{1}{2} a(\hat{\mathbf{x}} + \hat{\mathbf{z}}). \quad (5.2)$$

Here, a is the edge length of the conventional cubic cell (or the lattice constant), and $\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$ are the Cartesian unit vectors. Hence, the primitive cell volume is $a^3 / 4$. The

angle between any two primitive translation vectors is 60° . In Table 5.1, we list the used experimental lattice constants for the studied systems.

Table 5.1: Lattice constants (a) in Bohr units.

System	GaAs	AlAs	GaN	AlN	SiC	Si	Ge	C
a	10.680	10.680	8.532	8.249	8.230	10.260	10.680	6.739

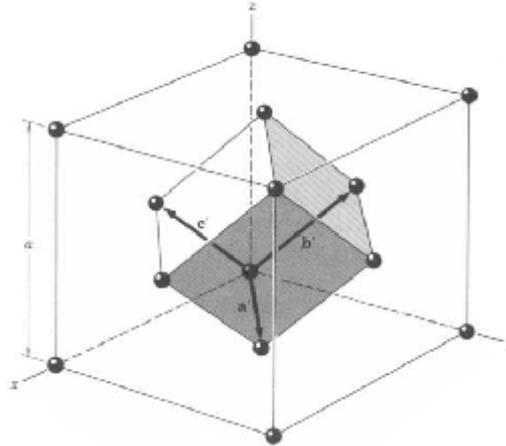


Fig. 5.1: FCC lattice. The bigger cell is the conventional unit cell and the other one is the primitive unit cell. Here a' , b' , c' are the primitive translation vectors.

5.3 Structures

Si, Ge and C crystallize in the diamond structure, while GaAs, AlAs, GaN, AlN and SiC all crystallize in the zincblende structure. In the diamond structure, the primitive unit cell has two identical atoms at $(0,0,0)$ and $(\frac{1}{4},\frac{1}{4},\frac{1}{4})a$ associated with each lattice point (of the FCC lattice), as shown in Fig. 5.2. Hence, the conventional cell contains eight atoms. The tetrahedral bonds in diamond structure are also clear in Fig. 5.2. Each atom has four nearest neighbors. The zincblende structure has equal number of two types of atoms distributed on the atomic sites of the diamond structure so that each atom has four nearest neighbors of the other kind, as shown in Fig. 5.3. Note that the diamond structure allows a center of inversion symmetry operation (which maps an atom at \mathbf{r} into an atom at $-\mathbf{r}$) at each midpoint between nearest neighboring atoms, while zincblende structure does not have inversion symmetry.

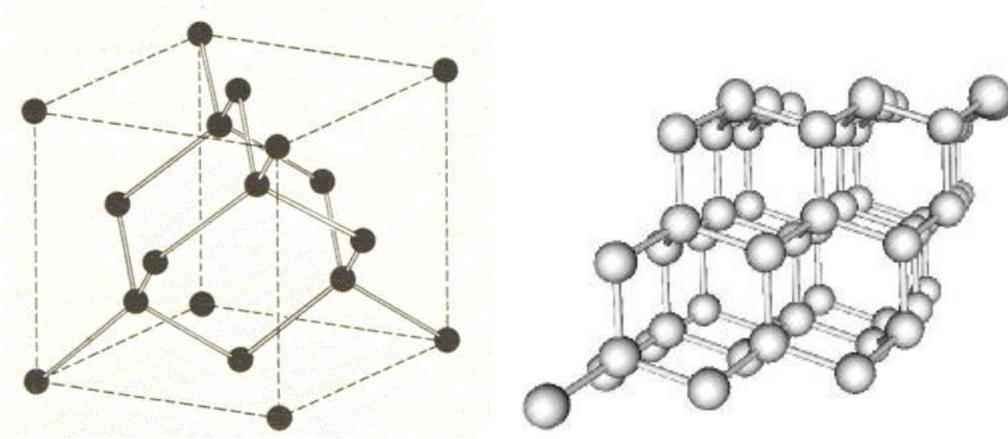


Fig. 5.2: Diamond crystal structure. The tetrahedral bonds are clear in both figures.

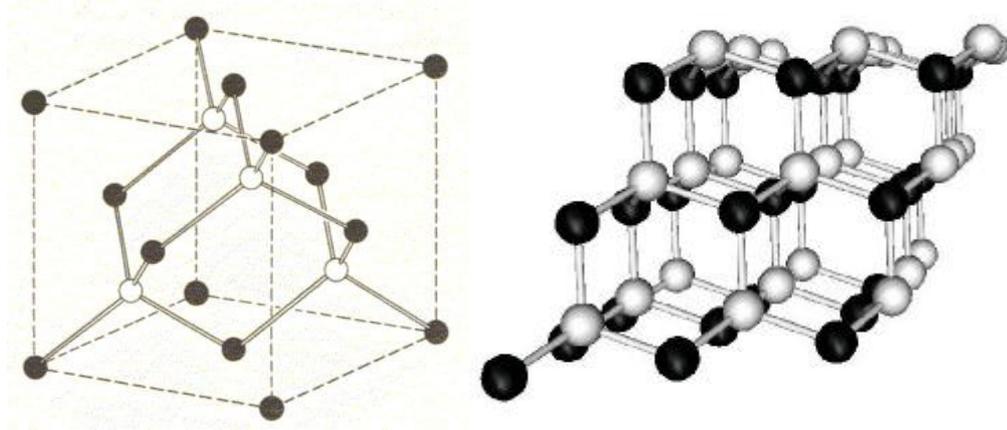


Fig. 5.3: Zinblende crystal structure.

5.4 Reciprocal Lattice

The reciprocal lattice vectors are defined via

$$\mathbf{G} = v_1 \mathbf{b}_1 + v_2 \mathbf{b}_2 + v_3 \mathbf{b}_3, \quad (5.3)$$

where v_1, v_2, v_3 are integers, and $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ are the primitive vectors in the reciprocal space. The primitive reciprocal lattice vectors to the FCC lattice are:

$$\mathbf{b}_1 = \frac{2p}{a}(-\hat{x} + \hat{y} + \hat{z}); \quad \mathbf{b}_2 = \frac{2p}{a}(\hat{x} - \hat{y} + \hat{z}); \quad \mathbf{b}_3 = \frac{2p}{a}(\hat{x} + \hat{y} - \hat{z}). \quad (5.4)$$

This means that the reciprocal to the FCC lattice is a Body-Centered Cubic (BCC) lattice. The shortest \mathbf{G} vectors are the eight vectors:

$$\frac{2p}{a}(\pm\hat{x} \pm \hat{y} \pm \hat{z}). \quad (5.5)$$

The first Brillouin Zone (BZ) is the Wigner-Seitz primitive unit cell of the reciprocal lattice. The Wigner-Seitz cell about a lattice point is the region of space that is closer to that point than to any other lattice point. The first Brillouin zone for the FCC lattice is

shown in Fig. 5.4. The \mathbf{k} -points shown are called high symmetry \mathbf{k} -points, which lie along high symmetry directions of the first BZ.

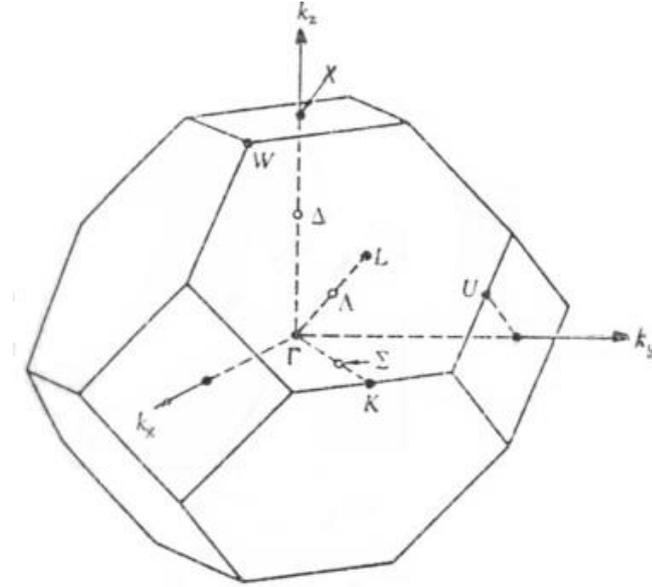


Fig. 5.4: Brillouin zone of the FCC lattice. The three axes intersect the BZ at $(2\pi/a)(100)$, $(2\pi/a)(010)$, and $(2\pi/a)(001)$.

5.5 Self-Consistent Calculations

Our calculations are performed using a standard PP-PW approach (see chapter 3). We have used norm-conserving pseudopotentials in the Kleinman-Bylander [36] representation, generated by using the Troullier and Martins scheme [40]. The used electronic configuration, core radii and local potentials (see chapter 3) are listed in Table 5.2. These pseudopotentials were carefully checked, and they are found to have very good transferability and free from ghost states [47]. For each of the atoms considered, except for N, a d-component is included which corresponds to the lowest unoccupied d-state. In order to be able to use reasonable core radii for this component for Ge and Ga we used reference energies of 15 and 20 eV, respectively.

The self-consistent calculations for the different systems were performed using a regular $4 \times 4 \times 4$ Monkhorst-Pack (MP) mesh [45] of \mathbf{k} -points (see Appendix A). For the FCC lattice, this mesh of \mathbf{k} -points is reduced by symmetry to only ten special ones.

The PW energy cutoffs used for the systems considered are shown in Table 5.3. These cutoffs are found to give excellent convergence. One should note that the occupied p-states of the first row elements (C and N) are highly localized. Hence, systems

involving such atoms require quite large cutoff energy. For example, we show in Fig. 5.5 the used LDA-PP's for C and Si. It is worth noting that the p -component of the ionic pseudopotential of C is much deeper than that of Si.

Table 5.2: Data used for the pseudopotential generation for the atoms involved. Here, Z is the atomic number, N_v is the number of valence electrons, and l_{local} is the angular momentum for the local components of the ionic pseudopotentials. r_{cl} are in Bohr units.

atom	Z	valence electronic configuration	N_v	core radii r_{cl}			l_{local}
				$l=0$	$l=1$	$l=2$	
C	6	$2s^2 2p^2$	4	1.5	1.5	1.5	2
N	7	$2s^2 2p^3$	5	1.5	1.5		1
Al	13	$3s^2 3p^1$	3	1.9	1.9	1.9	2
Si	14	$3s^2 3p^2$	4	2.1	2.1	2.1	2
Ga	31	$4s^2 4p^1$	3	2.1	2.1	2.2	0
Ge	32	$4s^2 4p^2$	4	2.2	2.2	2.2	2
As	33	$4s^2 4p^3$	5	2.1	2.1	2.1	2

Table 5.3: Plane wave energy cutoffs (E_{cut}) in Rydberg units (1 Hartree = 2 Rydberg).

System	GaAs	AlAs	GaN	AlN	SiC	Si	Ge	C
E_{cut}	25	25	60	60	60	25	25	60

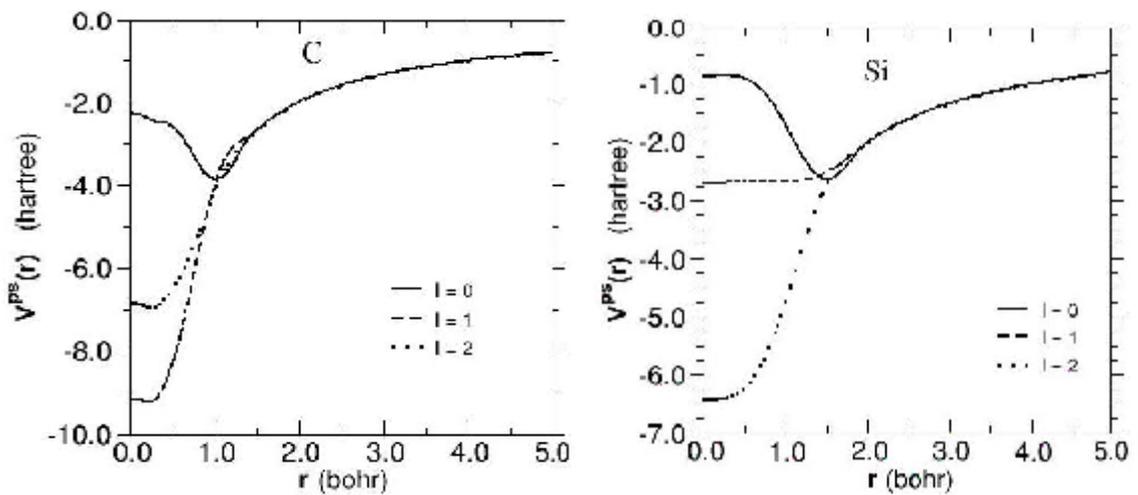


Fig. 5.5: Ionic pseudopotentials of C and Si.

5.6 Construction of the MLWF's

5.6.1 Reciprocal Space Grids

In the calculation of the overlap matrix elements, given by Eq. (4.35), the full sets (not the reduced ones) of MP meshes were used. The Bloch eigenfunctions and eigenvalues are obtained via band structure calculation, at fixed (self-consistent) KS potential. Moreover, as suggested by Marzari and Vanderbilt [12], the MP meshes were shifted in order to include the Γ point ($\mathbf{k}=0$). Since the lattice is FCC in real space, the grid is BCC in reciprocal space, and we have used the simplest possible representation of $\nabla_{\mathbf{k}}$ using only the first nearest neighbors ($Z=8$) of each \mathbf{k} -point [12]. However, for GaAs, we have used two other kinds of meshes. The first one is a uniform MP mesh that was shifted such that for each \mathbf{k} there is a $-\mathbf{k}$, without the inclusion of the Γ point. The second one is an equi-spaced cubic mesh [46], in this case we have 6 nearest neighbors for each \mathbf{k} -point. Note that all the reciprocal space grids, which are used in the calculation of the overlap matrix elements, satisfy the following condition:

$$\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}} = Nd_{\mathbf{R},0}, \quad (5.6)$$

which is clearly the same as equation (4.3) but with a discretized \mathbf{k} -space.

