

The pseudopotential-density-functional method applied to semiconducting crystals

Citation for published version (APA):

Denteneer, P. J. H. (1987). *The pseudopotential-density-functional method applied to semiconducting crystals*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Applied Physics and Science Education]. Technische Universiteit Eindhoven. <https://doi.org/10.6100/IR263952>

DOI:

[10.6100/IR263952](https://doi.org/10.6100/IR263952)

Document status and date:

Published: 01/01/1987

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

THE PSEUDOPOTENTIAL-DENSITY-FUNCTIONAL METHOD
APPLIED TO
SEMICONDUCTING CRYSTALS



P.J.H. DENTENEER

**THE PSEUDOPOTENTIAL-DENSITY-FUNCTIONAL METHOD
APPLIED TO
SEMICONDUCTING CRYSTALS**

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag
van de rector magnificus, Prof. dr. F.N. Hooge,
voor een commissie aangewezen door het college
van dekanen in het openbaar te verdedigen op
vrijdag 5 juni 1987 te 16.00 uur

door

PETER JAN HENDRIK DENTENEER
geboren te Brunssum

Dit proefschrift is goedgekeurd door de promotoren:

Prof.dr. W. van Haeringen

en

Prof.dr. J.T.L. Devreese

Aan mijn ouders

*While timorous knowledge stands considering,
audacious ignorance hath done the deed.*

Samuel Daniel

ABSTRACT

A detailed description is given of the pseudopotential-density-functional method to accurately calculate from first principles the electronic and atomic structure of the ground state of crystals. Density-functional theory necessitates the self-consistent solution of the one-electron Schrödinger equation, whereas pseudopotentials allow for the inclusion in the calculation of valence electrons only and for the expansion of the functions of interest in plane waves. All necessary formulae are given to obtain the self-consistent density of valence electrons, screening potential, and energy of the ground state.

Particular emphasis is placed on the application of the technique of "special points in the first Brillouin zone" to perform necessary integrations over reciprocal space. The exploitation of space-group symmetry in the solution of the Schrödinger equation is discussed and illustrated for the case of expansion of the wave function in plane waves. Furthermore, characteristic features of the calculational scheme connected with self-consistency and finite cutoffs are pointed out and utilized to reduce the computational work.

Results of calculations for silicon, diamond, and two structurally extreme polytypes of silicon carbide illustrate the method and techniques described. Finally, the applicability of the method to surfaces, interfaces, superlattices, and polytypes is briefly discussed.

CONTENTS

	page
Chapter 1 Introduction	1
Chapter 2 The pseudopotential-density-functional method in momentum space	7
2.1 Density-functional theory	8
2.2 Pseudopotential theory	13
2.3 Momentum-space formalism for self-consistent pseudopotential calculations	20
2.3.1 Total energy in direct space	21
2.3.2 Total energy in momentum space	24
2.3.3 Self-consistent solution of Kohn-Sham equations in momentum space	31
2.4 Matrix elements of norm-conserving pseudopotentials	36
2.5 Cutoff parameters	40
Chapter 3 Special points in the first Brillouin zone	45
3.1 General theory and application to charge-density calculations	46
3.2 Description and computerization of the Monkhorst-Pack scheme	53
3.2.1 MP-sets for face-centred cubic lattices	54
3.2.2 MP-sets for hexagonal lattices	57
3.3 Convergence of energy-band integrations using special points	60
3.4 Equivalent special-point sets for structurally different crystals	64
Chapter 4 Exploitation of crystal symmetry for electronic energy bands and states	71
4.1 Construction of symmetrised plane waves: theory	72
4.2 Construction of symmetrised plane waves: illustrative examples	78
4.3 Unfolding of symmetrised electron states	85

Chapter 5	Secrets de cuisine: capita selecta	89
5.1	Non-self-consistency correction	89
5.2	Numerical noise on total-energy-versus-volume curve	93
5.3	Accuracy of energy-band integrations using special points	98
Chapter 6	Application to silicon, diamond, and silicon carbide	101
6.1	Self-consistent valence-charge density of silicon and diamond	104
6.2	Ground-state properties of silicon and cubic SiC	110
6.3	Valence-charge density and band structure of cubic SiC	120
6.4	Accurate energy differences and equivalent special-point sets	124
6.5	Wurtzite SiC: mapping and relaxation	128
Chapter 7	Outlook: Towards a fundamental description of crystals with limited periodicity	135
References		141
Samenvatting		146
Curriculum Vitae		148

CHAPTER 1

INTRODUCTION

The study of the condensed state of matter -solids and liquids- constitutes one of the largest subfields of modern physics. In view of its link to society (materials science), the importance of this field is obvious. From a more scientific point of view the purpose of solid-state physics is, of course, to understand the properties of solids starting from basic notions; Why is one solid different from another? An increased understanding of the properties of solids immediately leads to a more systematic search for materials that have desirable properties. There is interest, for instance, in (i) solids that are as ductile and malleable as common metals, but are corrosion-resistant, (ii) solids with the hardness and chemical inertness of diamond, but not as costly, (iii) semiconductors with a band gap that is direct and corresponds to a desirable frequency (color), for use in light-emitting diodes, lasers, and photo-detectors, (iv) semiconductors with high electron mobilities, which have a higher potential operating speed in electronic devices, to mention a few.

To this end, experiments are needed to determine the properties of solids. We also need theories that tell us why solids have the properties they have. These theories should preferably start from elementary ingredients. Regarding solids, these elementary ingredients are the properties of the nuclei, the electrons, and their interactions. The latter category of theories are called *first-principles theories* or *ab-initio theories*. Quantum mechanics and statistical mechanics are such theories, which should in principle suffice to determine the properties of solids from first-principles. In practice, however, these general theories alone almost invariably generate a calculational scheme that is too complex to actually carry out. By making approximations that are not too drastic, it is possible to obtain theories that may still be called first-principles theories, but lead to practical schemes of calculation. The approximations of course must be carefully investigated for their appropriateness and should not violate the basic laws of quantum mechanics and statistical mechanics. Theories that need experimental data as input, e.g., in

order to determine the values of parameters in the theory, are called *empirical theories*. Such theories are in fact less fundamental with regard to predictive purposes. In what is called the scientific method, theories, irrespective of whether they are empirical or start from first principles, are first tested to *reproduce* the results of experiments and are subsequently tested to *predict* the results of experiments.

Only in the last decade it has become possible to employ first-principles theories in the computation of solid-state properties and to reliably predict experiments. This is partly due to the steady advance made in the development of theories, the most important reason, however, lies in the increase in computing power of the generations of digital computers that rapidly succeed each other. The latter development has led some people to discern a third way to study physics, in-between experimental and theoretical physics, namely that of *computational physics* [2]. Although computational physics has descended from theoretical physics historically, its approach is more akin to that of experimental physics. A large computer code must be designed and tested part by part just as careful as an experimental set-up. Both computer code and experimental set-up can be used to perform experiments, be it of a different, possibly supplementary, kind.

In a first-principles theory for solids, it appears to be necessary to solve the *Schrödinger equation* of quantum mechanics for electrons *self-consistently* (if the electrons may be treated non-relativistically). The need for self-consistency is caused by the fact that the electrons interact with each other. Therefore, in the Schrödinger equation for the individual electrons, the effective potential, describing the interactions the electrons experience, depends on the solutions of the equation itself. Only if the Schrödinger equation is solved in this self-consistent way, the electronic structure of the solid follows in a reliable way. Subsequently the total energy of the solid can be calculated, as well as first-order derivatives of the total energy with respect to changes in the atomic positions, providing forces, stresses and pressure. Assuming that the solid strives for the situation of minimum energy (and zero force), the equilibrium positions of the atoms can be found. Many other properties of the solid may also be found. In this way a microscopic description

-on the level of atoms- based on quantum mechanics is obtained.

Demanding a solid to be *periodic* -we call such a solid a crystal-, simplifies the above task of calculating electronic properties considerably. The attention can be confined to a unit cell, usually containing between 1 and 10 atoms, which is representative for the whole crystal when repeated in three independent directions.

In this work we give a reasonably complete description of a method, called the pseudopotential-density-functional method, by which the electronic-structure problem may self-consistently be solved without parameters determined from experiments. Since the growing ability of computational physics has made it a discipline in its own right, the study of its methods is appropriate. In the present method a distinction is made in the solid between electrons that are so tightly bound to the nuclei as to be negligibly perturbed from their behaviour in the atom (*core* electrons), and electrons that adjust themselves to the different environment in the solid (*valence* electrons). The latter electrons have appeared to be responsible for a majority of solid-state properties. *Pseudopotential theory* (see also chapter 2) assumes that the cores, i.e., nuclei plus core electrons, interact in the same way with available other electrons in the cases of both large and small separation of the atoms, as in a gas and a solid, respectively. In this theory the energy of the interactions within the core is not taken into account. This implies that the total energy we calculate, which will nevertheless be called "the total energy" in the remainder of this work, is really the difference between the actual total energy of the crystal and the energy of isolated cores. Pseudopotential theory combined with the periodicity of crystals allows for a convenient, Fourier analysed version of the calculational scheme. In this version no assumptions have to be made a priori about the form of the electron density. The latter fact makes the method particularly suited for calculations on crystals with *covalent bonds*, where the electron density accumulates in bonds between nearest-neighbour atoms. This in contrast to the more simple cases of metals, where the electrons are nearly free and their density consequently is fairly constant over the crystal, and ionic crystals, where the electrons can all be seen as belonging to one atom, resulting in largely spherical electron densities centred on the atomic positions.

Crystals with covalent bonds form a highly interesting class of materials. Nearly all semiconductors belong to this class and also diamond, which is an insulator. The physics of semiconductors lies at the basis of modern electronics, computers, and information handling hardware. To this day semiconductor technology is largely based on the semiconductor silicon. Although silicon has superb chemical and mechanical properties, it certainly is not an ideal choice regarding its electronic properties. The electron mobility in silicon is only average and its indirect band gap limits many optical applications. Currently, theory, experiment, and technology are joining hands to find out what materials may be suitable, as well as technologically feasible, to replace silicon. Candidates are, gallium arsenide, possibly combined with silicon, and germanium-silicon systems. Very recently, progress in developing diamond-transistors was reported [3]. Because of the whole of its natural properties, diamond is considered to be a material superior to silicon for this application. The method described in this work can prove useful in the undertaking of finding new materials.

In the following chapters the focus is more on the method itself than on underlying theories or calculated properties, although both of these subjects are also addressed. Consequently, some chapters are of a technical nature. We think this full exposition of the method is justified because of the gratifying results already attained with this method. It is furthermore useful in closing a gap between present and future practitioners of the method.

Chapter 2 starts with a general introduction to the two basic theories on which the pseudopotential-density-functional method is based: *density-functional theory* describes a system of many interacting electrons in an external potential in terms of the electron density, while *pseudopotential theory* describes the behaviour of valence electrons in a solid. The rest of the chapter provides a self-contained treatment of the calculational scheme that emerges from the combination of the two theories. Fourier analysis or, equivalently, expansion of the functions of interest in *plane waves* exploits the translational symmetry of the crystal and provides a transparent calculation scheme in Fourier -or *reciprocal*- space. Furthermore, a discussion is given of technical approximations that must be made to make calculations feasible.

In the method of chapter 2 frequently integrations over reciprocal space have to be performed. In chapter 3 a technique to this end is treated that makes explicit use of the symmetry of the crystal, the technique of *special points in the Brillouin zone*. This technique is especially suited for application to semiconductors and insulators.

The method of chapter 2 also results in the necessity of solving large sets of linear equations. In chapter 4 we show how the symmetry of the crystal and results from group theory are exploited to subdivide these large sets into a number of smaller sets of linear equations. This chapter is not essential to the method, but merely allows for considerable reduction in computing times. The chapter is reasonably self-contained and can properly be skipped on a first reading.

In chapter 5 some characteristic properties of the calculational scheme are discussed. These properties enable one to make a more convenient use of the method.

Chapter 6 includes a number of applications of the method to silicon, diamond, and two -semiconducting- modifications of silicon carbide. These applications serve as illustrations of the method and techniques of the preceding chapters, as illustrations of the potentialities of the method, and serve also as presentation of results for silicon carbide, for which until recently no such calculations had been performed.

As an outlook to the future, we end by sketching in chapter 7 the applicability of the pseudopotential-density-functional method to interfaces, surfaces, stacking faults, and superlattices of semiconductors. Such systems are becoming increasingly important in technology and the desire to study such systems motivated the present study.

CHAPTER 2

THE PSEUDOPOTENTIAL-DENSITY-FUNCTIONAL METHOD IN MOMENTUM SPACE

In this chapter a detailed description is given of a method by which ground-state properties of a solid may be calculated. This method, which we call the pseudopotential-density-functional method, finds its origin in two basic theories: density-functional theory (DFT) and pseudopotential theory. The method combines both theories in such a way that ground-state properties of a large class of solids -among which the semiconductors to which we will apply the method- may be determined. DFT is a theory describing a system of many electrons with mutual interactions in an external potential and is discussed in section 2.1. Pseudopotential theory, which will be discussed in section 2.2, deals with the behaviour of valence electrons in a solid. Valence electrons are electrons that originate from not completely filled shells of the atoms constituting the solid and may be held responsible for most of the properties of interest of a solid. The other electrons are called core electrons. Recent advances in especially pseudopotential theory, namely, the construction of so-called norm-conserving pseudopotentials, have made possible the accurate calculation of properties of solids without the need of any empirical or adjustable parameters.

Whereas sections 2.1 and 2.2 have a global character and may be regarded as a general introduction to the basic theories, sections 2.3 to 2.5 are more specific and detailed. In section 2.3 it is shown that when the problem is treated in momentum space, which is just another way of saying that all functions of interest (wave functions, charge densities, potentials,...) are expanded in plane waves (or -put differently- Fourier analysed), the calculational scheme becomes very transparent. In section 2.4 useful formulae are given that enable one to make almost direct use of tabulated versions of norm-conserving pseudopotentials for all elements in the periodic table. Finally, in section 2.5 the inevitable technical approximations are discussed that one is forced to make in order to make calculations feasible. In this connection the relevant cutoff parameters are introduced.

2.1 Density-functional theory

Density-functional theory (DFT) is an approach to describe a system of many interacting electrons and may as such be considered an alternative for the Hartree-Fock method. When applied to solids, DFT has definitely shown to be more practical and successful than the Hartree-Fock method. The theory was formulated first by Hohenberg and Kohn [4] and Kohn and Sham [5]. Since the density of particles plays a central role in the theory, DFT can be regarded as the direct descendant of the more intuitive theory of Thomas and Fermi [6]. Although DFT can be presented in terms of well-defined concepts, its application to actual solids still suffers from uncertainties. The most important uncertainties are: (i) Are the one-particle equations that emerge adequate to approximate the many-body problem?, (ii) What is the exact form of the exchange-correlation functional (to be introduced below)? Since our goal in this section is to only give a brief discussion of this theory, the reader is referred to the original papers and more recent reviews [7,8] for more elaborate discussions.

The theory finds its formal justification in the Hohenberg-Kohn (HK) theorem, which in its original form is applicable to the ground state of a system of spinless fermions (i.e., particles that obey Fermi-Dirac statistics) in an external potential. In this original form the theorem therefore applies to systems of electrons for which interactions connected with their spin are absent or may be neglected. The theorem may be summarized as follows:

- (i) The ground-state energy of a system of identical spinless electrons is a unique functional of the particle density. (The ground state is assumed to be nondegenerate).
- (ii) This functional has its minimum value for the correct ground-state density, when particle-number-conserving variations of the density are considered.

The ground-state-energy functional is written as:

$$E_v[n] = \int \psi_{\text{ext}}(\mathbf{r})n(\mathbf{r})d^3r + F[n]. \quad (2.1)$$

In (2.1) ψ_{ext} is the external potential, which is also a unique

functional (neglecting a possible additive constant) of the particle density $n(\mathbf{r})$. This is in fact the central and remarkable feature of the theorem: the fact that the external potential determines the particle density is obvious, the converse, however, is surprising and initially was greeted with some scepticism. The functional $F[n]$ includes all kinetic energy and electron-electron interaction terms. It is convenient to split off from $F[n]$ the energy due to the Coulomb interaction, also called Hartree energy:

$$F[n] = \frac{e^2}{8\pi\epsilon_0} \iint \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d^3r d^3r' + G[n]. \quad (2.2)$$

Here and everywhere else in this work, e is the charge of the electron (negative) and ϵ_0 the electric permittivity of the vacuum. It is important to note that the exact form of $G[n]$ is unknown. The ground state of the system is formally obtained by minimizing $E_v[n]$ with respect to density variations that conserve the number of particles N :

$$\int n(\mathbf{r})d^3r = N. \quad (2.3)$$

This leads to the variational equation:

$$\delta\left\{E_v[n] - \mu \int n(\mathbf{r})d^3r\right\} = 0, \quad (2.4)$$

in which a Lagrange multiplier μ is introduced due to the constraint (2.3). Applying (2.1), (2.2), and (2.4) the ground-state-density-determining equation is found:

$$\psi_{\text{ext}}(\mathbf{r}) + \frac{e^2}{4\pi\epsilon_0} \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d^3r' + \frac{\delta G[n]}{\delta n(\mathbf{r})} = \mu, \quad (2.5)$$

where the last term in the left-hand side of (2.5) is the functional derivative of $G[n]$ with respect to $n(\mathbf{r})$. Even if the functional form $G[n]$ were known, eq. (2.5) would still not give us a procedure to actually calculate the correct $n(\mathbf{r})$. Kohn and Sham [5] however supply a procedure that results in one-particle equations (the so-called Kohn-Sham (KS) equations), that we do know how to solve. Their line of reasoning runs as follows: consider a system of N non-interacting electrons in some external potential $\psi_{\text{ext},s}(\mathbf{r})$. The ground-state

density of this system is called $n(\mathbf{r})$. The functional $F[n]$ in (2.2) reduces to $T_s[n]$, the kinetic-energy functional of non-interacting electrons, and the equation determining $n(\mathbf{r})$ is given by (cf. (2.5)):

$$\psi_{\text{ext},s}(\mathbf{r}) + \frac{\delta T_s[n]}{\delta n(\mathbf{r})} = \mu_s. \quad (2.6)$$

The general form of $T_s[n]$ is again unknown (μ_s is determined by (2.3)), but now there is an alternative way to obtain $n(\mathbf{r})$: for non-interacting electrons the many-particle ground-state wave function is simply a completely anti-symmetrized product of one-electron wave functions $\varphi_i(\mathbf{r})$ (Slater determinant), each of which obeys the Schrödinger equation:

$$\left\{ -\frac{\hbar^2}{2m} \nabla^2 + \psi_{\text{ext},s}(\mathbf{r}) \right\} \varphi_i(\mathbf{r}) = \epsilon_i \varphi_i(\mathbf{r}), \quad i = 1 \dots N. \quad (2.7)$$

Planck's constant divided by 2π is denoted by \hbar and m is the electron mass. The prescription is to select those N states $\varphi_i(\mathbf{r})$ that have lowest energy ϵ_i . The density for this system of electrons is then given by:

$$n(\mathbf{r}) = \sum_{i=1}^N |\varphi_i(\mathbf{r})|^2. \quad (2.8)$$

So for this particular system of non-interacting electrons, there is indeed a way of finding the solution to equation (2.6). Kohn and Sham now show how this procedure may be used in the case of interacting electrons as well. The functional $G[n]$ is split up in two terms:

$$G[n] = T_s[n] + E_{xc}[n], \quad (2.9)$$

in which $T_s[n]$ is the kinetic energy of a system of non-interacting electrons with a density $n(\mathbf{r})$ and in which the remaining term $E_{xc}[n]$ by definition is called the exchange and correlation energy of the interacting system with density $n(\mathbf{r})$. Equation (2.5) now becomes:

$$\psi_{\text{ext}}(\mathbf{r}) + \frac{e^2}{4\pi\epsilon_0} \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d^3\mathbf{r}' + \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} + \frac{\delta T_s[n]}{\delta n(\mathbf{r})} = \mu. \quad (2.10)$$

This equation has the form of (2.6); the only difference is that $\psi_{\text{ext}}(\mathbf{r})$ is replaced by an "effective" potential $\psi_{\text{eff}}[n]$:

$$\psi_{\text{eff}}[n] = \psi_{\text{ext}}(\mathbf{r}) + \frac{e^2}{4\pi\epsilon_0} \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d^3r' + \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})}. \quad (2.11)$$

By analogy with the non-interacting case, the correct ground-state density of the interacting system is found by the self-consistent solution of the following set of one-particle equations (the KS-equations):

$$\left\{ -\frac{\hbar^2}{2m} \nabla^2 + \psi_{\text{eff}}[n] \right\} \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}), \quad i = 1 \dots N, \quad (2.12)$$

$$n(\mathbf{r}) = \sum_{i=1}^N |\psi_i(\mathbf{r})|^2. \quad (2.13)$$

Note that the self-consistency requirement is caused by the functional dependence of ψ_{eff} on $n(\mathbf{r})$. The total ground-state energy of the electron system is then given by:

$$E_v[n] = T_s[n] + \int \psi_{\text{ext}}(\mathbf{r})n(\mathbf{r})d^3r + \frac{e^2}{8\pi\epsilon_0} \iint \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d^3r d^3r' + E_{xc}[n]. \quad (2.14)$$

We have:

$$T_s[n] = \sum_{i=1}^N \int \psi_i^*(\mathbf{r}) \left[-\frac{\hbar^2}{2m} \nabla^2 \right] \psi_i(\mathbf{r}) d^3r. \quad (2.15)$$

In order to be able to find the self-consistent solution of (2.12)-(2.13) and to calculate $E_v[n]$, it is necessary to adopt some explicit form for $E_{xc}[n]$. A very useful approximation has proven to be:

$$E_{xc}[n] = \int \epsilon_{xc}(n(\mathbf{r}))n(\mathbf{r})d^3r. \quad (2.16)$$

where $\epsilon_{xc}(n)$ is the exchange and correlation energy of an interacting electron gas with uniform density n . For $\epsilon_{xc}(n)$ several useful

approximate expressions are known (see section 2.3). The approximation (2.16) is called the local-density-approximation (LDA), since the exchange and correlation energy density at position r is assumed to depend on the density at point r only. This assumption is valid if $n(r)$ is constant and the approximation can be considered acceptable for electron systems with almost constant or slowly varying $n(r)$. The approximation is in fact not justified for systems with large density gradients such as semiconductors. The apparent success of the approximation (2.16) in such cases is even more remarkable if one notes that, because of the definition implied in (2.9), $E_{xc}[n]$ must also contain some kinetic-energy contribution apart from "real" exchange and correlation energy contributions as in (2.16); this is due to the fact that $T_s[n]$ represents only part of the kinetic energy of the interacting system. By "real" exchange and correlation energy we mean the remaining energy of an electron gas when the kinetic and Hartree energies (and the energy due to a possible external potential) have been subtracted from the total energy. In view of the successful application of (2.16), the latter feature is either of minor importance or its effect is washed out by adopting approximate forms for $\epsilon_{xc}(n)$. Many more fundamental questions can be asked in connection with the KS-equations and their interpretation [7,8], but we will not go into these here.

The calculational scheme presented in section 2.3 is partly based on the equations (2.11)-(2.16). From what is put forward above this requires at least some justification. One can argue that calculations for real materials so far always have been more or less successful by employing the idea of an effective potential. Therefore there is a lot of faith in the resulting one-particle equations and a lot of experience in solving them. The attractive feature of this new scheme is that it promises to give the correct ground-state density and from that other ground-state properties. Therefore this new scheme is -even in some approximate fashion- worth exploring. In this connection it seems justified to say that an actual calculational scheme using the LDA has a somewhat less firm foundation in DFT than is usually suggested. On the other hand, DFT has been a strong motivation for such calculations and to a large degree these calculations can be justified by their success [9]. The major step forward with respect to Hartree-Fock theory is the inclusion -although in an approximate

manner- of extra interaction effects (these extra effects are conventionally denoted by "correlation" effects. This name, however, is misleading, since essentially other interactions -e.g., exchange- also result from correlations between the electrons).