In the calculation of the overlap matrix elements, one shall specify the set of nearest neighbors (or their equivalents) for each \mathbf{k} -point among the whole set. In the case of FCC lattice, we have eight nearest neighbors for each \mathbf{k} -point (since we have BCC lattice in reciprocal space). Fig. 5.6 demonstrates finding the set of nearest neighbors (or their equivalents) for three different \mathbf{k} -points in a $4\times 4\times 4$ MP mesh for the FCC lattice. In Fig. 5.6a, the specified \mathbf{k} -point (the larger dark point) has eight nearest neighbors (the small dark points) that all lie inside the same cell. This is not always the case. In Fig. 5.6b, the specified \mathbf{k} -point (located on the upper face) has six nearest neighbors inside the same cell, and two nearest neighbors outside it (not shown). However, when folded back into the considered cell, the two neighbors outside it map onto two existing \mathbf{k} -points in the shown mesh, which lie on the opposite face. These two points are marked by dark points. The mapping is done by adding a certain \mathbf{G} -vector, such that $\mathbf{k} = \mathbf{k}' + \mathbf{G}$, where \mathbf{k}' is outside the specified cell and \mathbf{k} is inside it. The situation becomes a bit more complicated if the considered point lies at the corner of the mesh, as shown in Fig. 5.6c. Each of such

points has four nearest neighbors inside and four outside the cell. Similar arguments are applied for the other \mathbf{k} -points.

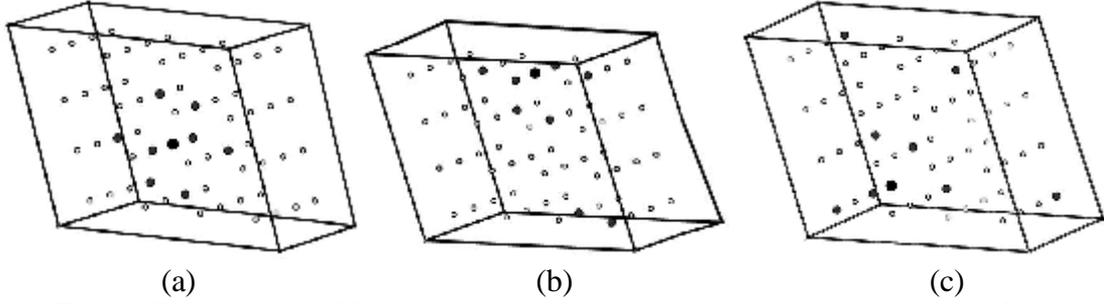


Fig. 5.6: Demonstration of finding the set of nearest neighbors (or their equivalents) for 3 different \mathbf{k} -points in a $4 \times 4 \times 4$ MP mesh for the FCC lattice. The specified point is the bigger dark one, and its nearest neighbors (or their equivalents) are the marked by smaller dark ones.

In order to obtain the periodic part of the Bloch's orbitals of the neighboring \mathbf{k} -points that lie outside the considered mesh, we note that

$$y_{n\mathbf{k}'}(\mathbf{r}) = y_{n\mathbf{k}}(\mathbf{r}). \quad (5.7)$$

Here, \mathbf{k} and \mathbf{k}' are as defined above. This is because they are identical points, ensured by the periodicity of the unit cell in reciprocal space. This implies that

$$u_{n\mathbf{k}'}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r}) e^{i\mathbf{G}' \cdot \mathbf{r}}. \quad (5.8)$$

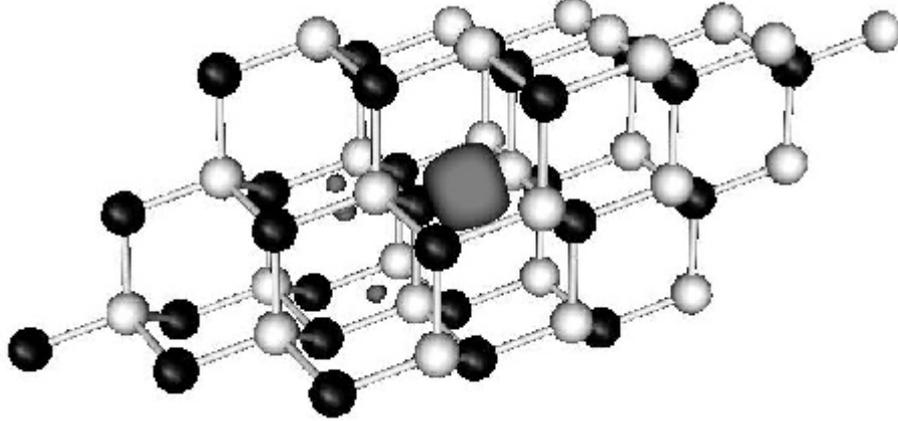
Therefore, $u_{n\mathbf{k}}$ are only explicitly calculated for the considered MP mesh, and that of the \mathbf{k}' -points are obtained from them by using the above relation. The calculations are performed in reciprocal space, and thus, the PW's coefficients of the Bloch's orbitals at the \mathbf{k}' -points are obtained according to the relation

$$C_{n\mathbf{k}'}(\mathbf{G}') = C_{n\mathbf{k}}(\mathbf{G}' - \mathbf{G}). \quad (5.9)$$

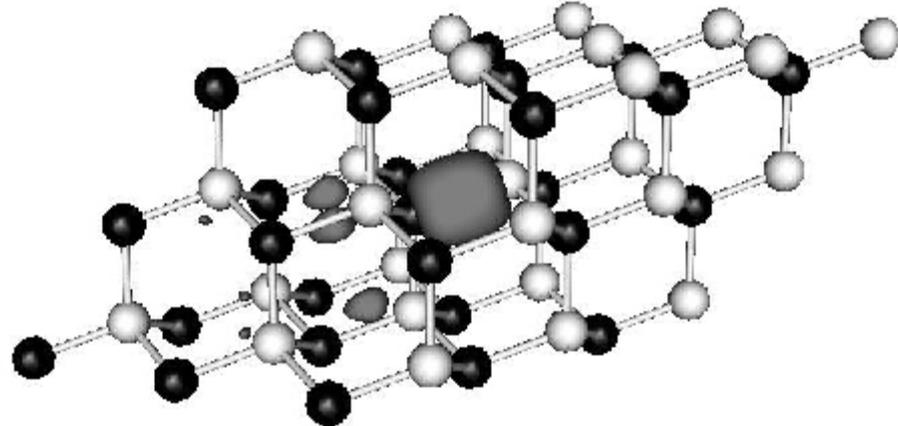
In order to investigate the effect of choosing a certain kind of the different \mathbf{k} -space meshes, we show in Figs. 5.7 one of the MLWF's in GaAs obtained by using three different kinds of \mathbf{k} -space meshes. Fig. 5.7a shows MLWF obtained using a uniform $8 \times 8 \times 8$ MP mesh (i.e. contains 512 \mathbf{k} -points) that was shifted in order to include the Γ -point. In this case, the minimized spread functional equals to 8.871 \AA^2 . Fig. 5.7b shows the corresponding one obtained by using the same MP mesh but was shifted in order to have, for each \mathbf{k} , a $-\mathbf{k}$. The minimized spread functional is now 8.981 \AA^2 . Fig. 5.7c shows the corresponding MLWF obtained by using a uniform cubic mesh of \mathbf{k} -points. This mesh contains 500 \mathbf{k} -points that lie inside the first BZ of the FCC lattice. Using this mesh, the minimized spread functional was 9.208 \AA^2 . Thus, although all of the

above meshes contain comparable number of \mathbf{k} -points, the spread functional depends on the kind of the set \mathbf{k} -points used. Nevertheless, we note that all of the MLWF's obtained above have the same location of the Wannier centers, and all of them have similar shape and confinement in the primitive unit cell. Thus, the results reported in the next chapter are constructed by using \mathbf{k} -meshes of the first kind.

a)



b)



c)

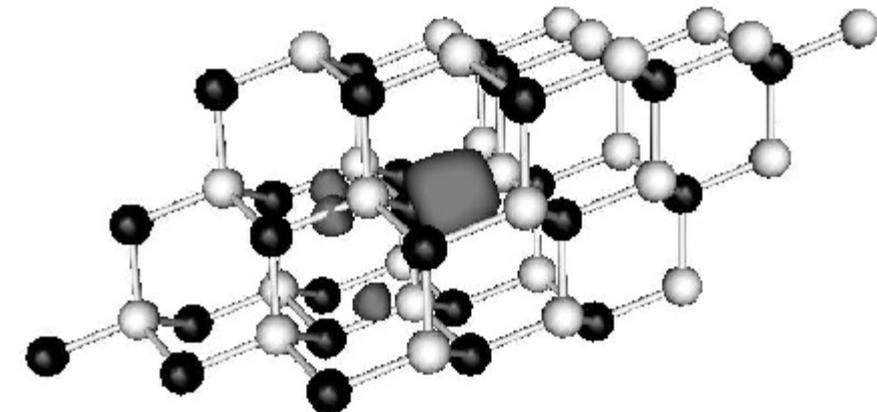


Fig. 5.7: a) One of the MLWF's in GaAs, obtained by using $8 \times 8 \times 8$ MP mesh that was shifted in order to include the Γ -point. b) The same MLWF obtained after using the same MP mesh but was shifted in order to have, for each \mathbf{k} , a $-\mathbf{k}$. c) The same MLWF obtained by using a uniform cubic mesh containing 500 \mathbf{k} -points.

5.6.2 Wavefunctions Initialization

As we have mentioned in chapter 4, the reference Bloch orbitals are prepared in order to make the starting Wannier functions somewhat localized. We do this by projecting Gaussian functions, centered in the middles of the tetrahedral bonds, on the Bloch orbitals. We have studied the effect of the choice of the standard deviation (spread) of these Gaussians on the required number of iterations to reach the convergence of the spread functional (Ω), and on Ω itself. This is done for GaAs, which has the largest lattice constant among the studied compounds, and for C, which has the smallest lattice constant among the studied compounds. This test is done using a regular $4 \times 4 \times 4$ MP mesh of \mathbf{k} -points.

The upper parts of figures 5.8 and 5.9 show the relation between the Gaussians standard deviation and the required number of iterations to reach convergence for C and GaAs, respectively. The lower parts show the relation between Gaussians spread and Ω . From Fig. 5.8 we see that the number of iterations increases slowly when we increase the Gaussians spread until it suddenly increases when we use Gaussians of spread more than, roughly speaking, about 0.9 times the C-C bond length. However, Ω remains unchanged. GaAs (Fig. 5.9) show a similar behavior. The number of iterations increases slowly when we increase the Gaussians spread until it increases rapidly when we use Gaussians of spread more than about 0.9 times the Ga-As bond length. Thus, we conclude that it is better to choose the Gaussians spread to be less than the bond length. For example choosing Gaussians spread to be half the bond length is expected to be a good choice. In our calculations, for all compounds considered, we have used Gaussians spread of 1 Å.

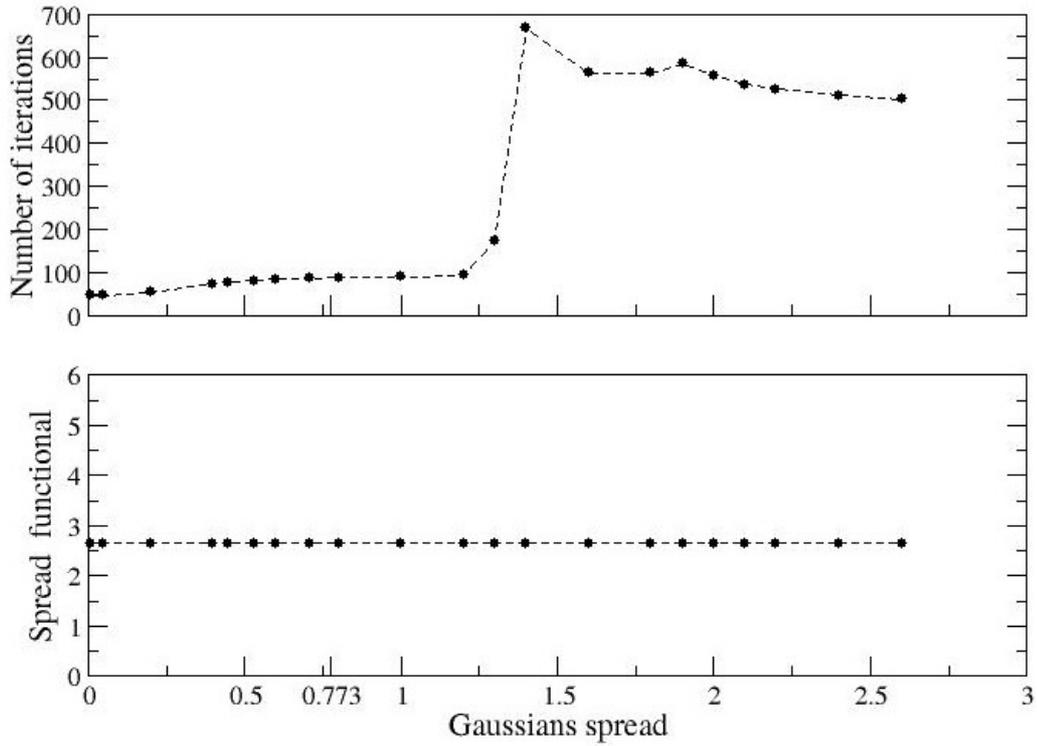


Fig. 5.8: Top: Gaussians spread (in \AA°) versus the required number of iterations for the spread functional to reach convergence, for C. Bottom: Gaussians spread (in \AA°) versus the spread functional (in \AA^2), also for C. Note that we have marked half the C-C bond length, which equals to $a\sqrt{3}/8$.

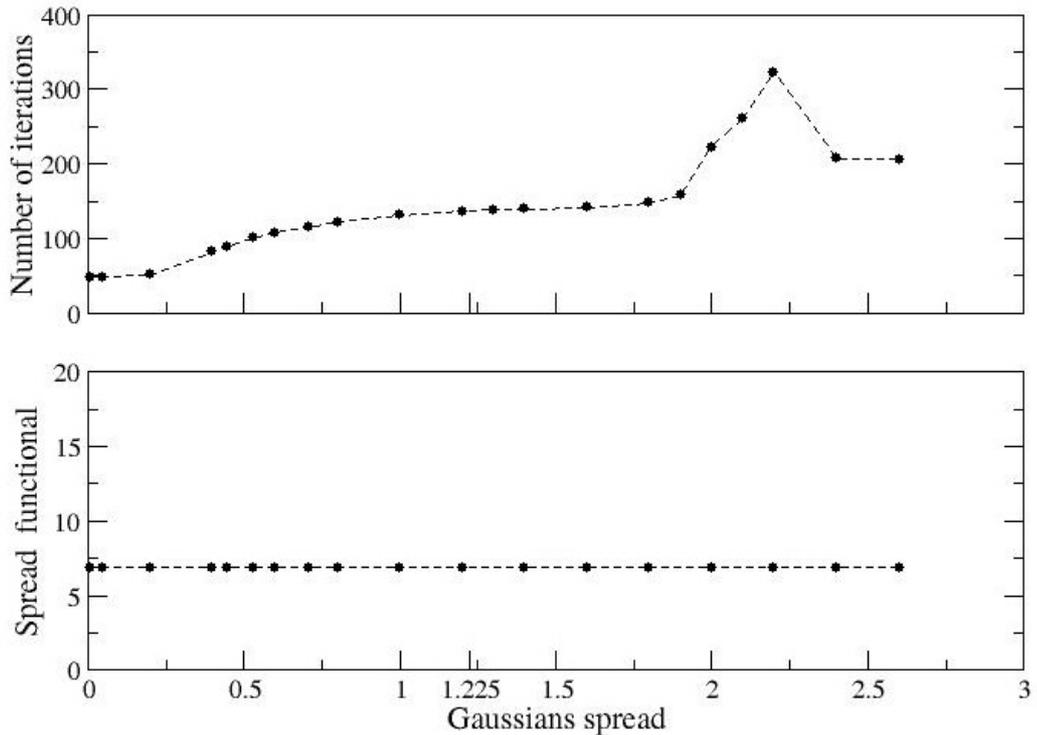


Fig. 5.9: Top: Gaussians spread (in \AA°) versus the required number of iterations for the spread functional to reach convergence, for GaAs. Bottom: Gaussians spread (in \AA°) versus the spread functional (in \AA^2), also for GaAs. Note that we have marked half the Ga-As bond length, which equals to $a\sqrt{3}/8$.

To show the importance of the wavefunctions initialization procedure, we show in Fig. 5.10a one of the calculated WF's of Si, using the $8 \times 8 \times 8$ MP \mathbf{k} -mesh, without any initialization and without using the minimization scheme of Ω introduced by Marzari and Vanderbilt. In this case, the unitary matrices $U^{(\mathbf{k})}$ are equal to the identity matrix. It is clear from the figure that the WF's are of widespread ($\Omega = 751 \text{ \AA}^2$) and are not confined in one unit cell. On the other hand, Fig 5.10b displays the same WF obtained after using the initialization procedure and again without using the minimization scheme of the spread functional. In this case, the WF is much more localized, with $\Omega = 68 \text{ \AA}^2$, and its shape is very similar to the maximally localized one which is shown in Fig. 5.10c. The Ω of this MLWF is equal to 8.232 \AA^2 . Thus, we can conclude that this initialization procedure provides a very good starting point.

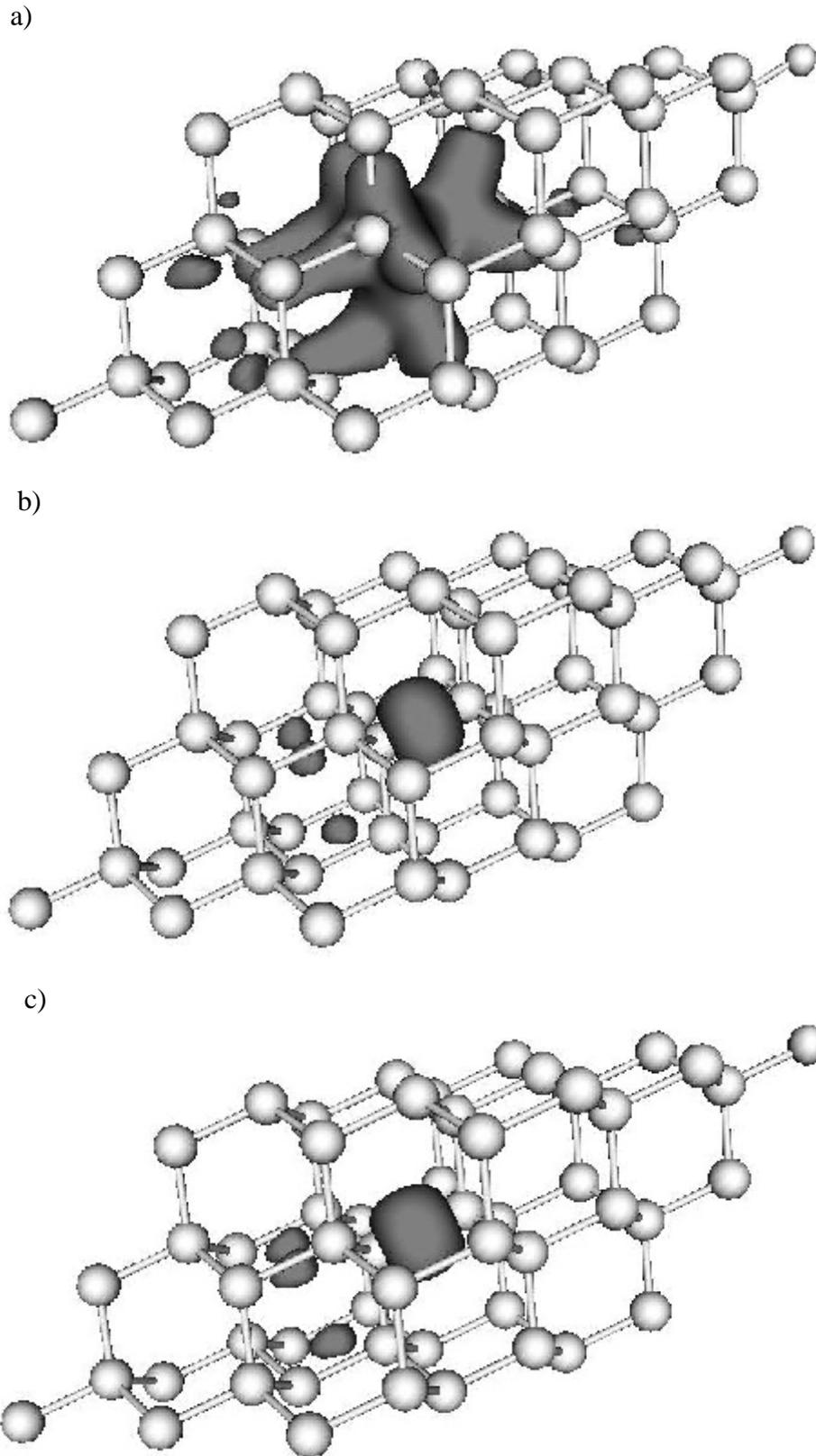


Fig. 5.10: a) One of the WF's of Si, obtained by using $8 \times 8 \times 8$ MP mesh, without any initialization or minimization of the spread functional. b) The same WF obtained after using the initialization procedure, without minimization of the spread functional. c) The same corresponding maximally localized one.

5.6.3 Plotting the MLWF's

The above displayed WF's and that shown in the next chapter are obtained as follows:

- 1- After finding the proper unitary matrix, a new set of the periodic part of the Bloch's wavefunction (u'_{nk}) are constructed, using the relation

$$u'_{nk}(\mathbf{r}) = \sum_m U_{mn}^{(k)} u_{mk}(\mathbf{r}). \quad (5.10)$$

Note that $u'_{nk}(\mathbf{r})$ obtained by using the inverse Fast Fourier Transform (FFT) are given at real space FFT grid, inside only one unit cell.

- 2- In order to have u'_{nk} in a large unit cell, we have carried out a $3 \times 3 \times 3$ repetition. This means that the size of the new (big) unit cell is 27 times larger than the original one.
- 3- The MLWF's are calculated in this big unit cell by multiplying by the proper phase factor and integrating over the first BZ

$$w_n(\mathbf{r} - \mathbf{R}) = \frac{1}{N} \sum_{\mathbf{k}} u'_{nk}(\mathbf{r}) e^{i(\mathbf{r}-\mathbf{R}) \cdot \mathbf{k}}. \quad (5.11)$$

The \mathbf{R} vector is chosen such that the WF's are centered at about the center of the big unit cell.

- 4- Both the surface and contour plots of the MLWF's are done, by using the pxviewer code [48]. As input to the program we supply the data as they are in the FFT grid points.

CHAPTER 6

RESULTS AND DISCUSSION

6.1 Introduction

In this chapter we will present and discuss the Maximally-Localized Wannier Functions (MLWF's), for four III-V semiconductors. These semiconductors are GaAs, AlAs, GaN and AlN. For comparison with other systems, we will also present and discuss the MLWF's of C, Si, Ge and SiC. In particular, we will focus on the spread of these functions and their centers (known as Wannier centers).

Although the accurate self-consistent calculations are nowadays a common practice, there are, however, cases where such calculations are extremely difficult (because of the complexity of the system) or the high accuracy they provide is not necessary (when studying broad trends is required). In such cases, concepts such as ionicity, electronegativity, covalent radii and others are of particular importance. It has been noted by Marzari and Vanderbilt that the Wannier centers can be used as a measure of the bond ionicity. However, this connection has not been exploited. In this chapter we will introduce a bond ionicity measure based on the Wannier center, and compare it with the empirical Phillips ionicity [49], and the self-consistent ionicity scale of Garcia and Cohen [50]. A more complete account of the new ionicity scale, applied to a large number of semiconductors, is currently under investigation.

The rest of this chapter is organized as follows. In section 6.2 we report and discuss the MLWF's of the eight semiconductors considered. Section 6.3 is devoted to the new ionicity scale. Finally, section 6.4 contains our main results and conclusions.

6.2 Maximally Localized Wannier Functions

6.2.1 For C, Si and Ge

As an example, we show in Fig. 6.1 the band structure of Si. In these systems, the upper three valence bands are not separated from the lowest energy one by a gap, as it is the case in the polar semiconductors. Therefore, the four valence bands of the above semiconductors have to be considered as a composite group of bands in the construction of MLWF's.

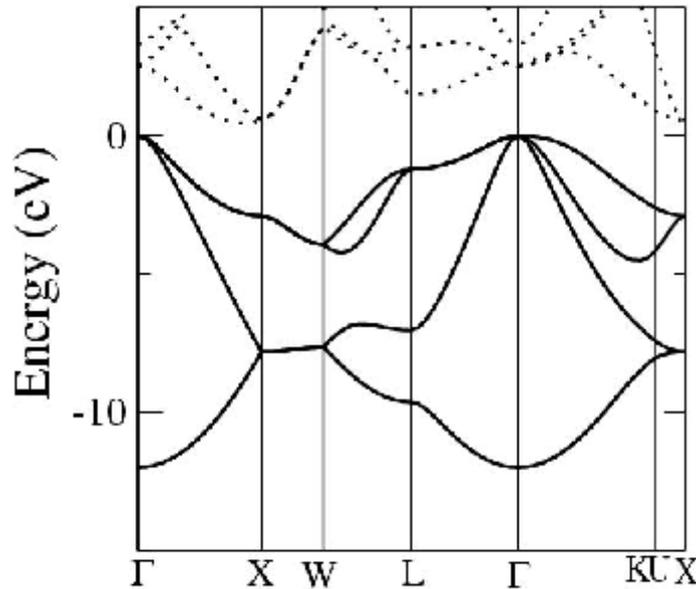


Fig. 6.1: The Band structure of Si. Solid curves: valence bands. Dotted curves: conduction bands.

The MLWF's are obtained as described in chapters 4 and 5, with $J = 4$. Because of the inversion symmetry, we expect the four MLWF's per primitive unit cell are centered exactly at the centers of the tetrahedral bonds. As noted by Marzari and Vanderbilt, the inversion symmetry allows for real coefficients of the PW's expansion (by taking the origin to be at the bond center) and so also the unitary matrices $U^{(k)}$. This yields zero Ω_D . Since Ω_I is gauge invariant, only Ω_{OD} needs to be minimized.

In Tables 6.1, 6.2 and 6.3 we show the variation of the spread functional, Ω , and its various parts (Ω_I , Ω_D and Ω_{OD}) for C, Si and Ge, respectively, with respect to the \mathbf{k} -space grids. These grids vary from $2 \times 2 \times 2$ to $12 \times 12 \times 12$ MP meshes. Also listed, in the same tables, is the parameter b which gives the relative distance (with respect to the bond length) of the Wannier center along the bond from one of its end atoms.