2.2 Pseudopotential theory

Pseudopotential theory is a theory that circumvents the need of an accurate description of the core electrons, i.e., the electrons occupying the completely filled shells of atoms. In a solid these electrons remain very localized around the atom, whereas the remaining electrons called valence electrons determine the majority of the properties of the solid. This is why pseudopotential theory is useful: it provides a simpler approach to the properties of solids. The first pseudopotential theory was formulated by Phillips and Kleinman [10] based on the orthogonalized-plane-wave (OPW) method of Herring [11].

In this section we briefly discuss the general concepts in pseudopotential theory. We also present the class of *ab initio* pseudopotentials called norm-conserving pseudopotentials that were put forward by Hamann et al. [12] not very long ago. These have put pseudopotential theory on a new level of sophistication, because self-consistent calculations with the necessary accuracy became possible. The discussion here is based on the more extensive reviews in ref.[13] regarding concepts and refs.[14] and [15] regarding norm-conserving pseudopotentials. We refer to these papers if not all details are given here (see also the review in ref.[16]).

The electrons in a solid move in the Coulomb field of the fixed nuclei and have their mutual interactions. All these interactions are assumed to be taken into account by adopting a one-electron picture, in which the electrons experience an effective potential V . These electrons now all obey a Schrödinger equation given by:

$$H|\psi\rangle = E|\psi\rangle, \quad (2.17a)$$

$$H = T + V. \quad (2.17b)$$

Here, T is the kinetic-energy operator $-(\hbar^2/2m)\nabla^2$. In (2.17a) we use

Dirac notation. In this notation a wave function $\psi(\mathbf{r})$ is denoted by a ket $|\psi\rangle$, its complex conjugate $\psi^*(\mathbf{r})$ by a bra $\langle\psi|$, and putting them together face to face implies integration, $\langle\varphi|\psi\rangle = \int \varphi^*(\mathbf{r})\psi(\mathbf{r})d^3\mathbf{r}$, so that the bra-ket combination defines a hermitian inner product (with the property: $\langle\varphi|\psi\rangle^* = \langle\psi|\varphi\rangle$). The matrix element of an operator O between two functions f_1 and f_2 is written as $\langle f_1|O|f_2\rangle$, meaning $\int f_1^*Of_2(\mathbf{r})d^3\mathbf{r}$.

Imagine the valence-electron state $|\psi\rangle$ to be written as a smooth pseudo-wave-function $|\varphi\rangle$ corrected to be orthogonal to all core-electron states $|c\rangle$:

$$|\psi\rangle = |\varphi\rangle - \sum_c |c\rangle\langle c|\varphi\rangle. \quad (2.18)$$

By smooth we mean expandable in few plane waves. It is reasonable to expect that $|\varphi\rangle$ will be smooth, as the effective potential outside the core regions is expected to be much smoother than inside the core regions. Note that we have not made any approximation yet, we only have made explicit the orthogonality of all core and valence states. If $|\varphi\rangle$ in (2.18) would be replaced by a single plane wave, eq. (2.18) represents a so-called OPW. OPW's appear to form a suitable basis set for calculations in solids, implying that the $|\varphi\rangle$ in (2.18) are indeed smooth. Substitution of (2.18) into (2.17a) gives an equation for $|\varphi\rangle$:

$$T|\varphi\rangle + V^{ps}|\varphi\rangle = E|\varphi\rangle, \quad (2.19)$$

where the pseudopotential V^{ps} is defined by combining the true potential V and the orthogonality terms (we use that $|c\rangle$ is an eigenstate of H with eigenvalue E_c):

$$V^{ps}|\varphi\rangle = V|\varphi\rangle + \sum_c (E - E_c)|c\rangle\langle c|\varphi\rangle. \quad (2.20)$$

We observe from (2.19) that for valence states -which are the ones we are interested in- the energy eigenvalues of the Hamiltonian H with the real potential are identical to those of the pseudo-Hamiltonian $H^{ps} = T + V^{ps}$. One easily verifies that this remains true if $E - E_c$ in (2.20) is replaced by a constant λ_c . This demonstrates the so-called non-uniqueness of the pseudopotential. It can be exploited to make the

pseudo-problem (2.19) as easy to solve as possible. We also note that V^{ps} in (2.20) is weaker than V , since the core energies E_c are lower than the valence energies E and the orthogonality term in (2.20) therefore is repulsive, partly cancelling the attractive potential V . This is consistent with the expected smoothness of the solution $|\varphi\rangle$ of (2.19). The non-uniqueness of V^{ps} motivates the operator approach to pseudopotentials [13]: we may define an operator V^{ps} in many ways as long as it gives the correct energy levels for valence states.

Another approach to pseudopotentials is possible, which is more closely related to scattering theory. A pseudopotential is now defined as one that gives the same scattering amplitudes [17] as the real potential for an incident plane wave with some reference energy E . The pseudopotential is allowed to differ from the real potential within a certain core sphere. It can be shown that a pseudopotential thus defined will also give the same band energy E for valence states. We will assume the pseudopotential to be spherically symmetric inside the core sphere. Therefore, in this scattering approach, the pseudopotential will depend on the angular momentum ℓ of the incident wave only and may generally be written as:

$$V^{ps} = \sum_{\ell} V_{\ell}^{ps} = \sum_{\ell} f_{\ell}(r) \mathcal{P}_{\ell}. \quad (2.21)$$

Here, \mathcal{P}_{ℓ} is a projection operator that picks out a specific angular-momentum component of the function that V^{ps} operates on. The functions $f_{\ell}(r)$ can be constructed such as to give the correct scattering properties of the core in a certain energy range (since the pseudopotential defined in this way is valid for one E only). From (2.21) (and also from (2.20)) it is clear that V^{ps} in general is a nonlocal operator, i.e., not a mere multiplication operator. The scattering approach -just as the former operator approach- allows to describe valence states by smooth and nodeless wave functions inside the core. The strong oscillations inside the core are eliminated by letting the pseudopotential reproduce the *reduced phase shifts* of the real potential instead of the *complete phase shifts*: this makes no difference for the scattering amplitudes.

So both approaches make it possible to replace the problem of finding the energies of the valence electrons via (2.17) by a problem of the type (2.19). The additional freedom of choice is exploited in

the operator approach to make the solutions $|\varphi\rangle$ as smooth as possible and in the scattering approach, where the solutions $|\varphi\rangle$ are smooth by construction, to maximize the energy range for which the pseudo-potential is valid.

If we would solve eq. (2.19) with V^p given by (2.20), we have to realize that this does not immediately give us the true valence states $|\psi\rangle$. Yet these are needed for almost all other properties one would want to calculate. In view of section 2.1, especially the charge density is of interest. One could of course use eq. (2.18) to construct $|\psi\rangle$, by using core states obtained from atomic calculations. Apart from the objection that the states $|c\rangle$ in principle should come out of the same calculation as the $|\varphi\rangle$ for use in eq. (2.18), there is another problem with eq. (2.18): the so-called orthogonality-hole problem. This is essentially a normalization problem. It occurs because with eq. (2.18) the true valence state $|\psi\rangle$ and the pseudo-wave-function $|\varphi\rangle$ cannot be normalized simultaneously. This is seen by multiplying (2.18) by bra's $\langle\psi|$ and $\langle\varphi|$, respectively, and combining the resulting equations:

$$\langle\psi|\psi\rangle = \langle\varphi|\varphi\rangle - \sum_c |\langle\varphi|c\rangle|^2. \quad (2.22)$$

The (positive) term $\sum_c |\langle\varphi|c\rangle|^2$ is called the orthogonality hole.

The solution $|\varphi\rangle$ to our substitute problem (2.19) is determined to within a constant factor. Suppose that this constant is chosen such that $|\varphi\rangle$ is normalized. The *normalized* true wave function $|\psi'\rangle$ with the same shape as $|\psi\rangle$ in (2.18) is then obtained by means of an additional factor γ :

$$|\psi'\rangle = \gamma(|\varphi\rangle - \sum_c |c\rangle\langle c|\varphi\rangle), \quad \gamma = (1 - \sum_c |\langle\varphi|c\rangle|^2)^{-1/2} \quad (2.23)$$

From (2.23) we see that $|\psi'\rangle$ has a larger amplitude than $|\varphi\rangle$ outside the core region (where $|c\rangle$ is negligible). Therefore, if we would use $|\varphi\rangle$ instead of $|\psi'\rangle$ for the determination of charge densities, too much of the total charge is put in the core region. So even outside the core one does not find the correct valence-charge density by using $|\varphi\rangle$. This is a serious problem in a self-consistent calculation via KS-equations, where the density is the crucial quantity (see section

2.1). Of course $|\varphi\rangle$ could be orthogonalized to all core states, but this is not easy and bypasses the idea behind and the advantage of pseudopotentials.

At first this problem blocked the way to the construction of *ab initio* pseudopotentials, i.e., pseudopotentials that correctly reproduce the energy eigenvalues and wave functions (outside some core radius) of valence states that are found in atomic all-electron calculations. Such pseudopotentials are required for self-consistent calculations in solids. The problem is caused by the construction of pseudopotentials implied in (2.20) (Phillips-Kleinman construction), but this way of constructing is not obligatory. It has been shown that the scattering approach enables one to overcome the problems with pseudo-wave-functions.

The most popular scheme to construct *ab initio* pseudopotentials was devised in ref.[12] (another scheme is given in ref.[18]), and starts with the construction from all-electron atomic calculations in the density-functional scheme of angular-momentum-dependent pseudopotentials, V_ℓ^{ps} , which by construction have the property that:

- (1) Energy eigenvalues for valence states in the all-electron calculation and in the pseudopotential calculation agree exactly for some chosen prototype configuration.
- (2) The wave functions of valence electrons in the all-electron calculation and in the pseudopotential calculation agree exactly outside a chosen core radius r_c . The pseudo-wave-function is chosen nodeless inside r_c .

Globally speaking, the construction of V_ℓ^{ps} is achieved by making some choice for the wave function within r_c (this can be done in arbitrarily many ways) and inverting the radial Schrödinger equation (there is no problem here because of the nodeless property of the adapted wave function). When pseudopotentials are constructed in this way they have two properties that make them *transferable*, i.e., useful in other situations than the one in which they are generated:

- (i) They yield the correct amount of charge inside the core radius r_c , so that the electrostatic potential outside r_c is the same for real and pseudo-charge-densities (*norm conservation*).
- (ii) The scattering amplitudes of the real ion cores are reproduced with minimum error as the energy starts deviating from the energy for which the pseudopotential was constructed. About the range of energies

for which this "minimum error" is acceptable nothing can be said a priori, but in practice these pseudopotentials turn out to be very satisfactory. In addition to this energy-"independence" of the pseudopotentials, it is shown in ref.[15] that there is also a perturbation-"independence", i.e., scattering amplitudes are not very sensitive to small variations in the potential.

To obtain ion-core pseudopotentials $V_{\ell}^{i\circ n}$ for the angular-momentum components ℓ , which are to be used in solid-state or molecular calculations, we must unscreen the pseudopotentials $V_{\ell}^{p\circ n}$, i.e., subtract the effect of the potential caused by all valence electrons of the atomic configuration. This potential is the sum of the Hartree and the exchange-correlation (XC) potentials due to the valence-(pseudo-)charge density (see section 2.1). By the core we always mean the nucleus plus the core electrons. Since the core is an ion, it is also called ion-core or ion. Note that both in the original atomic calculation and in the unscreening calculation the KS-equations of DFT with the LDA for exchange and correlation are employed. In fact, in the unscreening act it is furthermore assumed that the core- and valence-charge densities may be decoupled to calculate the XC-potential. This is clearly an approximation, since every useful form for the XC-potential (see section 2.3) is an explicitly nonlinear functional of the density. If this decoupling is not allowed, e.g., when core- and valence-charge densities overlap substantially, there are methods to correct for this [19].

Norm-conserving ion-core pseudopotentials, $V_{\ell}^{i\circ n}$, can be constructed for any element in the periodic table and for various ℓ [14].

From now on the assumption will be made -this is called the pseudopotential approximation- that these $V_{\ell}^{i\circ n}$ correctly represent the complete potential the valence electrons feel from nuclei plus core electrons and that this complete potential is not affected by using it in other environments than the one in which it was generated. This approximation is also called the "frozen-core approximation", because the interaction between the core and the valence electrons is assumed to be frozen (such as to lead to smooth valence states). As noted in ref.[15], this approximation is not identical to the approximation of the same name in which the core states are frozen (but the valence states still have the strong oscillations in the core region). It was proved that to first order in the error in the core-charge density the

total energy is exact in both approximations [20]. The frozen-core-approximation error may be estimated to be 0.1 eV/atom in the very worst cases and about 0.02 eV/atom for silicon and carbon (i.e., less than 0.5 % of the cohesive energies of their crystallized forms) to which we will apply our method. This implies that, as stated in chapter 1, the total energy in pseudopotential theory may indeed be seen as the difference between the actual total energy of the solid and the sum of energies of isolated cores.

Together with the theory of section 2.1 a scheme can now be put forward in which a crystal is seen as a many-electron system in which (valence) electrons move in the external potential formed by a periodic arrangement of ion-cores. The ground-state density and from that all ground-state properties are found by self-consistently solving the KS-equations (2.11)-(2.13), where v_{ext} is the sum of all ion-core pseudopotentials in the crystal. This scheme as well as its computational implications will be extensively discussed in the rest of this chapter. A conceptual difficulty with this scheme is that v_{ext} now is a nonlocal operator of the form (2.21). For general nonlocal v_{ext} the HK theorem no longer holds: the total energy of the ground state is then a unique functional of the density matrix rather than just its trace (i.e., the density [21]). To the author no rigorous justification is known to proceed with the density as crucial quantity in a self-consistent pseudopotential theory, where one uses nonlocal ion-core pseudopotentials in combination with the KS-equations of DFT. In practice, however, the results of such a procedure are very good and agree with those from all-electron calculations in the local-density-functional scheme, where the external potential is local, viz., the superposition of the Coulomb potentials of the nuclei. Perhaps a justification can be derived from the special form of nonlocality (eq. (2.21)) resulting from the assumed spherical symmetry of the core potentials.

2.3 Momentum-space formalism for self-consistent pseudopotential calculations

In this section we combine the theories of sections 2.1 and 2.2 to obtain a calculational scheme for the total ground-state energy per unit cell of an arbitrary periodic solid. This combination implies that the electrons in DFT will be the valence electrons only, which move in a (nonlocal) external potential given by the superposition of norm-conserving ion-core pseudopotentials. We furthermore impose periodicity upon the solid and call this a *crystal*. This implies that the wave functions $\psi_i(\mathbf{r})$ in (2.12) and (2.13) are replaced by pseudo-Bloch-functions $\psi_{n,\mathbf{k}}^{\text{ps}}(\mathbf{r})$, where n is the band index and \mathbf{k} a wave vector in the first Brillouin zone (1BZ) (reduced wave vector). We will immediately drop the superscript "ps" for the wave function, since from now on we will only consider pseudo-wave-functions.

In section 2.3.1 the KS-equations and a total-energy expression in \mathbf{r} -space (direct space) are given, whereas in section 2.3.2 these are Fourier transformed so that a formalism in momentum space results. In section 2.3.3 the necessary formulae and steps are discussed to obtain a self-consistent solution to the equations of section 2.3.2.

All formulae will be given in MKS-units contrary to popular practice in the literature. Using MKS-units is the best way to keep track of the dimension of all quantities appearing in the formulae. For use on computers a transition to some system of atomic units has to be made. In these atomic units (a.u.) all quantities in the computer program have a convenient order of magnitude. For these a.u. two possibilities are in general use: Rydberg atomic units and Hartree atomic units, named after the unit of energy resulting from these sets of units. In all our formulae the transition to these units is easily made as follows ("→" means "replace by"):

$$(i) \text{ Rydberg a.u.: } \epsilon_0 \rightarrow 1/4\pi, \hbar \rightarrow 1, m \rightarrow \frac{1}{2}, e^2 \rightarrow 2, \quad (2.24a)$$

$$(ii) \text{ Hartree a.u.: } \epsilon_0 \rightarrow 1/4\pi, \hbar \rightarrow 1, m \rightarrow 1, e \rightarrow 1. \quad (2.24b)$$

In both cases all quantities with the dimension $[\text{length}]^p$ must be expressed in units of a_0^p ($p \in \mathbb{R}$ and a_0 is the Bohr radius: $a_0 = 0.052917715 \text{ nm}$ [22]). Then all quantities with the dimension [energy] will be in units of 1 Rydberg ($= 13.605826 \text{ eV} = 21.799118 \times 10^{-19} \text{ J}$) for (i) and in units of 1 Hartree ($= 2 \text{ Rydberg}$) for (ii). Our motivation

for using MKS-units in our formulae is that we feel that most formulae in the literature are a confusing mixture of formulae that make sense regarding dimensions and formulae that are expressions directly programmable on a computer. We will give examples of this below (see below (2.61a)).

Another clarifying remark regarding dimensions we wish to make concerns the use of the word "potential": in the literature it is invariably used instead of "potential energy", although these quantities have different dimension. This usually causes no confusion, because there is only one type of particles, namely, electrons that experience potentials V_0 and having a potential energy eV_0 on account of that. In this work, we will also use the word "potential" as a short hand for "potential energy" and hence all potentials will have dimension [energy] (implying that the real potential has already been multiplied by e , the charge of the electron).

2.3.1 Total energy in direct space

The crystal is defined by giving three basis vectors $t_i (i=1,2,3)$ spanning the unit cell and the positions $t^{(j)}$ of atoms j within the unit cell. Vectors $R = n_1 t_1 + n_2 t_2 + n_3 t_3$ with integers n_1, n_2, n_3 are called Bravais-lattice vectors or just lattice vectors. The crystal volume is denoted by Ω and the crystal is considered to be composed of a large number of concatenated unit cells (with volume Ω_c). The unit cells the crystal is composed of are shifted with respect to each other over lattice vectors.

Following the prescription given in the beginning of section 2.3, the KS-equations for a crystal may be written as:

$$\left\{ -\frac{\hbar^2}{2m} \nabla^2 + \hat{V}_{eff}(r) - E_n(\mathbf{k}) \right\} \psi_{n,\mathbf{k}}(r) = 0, \quad (2.25)$$

$$\hat{V}_{eff}(r) = \sum_{R,j} \hat{V}_j^{pot}(\mathbf{r}-R-t^{(j)}) + \frac{e^2}{4\pi\epsilon_0} \int_{\Omega} \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d^3r' + V_{xc}(r), \quad (2.26)$$

$$n(r) = \frac{\Omega_c}{(2\pi)^3} \int_{1BZ} n_{\mathbf{k}}(r) d^3k, \quad n_{\mathbf{k}}(r) = \sum_{m,occ} |\psi_{m,\mathbf{k}}(r)|^2. \quad (2.27)$$

In (2.25)-(2.26) $E_n(\mathbf{k})$ is a band energy and $\hat{V}_j^{\text{ps.ion}}(\mathbf{r})$ the nonlocal pseudopotential operator for ion j , which is summed over lattice vectors \mathbf{R} and ions j in the unit cell. The second term in (2.26) (Hartree potential) will also be denoted by $V_H(\mathbf{r})$. $V_{xc}(\mathbf{r})$ is called the exchange-correlation (XC) potential and is defined by (cf. (2.11)):

$$V_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} = \left. \frac{d}{dn} [n \epsilon_{xc}(n)] \right|_{n=n(\mathbf{r})} \quad (2.28)$$

Expressions for $\epsilon_{xc}(n)$, the exchange and correlation energy of a homogeneous electron gas, are given in section 2.3.3. In (2.27) the sum over m is over occupied states. It is understood that states that are doubly occupied -as all electron states are with our spin-independent Hamiltonian- must be counted double. We remark that $n(\mathbf{r})$ is a particle density with dimension $[\text{volume}]^{-1}$ and the dimension of $\psi_{m,\mathbf{k}}(\mathbf{r})$ consequently is $[\text{volume}]^{-1/2}$.

We will now give the corresponding expression for the *total energy of the crystal*, E_{total} , which was defined as the difference between the actual total energy of the crystal and the energy of isolated cores (see chapter 1). We introduce the short-hand notation:

$$\sum_{\mathbf{k}} f(\mathbf{k}) = \frac{\Omega_c}{(2\pi)^3} \int_{\text{IBZ}} f(\mathbf{k}) d^3\mathbf{k} \quad (2.29)$$

We then have (cf. (2.14)):

$$E_{\text{total}} = E_{\text{kin}} + E_H + E_{xc} + E_{ec} + E_{cc}, \quad (2.30)$$

where E_{kin} is the kinetic energy of the electrons, E_H the Coulomb electron-electron interaction energy, E_{xc} the exchange and correlation energy of electrons, E_{ec} the interaction energy between electrons and cores, and E_{cc} the Coulomb core-core interaction energy. The addition of the latter term to expression (2.14), which gives the total energy of the electronic system, is necessary, because in our definition of total energy also the interaction between cores is included. We have:

$$E_{kin} = \int_{\Omega} \sum_{n,k} \psi_{n,k}^*(r) \left[-\frac{\hbar^2}{2m} \nabla^2 \right] \psi_{n,k}(r) d^3r, \quad (2.31a)$$

$$E_H = \frac{1}{2} \int_{\Omega} V_H(r) n(r) d^3r, \quad (2.31b)$$

$$E_{xc} = \int_{\Omega} \epsilon_{xc}(n(r)) n(r) d^3r, \quad (2.31c)$$

$$E_{ec} = \int_{\Omega} \sum_{n,k,R,j} \psi_{n,k}^*(r) \hat{V}_j^{ps-ion}(r-R-t^{(j)}) \psi_{n,k}(r) d^3r, \quad (2.31d)$$

$$E_{cc} = \frac{e^2}{8\pi\epsilon_0} \sum'_{R,j,R',j'} \frac{Z_j Z_{j'}}{|R+t^{(j)} - R' - t^{(j')}|}. \quad (2.31e)$$

The prime in (2.31e) excludes the term $R+t^{(j)} = R'+t^{(j')}$ and Z_j is the number of valence electrons of atom j . Expression (2.31e) is only correct for spherically symmetric and non-overlapping cores.

Since we will use norm-conserving ion-core pseudopotentials, the total pseudopotential operator is decomposed in its ℓ -dependent components:

$$\sum_{R,j} \hat{V}_j^{ps-ion}(r-R-t^{(j)}) = \sum_{R,j} \sum_{\ell=0}^{\infty} V_{\ell,j}^{ion}(r-R-t^{(j)}) \mathcal{P}_{\ell}(r,R+t^{(j)}), \quad (2.32)$$

where $\mathcal{P}_{\ell}(r,r')$ is an operator projecting r -dependent functions on eigenfunctions of the angular-momentum operator with quantum number ℓ centred around position r' . It is convenient to split off from the ionic pseudopotential a local (ℓ -independent) part:

$$V_{\ell,j}^{ion}(r-R-t^{(j)}) = V_{loc,j}(|r-R-t^{(j)}|) + \Delta V_{\ell,j}(|r-R-t^{(j)}|). \quad (2.33)$$

$V_{loc,j}(r)$ is chosen such as to contain the Coulomb tail $-Z_j e^2/4\pi\epsilon_0 r$ for $r \rightarrow \infty$, but no singularity for $r = 0$.

2.3.2 Total energy in momentum space

In this section the expressions of the preceding section are Fourier analysed, which will lead to a transformation of a set of differential equations into a set of linear equations, while volume integrals are replaced by summations over reciprocal-lattice vectors [23]. A function $f(\mathbf{r})$ that is periodic, i.e., $f(\mathbf{r}) = f(\mathbf{r} + \mathbf{R})$ with \mathbf{R} any lattice vector, can be expanded in plane waves (PW's) $\exp(i\mathbf{G}\cdot\mathbf{r})$, where \mathbf{G} is a reciprocal-lattice vector. A reciprocal-lattice vector is given by $\mathbf{G} = m_1\mathbf{b}_1 + m_2\mathbf{b}_2 + m_3\mathbf{b}_3$ with integers $m_i (i=1,2,3)$ and basis vectors of the reciprocal lattice $\mathbf{b}_i (i=1,2,3)$, that are related to the basis vectors of the Bravais lattice $\mathbf{t}_i (i=1,2,3)$ through: $\mathbf{b}_i \cdot \mathbf{t}_j = 2\pi\delta_{ij}$ ($i, j = 1,2,3$). So we have:

$$f(\mathbf{r}) = \sum_{\mathbf{G}} f(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}} \quad (2.34)$$

The Fourier components are given by:

$$f(\mathbf{G}) = \frac{1}{\Omega} \int_{\Omega} f(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}} d^3\mathbf{r}. \quad (2.35)$$

Equations (2.34) and (2.35) imply the following identity:

$$\frac{1}{\Omega} \int_{\Omega} e^{i(\mathbf{G}-\mathbf{G}')\cdot\mathbf{r}} d^3\mathbf{r} = \delta_{\mathbf{G},\mathbf{G}'}. \quad (2.36)$$

Because of the periodicity of the integrands in (2.35) and (2.36) Ω may be replaced by Ω_c .

The wave functions are not periodic, but can be chosen such as to obey the Bloch condition:

$$\psi_{n,\mathbf{k}}(\mathbf{r}+\mathbf{R}) = e^{i\mathbf{k}\cdot\mathbf{R}} \psi_{n,\mathbf{k}}(\mathbf{r}). \quad (2.37)$$

A convenient expansion of the wave function in PW's therefore is:

$$\psi_{n,k}(\mathbf{r}) = \sum_{\mathbf{G}} C_{n,k}(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}. \quad (2.38)$$

The functions $n(\mathbf{r})$, $V_H(\mathbf{r})$, $V_{xc}(\mathbf{r})$, and $\epsilon_{xc}(\mathbf{r})$ are periodic and can be expanded as in (2.34) with Fourier components $n(\mathbf{G})$, $V_H(\mathbf{G})$, $V_{xc}(\mathbf{G})$, and $\epsilon_{xc}(\mathbf{G})$, respectively. Note that by definitions (2.34)–(2.35) and (2.38) functions and their Fourier components have the same dimension.

Although one sees that PW's form a very natural basis set for the expansion of periodic functions in crystals, other choices can be made: another popular basis set in combination with pseudopotentials is the set consisting of Linear Combinations of Atomic Orbitals (LCAO's) [24]. The latter choice leads to a much more complicated calculational scheme than the one to be presented in this section, but has the advantage that fewer basis functions are needed. In all-electron calculations (i.e., no pseudopotentials are used) the basis sets used, such as LAPW's (Linearized Augmented Plane Waves [25]) and LMTO's (Linearized Muffin Tin Orbitals [26]), are usually accompanied by extra shape constraints on the functions of interest, whereas the PW-expansion is completely general. However, in an all-electron calculation PW's are not suitable basis functions: approximately 10^6 of them would be needed to describe the strong oscillations in the core region [13]. The pseudopotential-plane-wave method is considered to be best suited for calculations on open structures -i.e., solids with regions of negligible electron density- if the condition is fulfilled that the expansion (2.38) does not need to include so many PW's to become unmanageable. The prototypic example is silicon (Si).

The Fourier analysed version of (2.25) is a set of linear equations for the Fourier components $C_{n,k}(\mathbf{G})$:

$$\left[\frac{\hbar^2}{2m}(\mathbf{k}+\mathbf{G})^2 - E_n(\mathbf{k}) \right] C_{n,k}(\mathbf{G}) + \sum_{\mathbf{G}'} V_{eff}(\mathbf{k}+\mathbf{G}, \mathbf{k}+\mathbf{G}') C_{n,k}(\mathbf{G}') = 0, \quad (2.39)$$

where

$$V_{eff}(\mathbf{k}+\mathbf{G}, \mathbf{k}+\mathbf{G}') = \frac{1}{\Omega} \int_{\Omega} e^{-i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \hat{V}_{eff}(\mathbf{r}) e^{i(\mathbf{k}+\mathbf{G}')\cdot\mathbf{r}} d^3\mathbf{r}. \quad (2.40)$$

$V_{eff}(\mathbf{k}+\mathbf{G}, \mathbf{k}+\mathbf{G}')$ is composed of three terms:

$$V_{eff}(k+G, k+G') = V^{ps-ion}(k+G, k+G') + V_H(G-G') + V_{xc}(G-G'), \quad (2.41)$$

with

$$V^{ps-ion}(k+G, k+G') = \sum_j S_j(G-G') \left[V_{loc,j}(G-G') + \sum_{\ell} \Delta V_{\ell,j}(k+G, k+G') \right], \quad (2.42a)$$

$$V_{loc,j}(G-G') = \frac{1}{\Omega_{at}} \int_{\Omega} e^{-i(G-G') \cdot r} V_{loc,j}(r) d^3r, \quad (2.42b)$$

$$S_j(G-G') = \frac{\Omega_{at}}{\Omega_c} e^{-i(G-G') \cdot t(j)}, \quad (2.42c)$$

$$\Delta V_{\ell,j}(k+G, k+G') = \frac{1}{\Omega_{at}} \int_{\Omega} e^{-i(k+G) \cdot r} \Delta V_{\ell,j}(r) \phi_{\ell}(r, \mathbf{o}) e^{i(k+G') \cdot r} d^3r. \quad (2.42d)$$

Ω_{at} denotes the volume per atom. Since $V_{loc,j}(r)$ depends only on $r = |r|$, $V_{loc,j}(G-G')$ depends only on $q = |q| = |G-G'|$. Explicit expressions for $V_{loc,j}(q)$ and $\Delta V_{\ell,j}(k+G, k+G')$ for tabulated versions of norm-conserving ion-core pseudopotentials are given in section 2.4.

We now proceed to the expression for E_{total} to be obtained if PW-expansions of the various quantities in (2.30) are substituted. In this connection it has to be realized that due to the long-range nature of the Coulomb interaction the terms E_{ec} , E_H , and E_{xc} diverge, i.e., their Fourier terms for $G=0$ are infinite. These infinite terms, however, can be summed to give a finite contribution to the energy per cell as will be shown below. Substituting the PW-expansions in (2.31a) to (2.31d) and using (2.36) we obtain:

$$E_{kin} = \Omega \sum_{n, k, G} |C_{n, k}(G)|^2 \frac{\hbar^2}{2m} (k+G)^2, \quad (2.43a)$$

$$E_H = \frac{1}{2} \Omega \sum_G V_H(G) n^*(G), \quad (2.43b)$$

$$E_{xc} = \Omega \sum_G \epsilon_{xc}(G) n^*(G), \quad (2.43c)$$

$$E_{ec} = \Omega \sum_{\mathbf{G}} n^*(\mathbf{G}) \sum_{\mathbf{j}} S_{\mathbf{j}}(\mathbf{G}) V_{loc, \mathbf{j}}(\mathbf{G}) +$$

$$\Omega \sum_{\mathbf{n}, \mathbf{k}, \mathbf{G}, \mathbf{G}'} C_{\mathbf{n}, \mathbf{k}}^*(\mathbf{G}) C_{\mathbf{n}, \mathbf{k}}(\mathbf{G}') \sum_{\mathbf{j}} S_{\mathbf{j}}(\mathbf{G}-\mathbf{G}') \sum_{\ell} \Delta V_{\ell, \mathbf{j}}(\mathbf{k}+\mathbf{G}, \mathbf{k}+\mathbf{G}'). \quad (2.43d)$$

where

$$n(\mathbf{G}) = \sum_{\mathbf{n}, \mathbf{k}, \mathbf{G}''} C_{\mathbf{n}, \mathbf{k}}^*(\mathbf{G}'') C_{\mathbf{n}, \mathbf{k}}(\mathbf{G}''+\mathbf{G}). \quad (2.44)$$

We will now address the question what contribution to the total energy remains if the individually divergent $\mathbf{G}=0$ -terms in E_{ec} , E_H , and E_{cc} are summed. Note that the second term in the right-hand side (RHS) of (2.43d) is not involved in this discussion, since the local part of the pseudopotential was chosen to contain the Coulomb tail and so the nonlocal parts are short ranged and cause no divergencies. Our discussion is based on a similar discussion in ref.[27]. The first term in the RHS of (2.43d) will be denoted by E_{ec}^L . We start by splitting off from E_{ec}^L the Coulomb tail, so that a finite part FP remains:

$$E_{ec}^L = FP + E_{ec}^{Coul}, \quad (2.45a)$$

$$E_{ec}^{Coul} = \frac{e^2}{4\pi\epsilon_0} \int_{\Omega} n(\mathbf{r}) \sum_{\mathbf{R}, \mathbf{j}} \frac{-Z_{\mathbf{j}}}{|\mathbf{r}-\mathbf{R}-\mathbf{t}^{(\mathbf{j})}|} d^3\mathbf{r}. \quad (2.45b)$$

The sum $E_{ec}^{Coul} + E_H + E_{cc}$ should now be recognized as the electrostatic energy E_{es} of a lattice of point ions \mathbf{j} with charge $-eZ_{\mathbf{j}}$ in a periodic neutralizing background distribution $n(\mathbf{r})$ of electrons. The neutrality of this system implies:

$$\int_{\Omega_c} n(\mathbf{r}) d^3\mathbf{r} = \sum_{\mathbf{j}} Z_{\mathbf{j}}. \quad (2.46)$$

We further have from electrostatic theory [28] that:

$$E_{es} = \frac{1}{2}\epsilon_0 \int_{\Omega} |\mathbf{E}(\mathbf{r})|^2 d^3\mathbf{r}. \quad (2.47)$$

where $E(\mathbf{r})$ is the electric field at position \mathbf{r} in Ω . Note that since the whole system is periodic, $E(\mathbf{r})$ is also periodic and has a PW-expansion as in (2.34). We split up this system into two systems that are each periodic as well: system I is a periodic arrangement of positive point ions in a uniform neutralizing (therefore negatively charged) background with density $n_o^- = (1/\Omega_c) \sum_{(j)} Z_j$ and system II is a periodic distribution $n(\mathbf{r})$ of electrons in a uniform neutralizing (therefore positively charged) background with density $n_o^+ = n_o^-$. The electrostatic energy of system I can be expressed as:

$$E'_{cc} = \frac{e^2}{4\pi\epsilon_0} \left[\frac{1}{2} \sum_{\mathbf{R}, \mathbf{J}, \mathbf{R}', \mathbf{J}'} \frac{Z_j Z_{j'}}{|\mathbf{R}+\mathbf{t}^{(j)} - \mathbf{R}'-\mathbf{t}^{(j')}|} - \sum_{\mathbf{R}, \mathbf{J}} \int_{\Omega} \frac{Z_j n_o^-}{|\mathbf{r}-\mathbf{R}-\mathbf{t}^{(j)}|} d^3\mathbf{r} \right. \\ \left. + \frac{1}{2} \int_{\Omega} \int_{\Omega} \frac{(n_o^-)^2}{|\mathbf{r}-\mathbf{r}'|} d^3\mathbf{r} d^3\mathbf{r}' \right]. \quad (2.48)$$

E'_{cc} is related to the well-known Ewald energy γ_{Ewald} , which can be calculated by well-defined and fastly converging procedures [27,29]:

$$\gamma_{\text{Ewald}} = E'_{cc} \frac{\Omega a t}{\Omega}. \quad (2.49)$$

The Ewald energy thus is the interaction energy of system I per atom.

The electric field $E(\mathbf{r})$ of the original system is the sum of the electric fields $E_I(\mathbf{r})$ and $E_{II}(\mathbf{r})$ of the two subsystems. We may substitute a PW-expansion for E_I and E_{II} to obtain for E_{es} :

$$E_{es} = E'_{cc} + \frac{1}{2}\epsilon_0\Omega \sum_{\mathbf{G}} |E_{II}(\mathbf{G})|^2 + \epsilon_0\Omega \sum_{\mathbf{G}} E_I(\mathbf{G})^* \cdot E_{II}(\mathbf{G}). \quad (2.50)$$

We now are allowed to demand that $E_I(\mathbf{G})$ and $E_{II}(\mathbf{G})$ both equal zero for $\mathbf{G}=0$, which is equivalent to demanding the cell average of $E_I(\mathbf{r})$ and $E_{II}(\mathbf{r})$ to equal zero. If this were not so, we would -because of the periodicity of both systems- build up an electric field over the whole of system I and II. So we conclude that we may replace $E_{ec}^{\text{coul}} + E_H + E_{cc}$ by $E_{ec}^{\text{coul}'} + E_H' + E_{cc}'$, where E_{cc}' is given by (2.48) and the primes in $E_{ec}^{\text{coul}'}$ and E_H' denote that the $\mathbf{G}=0$ -term is put equal to zero.

To complete the discussion of the $\mathbf{G}=0$ -term in E_{total} , we now deal

with the $G=0$ -term in FP in (2.45a). From (2.45) and (2.43d) we have that this $G=0$ -term equals (substitute Fourier expansions of the r -dependent functions in (2.45)):

$$\Omega n^*(G=0) \sum_j \frac{\Omega_{at}}{\Omega_c} \frac{1}{\Omega_{at}} \int_{\Omega} V_{loc,j}(r) d^3r + \frac{e^2}{4\pi\epsilon_0} \frac{\Omega}{\Omega_c} n^*(G=0) \sum_j \int_{\Omega} \frac{Z_j}{r} d^3r.$$

From (2.46) we have $n^*(G=0) = n(G=0) = (1/\Omega_c) \sum_j Z_j$. So the $G=0$ -term of FP is given by:

$$\frac{\Omega}{\Omega_c} \sum_j \alpha_j \sum_j Z_j, \quad (2.51)$$

where

$$\alpha_j = \frac{1}{\Omega_c} \int_{\Omega} (V_{loc,j}(r) + \frac{Z_j e^2}{4\pi\epsilon_0 r}) d^3r. \quad (2.52)$$

Taking all this into account the total energy per unit cell E_{tot} is given by a sum of individually finite terms (note that from now on E_{kin} , E_{xc} , E_H^1 , and E_{ec}^1 have their previous value divided by N , the number of cells in the crystal):

$$E_{tot} = E_{kin} + E_H^1 + E_{xc} + E_{ec}^1 + \gamma_{Ewald} \frac{\Omega_c}{\Omega_{at}} + \sum_j \alpha_j \sum_j Z_j, \quad (2.53)$$

$$E_{kin} = \Omega_c \sum_{n,k,G} |C_{n,k}(G)|^2 \frac{\hbar^2}{2m} (k+G)^2, \quad (2.54a)$$

$$E_H^1 = \frac{1}{2} \Omega_c \sum_G V_H(G) n^*(G), \quad (2.54b)$$

$$E_{xc} = \Omega_c \sum_G \epsilon_{xc}(G) n^*(G), \quad (2.54c)$$

$$E_{ec}^1 = \Omega_c \sum_G n^*(G) \sum_j S_j(G) V_{loc,j}(G) +$$

$$\Omega_c \sum_{n,k,G,G'} C_{n,k}^*(G) C_{n,k}(G') \sum_j S_j(G-G') \sum_{\ell} \Delta V_{\ell,j}(k+G, k+G'), \quad (2.54d)$$

where the prime in the summation over \mathbf{G} denotes that the $\mathbf{G}=0$ -term is left out. We remark that our second term in the RHS of (2.54d) differs slightly from the expressions in refs.[27,30] -note in this connection that the structure factor $S_j(\mathbf{G})$ is generally complex and obeys $S_j^*(\mathbf{G}) = S_j(-\mathbf{G})$, but agrees with the expression in ref.[23] if one notices the different (unconventional) definition of the structure factor used there (see (9) of ref.[23]).

The last term in (2.54d) is the most difficult one to compute, since it contains a double summation over \mathbf{G} vectors and, furthermore, the Fourier components of the nonlocal part of the pseudopotential, which -as we will see in section 2.4- are complicated expressions. However, using (2.39) we can derive an alternative expression for E_{tot} , which does not include the term E_{ec} . Formula (2.39) can be rewritten as:

$$E_n(\mathbf{k})C_{n,\mathbf{k}}(\mathbf{G}) = \frac{\hbar^2}{2m}(\mathbf{k}+\mathbf{G})^2 C_{n,\mathbf{k}}(\mathbf{G}) + \sum'_{\mathbf{G}'} V_{eff}(\mathbf{k}+\mathbf{G},\mathbf{k}+\mathbf{G}') C_{n,\mathbf{k}}(\mathbf{G}'), \quad (2.55)$$

where the prime in the summation denotes that $V_{loc,j}(\mathbf{G}=\mathbf{G}')$ and $V_H(\mathbf{G}=\mathbf{G}')$ are set equal to zero; for the solutions $C_{n,\mathbf{k}}(\mathbf{G})$ this makes no difference, only the eigenvalues $E_n(\mathbf{k})$ are shifted by a constant. This convention is consistent with our analysis of the $\mathbf{G}=0$ -term in E_{tot} . Multiplying by $C_{n,\mathbf{k}}^*(\mathbf{G})$ and summing over \mathbf{G} gives:

$$E_n(\mathbf{k}) = \Omega_c \sum_{\mathbf{G}} |C_{n,\mathbf{k}}(\mathbf{G})|^2 \frac{\hbar^2}{2m}(\mathbf{k}+\mathbf{G})^2 + \Omega_c \sum_{\mathbf{G},\mathbf{G}'} V_{eff}(\mathbf{k}+\mathbf{G},\mathbf{k}+\mathbf{G}') C_{n,\mathbf{k}}(\mathbf{G}') C_{n,\mathbf{k}}^*(\mathbf{G}). \quad (2.56)$$

We may now sum over n and \mathbf{k} and use (2.41)-(2.42), (2.44), and (2.54) to obtain:

$$\sum_{n,\mathbf{k}} E_n(\mathbf{k}) = E_{kin} + 2E_H + \Omega_c \sum_{\mathbf{G}} V_{xc}(\mathbf{G}) n^*(\mathbf{G}) + E_{ec}. \quad (2.57)$$

Using this in (2.53) we find:

$$E_{\text{tot}} = \sum_{n,\mathbf{k}} E_n(\mathbf{k}) - E_H^1 + \Delta E_{xc} + \gamma_{\text{Ewald}} \frac{\Omega_c}{\Omega_{\text{at}}} + \sum_j \alpha_j \sum_j Z_j, \quad (2.58)$$

where

$$\Delta E_{xc} = \Omega_c \sum_{\mathbf{G}} \left[\epsilon_{xc}(\mathbf{G}) - V_{xc}(\mathbf{G}) \right] n^*(\mathbf{G}). \quad (2.59)$$

2.3.3 Self-consistent solution of Kohn-Sham equations in momentum space

If we have self-consistent solutions $C_{n,\mathbf{k}}(\mathbf{G})$ and $E_n(\mathbf{k})$ of the KS-equations in momentum space to our disposal, we are in the position to calculate E_{tot} from (2.58) if we know how to calculate $V_H(\mathbf{G})$, $V_{xc}(\mathbf{G})$, and $\epsilon_{xc}(\mathbf{G})$ from these self-consistent solutions. We proceed by (2.44) to calculate $n(\mathbf{G})$; since $V_H(\mathbf{r})$ and $V_{xc}(\mathbf{r})$ are all functionals of the density $n(\mathbf{r})$, it is straightforward in principle to calculate $V_H(\mathbf{G})$ and $V_{xc}(\mathbf{G})$. Here we will give details of this procedure.

Suppose the $n(\mathbf{G})$ are all known. The Hartree potential $V_H(\mathbf{r})$ is related to $n(\mathbf{r})$ by Poisson's equation:

$$\Delta V_H(\mathbf{r}) = - \frac{e^2 n(\mathbf{r})}{\epsilon_0}. \quad (2.60)$$

Remember that $V_H(\mathbf{r})$ has dimension [energy] and $n(\mathbf{r})$ dimension [volume]⁻¹. Substituting both PW-expansions the relation between $V_H(\mathbf{G})$ and $n(\mathbf{G})$ is a simple linear one:

$$V_H(\mathbf{G}) = \frac{e^2 n(\mathbf{G})}{\epsilon_0 |\mathbf{G}|^2}. \quad (2.61a)$$

For $\mathbf{G}=0$ $V_H(\mathbf{G})$ is put equal to zero, as was argued already in section 2.3.2. In the literature [31,23,32] we may find for $V_H(\mathbf{G})$:

- (1) $4\pi e^2 \rho(\mathbf{G}) / |\mathbf{G}|^2$, where apparently CGS-units are used and $\rho(\mathbf{G})$ has dimension [volume]⁻¹,
- (2) $8\pi \rho(\mathbf{G}) / |\mathbf{G}|^2$, where apparently Rydberg atomic units are used and $\rho(\mathbf{G})$ has dimension [volume]⁻¹, and
- (3) $4\pi e^2 \rho(\mathbf{G}) / (\Omega_c |\mathbf{G}|^2)$, where again CGS-units are used and $\rho(\mathbf{G})$ of (1) and (2) is now replaced by $\rho(\mathbf{G}) / \Omega_c$. Apparently $\rho(\mathbf{G})$ in this formula is

dimensionless.

When programming (2.61a) and going over to Rydberg atomic units, we -like in (3) above- introduce a dimensionless $n'(G)$, which is related to the distribution of electrons in a unit cell by:

$$n'(G) = \int_{\Omega_c} n(r) e^{-iG \cdot r} d^3r.$$

This results in the formula:

$$V_H(G) = \frac{8\pi}{\Omega_c} \frac{n'(G)}{|G|^2}, \quad (2.61b)$$

where $|G|$ should be expressed in a.u. and $V_H(G)$ is in Rydberg.

For the XC-potential and -energy the situation is more complicated. We first give some commonly used expressions for $\epsilon_{xc}(n)$, the exchange and correlation energy of a homogeneous interacting electron gas. These are given in terms of the Wigner-Seitz radius r_s , related to n by:

$$\frac{1}{n} = \frac{4\pi}{3} r_s^3. \quad (2.62)$$

$\epsilon_{xc}(n)$ can be seen as sum of an exchange part $\epsilon_x(n)$ and a correlation part $\epsilon_c(n)$. The Kohn-Sham form for $\epsilon_x(n)$ is given by [5]:

$$\epsilon_x(n) = - \frac{3e^2}{16\pi\epsilon_0} \left[\frac{9}{4\pi^2} \right]^{\frac{1}{3}} \frac{1}{r_s} = - \frac{e^2}{4\pi\epsilon_0} \frac{0.4582}{r_s}. \quad (2.63)$$

The Wigner interpolation form for $\epsilon_c(n)$ interpolates between known high- and low-density limits [33,34] for the correlation energy:

$$\epsilon_c(n) = - \frac{e^2}{4\pi\epsilon_0} \frac{0.44}{7.8a_0 + r_s}. \quad (2.64)$$

Since valence-charge densities generally have r_s around $2a_0$, it is clear from (2.63)-(2.64) that the correlation contribution to $\epsilon_{xc}(n)$ is very small.

Another expression -less popular nowadays- is given by the Slater $X\alpha$ -approximation to $\epsilon_{xc}(n)$:

$$\epsilon_{xc}(n) = -\frac{9\alpha e^2}{32\pi\epsilon_0} \left[\frac{9}{4\pi^2} \right]^{\frac{1}{3}} \frac{1}{r_s} = -\frac{\alpha e^2}{4\pi\epsilon_0} \frac{0.6872}{r_s}. \quad (2.65)$$

Here, α is an adjustable parameter to be chosen between $2/3$ (no correlation) and 1 in order to add some correlation to (2.63). For instance, for the r_s belonging to the average valence-charge density in Si, $r_s = 2a_0$, the sum of (2.63) and (2.64) intersects (2.65) for $\alpha = 0.8$.

The most popular form for $\epsilon_c(n)$ is the one parametrized by Perdew and Zunger [35] from the quantummechanical Monte Carlo simulations of the homogeneous interacting electron gas by Ceperley and Alder [36]. This parametrized form also gives the correct high- and low-density limits and is given by:

$$\epsilon_c(n) = \begin{cases} -\frac{e^2}{4\pi\epsilon_0 a_0} \frac{0.1423}{1 + 1.0529 \sqrt{R_s} + 0.3334 R_s} & (R_s \geq 1) \\ -\frac{e^2}{4\pi\epsilon_0 a_0} \left[0.0480 - 0.0311 \ln(R_s) + 0.0116 R_s - \right. \\ \left. 0.0020 R_s \ln(R_s) \right] & (R_s < 1) \end{cases} \quad (2.66)$$

Here $R_s = r_s/a_0$. This is the spin-unpolarized form; a spin-polarized form also exists [35]. Note that this $\epsilon_c(n)$ intersects the Wigner form for $R_s = 2.1$, i.e., again for about the average valence-charge density in silicon.