Table 6.1: Minimized spread functional and its various parts (invariant, off-diagonal, and diagonal) in C for different \mathbf{k} -meshes (in units of \AA^2), together with the relative position of the centers along the C-C bonds.

k set	Ω	Ω_I	Ω_{OD}	Ω_D	b
$2 \times 2 \times 2$	1.844	1.620	0.224	0	0.5
$4 \times 4 \times 4$	2.642	2.306	0.336	0	0.5
$8 \times 8 \times 8$	3.045	2.700	0.345	0	0.5
$12 \times 12 \times 12$	3.133	2.794	0.340	0	0.5

Table 6.2: Minimized spread functional and its various parts (invariant, off-diagonal, and diagonal) in Si, for different \mathbf{k} -meshes (in units of \AA^2), together with the relative position of the centers along the Si-Si bonds.

k set	Ω	Ω_I	Ω_{OD}	Ω_D	b
$2 \times 2 \times 2$	4.039	3.660	0.380	0	0.5
$4 \times 4 \times 4$	6.433	5.866	0.566	0	0.5
$8 \times 8 \times 8$	8.232	7.716	0.516	0	0.5
$12 \times 12 \times 12$	8.730	8.280	0.450	0	0.5

Table 6.3: Minimized spread functional and its various parts (invariant, off-diagonal, and diagonal) in Ge, for different \mathbf{k} -meshes (in units of \AA^2), together with the relative position of the centers along the Ge-Ge bonds.

k set	Ω	Ω_I	Ω_{OD}	Ω_D	b
$2 \times 2 \times 2$	5.320	4.507	0.813	0	0.5
$4 \times 4 \times 4$	7.713	7.001	0.712	0	0.5
$8 \times 8 \times 8$	10.116	9.485	0.631	0	0.5
$12 \times 12 \times 12$	11.041	10.462	0.578	0	0.5

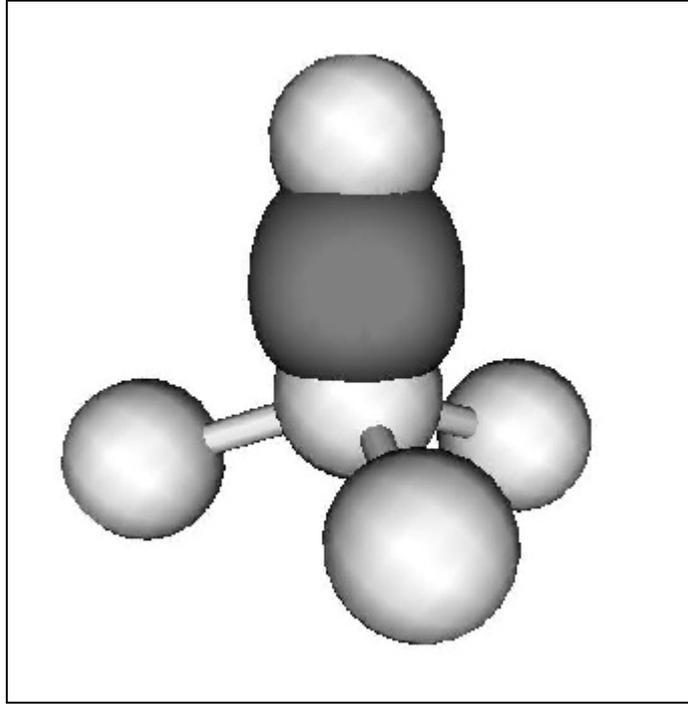
In Figs. 6.2, 6.3 and 6.4 we show one of the MLWF's of C, Si and Ge, respectively, obtained by using an $8 \times 8 \times 8$ MP mesh. The other three MLWF's, for each system, lie on the other three tetrahedral bonds. All of the obtained MLWF's are real, and their shape resembles that of the sp^3 hybrids combined to form s -bond orbitals: the MLWF's are located mainly in the bond region. The remarkable features to note from the above results are as follows:

- (i) As expected, Ω_D is identically zero for the above three systems, which crystallize in the diamond structure, and the Wannier centers lie exactly at the centers of the bonds. This reflects the accuracy and reliability of our calculations.
- (ii) The MLWF's are mainly confined in the primitive unit cell. This feature is of particular importance for the practical applications of the MLWF's, since it reduces to a large extent the number of matrix elements required. In fact, it has already been noted by Schnell, Czycholl and Albers [16] that matrix elements for only few neighbor shells are sufficient, when they used the MLWF's as basis to construct the Coulomb matrix elements. The behavior is also expected when the Fock operator matrix elements, required in the EXX calculations [7,8], are calculated by using the MLWF's basis set.
- (iii) The spread functional is a slowly converging function with respect to the density of the \mathbf{k} -mesh used. However, its slow convergence is mainly due to Ω_I . The Ω_{OD} is found to show, relatively, a very small variation. Similar behavior has been also observed by Marzari and Vanderbilt [12], see Table 6.4.
- (iv) Our results for Ω of Si, its various contributions, and their variation with respect to the density of the \mathbf{k} -space mesh agree nicely with the results of Marzari and Vanderbilt. The very small differences can be attributed to the differences in the used computational ingredients in the two calculations.
- (v) In spite of the fact that the Wannier centers of the MLWF's for the above three systems considered are located exactly at the bond centers and the very nice agreement between our results and that of Marzari and Vanderbilt, for Si, the displayed MLWF's (Figs. 6.2 to 6.4) don't show the expected inversion symmetry. In these plots, which have been done using the pxviewer [48], it appears that the center of such MLWF's is deviated slightly from the bond center toward one of the atoms. This is mainly because we have chosen the origin to be on the atom, not on the inversion center in the middle of the bond.

Table 6.4: Marzari and Vanderbilt results [12] for the minimized spread functional and its various parts in Si for different \mathbf{k} -meshes (in units of \AA^2), together with the relative position of the centers along the Si-Si bonds.

k set	Ω	Ω_I	Ω_{OD}	Ω_D	b
$2 \times 2 \times 2$	4.108	3.707	0.401	0	0.5
$4 \times 4 \times 4$	6.447	5.870	0.577	0	0.5
$8 \times 8 \times 8$	8.192	7.671	0.520	0	0.5

(a)



(b)

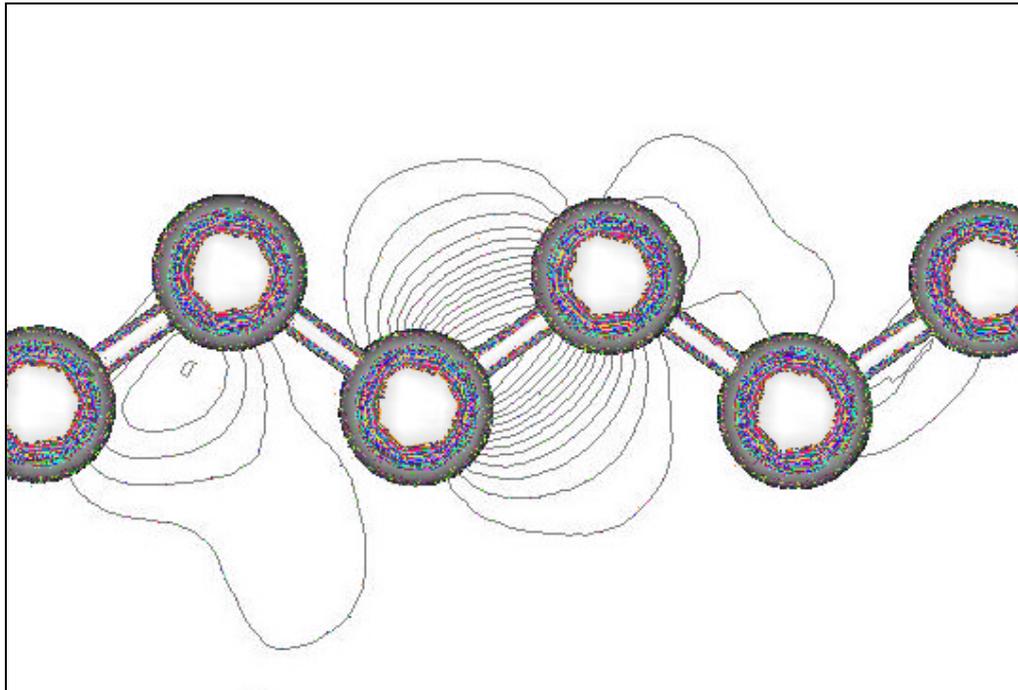
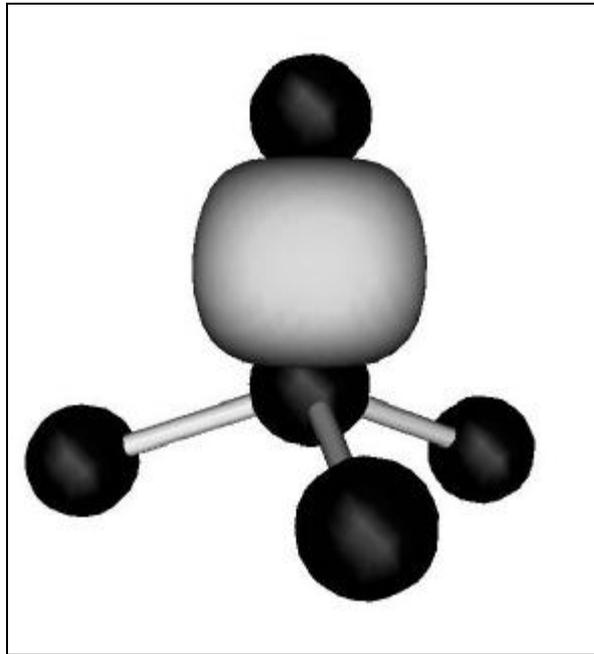


Fig. 6.2: One of the maximally-localized Wannier functions in C, for the $8 \times 8 \times 8$ \mathbf{k} -point sampling. The other Wannier functions lie on the other three tetrahedral bonds. (a) Isosurface plot (in gray), showing the tetrahedral bonds. (b) Contour plot in the (110) plane of the bond chains.

(a)



(b)

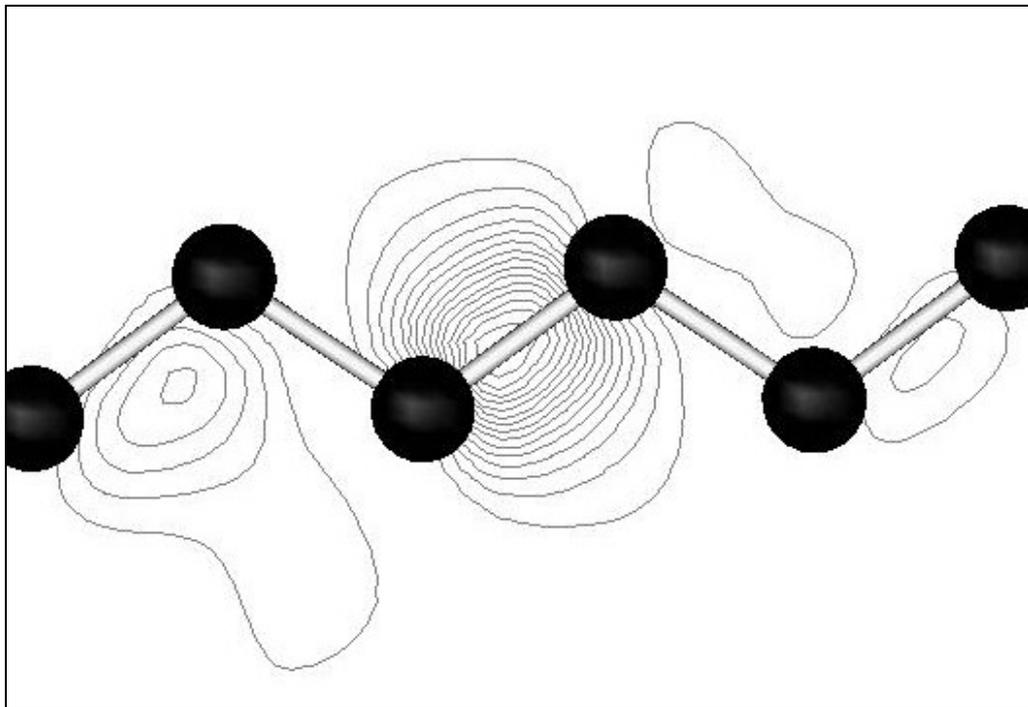
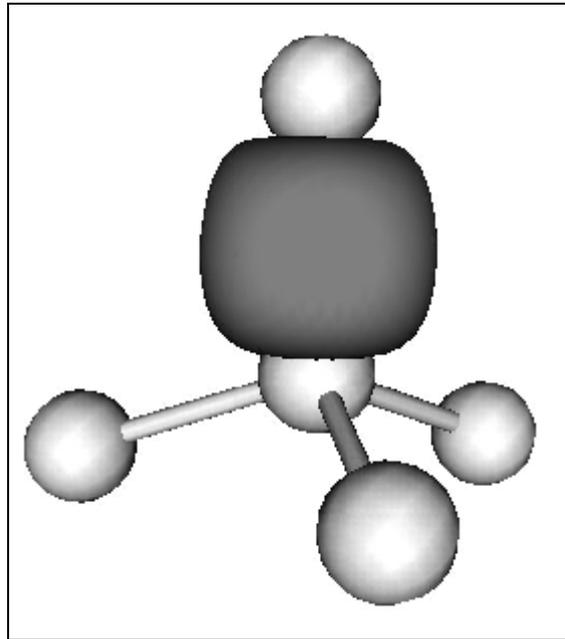


Fig. 6.3: As in Fig. 6.2, but for Si.

(a)



(b)

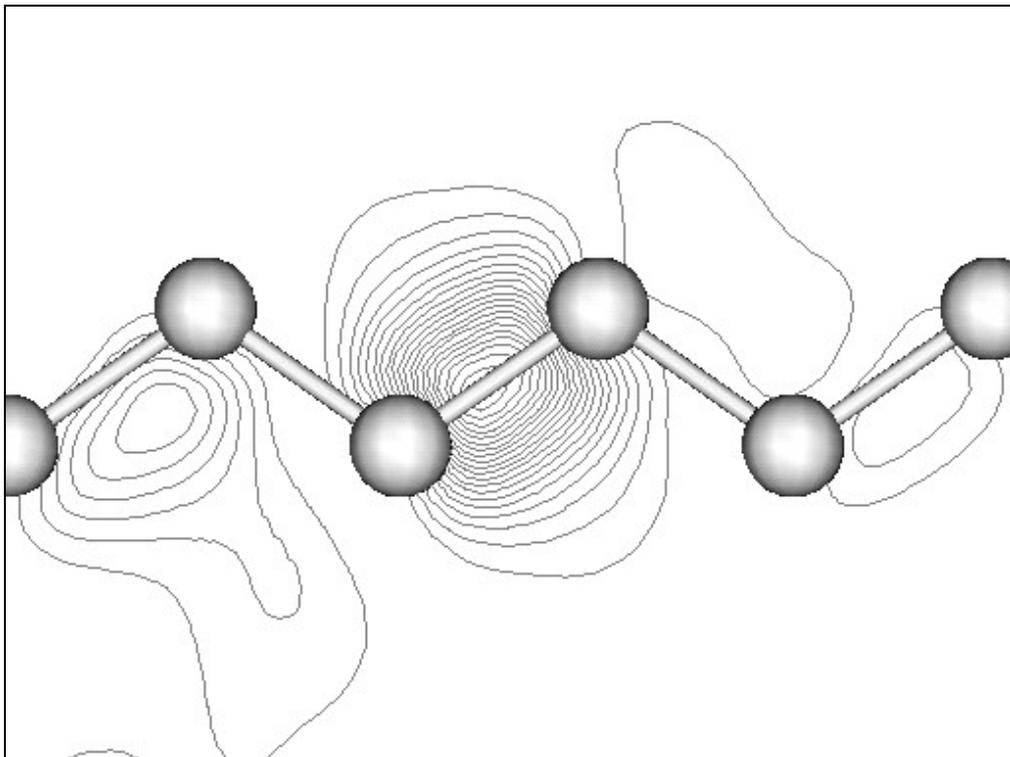


Fig. 6.4: As in Fig. 6.2, but for Ge.