We have used (2.63)-(2.64) in all our calculations. In that case the XC-potential is found via:

$$V_{xc}(n) = \frac{d}{dn} \left[n \epsilon_{xc}(n) \right] = \epsilon_{xc} - \frac{r_s}{3} \frac{d\epsilon_{xc}}{dr_s}. \quad (2.67)$$

We find for (2.63)-(2.64):

$$V_{xc}(n) = -\frac{e^2}{4\pi\epsilon_0} \left[\left[\frac{9}{4\pi^2} \right]^{\frac{1}{3}} \frac{1}{r_s} + \frac{0.44}{7.8a_0 + r_s} \left[1 + \frac{r_s}{3(7.8a_0 + r_s)} \right] \right]. \quad (2.68)$$

Using the LDA we have: $V_{xc}(\mathbf{r}) = V_{xc}(n(\mathbf{r}))$ and $\epsilon_{xc}(\mathbf{r}) = \epsilon_{xc}(n(\mathbf{r}))$. So exchange and correlation energy and potential are now expressed as local functionals of $n(\mathbf{r})$. The problem is that we do not have a functional relationship between their Fourier components like in

(2.61) for the Hartree potential. Although it is a diversion, the most efficient way has proven to be to calculate $n(\mathbf{r})$ on a fine mesh of points \mathbf{r} in the unit cell from the known $n(\mathbf{G})$, calculate $V_{xc}(\mathbf{r})$ and $\epsilon_{xc}(\mathbf{r})$ via (2.68) and (2.63)-(2.64), and subsequently make a Fourier inversion to obtain $V_{xc}(\mathbf{G})$ and $\epsilon_{xc}(\mathbf{G})$; this is discussed in a little more detail in section 2.5.

Now we can calculate E_{tot} from the self-consistent $C_{n,\mathbf{k}}(\mathbf{G})$ and $E_n(\mathbf{k})$ of (2.39), but to arrive at this self-consistent solution some steps still have to be discussed. The first item is the potential $V_{eff}(\mathbf{k}+\mathbf{G}, \mathbf{k}+\mathbf{G}')$ with which one starts. One possibility is to do a calculation with the form factors from the Empirical-Pseudopotential Method (EPM) first [37] and to use the density $n(\mathbf{G})$ from this calculation to calculate a first guess for the screening potential $V_{scr}(\mathbf{G})$:

$$V_{scr}(\mathbf{G}) = V_{xc}(\mathbf{G}) + V_H(\mathbf{G}). \quad (2.69)$$

In the EPM -which is not self-consistent- a set of equations of the form (2.39) is solved with a local potential of the form:

$$V(\mathbf{G}-\mathbf{G}') = \sum_j S_j(\mathbf{G}-\mathbf{G}') V_j^{ps}(|\mathbf{G}-\mathbf{G}'|), \quad (2.70)$$

which replaces the complete $V_{eff}(\mathbf{k}+\mathbf{G}, \mathbf{k}+\mathbf{G}')$ in (2.39). In (2.70) the $V_j^{ps}(q)$ are called form factors and these are chosen unequal to zero for just a few values of q lower than some q_{max} . These form factors are then considered as adjustable to obtain a band structure $E_n(\mathbf{k})$ in agreement with experiment. In such a calculation moderate numbers of PW's are usually sufficient. Other possibilities are to use a superposition of atomic charge densities, from which a V_{scr} may be calculated, or to screen a (local or nonlocal) ion-pseudopotential by a model dielectric function [38].

So far we have not discussed one of the main problems of this calculational scheme: it seems that (2.39) will have to be solved for very many \mathbf{k} points in order to approximate the integrations over 1BZ occurring frequently, e.g., in the calculation (2.44) of $n(\mathbf{G})$. In practice one can -especially for semiconductors and insulators- suffice with calculations for just a few special points in 1BZ. This is explained in chapter 3, where a modified form for (2.44) is derived

to be used in connection with special points.

Now we have all the equipment to determine self-consistent solutions of the KS-equations and to calculate the total energy: with some screening potential to start with ($V_{\text{scr}}^{(0)}(\mathbf{G}-\mathbf{G}')$), which is added to the ion-core pseudopotentials, we must solve the set of equations (2.39) for a few \mathbf{k} points. By the formula to be derived in chapter 3 we calculate the $n(\mathbf{G})$ from which a new screening potential $V_{\text{scr}}^{(1)}(\mathbf{G}-\mathbf{G}')$ is calculated as described above. $V_{\text{scr}}^{(1)}(\mathbf{G}-\mathbf{G}')$ is used as input for the next cycle and all steps are repeated. This whole procedure is repeated until self-consistency is achieved, i.e., $V_{\text{scr}}^{(n)}$ and $V_{\text{scr}}^{(n-1)}$ are equal to within some prescribed accuracy. In practical calculations the input and output screening potentials are mixed to accelerate convergence:

$$V_{\text{in}}^{(n+1)} = \alpha V_{\text{out}}^{(n)} + (1 - \alpha)V_{\text{in}}^{(n)}. \quad (2.71)$$

The mixing coefficient α is usually chosen between $\frac{1}{2}$ and 1 and may depend on the iteration number n . A more sophisticated convergence-acceleration scheme is the Broyden scheme introduced in solid-state calculations by Bendt and Zunger [39].

The whole calculational scheme is applicable, in principle, to any crystal. In practice, the applicability is limited to those crystals for which certain assumptions that were made hold, e.g., we should not construct a crystal in which cores overlap: this would make the pseudopotential approximation (see section 2.2) unreliable and contradicts the assumption made in connection with (2.31e). One is also limited by the question whether sensible results can be obtained with technical approximations that must be made to make calculations feasible; these technical approximations are discussed in section 2.5.

2.4 Matrix elements of norm-conserving pseudopotentials

In this section explicit matrix elements between PW's of tabulated versions of norm-conserving ion-core pseudopotentials are given. These are of almost direct use in (2.41)-(2.42). Two such tables exist: in ref.[40] and in ref.[14]. The first gives parameterized potentials in r -space for the elements Si, C, Ga, and P, whereas the second covers the whole periodic table from H to Pu.

In ref.[40] one of the ℓ -components (ℓ_o) of the pseudopotential is considered as the local part. In the notation of (2.33), the pseudopotentials are given by (we drop the subscript j because all formulae refer to only one -arbitrary- element):

$$V_{\text{loc}}(r) = V_{\ell_o}(r) = -\frac{e^2}{4\pi\epsilon_o} \left[\frac{Z}{r} \operatorname{erf}(\sqrt{\alpha_o}r) - \sum_{i=1}^5 c_i(\ell_o) e^{-\alpha_i(\ell_o)r^2} \right], \quad (2.72a)$$

$$\Delta V_{\ell_o}(r) = \frac{e^2}{4\pi\epsilon_o} (1 - \delta_{\ell, \ell_o}) \sum_{i=1}^5 c_i(\ell) e^{-\alpha_i(\ell)r^2}. \quad (2.72b)$$

The matrix elements of the local part between PW's $\exp(i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r})$ and $\exp(i(\mathbf{k}+\mathbf{G}')\cdot\mathbf{r})$, according to (2.42b), are easily shown to equal:

$$V_{\text{loc}}(\mathbf{G}-\mathbf{G}') = V_{\ell_o}(\mathbf{G}-\mathbf{G}') = -\frac{e^2}{\epsilon_o\Omega_{\text{at}}} \left[\frac{Z}{q^2} e^{-q^2/4\alpha_o} - \frac{\sqrt{\pi}}{4} \sum_{i=1}^5 c_i(\ell_o) \alpha_i^{-3/2}(\ell_o) e^{-q^2/4\alpha_i(\ell_o)} \right]. \quad (2.73)$$

where $q = |\mathbf{q}| = |\mathbf{G}-\mathbf{G}'|$. The quantity defined by (2.52), giving the finite part of the average of the local part of the pseudopotential, is for the form (2.72a) given by:

$$\alpha = \frac{e^2}{\epsilon_o\Omega_c} \left[\frac{Z}{4\alpha_o} + \frac{\sqrt{\pi}}{4} \sum_{i=1}^5 c_i(\ell_o) \alpha_i^{-3/2}(\ell_o) \right]. \quad (2.74)$$

To obtain the matrix elements of the nonlocal part one expands the PW's in Legendre polynomials P_ℓ and spherical Bessel functions of the first kind j_ℓ [41]:

$$e^{i\mathbf{K}\cdot\mathbf{r}} = \sum_{\ell=0}^{\infty} i^{\ell} j_{\ell}(Kr) (2\ell+1) P_{\ell}(\cos\beta), \quad (2.75)$$

where β is the angle between \mathbf{K} and \mathbf{r} and $K = |\mathbf{K}|$. Using the addition theorem and the orthonormality of spherical harmonics on the unit sphere [42], we obtain from (2.42d):

$$\Delta V_{\ell}(\mathbf{k}+\mathbf{G}, \mathbf{k}+\mathbf{G}') = \frac{4\pi}{\Omega_{\text{at}}} (2\ell+1) P_{\ell}(\cos\gamma) \int_0^{\infty} j_{\ell}(Kr) j_{\ell}(K'r) \Delta V_{\ell}(r) r^2 dr, \quad (2.76)$$

where $K = |\mathbf{K}| = |\mathbf{k}+\mathbf{G}|$, $K' = |\mathbf{K}'| = |\mathbf{k}+\mathbf{G}'|$ and $\cos\gamma = (\mathbf{K}\cdot\mathbf{K}')/KK'$ and $\cos\gamma = 1$ if K or K' equals zero. For functions $\Delta V_{\ell}(r)$ of the form $r^{2n} \exp(-\alpha r^2)$ ($\alpha > 0$, $n = 0, 1, 2, \dots$), the integral can be performed analytically [43], and for the form (2.72b) we obtain (We note that a factor of $(2\ell+1)$ is missing in eq. (19) of ref. [44]):

$$\Delta V_{\ell}(\mathbf{k}+\mathbf{G}, \mathbf{k}+\mathbf{G}') = \frac{e^2}{\epsilon_0 \Omega_{\text{at}}} (1 - \delta_{\ell, 1, \ell_0}) (2\ell+1) P_{\ell}(\cos\gamma) \sum_{i=1}^5 c_i(\ell) F_{\ell, \alpha_i}(\ell)(K, K'), \quad (2.77a)$$

with

$$F_{\ell, \alpha}(K, K') = \frac{\sqrt{\pi}}{4\alpha^{3/2}} h_{\ell} \left[\frac{KK'}{2\alpha} \right] e^{-\frac{K^2 + K'^2}{4\alpha}}, \quad (2.77b)$$

where $h_{\ell}(z)$ is a modified spherical Bessel function of the first kind, which for $\ell = 0, \pm 1, \pm 2, \dots$ only involves the calculation of functions not more complicated than one hyperbolic sine and/or cosine. In the pseudopotential approach only values $\ell = 0, 1, 2, \dots$ occur. We note that the values of the parameters α_0 , $\alpha_i(\ell)$, $c_i(\ell)$ ($\ell=0, 1, 2$, $i=1 \dots 5$) as given in table I and II in ref. [40] may only be used when Hartree a.u. are used. In the formulae (2.72)-(2.74) and (2.77) -where MKS units are used- the occurring parameters α_0 , α_i and c_i are those from the tables divided by the MKS-value of a_0^2 , a_0^2 , and a_0 , respectively.

In ref.[14] the pseudopotentials are parametrized as follows:

$$V_{\text{loc}}(r) = V_{\text{core}}(r) = -\frac{Ze^2}{4\pi\epsilon_0 r} \sum_{i=1}^2 c_i^{\text{core}} \text{erf}(\sqrt{\alpha_i^{\text{core}}} r), \quad (2.78a)$$

$$\Delta V_\ell(r) = \frac{e^2}{4\pi\epsilon_0} \sum_{i=1}^3 (A_i(\ell) + r^2 A_{i+3}(\ell)) e^{-\alpha_i(\ell)r^2}. \quad (2.78b)$$

Parameters are such that $c_1^{\text{core}} + c_2^{\text{core}} = 1$, so that V_{loc} has the Coulomb tail. The corresponding matrix elements between PW's are:

$$V_{\text{loc}}(\mathbf{G}-\mathbf{G}') = -\frac{Ze^2}{\epsilon_0 \Omega_{\text{at}}} \sum_{i=1}^2 c_i^{\text{core}} \frac{e^{-q^2/4\alpha_i^{\text{core}}}}{q^2}, \quad (2.79)$$

$$\Delta V_\ell(\mathbf{k}+\mathbf{G}, \mathbf{k}+\mathbf{G}') = \frac{e^2(2\ell+1)}{\epsilon_0 \Omega_{\text{at}}} P_\ell(\cos\gamma) \times \sum_{i=1}^3 \left[A_i(\ell) - A_{i+3}(\ell) \frac{d}{d\alpha_i}(\ell) \right] F_{\ell, \alpha_i}(\ell)(K, K'), \quad (2.80)$$

where all quantities are defined as before and

$$\frac{d}{d\alpha} \left[F_{\ell, \alpha}(K, K') \right] = \left[-\frac{3}{2\alpha} + \frac{K^2 + K'^2}{4\alpha^2} + \frac{\ell+1}{\alpha} \right] F_{\ell, \alpha}(K, K') - \frac{KK'}{2\alpha^2} F_{\ell-1, \alpha}(K, K') \quad (2.81)$$

In practice, it is better to use $V_{\text{core}}(r) + \Delta V_\ell(r)$ for one selected value of ℓ as local potential $V_{\text{loc}}(r)$. In this way both the local and the remaining nonlocal part of the pseudopotential are weaker [30]. From the above formulae one straightforwardly finds new expressions for local and nonlocal parts (2.79)-(2.80). Parameters c_i^{core} , α_i^{core} ($i=1,2$), $\alpha_j(\ell)$ ($j=1\dots6$, $\ell=0,1,2,3$) are given in table IV in ref.[14], whereas the $A_j(\ell)$ ($j=1\dots6$, $\ell=0,1,2,3$) must be obtained from tabulated parameters $C_j(\ell)$ by a numerically rather unstable but straightforward procedure. This instability, however, has not much consequence for the accuracy of $\Delta V_\ell(r)$ in (2.78b) [45]. Again we note that the values of all parameters in the table are to be used in connection with Hartree a.u. only and must be rescaled for use in other unit systems.

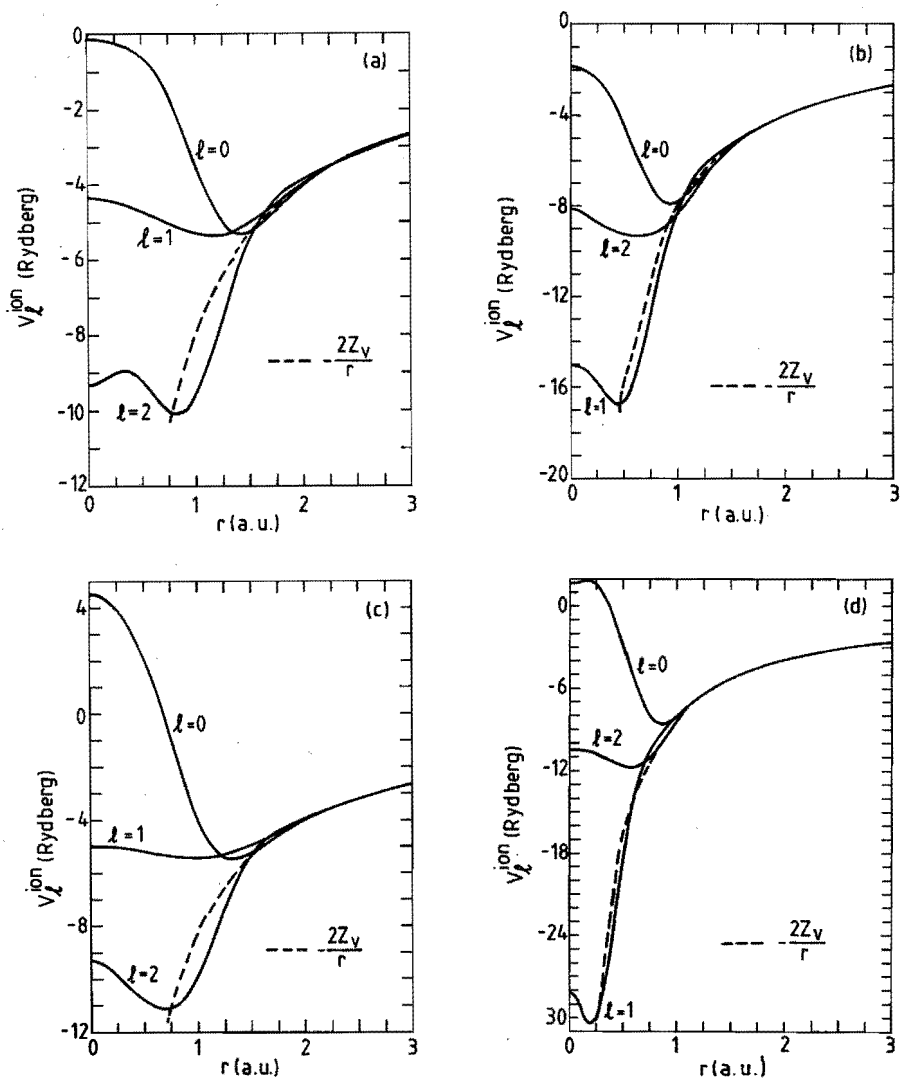


Fig. 2.1 Norm-conserving ion-core pseudopotentials:
 (a) silicon, ref.[40], (b) carbon, ref.[40],
 (c) silicon, ref.[14], (d) carbon, ref.[14].

In fig. 2.1 we show examples of both sets of pseudopotentials for Si and C as a function of r . It is observed, in accordance with the remark below (2.33), that the pseudopotentials V_l^{ion} have the Coulomb tail for large r and are finite for $r = 0$. The $l=1$ -component for C is

very deep, because the core of carbon does not contain p electrons, so there is practically no cancellation of the strong Coulomb potential (see (2.20)). As V_ρ^{ion} for $\ell=1$ deviates strongly from the $\ell=0$ and $\ell=2$ components, it is not preferable to take the $\ell=1$ component as the local potential, since this would lead to large (strong) $\Delta V_\rho(r)$ for $\ell=0,2$.

Figure 2.1 also suggests that the potentials from ref.[40] are probably more suited to be used with a PW-basis than those of ref.[14], because the deeper potentials require larger G - G' to describe the well for small r . This suggestion is confirmed by Nielsen and Martin [30]. We have used the pseudopotentials of ref.[40] in all our calculations reported in this work with $\ell_0=0$ (see (2.73) and (2.74)). As a final remark we note that fig. 2.1(a) deviates from the picture in ref.[40] for $\ell=2$. We speculate that the latter was obtained from the not yet parametrized (numerical) potential or with the parametrized form with non-rounded parameters, whereas we used the parameters from table I in ref.[40].

2.5 Cutoff parameters

The formalism in section 2.3 was derived assuming an expansion in an infinite number of PW's of all functions in r -space and consequently infinite summations over reciprocal-lattice vectors G in all formulae. In actual calculations these infinities of course are replaced by finite expansions and summations. This introduces what may be called technical or numerical approximations, in contrast to conceptual approximations like the one-electron picture or the LDA. These technical approximations are represented by so-called cutoffs. As was mentioned in section 2.3 also the integrations over 1BZ, occurring frequently, must be approximated, and so must the inverse Fourier transform to calculate $\epsilon_{xc}(r)$ and $V_{xc}(r)$.

In this section we will discuss the various cutoffs of our calculational scheme and their importance. The two most crucial cutoffs are the number of PW's into which the wave function is expanded in (2.38) and a cutoff related to the number of k points used to integrate over 1BZ. We will discuss two other cutoffs as well.

(A) The number of PW's N_{PW} in the expansion (2.38) is determined by a kinetic-energy cutoff E_{PW} : PW's are included in the expansion only if their kinetic energy $(\hbar^2/2m)(k+G)^2$ does not exceed E_{PW} . Put differently: a sphere of radius $G_{max} = (2mE_{PW})^{1/2}/\hbar$ and centre k is constructed in reciprocal space and all reciprocal-lattice vectors G within or on the sphere are included in the expansion. Clearly N_{PW} depends on k , though not very heavily, since k is a reduced wave vector; so in changing k the sphere is never shifted over a distance larger than a reciprocal-lattice vector. By this construction wave functions $\psi_{n,k}(r)$ are described with the same resolution in r -space for every k : variations of $\psi_{n,k}(r)$ on a length scale larger than $2\pi/G_{max}$ are correctly described. One could avoid the k -dependence of N_{PW} by including all G 's with $(\hbar^2/2m)G^2 \leq E_{PW}$, but in that case $\psi_{n,k}(r)$ may lose its symmetry properties. A reasonable estimate for N_{PW} is obtained by dividing the volume of a sphere of radius G_{max} by the volume of 1BZ of a Bravais lattice with unit-cell volume Ω_c (volume of 1BZ then is $(2\pi)^3/\Omega_c$):

$$N_{PW} \simeq \frac{\Omega_c (2mE_{PW})^{3/2}}{6\pi^2 \hbar^3} \quad (2.82)$$

This shows that N_{PW} scales linearly with Ω_c if the same resolution in r -space (same E_{PW}) is desired.

(B) As a consequence of (2.44), $n(r)$ will be expanded in many more PW's than the wave functions are, namely PW's $\exp(iG \cdot r)$, where G' is any difference vector of wave vectors $k+G$ included in the expansion (2.38). This implies that $n(G)$ will come out zero for very large G ($|G| > 2G_{max}$), whereas $n(G)$ for $|G|$ closely below $2G_{max}$ will be relatively less accurate, since a too little number of $k+G$ -pairs have this large difference. This is however not a problem, since for smooth wave functions the $C_{n,k}(G)$ fall off with increasing $|G|$. Therefore these relatively less accurate $n(G)$ are small anyway and unimportant in the calculation of $n(r)$ or in the summations in E_{tot} . The large and therefore important $n(G)$ with small $|G|$ are more accurate because many $k+G$ -pairs (especially the ones with small $|k+G|$, which have large $C_{n,k}(G)$) have this small difference. We may include in all summations over G only those with $|G| < \alpha G_{max}$, with $1.5 < \alpha < 2$. It is only for these G then that $n(G)$, $V_H(G)$, $V_{xc}(G)$, and $\epsilon_{xc}(G)$ need be calculated. This

may be done by introducing another cutoff N_{st} , being the number of stars of G vectors taken into account. A star is a set of G vectors transforming into each other under the operations of the point group of the crystal under consideration. N_{st} is a less important cutoff, since convergence of the summations in E_{t_0t} can be obtained to within sufficient accuracy without substantial increase in computational effort. To be on the safe side, one may choose N_{st} such that $\alpha=2$. Note that this choice increases the number of terms in a summation over G in E_{t_0t} with roughly a factor of 8 compared to the number of terms in the expansion (2.38) ($N_{PV} \sim G^3_{max}$).

(C) The approximation of the integral over 1BZ by a weighted sum over special points in the irreducible part of 1BZ can be represented by a cutoff limiting the number of special points used. This cutoff is discussed in chapter 3, where the theory of special points is presented together with examples and applications. Here, we suffice with the remark that with one or two special points already very accurate *energy differences* can be obtained, e.g., in the situation where the lattice parameters are varied, but E_{PV} is kept constant. If we retain the same crystal structure, this simulates putting a crystal under pressure, while describing the wave functions with constant resolution (see chapter 6).