6.2.2 For SiC

It is well known that the C pseudopotential is much stronger than that of the Si, which makes the Si-C bond partially ionic. In the case of the ZB structure of SiC, this leads to a splitting of the four valence bands (which were composite in Si and C) into three (upper) and one (lower) bands separated by a gap, as shown in Fig. 6.5.

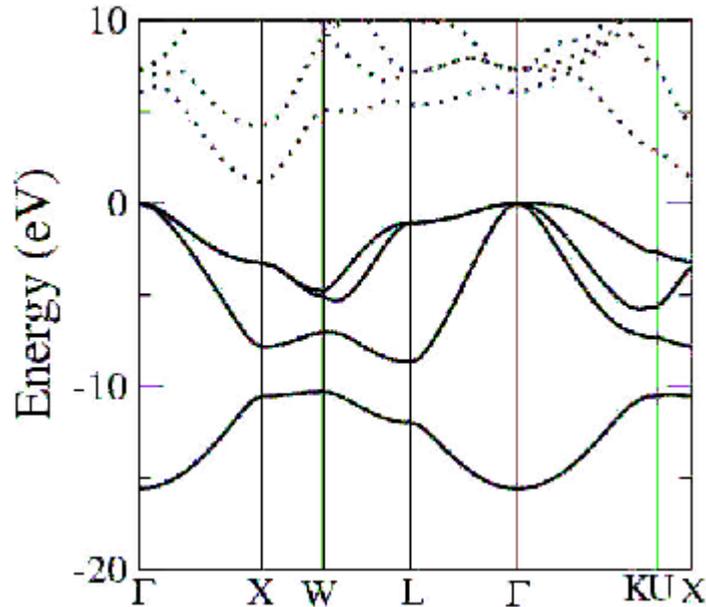


Fig. 6.5: The Band structure of SiC. Solid curves: valence bands. Dotted curves: conduction bands.

Thus, there are two options to construct the MLWF's: (i) Treating the lowest (isolated) valence band as a single band and construct its MLWF, and the upper three bands as a composite and construct their three MLWF's. (ii) To continue treating the four valence bands as a composite and construct their four MLWF's as done previously for group-IV semiconductors. In this work we have considered only the second option.

In Table 6.5 we show the variation of Ω for SiC, and its various components (Ω_1 , Ω_D and Ω_{OD}) with respect to the used density of \mathbf{k} -space mesh. This table also shows the similar variation of b . In Fig. 6.6 we show one of the MLWF's of SiC. The important features to note from these results are as follows:

- (i) Ω_D is not zero any more and b is larger than 0.5, because of the lack of inversion symmetry. The Wannier centers are shifted toward the C atoms, as expected. The large deviation from the center of the bond reflects a quite large

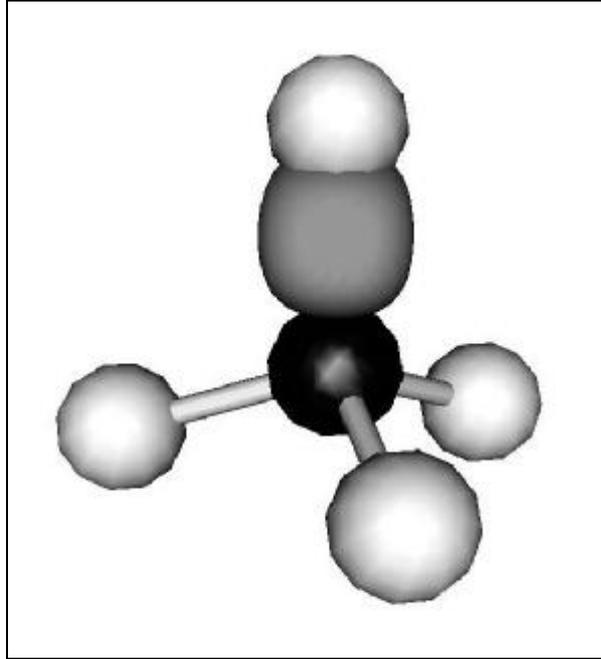
ionicity value for Si-C bond, despite the fact that both atoms have the same number of valence electrons.

- (ii) The convergence of Ω and its components with respect to density of \mathbf{k} -space mesh is very similar to that of group-IV semiconductors, and the slow convergence is mainly due to Ω_I .
- (iii) b shows a rapid convergence with respect to the density of the used \mathbf{k} -space mesh.
- (iv) As it was the case for group-IV semiconductors, the MLWF's are confined mainly inside the primitive unit cell.
- (v) The calculated value of Ω , obtained by using a certain MP mesh, is closer to that of bulk C more than that of Si. This can be understood as a consequence of the stronger C pseudopotential, see Sec. 5.5.

Table 6.5: Minimized spread functional and its various parts (invariant, off-diagonal, and diagonal) in SiC, for different k-meshes (in units of \AA^{-2}), together with the relative position of the centers from the Si atom along the Si-C bonds.

k set	Ω	Ω_I	Ω_{OD}	Ω_D	b
2×2×2	2.712	2.285	0.422	0.005	0.644
4×4×4	3.964	3.370	0.591	0.004	0.651
8×8×8	4.651	4.044	0.603	0.004	0.654
12×12×12	4.802	4.207	0.591	0.005	0.654

(a)



(b)

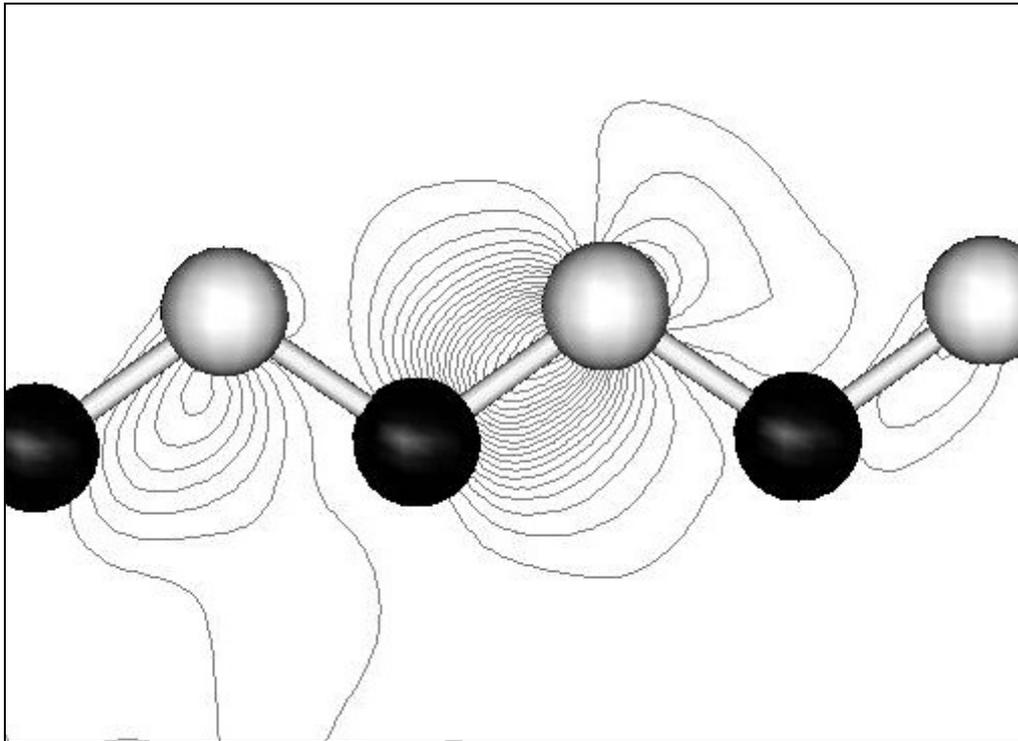


Fig. 6.6: One of the maximally-localized Wannier functions in SiC, for the $8 \times 8 \times 8$ k-point sampling. Si atoms are in black and C atoms are in white. The other Wannier functions lie on the other three tetrahedral bonds. (a) Isosurface plot (in gray), showing the tetrahedral bonds. (b) Contour plot in the (110) plane of the bond chains.

6.2.3 For GaAs, AlAs, GaN and AlN

As an example, we show in Fig. 6.7 the calculated band structure of GaAs. As in the case of SiC, the valence bands are separated into three composite bands (the upper three ones) and one isolated single band (the lowest one). Therefore, there are two options to construct the MLWF's, see Sec. 6.2.2. Here, we also opted for taking the four bands as a composite, $J = 4$. However, the other option has been considered, and we found that the MLWF of the lowest valence band is centered, in the case of GaAs, at the As atom. This is consistent with the well-known fact that this band has mainly an anion s-character. No attempt for constructing the MLWF's of the upper three bands has been made, in this work.

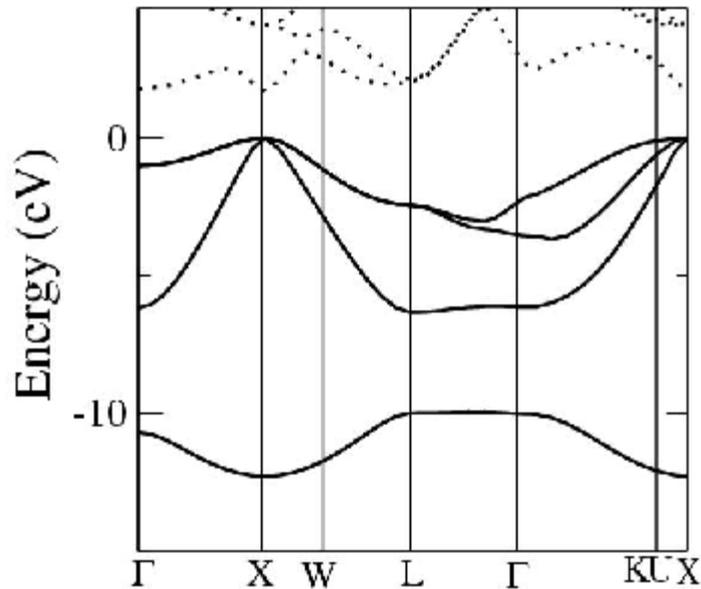


Fig. 6.7: The Band structure of GaAs. Solid curves: valence bands. Dotted curves: conduction bands.

In Tables 6.6 we show the variation of the spread functional Ω and its components (Ω_1 , Ω_D and Ω_{OD}) of GaAs, AlAs, GaN and AlN, respectively, with respect to the used MP mesh. In the same tables, we show the similar variation of b of the corresponding system. In Figures 6.8, 6.9, 6.10 and 6.11 we show one of the MLWF's of the above four semiconductors, respectively. The remarkable features to note are as follows.

- (i) Because of the lack of inversion symmetry, and as in the case of SiC, the obtained values of Ω_D are not zero, but they are very small. The quite strange

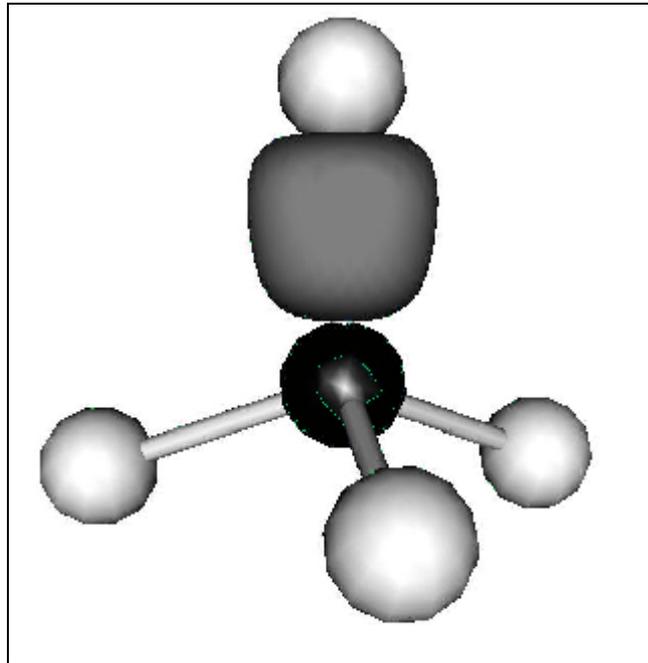
behavior is that the value of Ω_D decreases by going to systems with larger bond ionicity [49,50], and reaches a zero value for AlN.

- (ii) Because of the polar nature of the bonds of the above systems, the Wannier centers are shifted more towards the anion. As expected, the deviation from the bond center is found to be proportional to the bond ionicity, see the next section.
- (iii) Ω has a slow convergence with respect to the used density of the \mathbf{k} -space mesh, which originates mainly from the slow convergence of Ω_1 . These results and those obtained previously for C, Si, Ge and SiC show that this is a common behavior.
- (iv) b shows a rapid convergence with respect to the used density of the \mathbf{k} -space mesh, which also seems to be a common behavior.
- (v) The MLWF's are also confined in the primitive unit cell. The obtained value of Ω is found to depend strongly on the volume per atom: Ω of GaN and AlN are much smaller than that of GaAs and AlAs. Moreover, Ω is found to depend on the atomic volume: Ω is smaller in AlAs than in GaAs, and similarly in AlN and GaN. This is also reflected in the calculated values of b , which indicates that Ga is more electronegative than Al. This conclusion is consistent with the self-consistent ionicity scale of Garcia and Cohen [50], but not with the Phillips ionicity measure – for more details see the next section.

Table 6.6: Minimized spread functional and its various parts (invariant, off-diagonal, and diagonal) in for different k-meshes (in units of A°^2), together with the relative position of the centers along the bonds, for a) GaAs, b) AlAs, c) GaN and d) AlN.

a)					
k set	Ω	Ω_I	Ω_{OD}	Ω_D	b
2×2×2	4.424	3.908	0.507	0.009	0.611
4×4×4	6.898	6.282	0.610	0.006	0.616
8×8×8	8.871	8.287	0.578	0.007	0.618
12×12×12	9.499	8.953	0.539	0.007	0.618
b)					
k set	Ω	Ω_I	Ω_{OD}	Ω_D	b
2×2×2	4.409	3.850	0.554	0.004	0.635
4×4×4	6.634	5.931	0.701	0.002	0.649
8×8×8	8.090	7.436	0.652	0.003	0.654
12×12×12	8.456	7.847	0.606	0.003	0.655
c)					
k set	Ω	Ω_I	Ω_{OD}	Ω_D	b
2×2×2	2.547	2.142	0.400	0.004	0.723
4×4×4	3.611	3.112	0.497	0.002	0.726
8×8×8	4.109	3.612	0.494	0.003	0.728
12×12×12	4.216	3.726	0.487	0.003	0.729
d)					
k set	Ω	Ω_I	Ω_{OD}	Ω_D	b
2×2×2	2.414	1.972	0.440	0.001	0.740
4×4×4	3.343	2.755	0.587	0	0.752
8×8×8	3.711	3.103	0.608	0	0.757
12×12×12	3.789	3.178	0.611	0	0.760

(a)



(b)

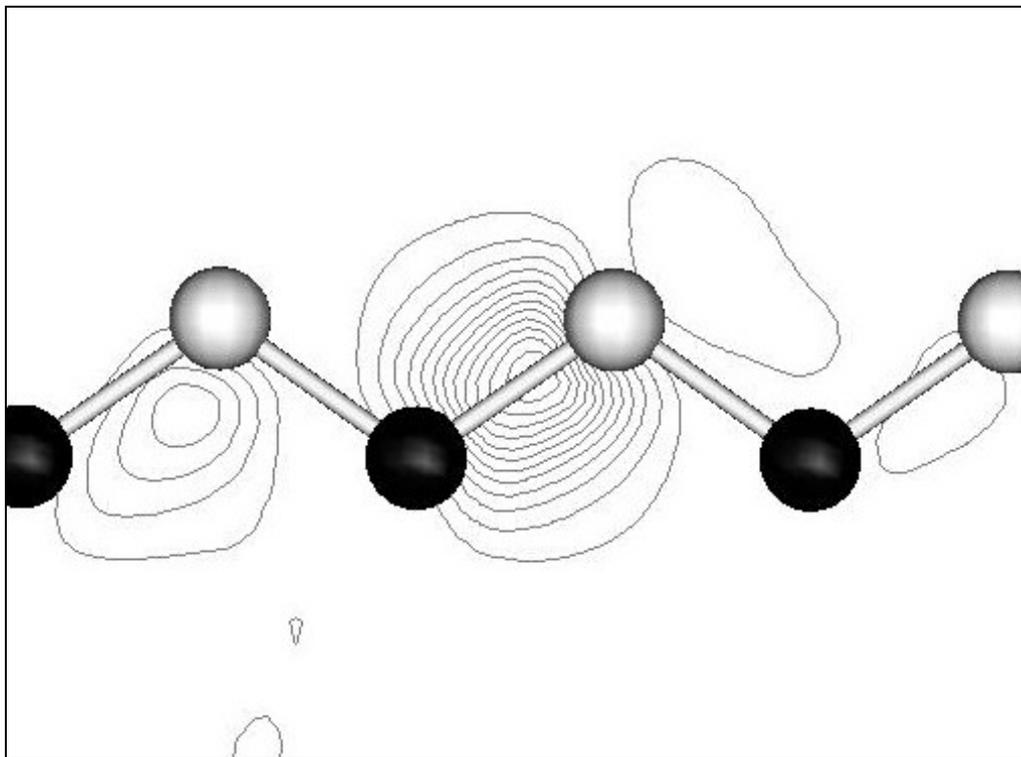
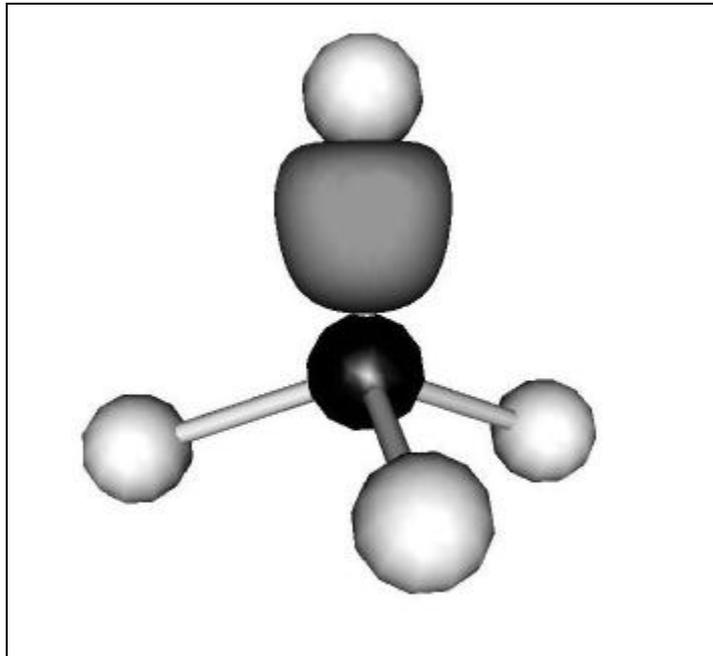


Fig. 6.8: One of the maximally-localized Wannier functions in GaAs, for the $8 \times 8 \times 8$ \mathbf{k} -point sampling. Ga atoms are in black and As atoms are in white. The other Wannier functions lie on the other three tetrahedral bonds. (a) Isosurface plot (in gray), showing the tetrahedral bonds. (b) Contour plot in the (110) plane of the bond chains.

(a)



(b)

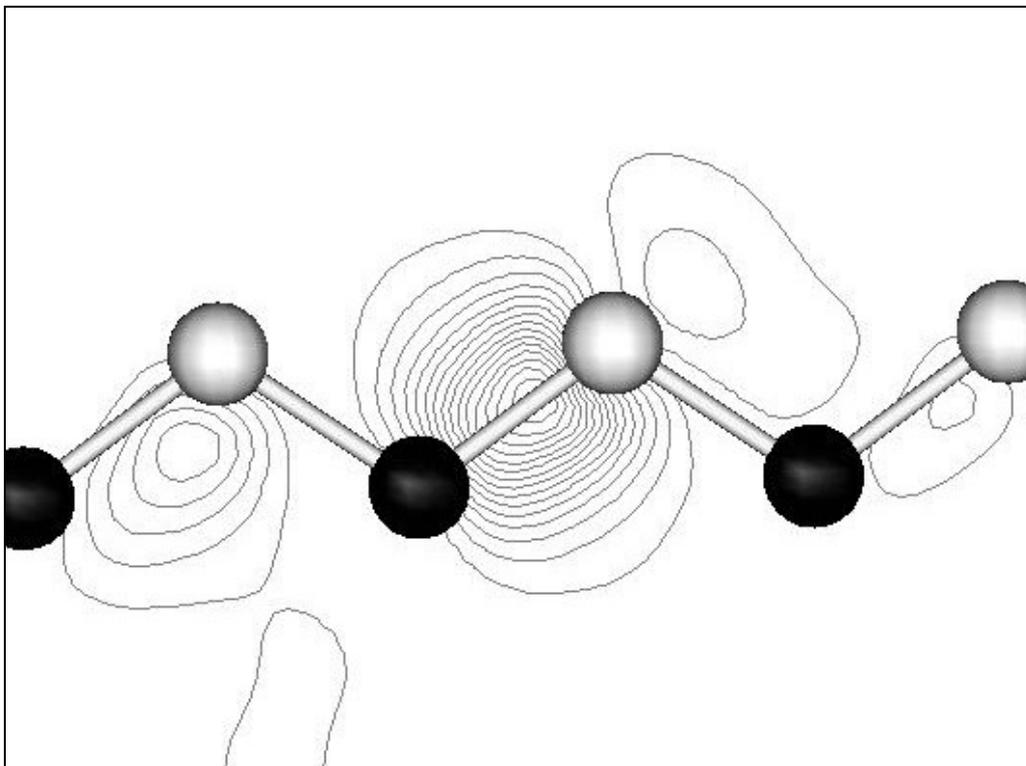
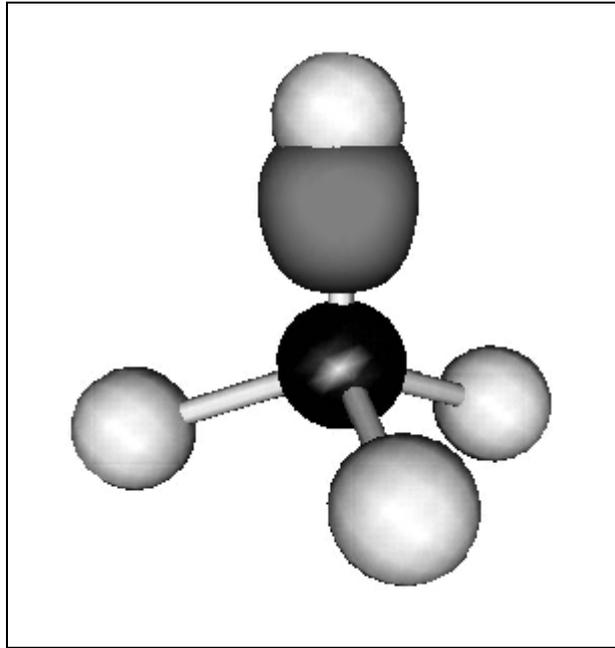


Fig. 6.9: One of the maximally-localized Wannier functions in AlAs, for the $8 \times 8 \times 8$ \mathbf{k} -point sampling. Al atoms are in black and As atoms are in white. The other Wannier functions lie on the other three tetrahedral bonds. (a) Isosurface plot (in gray), showing the tetrahedral bonds. (b) Contour plot in the (110) plane of the bond chains.

(a)



(b)

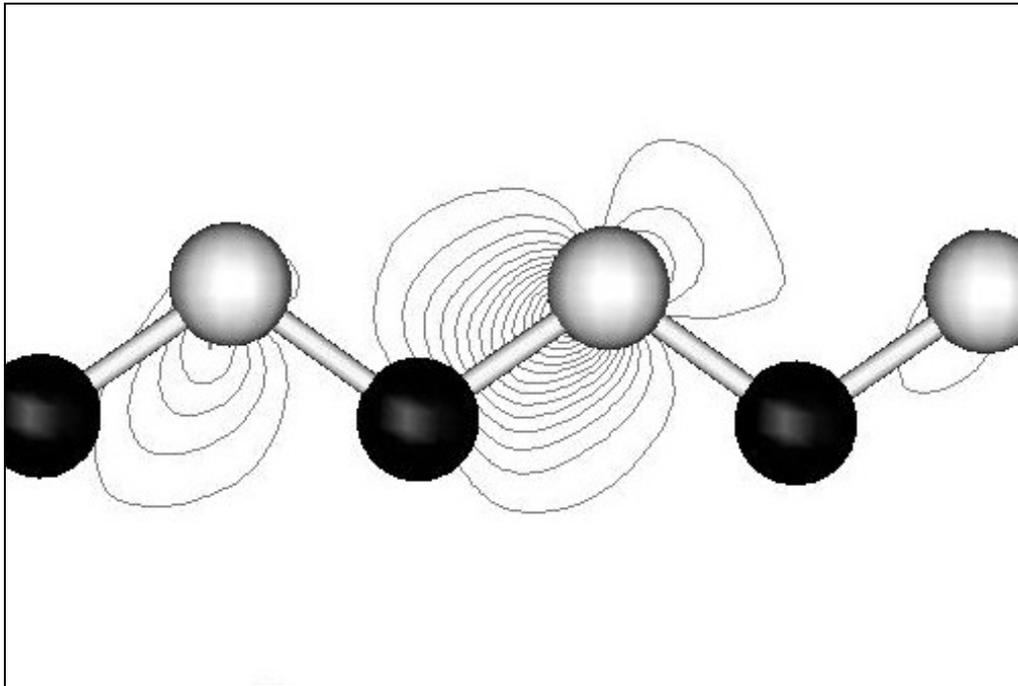
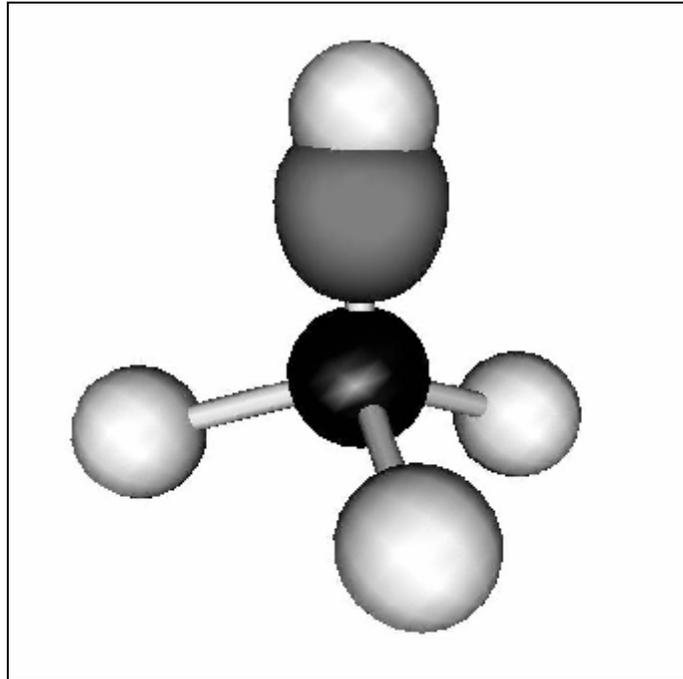


Fig. 6.10: One of the maximally-localized Wannier functions in GaN, for the $8 \times 8 \times 8$ \mathbf{k} -point sampling. Ga atoms are in black and N atoms are in white. The other Wannier functions lie on the other three tetrahedral bonds. (a) Isosurface plot (in gray), showing the tetrahedral bonds. (b) Contour plot in the (110) plane of the bond chains.