(D) As final cutoff parameter, we discuss the number of grid points in the unit cell of the crystal for which $n(\mathbf{r})$ and consequently $V_{xc}(\mathbf{r})$ and $\epsilon_{xc}(\mathbf{r})$ (see section 2.3) must be calculated before Fourier inverting to obtain $V_{xc}(G)$ and $\epsilon_{xc}(G)$. Both the calculation of $n(\mathbf{r})$ and the two Fourier inversions are performed by standard 3-dimensional discrete complex Fast Fourier Transform (FFT) [46]. To obtain Fourier components for reciprocal-lattice vectors G , $n(\mathbf{r})$ must be calculated for $\mathbf{r} = \xi\mathbf{t}_1 + \eta\mathbf{t}_2 + \zeta\mathbf{t}_3$, where $\{\mathbf{t}_i\}_{i=1,2,3}$ is the set of basis vectors spanning the unit cell in \mathbf{r} -space and $\xi=i/N_1$ ($i=1\dots N_1$), $\eta=j/N_2$ ($j=1\dots N_2$), $\zeta=k/N_3$ ($k=1\dots N_3$). The number of grid points is $N_1N_2N_3$. The numbers N_n ($n=1,2,3$) must be chosen at least so large that $N_n > \max\{m_n\}$, where $\max\{m_n\}$ is the maximum component along basis vector \mathbf{b}_n of the reciprocal lattice that occurs in the set of G for which the Fourier component is desired. This is an absolute minimal requirement in order not to get complete nonsense; for instance, if $N_n = \max\{m_n\}$ ($n=1,2,3$), the Fourier component for $G = \sum_{(n=1)}^{(3)} N_n \mathbf{b}_n$ would be equal to the $G=0$ Fourier component in a FFT, whereas the former should

in fact be orders of magnitude smaller for the smooth functions concerned. A very safe rule of thumb is to choose N_n equal to $2\max\{m_n\}$ and to disregard all Fourier components for G with components along basis vector b_n between $\max\{m_n\}$ and $2\max\{m_n\}$. The determination of satisfactory N_n can also be done by checking in how far symmetry properties of the Fourier components are satisfied: the crystal symmetry of functions in r -space causes some Fourier components to be equal to zero and some of the (generally complex) Fourier components to be real. The N_n can be adjusted to satisfy the requirements for these Fourier components to within some desired accuracy. We remark that this is a much more direct way of determining suitable N_n than the one that was used in ref.[47], where the effect of enlarging N_n on $E_{t,t}$ was studied.

Just as $N_{s,t}$ the N_n are relatively unimportant cutoffs because FFT is such an efficient technique that convergence with respect to increasing N_n to within sufficient accuracy can be achieved with very modest computational effort.

CHAPTER 3

SPECIAL POINTS IN THE FIRST BRILLOUIN ZONE

In the calculational scheme described in the preceding chapter, there is a need, among others, to accurately determine the valence-electron charge density $\rho(\mathbf{r})$ (see (2.27) and (2.29)):

$$\rho(\mathbf{r}) = e n(\mathbf{r}) = e \sum_{\mathbf{k}} \sum_{m, \text{occ}} |\psi_{m, \mathbf{k}}(\mathbf{r})|^2, \quad (3.1)$$

and the band-structure energy $\sum_{(n, \mathbf{k})} E_n(\mathbf{k})$ (see (2.58)). In these expressions the sum over \mathbf{k} represents in fact an average over 1BZ of the \mathbf{k} -dependent function. Note that in both cases the \mathbf{k} -dependent function is periodic in reciprocal space, i.e., $g(\mathbf{k}) = g(\mathbf{k}+\mathbf{G})$ for any reciprocal-lattice vector \mathbf{G} . If we were to evaluate the necessary 3-dimensional integrals by calculating the integrand at a fine mesh of \mathbf{k} points in 1BZ, we would have to solve the set of equations (2.39) for a very large number of \mathbf{k} points. From a computational point of view this should be considered as undesirable. It was an important discovery by Baldereschi around 1973 that the average is already very reasonably approximated by calculating the integrand at just a single \mathbf{k} point, the so-called mean-value point [48]. Based on this idea Chadi and Cohen gave a systematic way of choosing successively larger sets of so-called special points, resulting in successively better approximations for the average [49]. Other systematic approaches were later proposed by Monkhorst and Pack as well as Evarestov and Smirnov [50,51,52]. A review of the use of special points in solid state physics can be found in ref.[52].

The use of special points has proven to be essential to make the scheme of chapter 2 manageable for calculations on solids. Especially for semiconductors and insulators special-point approximations are very suitable. For these solids we deal with completely filled energy bands, implying that for every \mathbf{k} point the same number of states is occupied in the ground state. Therefore, the summation over occupied states in (2.58) and (3.1) can be performed before the average is evaluated. In the case of metals, for instance, not all bands are completely filled and we have to determine the Fermi level first to

find out how many bands are occupied for an arbitrary \mathbf{k} point. The values of $E_n(\mathbf{k})$ for just a few \mathbf{k} points do not suffice to determine the Fermi level. So the application to metals is not impossible, but less practical, since it requires sets with a large number of special points.

In section 3.1 we present the general theory of special points and derive a formula showing how to calculate for general crystal symmetry the Fourier components of the valence-charge density, $\rho(\mathbf{G})$, from the solutions $C_{n,\mathbf{k}}(\mathbf{G})$ of (2.39) for just a few \mathbf{k} points only. This formula replaces (2.44) in actual calculations. In section 3.2 the Monkhorst-Pack (MP) scheme for the generation of special-point sets is described; this scheme is formulated in a way suited for automatic generation. We will give very efficient automatic integration schemes for a subset of the MP-sets applicable in the respective cases of face-centred cubic (fcc) crystals with point group O_h , T_d or O and hexagonal crystals with point group D_{6h} , C_{6v} , D_6 or D_{3h} . For the notion of crystallographic point groups, for which we employ the Schönflies notation [53], we refer to chapter 4. We note here that in most of the pioneering work on special points the fact that the application of special point sets not only depends on the Bravais lattice, but also on the point group of the specific crystal, was not mentioned (see section 3.1). In section 3.3 the degree of accuracy of successively larger MP-sets is discussed in the application to the integration of energy bands. In section 3.4 we present prescriptions to select special-point sets for structurally different crystals the total energies of which one wishes to compare. A good procedure for an equivalent choice in both cases is essential for the accurate determination of energy differences between such crystals.

3.1 General theory and application to charge-density calculations

Any function $g(\mathbf{k})$ that is periodic in reciprocal space can be expressed as a Fourier series (cf. (2.34)):

$$g(\mathbf{k}) = \sum_{n=0}^{\infty} g_n e^{i\mathbf{k} \cdot \mathbf{R}_n}, \quad (3.2)$$

where the R_n are lattice vectors ($R_0 = 0$). We introduce the function:

$$f(\mathbf{k}) = \langle g(\mathbf{k}) \rangle_{\mathcal{G}_L} \equiv \frac{1}{N_\alpha} \sum_{\alpha} g(\alpha \mathbf{k}), \quad (3.3)$$

where the sum runs over all operations α in the lattice point group \mathcal{G}_L and where N_α is the number of elements in \mathcal{G}_L . By definition, \mathcal{G}_L is the largest crystallographic point group that leaves the Bravais lattice under consideration invariant; it therefore contains at least the identity E and the inversion I . The function $f(\mathbf{k})$ is clearly invariant under all operations in \mathcal{G}_L , i.e., $f(\mathbf{k}) = f(\alpha \mathbf{k})$, and is said to possess the complete symmetry of the lattice. The construction defined by (3.3) and leading to $f(\mathbf{k})$ is called "symmetrisation of $g(\mathbf{k})$ with respect to the point group \mathcal{G}_L ".

The function $f(\mathbf{k})$ can be expanded as follows:

$$f(\mathbf{k}) = f_0 + \sum_{m=1}^{\infty} f_m A_m(\mathbf{k}), \quad (3.4)$$

where

$$A_m(\mathbf{k}) = \sum_{\substack{|\mathbf{R}_m| \\ = C_m}} e^{i\mathbf{k} \cdot \mathbf{R}_m} \quad (m=1,2,\dots), \quad (3.5)$$

and where $f_0 = g_0$. The sum in (3.5) is over stars of lattice vectors, i.e., lattice vectors that transform into each other under operations of the lattice point group. The stars are ordered in such a way that $0 < C_m \leq C_{m+1}$. The $A_m(\mathbf{k})$ are called symmetrised plane waves, since they are sums of plane waves $\exp(i\mathbf{k} \cdot \mathbf{R}_m)$ having the complete symmetry of the lattice. From (3.5) it is clear that the $A_m(\mathbf{k})$ are periodic. Another important property is that their average over 1BZ vanishes:

$$\frac{\Omega_c}{(2\pi)^3} \int_{1BZ} A_m(\mathbf{k}) d^3\mathbf{k} = 0 \quad (m=1,2,\dots). \quad (3.6)$$

Equation (3.6) implies that the average of $f(\mathbf{k})$ over 1BZ is just f_0 .

Suppose a point \mathbf{k}_0 would exist such that:

$$A_m(\mathbf{k}_0) = 0 \quad (m=1,2,\dots,M), \quad (3.7)$$

for $M \rightarrow \infty$. In that case it is clear from (3.4) that the average f_0 equals $f(\mathbf{k}_0)$. There is no lattice for which such a point exists, but we may attempt to find the point that satisfies (3.7) for as large an M as possible. For a smooth function of \mathbf{k} the coefficients f_m in (3.4) fall off with increasing m and we find a reasonable approximation to f_0 with one point only. Chadi and Cohen have generalised this idea to obtain sets of N_{sp} special points \mathbf{k}_s with weighting factors w_s ($s = 1 \dots N_{sp}$) that satisfy:

$$\sum_s w_s A_m(\mathbf{k}_s) = 0 \quad (m = 1, 2 \dots M-1), \quad (3.8)$$

$$\sum_s w_s = 1. \quad (3.9)$$

where the sum is over special points ($s = 1 \dots N_{sp}$). In their scheme, it is possible to make M in (3.8) as large as one desires by carefully selecting more and more special points. The special points can always be chosen to lie in the irreducible wedge of 1BZ (IBZ), because the $A_m(\mathbf{k})$ have the complete symmetry of the lattice. The average over 1BZ, f_0 , according to (3.4), (3.8), and (3.9) can now be written as:

$$f_0 = \sum_s w_s f(\mathbf{k}_s) - \sum_{m=M}^{\infty} \sum_s w_s f_m A_m(\mathbf{k}_s). \quad (3.10)$$

The first coefficient f_m appearing in (3.10) is f_M . Therefore, for smooth functions $f(\mathbf{k})$ the approximation:

$$f_0 = \sum_s w_s f(\mathbf{k}_s). \quad (3.11)$$

improves, if the value of M is increased. This can be achieved by increasing the number of special points N_{sp} . The task is to find sets that lead to large values of M with N_{sp} values as small as possible. The efficiency of a special-point set is often denoted by the ratio $(M-1)/N_{sp}$.

A function $A_m(\mathbf{k})$ for which (3.8) holds is called: a symmetrised plane wave that is integrated exactly by the special-point set. We

introduce R_M as the length of the lattice vector R_M associated with the first symmetrised plane wave that is not integrated exactly by a special-point set. In general, many $A_m(k)$ with $m > M$ are also integrated exactly by such a special-point set. R_M is a slightly better quantity than M to characterize a special-point set, since in case of more stars with the same C_m the ordering is not unique. In the sense of section 2.5, R_M can be called: the cutoff parameter that is associated with the special-points approximation (3.11). At first sight, it might seem that R_M is a cutoff similar to E_{PV} (or G_{max} : see section 2.5). There is, however, an important difference: plane waves with kinetic energy larger than E_{PV} are really excluded from the expansion of the wave function, whereas an infinite number of $A_m(k)$ with $m > M$ may still be included in the expansion of the function to be averaged, because they are integrated exactly anyhow.

If we let all operations α of the lattice point group operate on the k_s in IBZ, we find a set of k points in 1BZ. The number of actually different points in the latter set we call N_{BZ} . This number is connected to the symmetry and thereby to the weights w_s of the special points: a special point of high symmetry is mapped onto fewer points in 1BZ and because of that has lower weight. Put differently: the weight of a special point k_s is proportional to the number of different k points obtained (note that two k points both situated at a BZ face and separated by a reciprocal-lattice vector are considered to be equal).

In order to characterize a given special-point set, each of the three above introduced parameters, N_{sp} , R_M , and N_{BZ} , has its own relevance. Though these parameters are strongly related, each of them emphasizes a specific aspect of the special-point set under consideration. Generally speaking, the integral or average of a function over 1BZ is better approximated the larger these quantities are.

We now come to the question of how to apply the special-point sets defined above to the determination of the average over 1BZ of periodic functions $g(k)$ in crystals. As is easily deduced from (3.2), this average equals g_0 and can be approximated by (use (3.3) and (3.11)):

$$g_0 = \sum_s w_s \frac{1}{N} \sum_{\alpha \in \mathcal{G}_L} g(\alpha k_s). \quad (3.12)$$

If no direct relation exists between $g(\mathbf{k})$ and $g(\alpha\mathbf{k})$ for any $\alpha \in \mathcal{G}_L$, one has to calculate $g(\alpha\mathbf{k}_S)$ for all N_{BZ} different \mathbf{k} points $\alpha\mathbf{k}_S$. The \mathbf{k} -dependent functions $E_n(\mathbf{k})$ and $\rho_{\mathbf{k}}(\mathbf{r})$, to which we wish to apply the special-point set procedure, however, are clearly different in this respect:

- (1) The function $E_n(\mathbf{k})$ is invariant under operations of the crystallographic point group \mathcal{G}_C . In general, \mathcal{G}_C is a subgroup of \mathcal{G}_L . Furthermore, because of invariance under time-reversal, the relation $E_n(\mathbf{k}) = E_n(-\mathbf{k})$ holds, even if the inversion I is not contained in \mathcal{G}_C .
- (2) The function $\rho_{\mathbf{k}}(\mathbf{r})$ also has inversion symmetry because of time-reversal symmetry. Furthermore, $\rho_{\mathbf{k}}(\mathbf{r})$ and $\rho_{\beta\mathbf{k}}(\mathbf{r})$ are related if $\beta \in \mathcal{G}_C$ (This relation will be derived below).

It is helpful to define the point group \mathcal{G}'_C , which coincides with \mathcal{G}_C if \mathcal{G}_C contains I and is the *direct-product group* $\mathcal{G}_C \times I$ if \mathcal{G}_C does not contain I (We define $\mathcal{G}_C \times I$ as follows: $\beta \in \mathcal{G}_C \times I$ only if $\beta = \alpha_1 \alpha_2$ with $\alpha_1 \in \mathcal{G}_C$ and α_2 is either E or I). From the above it is clear that, because of the above quoted symmetry properties (1) and (2), $E_n(\mathbf{k})$ and $\rho_{\mathbf{k}}(\mathbf{r})$ only have to be evaluated at points \mathbf{k}_S , that are inequivalent with respect to \mathcal{G}'_C . The weighting factor w_S , belonging to \mathbf{k}_S , is proportional to the number of different \mathbf{k} points in 1BZ obtained if the operations of \mathcal{G}'_C act on \mathbf{k}_S . The w_S again sum to unity. If \mathcal{G}'_C is identical to \mathcal{G}_L the special-point set $\{\mathbf{k}_S; w_S\}$ is recovered.

From now on we will consider only cases in which \mathcal{G}'_C is identical to \mathcal{G}_L as these cases are most common (and apply in all our examples). From the above discussion it is clear that special-point sets for other cases can straightforwardly be constructed as well. We note that the restriction to cases for which \mathcal{G}'_C equals \mathcal{G}_L still allows that the same special-point set $\{\mathbf{k}_S; w_S\}$, derived using \mathcal{G}_L , may be applied to crystal structures with different crystallographic point groups. More precisely, the same special-point set may be used for all crystallographic point groups \mathcal{G}_C which are either equal to \mathcal{G}_L or which fulfill the relation $\mathcal{G}_C \times I = \mathcal{G}_L$. For example, special-point sets for the fcc Bravais lattice ($\mathcal{G}_L = O_h$) may be applied to the diamond structure, with $\mathcal{G}_C = O_h$, the zincblende structure, with $\mathcal{G}_C = T_d$, and structures with $\mathcal{G}_C = O$, since $O_h = T_d \times I$ and $O_h = O \times I$. Similarly, special-point sets for hexagonal Bravais lattices ($\mathcal{G}_L = D_{6h}$) may be applied to crystal structures with $\mathcal{G}_C = D_{6h}, C_{6v}, D_6$ or D_{3h} . For both lattices the reason is the invariance of $E_n(\mathbf{k})$ and $\rho_{\mathbf{k}}(\mathbf{r})$ with respect to the

inversion $\mathbf{k} \rightarrow -\mathbf{k}$.

Since, for the cases to be considered from now on, $E_n(\mathbf{k})$ has the complete symmetry of the lattice, the average involved in (2.58) is, according to (3.11), approximated by:

$$\sum_{n, \mathbf{k}} E_n(\mathbf{k}) = \sum_s w_s \sum_{n, \text{occ}} E_n(\mathbf{k}_s). \quad (3.13)$$

Equation (3.13) shows that, much to our advantage, eqs. (2.39) have to be solved for the special points \mathbf{k}_s only. The situation is a bit more complicated for the calculation of $\rho(\mathbf{r})$, since $\rho_{\mathbf{k}}(\mathbf{r})$ is generally not invariant under operations of the crystal point group. However, $\rho_{\mathbf{k}}(\mathbf{r})$ and $\rho_{\beta\mathbf{k}}(\mathbf{r})$ are related for all $\beta \in \mathcal{G}_c$ and we furthermore have, $\rho_{\mathbf{k}}(\mathbf{r}) = \rho_{-\mathbf{k}}(\mathbf{r})$. This leads to the possibility to calculate Fourier components of $\rho(\mathbf{r})$ directly from the solutions $C_{n, \mathbf{k}}(\mathbf{G})$ of eqs. (2.39) for the special points $\mathbf{k}=\mathbf{k}_s$ only. We will now derive the formula that makes this explicit and that will replace the so far only formally useful formula (2.44).

The function $\rho_{\mathbf{k}}(\mathbf{r})$ is periodic in \mathbf{r} -space and may therefore be expanded in a Fourier series:

$$\rho_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} \rho_{\mathbf{k}}(\mathbf{G}) e^{i\mathbf{G} \cdot \mathbf{r}}. \quad (3.14a)$$

From eqs. (2.27) and (2.44) we infer:

$$\rho_{\mathbf{k}}(\mathbf{G}) = e \sum_{n, \text{occ}} \sum_{\mathbf{G}'} C_{n, \mathbf{k}}^*(\mathbf{G}') C_{n, \mathbf{k}}(\mathbf{G}' + \mathbf{G}). \quad (3.14b)$$

Equation (3.12) applied to $g(\mathbf{k}) = \rho_{\mathbf{k}}(\mathbf{r})$ leads to:

$$\rho(\mathbf{r}) = \sum_s w_s \frac{1}{N_\alpha} \sum_{\alpha \in \mathcal{G}_l} \rho_{\alpha\mathbf{k}_s}(\mathbf{r}). \quad (3.15a)$$

Because of time-reversal symmetry, we may always write:

$$\rho(\mathbf{r}) = \sum_s w_s \frac{1}{N_\beta} \sum_{\beta \in \mathcal{G}_c} \rho_{\beta\mathbf{k}_s}(\mathbf{r}), \quad (3.15b)$$

where the second summation is over N_β operations β in \mathcal{G}_c . Realizing

that \mathcal{G}_c is a group, so that a sum over β is also a sum over β^{-1} , and substituting Fourier series for $\rho(\mathbf{r})$ and $\rho_{\beta^{-1}\mathbf{k}_s}(\mathbf{r})$, we obtain via (3.15b):

$$\rho(\mathbf{G}) = \sum_s w_s \frac{1}{N_\beta} \sum_{\beta} \rho_{\beta^{-1}\mathbf{k}_s}(\mathbf{G}). \quad (3.15c)$$

We now proceed as follows to express $\rho_{\beta^{-1}\mathbf{k}}(\mathbf{G})$ in terms of $\rho_{\mathbf{k}}(\mathbf{G}')$: Let \mathcal{R}_β be an operation of the space group with point-group part β , associated nonprimitive translation $\tau(\beta)$, and no primitive translation (these notions are defined in chapter 4 together with a discussion of the effect of an operation of the space group on general \mathbf{r} -dependent functions). Its effect on a pseudo-Bloch-function $\psi_{\mathbf{n},\mathbf{k}}(\mathbf{r})$ is:

$$\mathcal{R}_\beta \psi_{\mathbf{n},\mathbf{k}}(\mathbf{r}) = \psi_{\mathbf{n},\mathbf{k}}(\beta\mathbf{r} + \tau(\beta)). \quad (3.16)$$

Since $\beta\mathbf{R}$ is a lattice vector if \mathbf{R} is one, the Bloch character of $\mathcal{R}_\beta \psi_{\mathbf{n},\mathbf{k}}(\mathbf{r})$ implies:

$$\begin{aligned} \psi_{\mathbf{n},\mathbf{k}}(\beta(\mathbf{r}+\mathbf{R}) + \tau(\beta)) &= e^{i\mathbf{k}\cdot\beta\mathbf{R}} \psi_{\mathbf{n},\mathbf{k}}(\beta\mathbf{r} + \tau(\beta)) \\ &= e^{i\beta^{-1}\mathbf{k}\cdot\mathbf{R}} \psi_{\mathbf{n},\mathbf{k}}(\beta\mathbf{r} + \tau(\beta)). \end{aligned} \quad (3.17)$$

Equation (3.17) shows that $\psi_{\mathbf{n},\mathbf{k}}(\beta\mathbf{r} + \tau(\beta))$ is a Bloch function with reduced wave vector $\beta^{-1}\mathbf{k}$ or:

$$\mathcal{R}_\beta \psi_{\mathbf{n},\mathbf{k}}(\mathbf{r}) = \psi'_{\mathbf{n},\beta^{-1}\mathbf{k}}(\mathbf{r}). \quad (3.18)$$

Substituting Fourier expansions for $\psi_{\mathbf{n},\mathbf{k}}(\mathbf{r})$ and $\psi'_{\mathbf{n},\beta^{-1}\mathbf{k}}(\mathbf{r})$ and using the fact that a sum over reciprocal-lattice vectors \mathbf{G} is also a sum over $\beta^{-1}\mathbf{G}$ for any β^{-1} in \mathcal{G}_c , we obtain from (3.18):

$$C_{\mathbf{n},\mathbf{k}}(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G})\cdot\tau(\beta)} = C_{\mathbf{n},\beta^{-1}\mathbf{k}}(\beta^{-1}\mathbf{G}). \quad (3.19)$$

Using this, we have from (3.14b):

$$\begin{aligned}
\rho_{\beta^{-1}\mathbf{k}}(\mathbf{G}) &= e \sum_{\mathbf{n}, \text{occ}} \sum_{\mathbf{G}'} C_{\mathbf{n}, \beta^{-1}\mathbf{k}}^*(\beta^{-1}\mathbf{G}') C_{\mathbf{n}, \beta^{-1}\mathbf{k}}(\beta^{-1}\mathbf{G}' + \mathbf{G}) \\
&= e \sum_{\mathbf{n}, \text{occ}} e^{i\beta\mathbf{G} \cdot \boldsymbol{\tau}(\beta)} \sum_{\mathbf{G}'} C_{\mathbf{n}, \mathbf{k}}^*(\mathbf{G}') C_{\mathbf{n}, \mathbf{k}}(\mathbf{G}' + \beta\mathbf{G}) \\
&= \rho_{\mathbf{k}}(\beta\mathbf{G}) e^{i\beta\mathbf{G} \cdot \boldsymbol{\tau}(\beta)}. \tag{3.20}
\end{aligned}$$

Using (3.20) in (3.15c), we obtain:

$$\rho(\mathbf{G}) = \sum_s w_s \frac{1}{N_\beta} \sum_{\beta} e^{i\beta\mathbf{G} \cdot \boldsymbol{\tau}(\beta)} \rho_{\mathbf{k}_s}(\beta\mathbf{G}). \tag{3.21}$$

With (3.14b) we finally obtain the desired expression:

$$\rho(\mathbf{G}) = \frac{e}{N_\beta} \sum_s w_s \sum_{\beta} e^{i\beta\mathbf{G} \cdot \boldsymbol{\tau}(\beta)} \sum_{\mathbf{n}, \text{occ}} \sum_{\mathbf{G}'} C_{\mathbf{n}, \mathbf{k}_s}^*(\mathbf{G}') C_{\mathbf{n}, \mathbf{k}_s}(\mathbf{G}' + \beta\mathbf{G}). \tag{3.22}$$

An expression like (3.22) was given before [47], but it contained at least one (printing) error and no derivation was given; this combined with the importance of (3.22) in actual calculations motivated the detailed derivation given above. Equation (3.21) shows that for space groups with all $\boldsymbol{\tau}(\beta) = 0$ (symmorphic space groups) the symmetrisation with respect to the point group \mathcal{G}_t reduces to taking the arithmetical mean over a star of reciprocal-lattice vectors.

3.2 Description and computerization of the Monkhorst-Pack scheme

A few schemes to generate special points as defined in section 3.1 have been proposed. Of these the most popular have been the scheme of Chadi and Cohen [49] and the scheme of Monkhorst and Pack [50] with extensions and generalisations [50-52,54]. The MP-scheme is more systematic and therefore easier to computerize. In this section we will describe this scheme.

The general procedure is to start with a set of \mathbf{k} points:

$$\mathbf{k} = u_p \mathbf{b}_1 + u_r \mathbf{b}_2 + u_s \mathbf{b}_3, \tag{3.23}$$

where $\mathbf{b}_1, \mathbf{b}_2$, and \mathbf{b}_3 are the basis vectors of the reciprocal lattice and the components u_p, u_r , and u_s along these basis vectors have the form (q is a positive integer):

$$u = (2j-q-1)/2q \quad (j=1,2\dots q). \quad (3.24)$$

We define \mathbf{R}_n to be the lattice vector with integer components n_1, n_2 , and n_3 along the basis vectors $\mathbf{t}_1, \mathbf{t}_2$, and \mathbf{t}_3 of the direct lattice, respectively. It can be shown that a sum over all q^3 points defined by (3.23) and (3.24) of the plane wave $\exp(i\mathbf{k}\cdot\mathbf{R}_n)$ equals zero for all \mathbf{R}_n for which holds:

$$|n_i| < q \quad \text{for } i = 1, 2 \text{ or } 3. \quad (3.25)$$

The set of q^3 points may be reduced to points lying in IBZ using operations of the lattice point group. The resulting set is called a set of N_{sp} special points \mathbf{k}_s with weighting factors w_s proportional to the number of points in IBZ they are equivalent with. The special-point set $\{\mathbf{k}_s; w_s\}$ also sums $\exp(i\mathbf{k}\cdot\mathbf{R}_n)$ to zero for all \mathbf{R}_n given above.

This completes the description of the MP-scheme applicable to any Bravais lattice. However, it appears to be possible for special cases to generate more efficient special-point sets, i.e., with larger R_M and/or smaller N_{sp} , by modifying or generalising the scheme described above. In the following subsections we give such generation schemes for the special cases of fcc and hexagonal Bravais lattices. We do not employ the generalisation proposed in ref.[51], that implies the addition of an arbitrary wave vector \mathbf{k}_0 to (3.23), which may be varied to optimize the efficiency.

3.2.1 MP-sets for face-centred cubic lattices

We start by choosing basis vectors for the reciprocal lattice of a fcc lattice : $\mathbf{b}_1 = (2\pi/a_c)(-\mathbf{e}_x + \mathbf{e}_y + \mathbf{e}_z)$, $\mathbf{b}_2 = (2\pi/a_c)(\mathbf{e}_x - \mathbf{e}_y + \mathbf{e}_z)$, and $\mathbf{b}_3 = (2\pi/a_c)(\mathbf{e}_x + \mathbf{e}_y - \mathbf{e}_z)$, where a_c is the lattice parameter of a cubic lattice (see fig. 3.1(a)) and $\mathbf{e}_x, \mathbf{e}_y$, and \mathbf{e}_z are unit vectors in three Cartesian directions. With respect to these unit vectors we

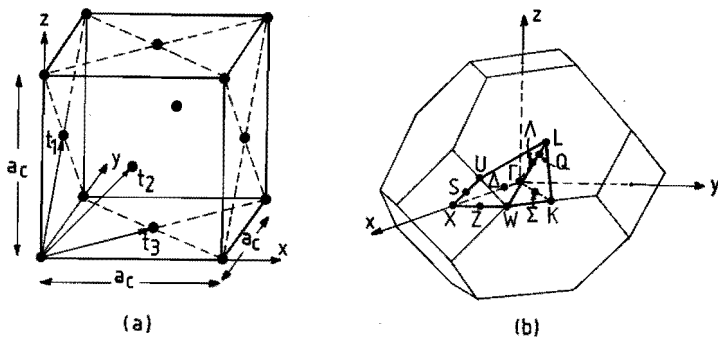


Fig. 3.1 (a) The primitive unit cell of a fcc Bravais lattice is spanned by three basis vectors: $t_1 = \frac{1}{2}a_c(e_y + e_z)$, $t_2 = \frac{1}{2}a_c(e_x + e_z)$, $t_3 = \frac{1}{2}a_c(e_x + e_y)$, where a_c is the lattice parameter of a cubic lattice. (b) First Brillouin zone of a fcc Bravais lattice with symmetry points in the irreducible part of the zone, following the notation of Bouckaert et al. [55].

write a k point as:

$$k = \frac{2\pi}{a_c} (k_x e_x + k_y e_y + k_z e_z) . \quad (3.26)$$

The modified MP-scheme for fcc lattices is based on the observation that half of the q^3 points defined by (3.24) for a simple cubic (sc) lattice with lattice parameter $\frac{1}{2}a_c$ lie in 1BZ of a fcc lattice with lattice parameter a_c . For this particular sc lattice the basis vectors of the reciprocal lattice are: $d_1 = (4\pi/a_c)e_x$, $d_2 = (4\pi/a_c)e_y$, and $d_3 = (4\pi/a_c)e_z$. The MP-points (3.24) are then given by (cfr. (3.26)):

$$k_x, k_y, k_z = (2j - q - 1)/q \quad (j = 1, 2, \dots, q). \quad (3.27)$$

In the appendix of ref.[50], it is shown that the points given by (3.27) that moreover lie in 1BZ of a fcc lattice with lattice parameter a_c form suitable points to integrate k -dependent functions in fcc lattices. The form of the MP-points given by (3.27) is convenient to immediately select the points lying in 1BZ of the fcc lattice, since 1BZ is determined by the conditions (see fig. 3.1(b)):

$$0 \leq k_z \leq k_y \leq k_x \leq 1, \quad (3.28a)$$

$$k_x + k_y + k_z \leq 3/2. \quad (3.28b)$$

The equations (3.27)-(3.28) generate MP-sets of special points k_s for fcc lattices. The accompanying weighting factors w_s are simply determined by the conditions that they (i) are inversely proportional to the number of operations in \mathcal{G}_L that leave k_s invariant and (ii) sum up to one.

As noted in ref.[50], it is advantageous to restrict q in (3.27) to even values. If now $q/2$ is odd, the scheme may generate different points in IBZ that are nevertheless equivalent, i.e., they transform into each other under operations of the point group accompanied by translations over reciprocal-lattice vectors. This can occur for points on that face of IBZ for which the equality in (3.28b) holds (LKUW-face; see fig. 3.1(b)). For instance, for $q=10$ both the points $(2\pi/a_c)(0.9,0.3,0.3)$ and $(2\pi/a_c)(0.7,0.7,0.1)$ are generated in IBZ, satisfying (3.27)-(3.28), but closer analysis teaches that these points are equivalent. Of these equivalent points only one is to be included in the special-point set, with w_s determined as explained before. The generation of inequivalent points in IBZ is completely unambiguous if the additional condition $k_y \leq 1/2$ is imposed on points on the LUKW-face.

If $q/2$ is even (or: q is 4-fold), no points are generated on the LUKW-face as may be checked from (3.27)-(3.28). Furthermore, k_x, k_y , and k_z do not take on values 0, $1/2$ or 1. This results in the occurrence of only three possible weighting factors w :

$$(1) w = 12/N \quad \text{if } k_x \neq k_y \text{ and } k_y \neq k_z,$$

$$(2) w = 6/N \quad \text{if } k_x = k_y \text{ or } k_y = k_z,$$

$$(3) w = 2/N \quad \text{if } k_x = k_y \text{ and } k_y = k_z.$$

Here $N = (\frac{1}{2}q)^3$. So if q is taken to be 4-fold, eqs. (3.27)-(3.28) generate sets of special points the weighting factors of which are immediately obtained by checking whether the point belongs to class (1), (2) or (3). All this leads to a very efficient computerization of special-point approximations to Brillouin-zone integrations. The scheme for $q/2$ odd can of course also be computerized, but in a more cumbersome fashion. We will use the computerization scheme for q

4-fold to study the convergence of special-points approximations with respect to increasing q (see section 3.3).

The important properties of MP-sets for fcc lattices can also be expressed in terms of the parameter q . It can be shown that:

$$N_{sp}(q) = \begin{cases} \frac{1}{96} q(q+2)(q+4) & q/2 \text{ even} \\ \frac{1}{96} (q+2)(q^2+4q+12) & q/2 \text{ odd} \end{cases} \quad (3.29a)$$

Since half of the points given by (3.27) are situated in 1BZ of the fcc lattice, we obtain:

$$N_{BZ}(q) = \frac{1}{2}q^3. \quad (3.29b)$$

One may prove for the cutoff R_M :

$$R_M(q) = \frac{1}{2}qa_c. \quad (3.29c)$$

From (3.29) we observe that for larger q more $A_m(\mathbf{k})$ are integrated exactly (larger R_M) and reciprocal space is sampled better (larger N_{BZ}) at the cost, however, of calculating the integrand at more \mathbf{k} points (larger N_{sp}).

3.2.2 MP-sets for hexagonal lattices

In this section the MP-scheme modified to treat hexagonal lattices is described [54]. For hexagonal lattices, we choose as basis vectors for the reciprocal lattice: $\mathbf{b}_1' = (2\pi/a)(2/3\sqrt{3}\mathbf{e}_x)$, $\mathbf{b}_2' = (2\pi/a)(1/3\sqrt{3}\mathbf{e}_x + \mathbf{e}_y)$, $\mathbf{b}_3' = (2\pi/c)\mathbf{e}_z$, with a and c the lattice parameters of a hexagonal lattice (see fig. 3.2(a)). Since in general a and c will be such that the length of \mathbf{b}_3' differs from the length of \mathbf{b}_1' and \mathbf{b}_2' , it is advisable to introduce two parameters q_a and q_c that determine the mesh in reciprocal space. The mesh of points in 1BZ is determined by:

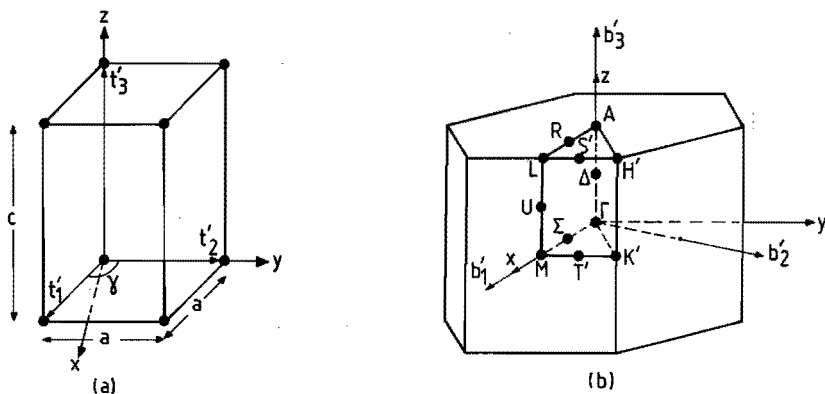


Fig. 3.2 (a) The primitive unit cell of a hexagonal Bravais lattice is spanned by three basis vectors: $t'_1 = \frac{1}{2}a\sqrt{3}e_x - ae_y$, $t'_2 = ae_y$, $t'_3 = ce_z$, where a and c are the lattice parameters of the hexagonal lattice and $\gamma = 2\pi/3$. (b) First Brillouin zone of a hexagonal Bravais lattice with symmetry points in the irreducible part of the zone, following the notation of Herring [56]. Basis vectors b'_i ($i=1,2,3$) are given in the text.

$$u_p, u_r = (p-1)/q_a \quad (p=1,2,\dots,q_a), \quad (3.30a)$$

$$u_s = (2s-q_c-1)/2q_c \quad (s=1,2,\dots,q_c). \quad (3.30b)$$

The special points, lying in IBZ (see fig. 3.2(b)), follow from the restrictions (apply point group D_{6h}):

$$0 \leq 2u_p + u_r \leq 1, \quad (3.31a)$$

$$0 \leq u_r \leq u_p \leq \frac{1}{2}, \quad (3.31b)$$

$$0 \leq u_s < \frac{1}{2}. \quad (3.31c)$$

Regarding the dependence of the important properties on q_a and q_c , it can be shown that:

$$N_{sp}(q_a, q_c) = P_a(q_a) P_c(q_c), \quad (3.32a)$$

where

$$P_a(q_a) = (\alpha+1)(3\alpha+\beta) + \delta_{\beta,0}, \quad (3.32b)$$

$$\alpha = \frac{1}{6}(q_a - \beta), \quad \beta = q_a \bmod 6, \quad (3.32c)$$

$$P_c(q_c) = \begin{cases} \frac{1}{2}q_c & q_c \text{ even} \\ \frac{1}{2}(q_c+1) & q_c \text{ odd} \end{cases} \quad (3.32d)$$

In table 3.1 we give some values for P_a and P_c . From table 3.1 it is clear that it is advantageous to choose q_c even. From (3.30) we have:

$$N_{BZ}(q_a, q_c) = q_a^2 q_c. \quad (3.32e)$$

The dependence of R_M on q_a and q_c is for these sets given by:

$$R_M(q_a, q_c) = a(\min(q_a, q_c/a)). \quad (3.32f)$$

With $\min(x,y)$ we denote the minimum of x and y .

q_a	$P_a(q_a)$	q_c	$P_c(q_c)$
1	1	1	1
2	2	2	1
3	3	3	2
4	4	4	2
5	5	5	3
6	7	6	3
7	8	7	4
8	10	8	4
9	12	9	5

Table 3.1. Dependence of the functions P_a and P_c determining the number of special points (see (3.30a)) on the parameters q_a and q_c , respectively.

We have not attempted to prove (3.32f), but it holds in all cases considered.

The sets determined by (3.30)-(3.31) do not have ambiguities like in the fcc case. Taking q_c to be even, u_s never assumes the value zero. With this restriction there are always exactly 12 D_{6h} -operations that leave the third component of a point $k = [u_p, u_r, u_s]$ invariant (numbers between square brackets are always components with respect to our choice of basis vectors). Therefore, the weighting factor, which is inversely proportional to the number of D_{6h} -operations that leave the complete k invariant, depends on the pair of components (u_p, u_r) only. For q_a and q_c even we have 7 distinct possibilities:

(1) $w = 2/N_{BZ}$ if $(u_p, u_r) = (0,0)$,

(2) $w = 6/N_{BZ}$ if $(u_p, u_r) = (\frac{1}{2}, 0)$,

(3) $w = 12/N_{BZ}$ if $(u_p, u_r) = (p, p)$ and $p \neq 1/3$, $0 < p < \frac{1}{2}$,

(4) $w = 4/N_{BZ}$ if $(u_p, u_r) = (1/3, 1/3)$,

(5) $w = 12/N_{BZ}$ if $(u_p, u_r) = (p, 0)$ and $0 < p < \frac{1}{2}$,

(6) $w = 12/N_{BZ}$ if $(u_p, u_r) = (p, p')$ and $p = -\frac{1}{2}p' + \frac{1}{2}$, $0 < p, p' < \frac{1}{2}$,

(7) $w = 24/N_{BZ}$ if $(u_p, u_r) = (p, p')$ and $p \neq p'$, $p \neq -\frac{1}{2}p' + \frac{1}{2}$, $0 < p, p' < \frac{1}{2}$.

N_{BZ} is given by (3.32e). Equations (3.30)-(3.31) together with the easily determinable weights can be considered as an efficient approximation scheme for integrals over 1BZ. We will use this scheme in section 3.3.

3.3 Convergence of energy-band integrations using special points

In general, the accuracy of averaging over 1BZ using special-point sets is hard to determine, since in most applications the evaluation of the integrand at one k point already is a hard problem (In our application the set of equations (2.39) must be solved). Furthermore, we do not know how smooth the function of k is that we wish to average or, equivalently, how rapidly the coefficients f_m in expansion (3.4) fall off. Based on global arguments and assuming smoothness of the function to be averaged, it was argued in refs.[49] and [50] that the error made by using a special-point set falls off proportional to R_M^{-3} . In terms of the parameter q defined for MP-sets for fcc lattices, this implies (see (3.29c)) that this error falls off proportional to q^{-3} .

In this section, we study as an illustrative example the average over 1BZ of the function $f(\mathbf{k})=|\mathbf{k}|^2$. This function has three nice features for our purpose of studying the convergence of energy-band integrations using special points: (i) it can be integrated exactly, (ii) its average can with low computational effort be calculated with very large numbers of special points, and (iii) it is an example of an energy band, viz., the lowest energy band in a lattice where the potential equals zero (the so-called empty lattice). We will study the average of $f(\mathbf{k})$ over 1BZ both for fcc and hexagonal lattices.

We start by evaluating the average I_{fcc} of $|\mathbf{k}|^2$ over 1BZ of a fcc lattice:

$$I_{\text{fcc}} = \frac{\Omega_c(\text{fcc})}{(2\pi)^3} \int_{1\text{BZ}(\text{fcc})} |\mathbf{k}|^2 d^3\mathbf{k}. \quad (3.33)$$

Substituting $\Omega_c(\text{fcc}) = \frac{1}{4}a_c^3$ and realizing that the integrand has the complete symmetry of the lattice point group O_h , so that the integration can be limited to 1BZ (defined by (3.28)), we may write:

$$I_{\text{fcc}} = 12 \left[\frac{2\pi}{a_c} \right]^2 T, \quad (3.34a)$$

where

$$T = \int_0^1 dx \int_0^{\min(x, 3/2-x)} dy \int_0^{\min(y, 3/2-y-x)} dz (x^2+y^2+z^2). \quad (3.34b)$$

After a straightforward but tedious calculation we find:

$$I_{\text{fcc}} = \frac{19}{32} \left[\frac{2\pi}{a_c} \right]^2 = 0.59375 \left[\frac{2\pi}{a_c} \right]^2. \quad (3.35)$$

We also approximated I_{fcc} by $I_{\text{fcc}}(q)$ using the MP-sets defined in section 3.2.1 for $q = 2, 4, \dots, 16$. The results are given in table 3.2, where the last column gives the relative difference with the exact result (3.35).

q	$N_{sp}(q)$	$I_{fcc}(q)$ (in $(2\pi/a_c)^2$)	δ (in %)
2	1	3/4=0.75	+26.32
4	2	9/16=0.5625	- 5.26
6	6	23/36=0.6388	+ 7.60
8	10	75/128=0.58593775	- 1.32
10	19	1491/2500=0.5964	+ 0.45
12	28	85/144=0.590277	- 0.58
14	44	5715/9604=0.59506456	+ 0.22
16	60	303/512=0.59179769	- 0.33

Table 3.2. Approximations $I_{fcc}(q)$ of the average of $|k|^2$ over 1BZ of a fcc lattice using progressively larger sets of special points characterized by q . $N_{sp}(q)$ is the number of special points in a set with parameter q and δ is the relative difference with the exact result $I_{fcc}=0.59375$. (A string of barred figures is recurring).

For 4-fold q , we used the computerization procedure described in section 3.2.1 to calculate $I_{fcc}(q)$ for very large values of q . The surprising results is that all results of calculations with q 4-fold are exactly given by:

$$I_{fcc}(q) = \left[\frac{2\pi}{a_c} \right]^2 \left[\frac{19}{32} - \frac{1}{2q^2} \right]. \quad (3.36)$$

Note that for $q \rightarrow \infty$ $I_{fcc}(q)$ converges to the exact result (3.35) as it should. More interesting is that we seem to have an analytical result for the approximation. Strictly speaking, we have not proven that (3.36) holds for all q that are 4-fold, but in our calculations we have found no exception in about 30 values of q the largest being $q=400$. Presently, there seems little use in trying to prove (3.36), since the formula has only illustrative purposes. It would however be the first result of this kind as far as we know. As an illustration of the efficiency of our computerization procedure, we mention that the calculation of $I_{fcc}(q)$ for $q=200$ ($N_{sp}=85850$) took only 1.6 seconds and for $q=400$ ($N_{sp}=676700$) only 12.0 seconds of processing time on a Burroughs 7900 computer.

Another important thing to notice concerning (3.36) is that the error falls off proportional to q^{-2} , which is much more unfavourable than q^{-3} as suggested earlier for smooth functions. Of course our form of an energy band is not very realistic: it is much sharper at the edges of 1BZ than a real band structure is. A more realistic example is given in chapter 5 (section 5.3); the example there also shows that special-point approximation schemes converge much slower when integrating energy bands $E_n(\mathbf{k})$ than in the case of integrating $\rho_{\mathbf{k}}(r)$.

We have also evaluated the average of $|\mathbf{k}|^2$ over 1BZ of a hexagonal lattice:

$$I_{\text{HEX}} = \frac{\Omega_c(\text{hex})}{(2\pi)^3} \int_{\text{1BZ(hex)}} |\mathbf{k}|^2 d^3\mathbf{k}, \quad (3.37)$$

where $\Omega_c(\text{hex}) = \frac{1}{2}\sqrt{3}a^2c$. The integration in the z-direction (see fig. 3.2(b)) can be performed first so that an integral over a hexagon remains. We eventually arrive at:

$$I_{\text{HEX}} = \left[\frac{2\pi}{a}\right]^2 \left[\frac{5}{27} + \frac{1}{12}\left[\frac{a}{c}\right]^2\right]. \quad (3.38)$$

We denote the approximation to I_{HEX} using a MP-set with parameters q_a and q_c (see section 3.2.2) by $I_{\text{HEX}}(q_a, q_c)$. Again the integration in the z-direction may be performed separately and we may write:

$$I_{\text{HEX}}(q_a, q_c) = \left[\frac{2\pi}{a}\right]^2 \left[I_a(q_a) + I_c(q_c) \right]. \quad (3.39)$$

We calculated $I_a(q_a)$ and $I_c(q_c)$ for a large number of values q_a and q_c (both chosen even) with the computerization scheme given in section 3.2.2. For $I_c(q_c)$ we find numerically *and* analytically that again an exact relation holds:

$$I_c(q_c) = \frac{1}{12}\left[\frac{a}{c}\right]^2 \left[1 - \frac{1}{q_c^2} \right]. \quad (3.40)$$

For $I_a(q_a)$ such a relation summarizing all calculations was not found. In table 3.3 we give results for $I_a(q_a)$ for several values of q_a . The results converge to the exact result $I_{\text{HEX}} = 5/27 (= 0.\overline{185})$.

q_a	$I_a(q_a)$
4	0.1875
6	0.1882716
8	0.1875
10	0.1860
20	0.1855
100	0.185196
200	0.185188
400	0.1851859
1000	0.1851853

Table 3.3. Results of calculations of $I_a(q_a)$ (defined in the text) for various values of q_a . Results are rounded to 7 figures behind the decimal point.

Table 3.3 and eq. (3.40) show that, just as in the fcc case, a reasonable approximation (i.e., at the percentage level) is obtained with small values of the q -parameter(s), but that the convergence to the exact result is rather slow. We finally mention that the calculation for $q_a=400$ took 0.4 seconds and for $q_a=1000$ took 1.6 seconds of processing time on a Burroughs 7900 computer. This means that our computerization procedure takes practically no time at all to integrate a simple analytic function approximately, even if a very dense set of points is used.