(a)



(b)

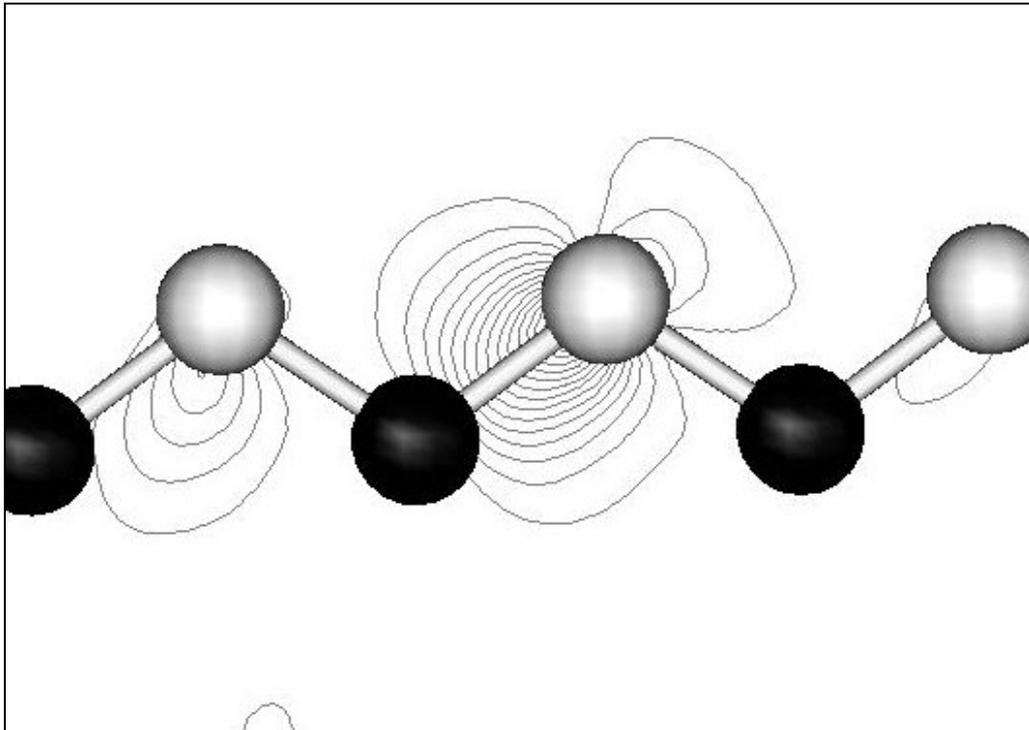


Fig. 6.11: One of the maximally-localized Wannier functions in AlN, for the $8 \times 8 \times 8$ \mathbf{k} -point sampling. Al atoms are in black and N atoms are in white. The other Wannier functions lie on the other three tetrahedral bonds. (a) Isosurface plot (in gray), showing the tetrahedral bonds. (b) Contour plot in the (110) plane of the bond chains.

6.3 Bond Ionicity

The bond ionicity is a very important concept, which helps in getting a better and better understanding of the behavior of matter. Several empirical ionicity scales have been suggested, and the most widely accepted one is that of Phillips [49]. A self-consistent ionicity scale has been introduced by Garcia and Cohen [50]. Important differences between these two scales have been found, especially for the systems involving first row elements. The bond ionicity in these systems is significantly underestimated by the Phillips scale. Another difference is in the bond ionicity of the common anion semiconductors. For example, according to the Phillips ionicity scale Ga-As bond has slightly larger ionicity than the AlAs one, while according to the self-consistent scale this behavior is reversed. Thus, a new and accurate ionicity scale is of special importance. In the following we provide such a scale, based on the deviation of the Wannier center from the bond center.

The Phillips ionicity scale [49] of the $A^N B^{8-N}$ crystals is based on two band parameters, C and E_h . The E_h is homopolar band gap results from the symmetric part of the potential, $(V_A + V_B)/2$, whereas C is the ionic or charge transfer gap results from the asymmetric part of the potential, $(V_A - V_B)/2$. From these two parameters the average valence-conduction gap, E_g , is defined according to the relation

$$E_g^2 = E_h^2 + C^2. \quad (6.1)$$

The Phillips ionicity scale is defined as

$$f_i = \frac{C^2}{E_h^2 + C^2} = \frac{C^2}{E_g^2}. \quad (6.2)$$

Now, E_h of a compound crystal of lattice parameter, a , is obtained from that of Si according to

$$E_h = E_h(\text{Si}) \left(\frac{a(\text{Si})}{a} \right)^{2.5}. \quad (6.3)$$

Thus, the estimation of E_h and the knowledge of E_g determine C and, hence, f_i .

The Garcai and Cohen [50] first-principles ionicity scale is defined as follows. The self-consistent charge density is separated into symmetric and asymmetric parts according to

$$n_s(\mathbf{r}) = \frac{1}{2}(n(\mathbf{r}) + n(-\mathbf{r})); n_A(\mathbf{r}) = \frac{1}{2}(n(\mathbf{r}) - n(-\mathbf{r})). \quad (6.4)$$

Then, the parameters S_S and S_A is defined as the cell average of the above parts, respectively, or

$$S_S = \frac{1}{V_c} \int n_s^2(\mathbf{r}) d\mathbf{r}; S_A = \frac{1}{V_c} \int n_A^2(\mathbf{r}) d\mathbf{r}. \quad (6.5)$$

Finally, the asymmetry coefficient, g , which is introduced as a measure of the ionicity is defined as

$$g = \sqrt{\frac{S_A}{S_S}}. \quad (6.6)$$

In this case g is obtained completely from the self-consistent $n(\mathbf{r})$.

To introduce a new ionicity scale based on the deviation of the Wannier center from the bond center, two guidelines have to be observed: (i) The ionicity should be between 0 and 1. (ii) The obtained ionicities should be somehow consistent with those of the above described scales. We found that a very reasonable choice is

$$w_i = (2b - 1)^{0.75}, \quad 0.5 \leq b \leq 1 \quad (6.7)$$

The obtained values of w_i for the studied systems, compared with those of f_i and g , are listed in Table 6.7. Furthermore, plots between these different ionicity scales are provided in Figs. 6.12, 6.13 and 6.14. It is clear, from both the table and the figures, that the new ionicity scale fits nicely with the above two scales. In particular, for GaN and AlN our results lie in between those of f_i and g . Thus we believe that the surprising large difference is due to underestimation of the ionicity by Phillips and overestimation by the self-consistent measures. The ionicity of the Si-C bond is also underestimated by the former approach, and our result is close to that of the latter. Finally, we note that the ionicity of the GaAs and AlAs as obtained by Phillips are reversed in our present results and those of Garcia and Cohen [50]. Similar behavior occurs in the case of GaN and AlN. This shows that the electronegativity of the Ga atom is larger than that of Al.

Table 6.7: Phillips ionicity (f_i), charge asymmetry coefficient (g) and our new ionicity scale w_i , for the studied compounds.

Compound	f_i	g	w_i
Si	0.000	0.000	0.000
Ge	0.000	0.000	0.000
GaAs	0.310	0.316	0.339
AlAs	0.274	0.375	0.415
SiC	0.177	0.475	0.413
GaN	0.500	0.780	0.557
AlN	0.449	0.794	0.612

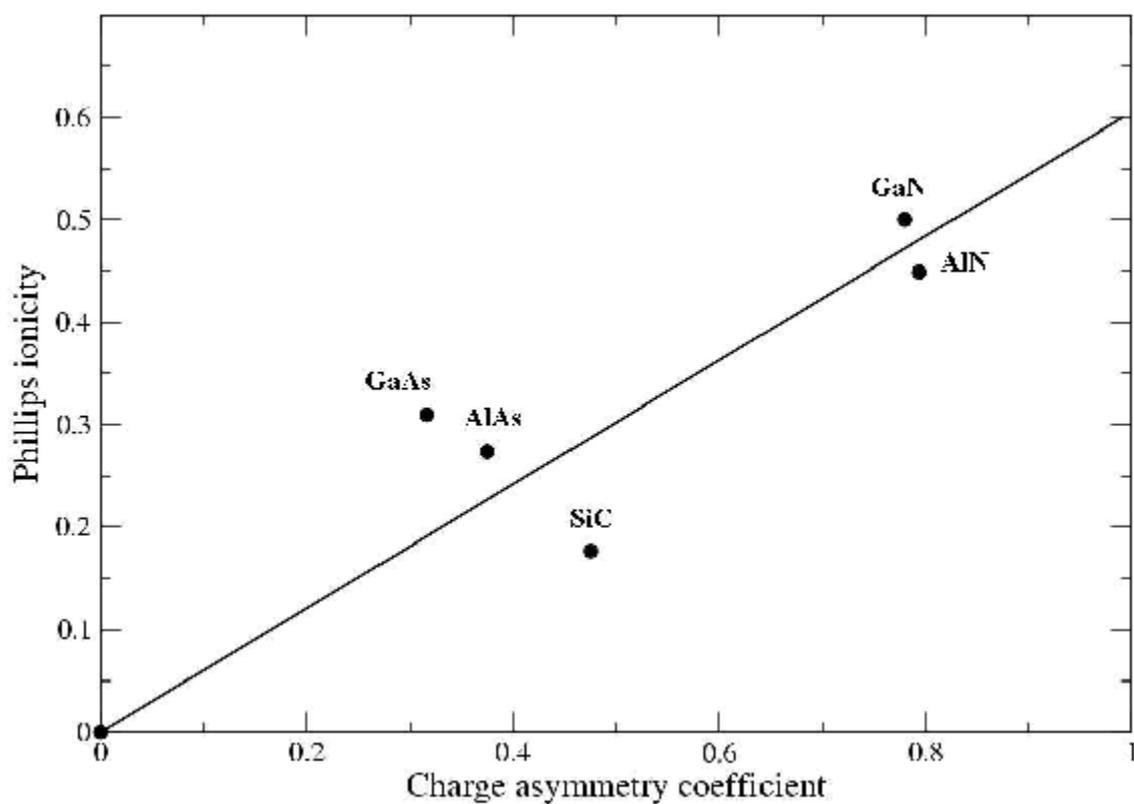


Fig. 6.12: Comparison of the Phillips ionicity (f_i) to the charge asymmetry coefficient (g) for the studied compounds.

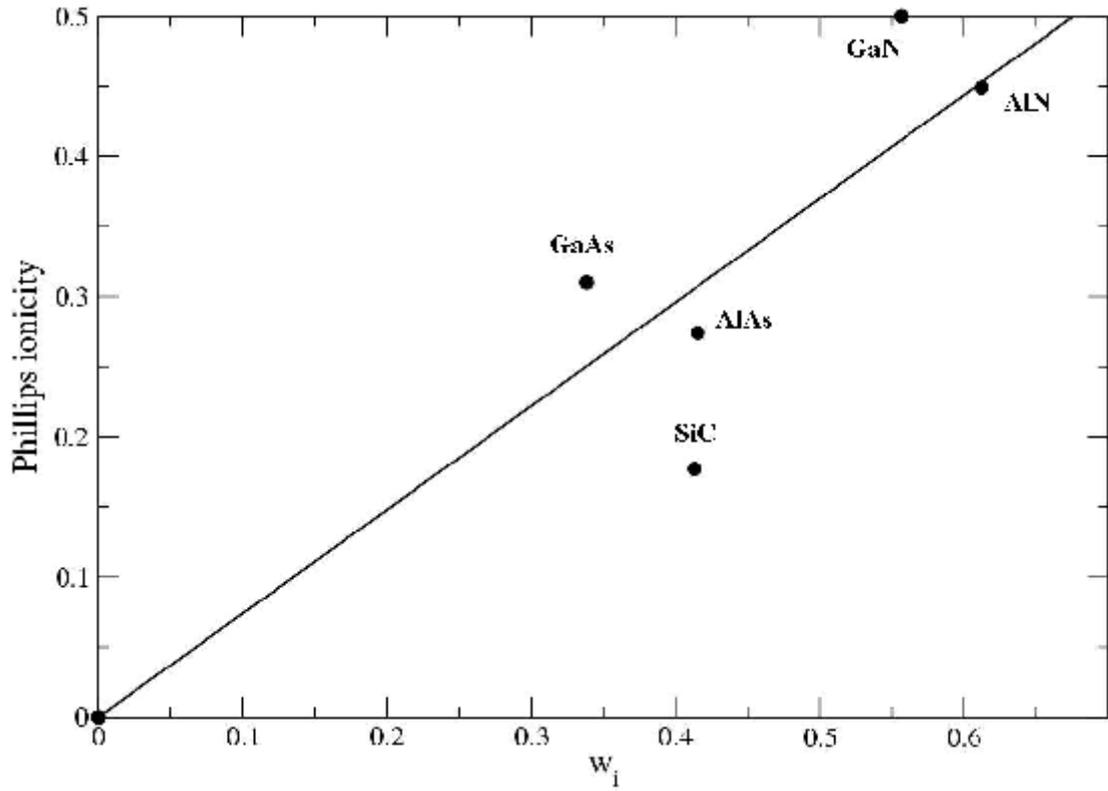


Fig. 6.13: Comparison of the w_i to the Phillips ionicity for the studied compounds.