3.4 Equivalent special-point sets for structurally different crystals

In the preceding sections we have seen that a reasonable approximation to an integral over IBZ may be obtained by calculating the integrand for just a few carefully selected points in IBZ. As was explained in section 2.5, it usually is not possible to obtain a fully converged value for the total energy E_{tot} -due to finite cutoffs that must be used-, but what can be obtained are accurate total-energy differences ΔE_{tot} . To this end one has to assure that all technical approximations are made in an equivalent way in the calculations of E_1 and E_2 . Here,

E_1 and E_2 are the total energies of crystals C_1 and C_2 , respectively, while $E_1 - E_2 = \Delta E_{tot}$. If approximations are made in an equivalent way, the systematic part of the error will cancel when calculating a total-energy difference. One may hope that this systematic part is the larger part. In this section we discuss how approximations using special points can be made in an equivalent way. We restrict ourselves to cases in which C_1 and C_2 contain the same type(s) of atom(s), but their crystal structure or lattice parameters are different.

In case the crystal structures of C_1 and C_2 are equal, but with different lattice parameters, the symmetry of the crystals C_1 and C_2 is clearly the same. Equivalent special-point approximations are then obtained by using special points for crystal C_2 which have the same coordinates with respect to the basis vectors of the reciprocal lattice of C_2 as the special points for crystal C_1 have with respect to the basis vectors of the reciprocal lattice of C_1 . This procedure is in fact a scaling of special points with the basis vectors. That such a prescription for equal crystal structures works well will be shown in chapter 6 (section 6.2).

If the crystal structures of C_1 and C_2 are different but the unit cell of C_2 can be obtained by small distortions of the unit cell of C_1 , a prescription similar to the one above is possible. Now the set of undistorted k points in 1BZ of C_1 must be distorted by the "reciprocal distortion" associated with the distortion in r -space. The set of resulting points can be reduced with the point group of the distorted crystal C_2 . If C_1 has higher symmetry than C_2 , the set of special points for C_2 will be larger than the one for C_1 . Obtaining equivalence regarding special-point approximations in this way has been successful in the calculation of phonon frequencies and elastic properties of semiconductors [30]. Proceeding in the way described above, energy differences have been determined with an accuracy of 1 in 10^5 using typically numbers of special points between 1 and 10.

This accuracy appears to be much harder to obtain if C_1 and C_2 are not related by a small distortion of the unit cell. Such crystals will be called *structurally distinct*.

In sections 3.1 and 3.2 we introduced three important parameters, two of which determine the accuracy with which a special-point formula like (3.11) approximates an average over 1BZ. These two parameters, R_M and N_{BZ} , are of course closely related to the third parameter, the

number of special points N_{sp} . The parameter R_M is especially relevant in the case of smooth functions: expansion coefficients f_m in (3.4) then fall off rapidly with increasing m , so the accuracy is not seriously affected if their associated symmetrised plane wave $A_m(\mathbf{k})$ ($m \geq M$) is not integrated exactly. For functions the smoothness of which is questionable, the relevance of R_M is less clear: one does not know a priori whether or not there are f_m for $m \geq M$ that are large and whose $A_m(\mathbf{k})$ is not integrated exactly. In such cases the parameter N_{BZ} seems to be the more relevant one; it measures the number of sampling points in 1BZ. In fact, this number does not guarantee a good sampling, since to that end the N_{BZ} points should furthermore be distributed evenly over 1BZ. The evenness of the distribution is related to the parameter R_M as we will now illustrate. Consider the MP-sets for hexagonal lattices from section 3.2.2. The parameters N_{BZ} and R_M are then given by (3.32e) and (3.32f), respectively. Let for some choice of q_a and q_c the numbers q_a and $q_c c/a$ be almost equal, and let us enlarge N_{BZ} by enlarging q_a but keeping q_c constant. It is then obvious that the distribution of points gets less even, while the parameter R_M soon will not increase anymore. To obtain a more even distribution one should also increase q_c and this in turn causes R_M to increase. As a general rule one should improve the sampling of 1BZ by enlarging N_{BZ} in such a way that also R_M gets larger. This is a natural way to obtain an even distribution. Note that for the fcc lattice an even distribution is already assured by construction: all directions are treated on an equal footing by using only one parameter q (see (3.27)). As a result we have that if N_{BZ} increases, R_M increases automatically (see (3.29)).

We will now propose a systematic way of choosing special-point sets for structurally distinct crystals C_1 and C_2 so that the special-point approximations are as equivalent as possible in the two cases. We call such a pair of special-point sets *equivalent special-point sets* (ESPS). From the discussion in the preceding paragraph the following prescription follows naturally: special-point sets for C_1 and C_2 should be chosen such that $R_M(C_1)$ and $R_M(C_2)$ as well as $N_{BZ}^*(C_1)$ and $N_{BZ}^*(C_2)$ are as close as possible. Here, N_{BZ}^* is the number of \mathbf{k} points in a unit volume in reciprocal space. By attempting to establish equivalence of the parameter N_{BZ}^* , we obtain a more or less equivalent treatment of crystals with different volumes of 1BZ. More explicitly

we express this as follows: suppose a special-point set for C_1 is chosen. We select a set for C_2 , out of sets available, by minimizing the following quantity:

$$\frac{|R_M(C_2) - R_M(C_1)|}{R_M(C_1)} + \frac{|N_{BZ}^*(C_2) - N_{BZ}^*(C_1)|}{N_{BZ}^*(C_1)}. \quad (3.41)$$

If we improve the approximation of the average for C_1 , the above prescription gives us a systematic way of improving the approximation for C_2 in an equivalent way. In this way we may hope -there is no guarantee- that the energy difference between C_1 and C_2 converges faster than the difference in individual total energies. If more is known about the smoothness of the function to be averaged, the expression (3.41) could be generalised to give more weight to one of the two terms.

fcc				hexagonal				
q	N_{sp}	N_{BZ}^*	R_M (in a)	q_a	q_c	N_{sp}	N_{BZ}^*	R_M (in a)
2	1	4	1.4	1	2	1	4	1.0
4	2	32	2.8	3	2	3	36	3.0
6	6	108	4.2	4	4	8	128	4.0
8	10	256	5.7	6	4	14	288	6.0
10	19	500	7.1	8	4	20	512	6.5
12	28	864	8.5	8	6	30	768	8.0
14	44	1372	9.9	10	6	42	1200	9.8
16	60	2048	11.3	12	8	76	2304	12.0

Table 3.4. Properties of equivalent special-point sets (ESPS) for fcc and hexagonal lattices. N_{sp} is the number of special points in IBZ, N_{BZ}^* the number of sampling points in a unit volume in reciprocal space, which was chosen to be the volume of 1BZ of a fcc lattice. R_M indicates the first symmetrised plane wave, in an expansion of a periodic function in symmetrised plane waves, that is not integrated exactly by the special-point set. The parameters q , q_a , and q_c characterize the sets completely (see section 3.2). R_M is always expressed in units of a , the hexagonal lattice parameter ($a_c = a\sqrt{2}$).

In table 3.4 we give examples of ESPS for fcc lattices and hexagonal lattices selected from the large class of MP-sets described in section 3.2. For the hexagonal lattice a c/a -ratio of $(2/3)\sqrt{6}$ was chosen. With this c/a -ratio the sets can be applied to the determination of energy differences between crystals with zincblende structure and wurtzite structure (such applications are found in chapter 6).

The zincblende structure consists of two fcc-sublattices with different types of atoms on each sublattice and one sublattice shifted over $[\frac{1}{4}, \frac{1}{4}, \frac{1}{4}]$ with respect to the other (see fig. 3.1(a)). If the atoms on the sublattices are the same, we have the diamond structure. In both the zincblende and diamond structure, there are two atoms in a primitive unit cell. The wurtzite structure has a hexagonal lattice (see fig. 3.2(a)) with two types of atoms in the unit cell at positions $[0,0,0]$ and $[2/3, 1/3, 1/2]$ for one type and at positions $[0,0,u]$ and $[2/3, 1/3, u+1/2]$ for the other. Usually u is around $3/8$ and the c/a -ratio is always close to $(2/3)\sqrt{6}=1.633$. If c/a and u have exactly these values, we speak of an "ideal" wurtzite structure. In that case, all atoms are at the centre of regular tetrahedra just as in the zincblende structure (If the cubic lattice parameter a_c equals $a\sqrt{2}$, these regular tetrahedra have the same size in both structures). If all 4 atoms in the primitive unit cell are of the same type, we have the hexagonal-diamond structure.

Sometimes an alternative exists to obtain equivalence in the treatment of structurally distinct crystals; this alternative requires the use of supercells for at least one of the crystals. Here we mean by a supercell a unit cell containing more atoms than the primitive unit cell but describing the same crystal. By using supercells, both crystals can be described with a unit cell of similar symmetry. We will illustrate this for the comparison of zincblende and wurtzite structures in which all atoms are at the centre of regular tetrahedra of the same size. We further assume that all bond lengths from atoms at the centre to atoms at the corners of the regular tetrahedra are equal. This implies: $a_c = a\sqrt{2}$, $c/a = (2/3)\sqrt{6}$ and $u = 3/8$. The zincblende structure can be described with a hexagonal unit cell containing 6 atoms. This is schematically depicted in fig. 3.3, where a plane is shown containing the z -axis and the atoms in a hexagonal

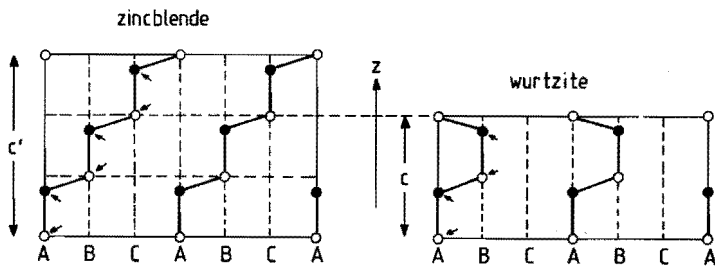


Fig. 3.3 Schematic representation of atomic positions and bonds in an equivalent plane for the zincblende and the wurtzite structure. The z -direction is along the $[1,1,1]$ -direction of the zincblende structure (cf. fig. 3.1(a)) and the usual z -direction of the wurtzite structure (cf. fig. 3.2(a)). If for both structures a unit cell with hexagonal symmetry is chosen, the little arrows denote the atoms in such a unit cell. The characters A, B, and C denote the type of position the atoms occupy in a plane perpendicular to the z -direction (see fig. 3.4).

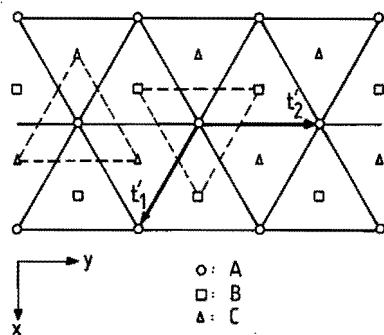


Fig. 3.4 Schematic representation of atomic positions in planes perpendicular to the z -direction in fig. 3.3. The vectors t_1' and t_2' are as in fig. 3.2(a). The atoms in one plane all occupy either A- or B- or C-type positions and are arranged in equilateral triangles.

unit cell (denoted by little arrows). The c' -axis points in the $[1,1,1]$ -direction of the zincblende structure (see fig. 3.1(a)). In fig. 3.3 an equivalent plane is also shown for the wurtzite structure. In planes perpendicular to the z -axis that go through the atomic positions shown in fig. 3.3 the atoms are all of the same type and are arranged in equilateral triangles with side a . The atoms in such a plane all occupy either A- or B- or C-type positions (fig. 3.4). From fig. 3.3 we see that $c' = 3c/2$, therefore the height of 1BZ for the "hexagonal-zincblende" structure will be $2/3$ times the height of 1BZ for the wurtzite structure (cf. fig. 3.2(b)). By choosing q_a equal for both structures and q_c 's with a ratio of $2/3$ (e.g., $q_c = 4$ for "hexagonal-zincblende" and $q_c = 6$ for wurtzite), R_M is exactly equal for both sets (see (3.32f)). Even more important for the equivalence of such sets is that not only N_{BZ}^* is also exactly equal, but that the sample k points in all of reciprocal space (to be found by repeating the k points in 1BZ along the basis vectors of the reciprocal lattice) are identical for both sets. Recently, this procedure was followed in ref. [57]. The disadvantage of this scheme is clear from eq. (2.82): by enlarging the unit cell with a factor of three, N_{PV} is multiplied by three (when keeping the resolution with which the wave functions are described constant, i.e., constant E_{PV}). Compared to using a primitive unit cell this approach increases the computational effort in solving the set of equations (2.39) by a factor of $3^3 = 27$ and the computational effort in calculating the (complicated) matrix elements of the nonlocal potential by a factor of $3^2 = 9$.

CHAPTER 4

EXPLOITATION OF CRYSTAL SYMMETRY FOR ELECTRONIC ENERGY BANDS AND STATES

In this chapter we present concepts and examples of how group theory may be used to reduce the size of secular equations of the type (2.39) if the reduced wave vector k occupies a symmetric position in the first Brillouin zone (1BZ). If there were no limits on the capacity of computers and the access users have to them, this chapter would be superfluous, but since these limits exist the material presented here can be of great help to reduce the computational effort.

We will not give a complete description of group theory as applied to the study of crystals - it can be found in a number of textbooks [58,59,60]-, but limit ourselves to the problem of simplifying the diagonalization of a Hamiltonian matrix expressed with respect to a basis of plane waves (PW's). We also pay attention to how the eigenvectors of this matrix should be handled in order to evaluate the valence-charge density (see chapter 2). Only concepts and results of group theory needed for this purpose will be presented, but no detailed discussions or proofs of these will be given.

The structure of a crystal is completely defined by specifying :

(i) 3 linearly independent basis vectors $t_1, t_2,$ and t_3 and (ii) the positions and types of atoms in a unit cell spanned by these vectors. The crystal structure can alternatively be classified by giving its space group \mathcal{G} . The symmetry operations of \mathcal{G} consist of primitive translations -i.e., translations over vectors $R = n_1 t_1 + n_2 t_2 + n_3 t_3$ with $n_1, n_2,$ and n_3 integers-, rotations around symmetry axes, reflections through planes, an inversion in the origin or combinations thereof. Some rotations, reflections or the inversion in the origin have a nonprimitive translation τ associated with them. Space groups containing these last type of operations are called *non-symmorphic space groups*, in contrast to *symmorphic space groups*, where none of the symmetry operations has a nonprimitive translation.

Following the notation of ref.[58], we denote an operation of the space group by $\{\mathcal{R}|R\}$, where R is the primitive translation and \mathcal{R} denotes an orthogonal, i.e., length-preserving operation, possibly

accompanied by a nonprimitive translation. The group of primitive translations \mathcal{T} consisting of operations $\{E|\mathbf{R}\}$, where E denotes the identity, is an invariant subgroup of \mathcal{G} . The group consisting of operations $\{\mathcal{R}|\mathbf{0}\}$, neglecting possible nonprimitive translations, is called the *crystallographic point group* \mathcal{G}_0 . In case of symmorphic space groups, there is a neat division of a symmetry operation $\{\mathcal{R}|\mathbf{R}\}$ into a purely orthogonal transformation $\{\mathcal{R}|\mathbf{0}\}$ and a pure translation $\{E|\mathbf{R}\}$. For a non-symmorphic space group this is not possible, since with some orthogonal transformations a (nonprimitive) translation is associated. In that case \mathcal{G}_0 is not a subgroup of \mathcal{G} .

In this work we will use the notation of Schönflies [53] to denote space groups and point groups. In this notation the zinblende structure has (symmorphic) space group T_d^2 and point group T_d , whereas the wurtzite and graphite structure have (non-symmorphic) space group C_{6v}^4 and point group C_{6v} .

4.1 Construction of symmetrised plane waves: theory

In accordance with the Bloch theorem, the wave function $\psi_{n,\mathbf{k}}(\mathbf{r})$ of an electron with reduced wave vector \mathbf{k} and bandindex n is expanded in plane waves as follows (see (2.38)):

$$\psi_{n,\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} C_{n,\mathbf{k}}(\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}. \quad (4.1)$$

The effect of a general operation of the space group $\{\mathcal{R}_j|\mathbf{R}\}$ on a single plane wave, $\varphi = \exp(i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r})$, is:

$$\begin{aligned} \{\mathcal{R}_j|\mathbf{R}\}\varphi &= e^{i(\mathbf{k}+\mathbf{G})\cdot\alpha_j\mathbf{r}} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{R}} e^{i(\mathbf{k}+\mathbf{G})\cdot\tau_j} \\ &= e^{i\alpha_j^{-1}(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} e^{i\mathbf{k}\cdot\mathbf{R}} e^{i(\mathbf{k}+\mathbf{G})\cdot\tau_j}. \end{aligned} \quad (4.2)$$

In (4.2) the vector τ_j is the nonprimitive translation associated with the operation $\{\mathcal{R}_j|\mathbf{R}\}$, while α_j^{-1} is the 3x3-matrix representing the orthogonal transformation $\{\mathcal{R}_j|\mathbf{0}\}$. The operation $\{\mathcal{R}_j|\mathbf{0}\}$ will be denoted by \mathcal{R}_j from now on. The factor $\exp(i\mathbf{k}\cdot\mathbf{R})$ in the right-hand side (RHS) of (4.2) reflects that the basis functions φ are chosen to be Bloch functions. By this choice we have fully exploited the translational

symmetry of the crystal. In order to investigate the exploitation of the remaining symmetry we may safely put R equal to 0 . Equation (4.2) reveals that the result of an operation of the space group \mathfrak{K}_j on a plane wave, apart from a possible phase factor, is again a plane wave, with a reduced wave vector that generally will be different from k . If apart from the identity operator $\{E|0\}$ none of the \mathfrak{K}_j operations result in a plane wave with the same reduced wave vector k , we speak of a *general k point*, and we cannot expect help from group theory. If on the other hand operations \mathfrak{K}_j exist, such that:

$$\underline{\alpha}_j^{-1}(k+G) = k+G', \quad (4.3)$$

thus producing a plane wave with the same reduced wave vector k , group theory may be of help as will be shown below. Note that there is no need for $\underline{\alpha}_j^{-1}$ to leave k invariant: changing by a reciprocal-lattice vector $G-G'$ is also allowed. From now on we will call an $\underline{\alpha}_j^{-1}$ satisfying (4.3): an operation that leaves k invariant. The operations in \mathcal{G} for which (4.3) holds form the *group of the wave vector k*: $\mathcal{G}(k)$. The operations $\underline{\alpha}_j^{-1}$ leaving k invariant form the *point group of the wave vector k*: $\mathcal{G}_o(k)$. This point group is isomorphic with one of the 32 possible crystallographic point groups.

At given reduced wave vector k , the operations $\underline{\alpha}_j^{-1}$ for which (4.3) holds subdivide the set of G vectors into shells; G and G' belong to the same shell if some $\underline{\alpha}_j^{-1}$ exists for which (4.3) is valid. Note that for a general k point each shell contains precisely one G vector. The associated set of PW's $\varphi = \exp(i(k+G) \cdot r)$ are similarly said to be subdivided in shells. The shells are denoted by the symbol $\mathcal{G}_s(k)$ ($s=1,2,\dots$), where the index s orders the shells by means of the non-decreasing magnitude of the vector $k+G$: $G \in \mathcal{G}_s(k)$ and $G' \in \mathcal{G}_{s+1}(k)$ implies $|k+G| \leq |k+G'|$. Since the $\underline{\alpha}_j^{-1}$ are length preserving, the length of $k+G$ for G 's within one shell is equal. It need, however, not be true that $k+G$ vectors with equal length are included in the same shell. In practical cases the number of shells is limited and will be denoted by the capital number S . We define $N(s)$ to be the number of members in shell s . The limited expansion of the wave function in (4.1) then consists of N terms:

$$N = \sum_{s=1}^S N(s) . \quad (4.4)$$

We note that $N(s)$ can never be larger than g_0 : the number of elements in the point group $\mathcal{G}_0(\mathbf{k})$ also called the *order* of $\mathcal{G}_0(\mathbf{k})$.

By way of the so-called projection-operator method [61], it is possible to linearly combine PW's in one shell to form symmetrised plane waves (SPW's), i.e., functions that transform according to the irreducible representations (irreps) of the group of the wave vector. These SPW's are given by:

$$f_{jl}^{(p)} = \sum_i \Gamma_p(\mathcal{R}_i)_{jl}^* \mathcal{R}_i \varphi , \quad (4.5)$$

where the \mathcal{R}_i denote the operations of $\mathcal{G}(\mathbf{k})$, φ is one of the PW's in the shell (since the sum is over all operations in $\mathcal{G}(\mathbf{k})$ the choice is arbitrary), and the $\Gamma_p(\mathcal{R}_i)_{jl}$ are matrix elements of irreducible representation p belonging to operation \mathcal{R}_i . The star in (4.5) denotes complex conjugation. The indices j and l range from 1 to n_p , where n_p is the dimension of irreducible representation p . We call $f_{jl}^{(p)}$ the j -th partner in a set of n_p basis functions for the p -th irreducible representation. From the range of indices j and l we see that we may find n_p sets of symmetrised functions each consisting of n_p partners. From the examples in the next section we will see that not always so many suitable symmetrised functions are found. Functions $f_{jl}^{(p)}$ with the same j and p but different l are also called: functions that transform according to the same row of the same irrep. A function transforms according to row j of irrep p if the following formula holds for all \mathcal{R}_i in $\mathcal{G}(\mathbf{k})$:

$$\mathcal{R}_i f_{jl}^{(p)} = \sum_{m=1}^{n_p} \Gamma_p(\mathcal{R}_i)_{mj} f_{ml}^{(p)} . \quad (4.6)$$

If p is a one-dimensional irrep, eq. (4.6) reduces to (we may omit indices j, l, m and the sum over m):

$$\mathcal{R}_i f^{(p)} = \Gamma_p(\mathcal{R}_i) f^{(p)} . \quad (4.7)$$

The proof that functions constructed according to (4.5) indeed

satisfy (4.6) is given in ref.[61] and rests upon the fact that the Γ 's are unitary matrix representations of $\mathcal{G}(\mathbf{k})$. The representations are unitary, since the unsymmetrised functions -the PW's φ_i - are orthonormal with respect to the hermitian inner product $\langle f | g \rangle$:

$$\langle f | g \rangle = \int f^*(\mathbf{r})g(\mathbf{r})d^3r, \quad (4.8)$$

where the integration is over all space.

We postpone the discussion of how to obtain the $\Gamma_p(\mathcal{R}_i)_{jl}$ needed to construct the SPW's until the next section and first mention the advantage of these newly formed functions. The advantage is that the Hamiltonian matrix calculated with respect to the basis of SPW's (to be called \underline{H}') has matrix elements equal to zero between SPW's belonging to different irreps of $\mathcal{G}(\mathbf{k})$. In case of higher-dimensional irreps, there are also no matrix elements unequal to zero between different partners in the same set of basis functions. This is expressed in the following formula, which is a slight extension of a formula in ref.[61] and equivalent to one in ref.[62]:

$$\langle f_{jl}^{(p)}(s) | H_{op} | g_{j'l'}^{(p)}(s') \rangle = H_{ll'}^{(p)}(s,s') \delta_{p,p} \delta_{j,j'} \quad (4.9)$$

In (4.9) δ is the usual Kronecker delta and the extra indices s and s' are added to denote the shell the symmetrised function is constructed for ($s, s' = 1 \dots S$). Note that there may be matrix elements unequal to zero between SPW's with the same index j but different index l . In words one may express (4.9) as: the only non-zero matrix elements of \underline{H}' occur between SPW's that transform according to the same row of the same irrep. In the practice of our plane-wave expansion, eq. (4.9) implies, that for every (s, s') -combination of shells one obtains a matrix $H_{ll'}$, independent of j for every irrep p . The indices l and l' can take on values from 1 to n_p , depending on the number of PW's in shell s and s' . For some shells some of the $f_{jl}^{(p)}(s)$ may be equal to each other or equal to zero and therefore they do not all represent suitable basis functions (see, e.g., example 2(A) below).