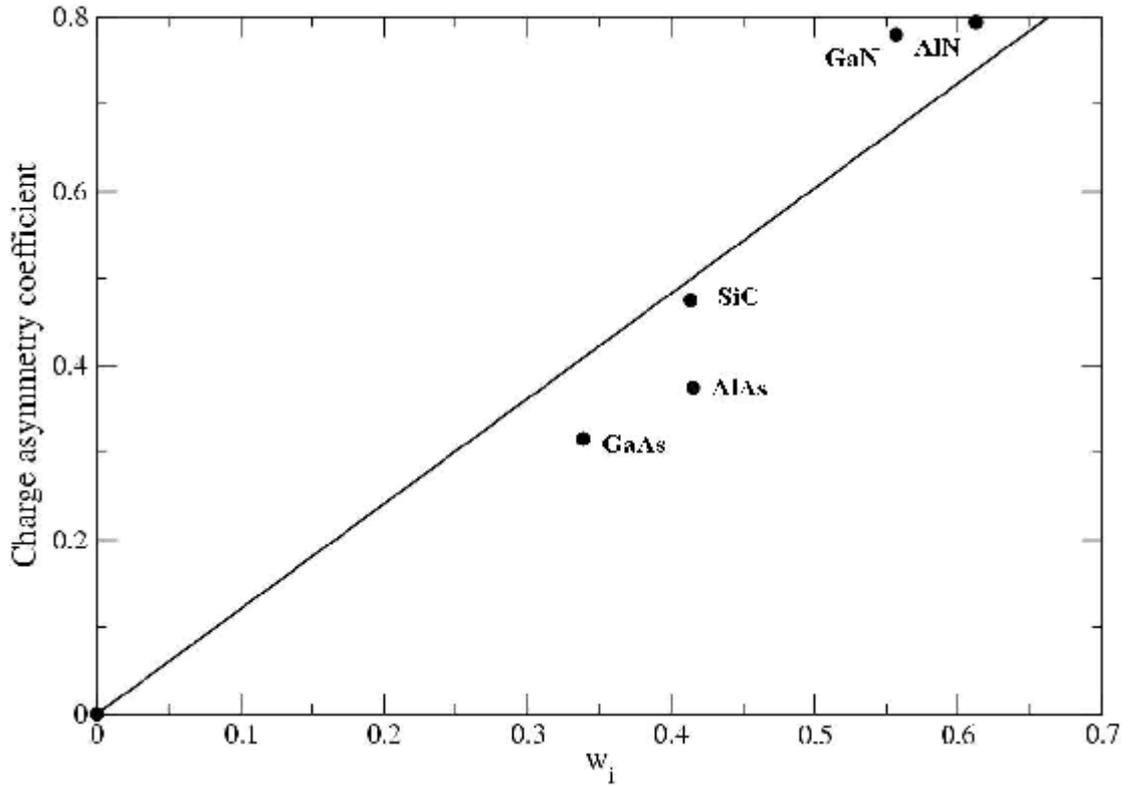


Fig. 6.14: Comparison of the w_i to the charge asymmetry coefficient for the studied compounds.

6.4 Conclusions

We have successfully developed FORTRAN77 and C++ codes to construct Maximally-Localized Wannier Functions (MLWF's) according to the scheme of Marzari and Vanderbilt. The latter code is used to calculate the MLWF's of eight semiconductors, namely, C, Si, Ge, SiC, GaAs, AlAs, GaN and AlN. Based on the obtained Wannier centers, a new bond ionicity scale has been introduced. In the following we draw our main results and conclusions.

- (i) The MLWF's are mainly confined to the primitive unit cell, for all the studied systems. This is a very important feature for their application as a very accurate minimal basis in a variety of theoretical approaches, such as the EXX method.
- (ii) The spread functional is found to converge slowly with respect to the density of \mathbf{k} -space mesh used, and this is mainly due to the gauge invariant part.
- (iii) The deviation of the Wannier center from the center of the bond is found to be proportional to the bond ionicity.
- (iv) Based on (iii), a new ionicity scale has been introduced. The obtained ionicities of the studied systems are in much better agreement with those obtained self-consistently (from the calculated charge density) than with the empirical Phillips ionicity scale.

APPENDIX A

SPECIAL POINTS

Many solid state calculations in the solids require averaging of a periodic function of the wave vector, \mathbf{k} , over the BZ. Such calculations are often complicated, time consuming, and, in principle, require the knowledge of the functional values at each \mathbf{k} -point in the BZ. In practice, these averages are determined by sampling the function at a discrete set of points and summing with an appropriate weight for each point. The problem is how to select the most efficient set, of the least number of \mathbf{k} -points, from which the average values of various quantities can be calculated to a sufficient accuracy.

Methods for finding such sets of “special” \mathbf{k} -points have been introduced by Baldereschi in 1973 [51], and developed further by Chadi and Cohen [52]. An alternative method was introduced by Monkhorst and Pack (MP) [45]. In our calculations, the special \mathbf{k} -points were generated using the MP scheme.

Monkhorst and Pack have defined a mesh of q^3 \mathbf{k} -points, which are uniformly spaced in the BZ, given by

$$\mathbf{k}_{p,r,s} = u_p \mathbf{b}_1 + u_r \mathbf{b}_2 + u_s \mathbf{b}_3, \quad (\text{A.1})$$

Where \mathbf{b}_1 , \mathbf{b}_2 and \mathbf{b}_3 are the primitive reciprocal lattice translation vectors, and

$$u_j = \frac{2j - q - 1}{2q}, \quad j = p, r, s = 1, 2, 3, \dots, q. \quad (\text{A.2})$$

This set of q^3 \mathbf{k} -points is then divided in symmetry related shells, and we choose per each shell one point in the irreducible part of the BZ (IBZ). The weight of this special \mathbf{k} -point is the ratio between the number of points in the shell and the total number of points.

For an illustration, we show in Fig. A.1 the $2 \times 2 \times 2$ MP cubic mesh of a simple cubic lattice. In this case, the cubic BZ is divided into 8 smaller cubes of equal sizes, and the MP mesh points are nothing but the centers of these 8 cubes.

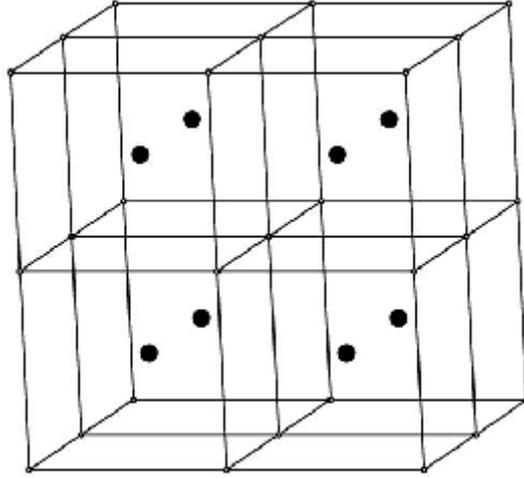


Fig. A.1: $2 \times 2 \times 2$ MP mesh of the simple cubic lattice.

The $4 \times 4 \times 4$ MP mesh of the FCC lattice is shown in Fig. A.2. In this case, the reciprocal space unit cell (defined by \mathbf{b}_1 , \mathbf{b}_2 and \mathbf{b}_3) is divided uniformly into 64 identical cells, and the MP mesh points are the centers of these small cells. These 64 \mathbf{k} -points are reduced by symmetry to the ten special \mathbf{k} -points of Chadi and Cohen [52] (used in the self-consistent calculations).

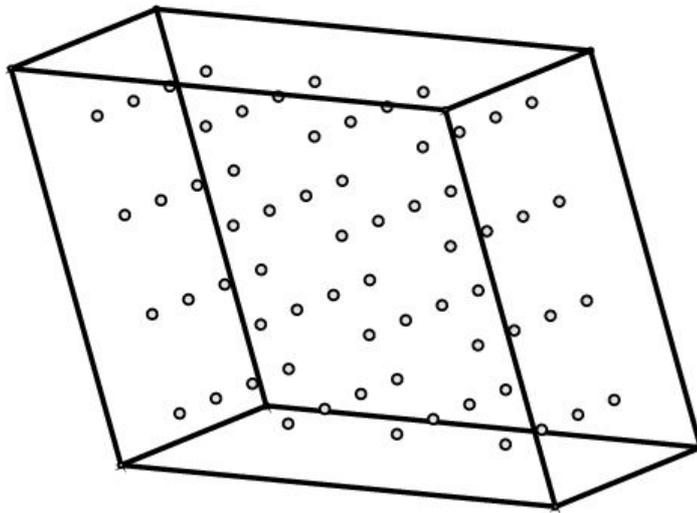


Fig. A.2: $4 \times 4 \times 4$ MP mesh of the FCC lattice.

REFERENCES

- [1] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
- [2] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
- [3] S. Lundqvist and N. H. March, *Theory of the Inhomogeneous Electron Gas* (Plenum, New York, 1983).
- [4] D. C. Langreth and M. J. Mehl, Phys. Rev. B **28**, 1809 (1993).
- [5] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
- [6] J. P. Perdew, K. Burke, and Y. Wang, Phys. Rev. B **54**, 16 533 (1996).
- [7] T. Kotani, Phys. Rev. B **50**, 14 816 (1994).
- [8] M. Städele, M. Moukara, J. A. Majewski, P. Vogl, and A. Görling, Phys. Rev. B **59**, 10 031 (1998).
- [9] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, 1976).
- [10] C. Kittel, *Introduction to Solid State Physics* (John Wiley & Sons, Inc., 1996).
- [11] G. H. Wannier, Phys. Rev. **52**, 191 (1937).
- [12] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).
- [13] I. Souza, N. Marzari, and D. Vanderbilt, Phys. Rev. B **65**, 035109 (2002).
- [14] G. Galli and M. Parrinello, Phys. Rev. Lett. **69**, 3547 (1992).
- [15] G. Galli, Curr. Opin. Sol. State Mater. Sci. **1**, 864 (1996).
- [16] I. Schnell, G. Czycholl, and R. C. Albers. To be published (2003).
- [17] R. D. King-Smith and D. Vanderbilt, Phys. Rev. B **47**, 1651 (1993).
- [18] R. Resta, Rev. Mod. Phys. **66**, 899 (1994).
- [19] M. D. Segall, P. J. D. Lindan, M. J. Probert, C. J. Pickard, P. J. Hasnip, S. J. Clark and M. C. Payne, J. Phys.: Condensed Matter **14**, 2717-2744 (2002).
- [20] <http://www.sfhingx.de>.
- [21] M. Born and J.R. Oppenheimer, Ann. Phys. **84**, 457 (1927).
- [22] W. Kohn, Rev. Mod. Phys. **71**, 1253 (1999).

- [23] J. A. Pople, *Rev. Mod. Phys.* **71**, 1267 (1999).
- [24] M. Levy, *Proc. Natl. Acad. Sci. USA* **76**, 6062 (1979).
- [25] M. Lüders, A. Ernst, W. M. Temmerman, Z. Szotek, and P. J. Durham, *J. Phys. Cond. Mat.* **13**, 8587 (2001).
- [26] J. Muskat, A. Wander, and N. M. Harrison, *Chem. Phys. Lett.* **342**, 397 (2001).
- [27] D. M. Cerperly and B. J. Alder, *Phys. Rev. Lett.* **45**, 566 (1980).
- [28] J. P. Perdew and A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).
- [29] R. O. Jones and O. Gunnarsson, *Rev. Mod. Phys.* **61**, 689 (1989).
- [30] P. Giannozzi, S. De Gironcoli, P. Pavone and S. Baroni, *Phys. Rev. B* **43**, 7231 (1991).
- [31] X. Gonze, D. C. Allan and M. P. Teter, *Phys. Rev. Lett.* **68**, 3603 (1992).
- [32] M. L. Cohen and T. K. Bergstresser, *Phys. Rev.* **141**, 789 (1966).
- [33] D. R. Hamann, M. Schlüter, and C. Chiang, *Phys. Rev. Lett.* **43**, 1494 (1979).
- [34] G. P. Kerker, *J. Phys. C* **13**, L189 (1980).
- [35] A. Zunger and M. L. Cohen, *Phys. Rev. B* **20**, 4082 (1979).
- [36] L. Kleinman and D. M. Bylander, *phys. Rev. Lett.* **48**, 1425 (1982).
- [37] D. Vanderbilt, *Phys. Rev. B* **41**, 7892 (1990).
- [38] A. M. Rappe, K. M. Rabe, E. Kaxiras, and J. O. Joannopoulos, *Phys. Rev. B* **41**, 1227 (1990).
- [39] J. S. Lin, A. Qteish, M. C. Payne, and V. Heine, *Phys. Rev. B* **47**, 4174 (1993).
- [40] N. Troullier and J. L. Martins, *Phys. Rev. B* **43**, 1993 (1991).
- [41] G. B. Bachelet, D. R. Hamann and M. Schlüter, *Phys. Rev. B* **26**, 4199 (1982).
- [42] J. Callaway, *Energy Band Theory* (New York: Academic Press, 1964) p99.
- [43] M. C. Payne, M. P. Teter, D. C. Allen, T. A. Arias, and J. D. Joannopoulos, *Rev. Mod. Phys.* **64**, 1045 (1992).
- [44] E. I. Blount, *Solid State Physics* **13**, 305 (1962).
- [45] H. J. Monkhorst and J. D. Pack, *Phys. Rev. B* **13**, 5188 (1976).

- [46] I. Schnell, *Ab-initio Wannier Functions, Coulomb Matrix Elements, Hartree (-Fock) and LSDA Claculations for the 3d Transition Elements Fe, Co, Ni and Cu*. Doctoral Thesis, University of Bremen, Germany (2002).
- [47] X. Gonze, P. Käckell and M. Scheffler, *Phys. Rev. B* **41**, 12264 (1990).
- [48] <http://www.phinax.de>.
- [49] J. C. Phillips, *Rev. Mod. Phys.* **42**, 317 (1970); *Bonds and Bands in Semiconductors* (Academic, New York, 1973).
- [50] A. Garcia and M. L. Cohen, *Phys. Rev. B* **47**, 4215 (1993).
- [51] A. Baldereschi, *Phys. Rev. B* **7**, 5212 (1973).
- [52] D. J. Chadi and M. L. Cohen, *Phys. Rev. B* **8**, 5747 (1973).