From formula (4.9) we infer that we can rearrange rows and columns of \underline{H}' such that a block form results. First one puts all rows and columns together that correspond to the same irrep p . Because of the

factor $\delta_{p,p}$, \underline{H} is now in block form with one block for every irrep p . The factor $\delta_{j,j}$, enables us to further break up blocks belonging to irreps with $n_p \geq 2$ into n_p blocks. Since in (4.9) the matrix element is independent of j , these n_p blocks are all identical. To obtain the energy levels or, equivalently, the band structure at the k point under consideration, we may diagonalize all blocks separately. Because for an irrep with dimension n_p we have n_p identical blocks, we find an n_p -fold degenerate energy level for such an irrep of $\mathcal{G}(k)$. This is called a degeneracy caused by symmetry in contrast to a so-called accidental degeneracy, which occurs if at a certain k point bands happen to cross.

The gain of calculating the Hamiltonian with respect to a basis of SPW's instead of with respect to a basis of PW's is obvious if one realizes that the computing time of solving an eigenvalue problem numerically is proportional to N^3 , where N is the dimension of the matrix to be diagonalized. In our problem N is the number of PW's into which $\psi_{n,k}$ is expanded (see (4.1) and (4.4)). Therefore, breaking up an eigenvalue problem into two eigenvalue problems of half the size of the original problem reduces computing time with a factor of 4. This example already applies if the k point is invariant under one extra operation of the point group α_j^{-1} only (apart from the (trivial) identity operation). The gain is usually a little less than a factor of 4, since in general the eigenvalue problem is not split up into two equal parts.

After having demonstrated the superiority of the basis of SPW's over the basis of PW's regarding computing time (in certain applications the symmetry analysis of eigenstates, that is allied with the procedures above, may also be of help. For our type of problem this is not the most important advantage), we discuss the basis transformation in a little more detail. The SPW's $f_{j1}^{(p)}$ constructed by (4.5) are not normalized with respect to the inner product (4.8), but since the PW's are normalized, we straightforwardly obtain the *normalized* SPW's $F_i^{(s)}$ ($i = 1 \dots N(s)$, $s = 1 \dots S$). We introduce the unitary matrix \underline{U} , which connects the two orthonormal bases. It is composed of matrices $u^{(s)}$ along the diagonal, that are defined through:

$$F_i^{(s)} = \sum_{j=1}^{N(s)} u_{ij}^{(s)} \varphi_j^{(s)} \quad (i=1 \dots N(s), s=1 \dots S). \quad (4.10)$$

In (4.10) $\varphi_j^{(s)}$ is the j-th member PW of shell s.

For a derivation of the connection between the Hamiltonian on a basis of PW's (\underline{H}) and on a basis of SPW's (\underline{H}'), we may condense the sub- and superscripts of F and φ into one index, such that $F_\alpha \equiv F_i^{(s)}$ with $\alpha = i + \sum_{(j=1)}^{(s)} N(j)$. Formula (4.10) can now be written as:

$$F_\alpha = \sum_{\beta=1}^N U_{\alpha\beta} \varphi_\beta \quad (\alpha=1\dots N), \quad (4.11)$$

where $U_{\alpha\beta}$ is the complete $N \times N$ -matrix with matrices $u^{(s)}$ along the diagonal. From the orthonormality of the φ we infer:

$$\langle \varphi_\beta | F_\alpha \rangle = U_{\alpha\beta}. \quad (4.12)$$

If $H_{\alpha\beta}$ is the matrix element of the Hamiltonian between PW's φ_α and φ_β .

$$H_{\alpha\beta} = \langle \varphi_\alpha | H_{op} | \varphi_\beta \rangle. \quad (4.13)$$

then the matrix element $H'_{\alpha\beta}$ between normalized SPW's F_α and F_β is:

$$\begin{aligned} H'_{\alpha\beta} &= \langle F_\alpha | H_{op} | F_\beta \rangle \\ &= \sum_{\gamma, \delta} \langle \varphi_\gamma | U_{\alpha\gamma}^* H_{op} U_{\beta\delta} | \varphi_\delta \rangle \\ &= \sum_{\gamma, \delta} U_{\alpha\gamma}^* H_{\gamma\delta} U_{\delta\beta} \end{aligned} \quad (4.14)$$

Formula (4.14) has the usual form of a unitary basis transformation. It suggests that two matrix multiplications with $N \times N$ -matrices have to be performed. For readers concerned about the computational demands of the present basis transformation, we remark that, because of the simple form \underline{U} has, only matrix multiplications with $N(s) \times N(s')$ -matrices -with $N(s)$ typically ranging between 1 and 12- have to be performed. These matrix multiplications are combined with the separation of \underline{H}' into blocks, as described above. The computational effort involved in the matrix transformation is negligible compared to the diagonalization of the blocks of \underline{H}' . The only input required is

the set of matrices $\underline{\alpha}_i^{-1}$ ($i=1\dots g_0$) that leave the \mathbf{k} point invariant, the associated nonprimitive translations τ_i , and the matrix elements $\Gamma_p(\mathcal{G}_i)_{jl}$.

4.2 Construction of symmetrised plane waves: illustrative examples

In this section the theory of section 4.1 is illustrated with some examples (For these we will choose some of the special \mathbf{k} points from chapter 3). We will not give well-specified procedures to construct what we will call Slater tables of matrix elements $\Gamma_p(\mathcal{G}_i)_{jl}$. One reason is that it is not easy to give a completely general procedure. Moreover, these Slater tables may be found for the most symmetric \mathbf{k} points in the Brillouin zones of the most commonly found crystal structures in Appendix 3 of ref.[58].

For symmorphic space groups Slater tables can easily be constructed; because of the possibility of a neat division of $\mathcal{G}(\mathbf{k})$ into an orthogonal and translational part, one only needs to deal with $\mathcal{G}_0(\mathbf{k})$. This $\mathcal{G}_0(\mathbf{k})$ is always isomorphic to one of the 32 possible point groups, the irreducible matrix representations of which may be found for instance in Appendix 1 of ref.[60]. With these matrix representations at hand the Slater table can immediately be assembled. We note that the character tables for point groups given in most books on group theory are not sufficient to write down the Slater tables, unless $\mathcal{G}_0(\mathbf{k})$ has one-dimensional irreps only (For one-dimensional irreps the character, which is the trace of the matrix representation, is equal to the (1×1) matrix representation). For non-symmorphic space groups the situation is less clear and the Slater tables are usually arrived at more or less by trial and error. Especially for symmetric points lying on the boundaries of 1BZ the matrix elements of the irreps of $\mathcal{G}(\mathbf{k})$ (even the dimensionalities of the irreps) can be quite different from what would be expected from the irreps of $\mathcal{G}_0(\mathbf{k})$. However, also in these cases well-defined procedures can be carried through to find irreps of $\mathcal{G}(\mathbf{k})$ with the translational part factored out [60].

Of course the correctness of Slater tables can always be checked by verifying whether SPW's constructed using the presumed Slater table and (4.5) indeed satisfy (4.6). With some experience in checking

Slater tables the procedure termed "trial and error" above usually converges rapidly to the correct table.

Example 1: One of the special points out of the two-point set of Chadi and Cohen (or: Monkhorst and Pack; see chapter 3) applicable to fcc-lattices is the one with coördinates $k_1 = [\frac{1}{4}, \frac{1}{4}, \frac{1}{4}]$ with respect to the basis $\{b_i\}_{i=1,2,3}$ in reciprocal space (see section 3.2.1). Coördinates between square brackets always denote coördinates with respect to the basis at hand. Suppose we want to calculate the band energies at this k point for a compound with zinblende structure (space group T_d^2 , point group T_d). The space group T_d^2 is symmorphic, so there are no nonprimitive translations. The point group T_d consists of 24 symmetry operations $\mathfrak{K}_1 \dots \mathfrak{K}_{24}$. (For the numbering and notation of the operations we use the convention of ref.[58]). Of these only \mathfrak{K}_1 (the identity) and \mathfrak{K}_{19} leave k_1 invariant. The matrices representing the operation of \mathfrak{K}_1 and \mathfrak{K}_{19} in 3-dimensional reciprocal space with respect to basis $\{b_i\}_{i=1,2,3}$ are:

$$\underline{\alpha}_1^{-1} = E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \underline{\alpha}_{19}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

The point group of k_1 , $\mathcal{G}_o(k_1)$, is C_s containing a reflection through a plane (σ) besides the identity (E). The character table of C_s is given in table 4.1(a). C_s has two one-dimensional irreps; a result from group theory is that the sum of squared dimensionalities of irreps of a group must equal the order of the group; so this is the only possibility.

	E	σ
1	1	1
2	1	-1

(a)

	\mathfrak{K}_1	\mathfrak{K}_{19}
1	1	1
2	1	-1

(b)

Table 4.1. (a) Character table for point group C_s .
 (b) Slater table for $k_1=[\frac{1}{4}, \frac{1}{4}, \frac{1}{4}]$ for space group T_d^2 .

The Slater table for k_1 is given in table 4.1(b) and is identical to table 4.1(a). We will now illustrate the construction of SPW's from PW's for a typical shell of PW's. Since in this example and the following examples we will be talking about one typical shell at a time, we will omit the superscript s denoting the shell from now on. The 9^{th} shell $\mathcal{S}_9(k_1)$ consists of 2 PW's, φ_1 and φ_2 , where $\varphi_j = \exp(i(k_1 + G_j) \cdot r)$, which have associated reciprocal-lattice vectors $G_1 = [0, 1, 0]$ and $G_2 = [0, 0, 1]$. From the matrices $\underline{\alpha}_{11}^{-1}$ and $\underline{\alpha}_{19}^{-1}$ it can be checked that this is indeed a shell for k_1 . Using table 4.1(b) and (4.5) we have (take $\varphi = \varphi_1$): $f^{(1)} = \varphi_1 + \varphi_2$ and $f^{(2)} = \varphi_1 - \varphi_2$. The $\mathcal{R}_1 \varphi$ introduce no additional phase factors since T_d^2 is symmorphic. Formula (4.7) is easily verified and so $f^{(1)}$ and $f^{(2)}$ are suitable SPW's. The normalized SPW's are: $F_1 = \frac{1}{\sqrt{2}}(\varphi_1 + \varphi_2)$ and $F_2 = \frac{1}{\sqrt{2}}(\varphi_1 - \varphi_2)$. The u -matrix for this shell as defined by (4.10) is:

$$\underline{u} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

Example 2: As an example for the same space group, but with higher symmetry, we treat the k point with A -symmetry: $k_2 = [p, p, p]$ with $0 \leq p \leq \frac{1}{2}$. The second special point in the two-point set of Chadi and Cohen [49] is an example of such a point ($p = \frac{1}{4}$). We note that the following symmetry analysis is also applicable for special values of p , where k_2 has higher symmetry, i.e., the Γ - and L -points ($p=0$ and $p=\frac{1}{2}$, respectively); there is no obligation to use the full symmetry group of the k point at hand. The point group $\mathcal{S}_0(k_2)$ consists of 6 symmetry operations $\mathcal{R}_1, \mathcal{R}_5, \mathcal{R}_9, \mathcal{R}_{19}, \mathcal{R}_{21}, \mathcal{R}_{23}$ and is isomorphic to the point group C_{3v} , which is a subgroup of T_d . The $\underline{\alpha}^{-1}$ -matrices are:

$$\begin{aligned} \underline{\alpha}_{51}^{-1} &= \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} & \underline{\alpha}_{91}^{-1} &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \\ \underline{\alpha}_{21}^{-1} &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} & \underline{\alpha}_{23}^{-1} &= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

($\underline{\alpha}_{11}^{-1}$ and $\underline{\alpha}_{19}^{-1}$ were already given with example 1).

In order to understand the character table (table 4.2(a)) for C_{3v} , we recall from standard group theory that the elements of a group can

	\mathcal{C}_1	\mathcal{C}_2	\mathcal{C}_3
1	1	1	1
2	1	1	-1
3	2	-1	0

(a)

	\mathcal{R}_1	\mathcal{R}_5	\mathcal{R}_9	\mathcal{R}_{19}	\mathcal{R}_{21}	\mathcal{R}_{23}
1	1	1	1	1	1	1
2	1	1	1	-1	-1	-1
$(3)_{11}$	1	$-\frac{1}{2}$	$-\frac{1}{2}$	1	$-\frac{1}{2}$	$-\frac{1}{2}$
$(3)_{21}$	0	$\frac{1}{2}\sqrt{3}$	$-\frac{1}{2}\sqrt{3}$	0	$-\frac{1}{2}\sqrt{3}$	$\frac{1}{2}\sqrt{3}$
$(3)_{12}$	0	$-\frac{1}{2}\sqrt{3}$	$\frac{1}{2}\sqrt{3}$	0	$-\frac{1}{2}\sqrt{3}$	$\frac{1}{2}\sqrt{3}$
$(3)_{22}$	1	$-\frac{1}{2}$	$-\frac{1}{2}$	-1	$\frac{1}{2}$	$\frac{1}{2}$

(b)

Table 4.2. (a) Character table for point group C_{3v} .
 (b) Slater table for point with Λ -symmetry for space group T_d^2 .

be subdivided into so-called classes and that the number of irreps of a group equals the number of classes. The characters for elements in one class are equal. For C_{3v} there are 3 classes. In terms of the elements of $\mathcal{G}_o(\mathbf{k}_2)$ these are: $\mathcal{C}_1 = \{\mathcal{R}_1\}$, $\mathcal{C}_2 = \{\mathcal{R}_5, \mathcal{R}_9\}$ and $\mathcal{C}_3 = \{\mathcal{R}_{19}, \mathcal{R}_{21}, \mathcal{R}_{23}\}$. Since there are 6 elements in $\mathcal{G}_o(\mathbf{k}_2)$, it must have -because of the rule quoted with example 1- 2 one-dimensional and 1 two-dimensional irreps.

In table 4.2(b) the Slater table for a point with Λ -symmetry for space group T_d^2 is given. One may check that the characters of the symmetry operations coincide with the ones for the corresponding class in table 4.2(a).

We now construct SPW's for two typical shells of PW's when $p=4$:

(A) The 5th shell for \mathbf{k}_2 consists of 3 PW's, φ_1, φ_2 , and φ_3 , with associated reciprocal-lattice vectors: $G_1 = [1,0,0]$, $G_2 = [0,1,0]$, $G_3 = [0,0,1]$. The SPW's constructed with table 4.2(b) and (4.5) are ($\varphi = \varphi_1$):

$$\begin{aligned}
 f^{(1)} &= \varphi_1 + \varphi_2 + \varphi_3, & f^{(2)} &= 0, \\
 f_{11}^{(3)} &= 2\varphi_1 - \varphi_2 - \varphi_3, & f_{21}^{(3)} &= \sqrt{3}(\varphi_2 - \varphi_3), \\
 f_{12}^{(3)} &= 0, & f_{22}^{(3)} &= 0.
 \end{aligned}$$

As was stated before, 3 PW's result in 3 SPW's. We may again check (4.6). For instance, $\mathcal{R}_5 f_{11}^{(3)} = 2\varphi_2 - \varphi_3 - \varphi_1$, which indeed equals:

$$\sum_{(m=1)}^{(2)} \Gamma_3(\mathcal{R}_5) f_{m1}^{(3)} f_{m1}^{(3)} = -\frac{1}{2} f_{11}^{(3)} + \frac{1}{2}\sqrt{3} f_{21}^{(3)} = -\varphi_1 + 2\varphi_2 - \varphi_3.$$

The normalized SPW's are:

$$F_1=(1/\sqrt{3})(\varphi_1+\varphi_2+\varphi_3), F_2=(1/\sqrt{6})(2\varphi_1-\varphi_2-\varphi_3), F_3=(1/\sqrt{2})(\varphi_2-\varphi_3).$$

The u-matrix for this shell follows immediately:

$$\underline{u} = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 2/\sqrt{6} & -1/\sqrt{6} & -1/\sqrt{6} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

(B) The 9^{th} shell for k_2 consists of 6 PW's $\varphi_1 \dots \varphi_6$ with associated reciprocal-lattice vectors : $G_1=[-1,0,1]$, $G_2=[1,-1,0]$, $G_3=[0,1,-1]$, $G_4=[-1,1,0]$, $G_5=[1,0,-1]$, $G_6=[0,-1,1]$. The SPW's are:

$$\begin{aligned} f^{(1)} &= \varphi_1+\varphi_2+\varphi_3+\varphi_4+\varphi_5+\varphi_6, & f^{(2)} &= \varphi_1+\varphi_2+\varphi_3-\varphi_4-\varphi_5-\varphi_6, \\ f_{11}^{(3)} &= \varphi_1-\frac{1}{2}\varphi_2-\frac{1}{2}\varphi_3+\varphi_4-\frac{1}{2}\varphi_5-\frac{1}{2}\varphi_6, & f_{21}^{(3)} &= \frac{1}{2}\sqrt{3}(\varphi_2-\varphi_3-\varphi_5+\varphi_6), \\ f_{21}^{(3)} &= -\frac{1}{2}\sqrt{3}(\varphi_2-\varphi_3+\varphi_5-\varphi_6), & f_{22}^{(3)} &= \varphi_1-\frac{1}{2}\varphi_2-\frac{1}{2}\varphi_3-\varphi_4+\frac{1}{2}\varphi_5+\frac{1}{2}\varphi_6. \end{aligned}$$

In contrast to example 2(A), we now find for irrep $p=3$ two sets of SPW's each consisting of two partners, namely $\{f_{11}^{(3)}, f_{21}^{(3)}\}$ and $\{f_{12}^{(3)}, f_{22}^{(3)}\}$. We need both sets since the shell consisted of 6 PW's. An important thing to note is that \underline{H}' (see section 4.1) will contain non-zero matrix elements between $f_{11}^{(3)}$ and $f_{12}^{(3)}$, because they transform according to the same row of the same irrep (cf. (4.9)). The same holds for $f_{21}^{(3)}$ and $f_{22}^{(3)}$. So this shell leads to 2 SPW's that transform according to the same row $j=1$ of the same irrep $p=3$. Therefore, the dimension of the block associated with this combination of p and j is enlarged by two. The same holds for the combination $j=2$ and $p=3$. For the shell of example 2(A) the dimension of both these blocks is only enlarged by one. This observation will be relevant for the treatment of eigenvectors resulting from the different blocks (section 4.3).

The normalized SPW's F_i ($i=1 \dots 6$) are again easily obtained and the u-matrix can be written down immediately from these. The dimension 6 for the u-matrix is the largest one can obtain for a general point with Λ -symmetry, since such a point has only 6 symmetry operations.

Example 3. As final example, we will treat a symmetric k point in 1BZ of a structure with the non-symmorphic space group C_{6v}^4 . This space group applies to, e.g., the wurtzite and graphite structures. The point group C_{6v} has 12 symmetry operations, the effect of which on a plane wave is given in table A3-1 of ref.[58]. The operations X_6 and Y_1 (in the notation of ref.[58]) leave the k point $k_3=[p,p,p_3]'$ invariant. Coördinates are given with respect to the basis $\{b_i'\}_{i=1,2,3}$

(see section 3.2.2). Points with the type of symmetry of k_3 occur frequently when generating special points with the Monkhorst/Pack-scheme for hexagonal lattices (see chapter 3). The matrices with respect to basis $\{b'_i\}_{i=1,2,3}$, which represent the effect of the operations of the space group on the wave vector, are the identity matrix for X_0 , and for Y_1 :

$$\underline{\alpha}^{-1}(Y_1) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The nonprimitive translation associated with Y_1 is $\tau = [0, 0, \frac{1}{2}]'$ (coordinates with respect to basis $\{t'_j\}_{j=1,2,3}$, which is connected to basis $\{b'_i\}_{i=1,2,3}$ through $b'_i \cdot t'_j = 2\pi\delta_{ij}$). So the effect of Y_1 on a PW is the transformation of the wave vector through $\underline{\alpha}^{-1}(Y_1)$ and multiplication by a phase factor $\exp(i(\mathbf{k}+\mathbf{G})\cdot\tau)$.

The point group $\mathcal{G}_0(k_3)$ is C_3 , the character table of which was given in table 4.1(a). In this case, however, the Slater table is not identical to the character table. Table 4.3 shows this Slater table.

	X_0	Y_1
1	1	α
2	1	$-\alpha$

Table 4.3. Slater table for point $[p, p, p_3]'$ for space group C_{6v}^4 . α stands for $\exp(\pi ip_3)$.

A typical shell of reciprocal-lattice vectors for k_3 is the one consisting of $G_1 = [1, 0, 1]'$ and $G_2 = [0, 1, 1]'$ with PW's φ_1 and φ_2 . With (4.5) the SPW's are: $f^{(1)} = \varphi_1 - \varphi_2$, $f^{(2)} = \varphi_1 + \varphi_2$. In the calculation of $f^{(1)}$ and $f^{(2)}$ the phase factors α^x resulting from the Slater table are partly compensated by the phase factors resulting from the nonprimitive translation. Using $Y_1\varphi_1 = -\alpha\varphi_2$ and $Y_1\varphi_2 = -\alpha\varphi_1$, one easily verifies that $f^{(1)}$ and $f^{(2)}$ indeed satisfy (4.7). So only by including the phase factors α in the Slater table, eq. (4.5) leads to SPW's that transform correctly. The u-matrix for this shell is given by:

$$\underline{u} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Apparently, the effect of nonprimitive translations on the construction of the Slater table is that the column of the character table of $\mathcal{G}_0(\mathbf{k}_3)$ associated with an operation with accompanying τ must be multiplied by $\exp(i\mathbf{k}_3 \cdot \tau) = \alpha$. This rule of thumb is not always applicable for \mathbf{k} points with higher symmetry; as was mentioned before Slater tables for points on the boundaries of 1BZ can be completely different from the Slater tables of their point group. The general observation, however, that nonprimitive translations introduce phase factors in the Slater table remains true.

From the above examples one might gain the impression that the matrices \underline{u} are always real, so that the basis transformation is simply orthogonal (instead of unitary). Although this is true in a great deal of situations, there are some exceptions. An example of such an exception is point W ($[1/4, 1/2, 3/4]$) in 1BZ of a zincblende structure: here the Slater table contains some purely imaginary characters, whereas there are no nonprimitive translations. Hence, the matrix \underline{U} contains purely imaginary elements.

We conclude this section with a number of practical remarks of general character:

(i) Besides the verification of Slater tables through (4.5) and (4.6) this verification can also be achieved by performing two calculations of the band structure at the \mathbf{k} point at hand; one using the presumed Slater table and one using no symmetry (i.e., a Slater table consisting of one element equal to unity for the identity operation). Using no symmetry, \underline{U} equals the unity matrix and the Hamiltonian matrix has to be diagonalized completely. The results should be identical to those obtained using the presumed Slater table. This can be done for small N , although N should be so large that at least one shell with the maximum number of members is included in the basis set of PW's.

(ii) The possibility of not using the complete symmetry of a highly symmetric \mathbf{k} point has the advantage that Slater tables are more easily constructed. Furthermore, the extra saving of computing time when using full symmetry is not always substantial. As an example, we discuss the Γ -point ($\mathbf{k}=0$) in the zincblende structure. $\mathcal{G}_0(\Gamma)$ contains 24 symmetry operations and has 2 one-dimensional, 1 two-dimensional

and 2 three-dimensional irreps. This implies (see section 4.1) that \underline{H}' is split into 10 blocks of which only 5 need be diagonalized (one for every irrep). If we suppose all blocks to be of equal dimension -for the ease of this argumentation-, use of symmetry reduces computing time by a factor of $1/5 \times (10)^3 = 200$ or, equivalently, by 99.5% compared to the case when no symmetry is used. If, however, we employ only Λ -symmetry (see example 2), we have 2 one-dimensional and 1 two-dimensional irreps, resulting in a break-up of \underline{H}' into 4 blocks of which only 3 need be diagonalized. Supposing all blocks to be of equal size, computing time now is reduced by a factor of $1/3 \times (4)^3 \approx 21$ or, equivalently, by 95.3%. The extra saving of 4.2% by using full symmetry instead of Λ -symmetry usually is not considered to balance the extra trouble and error-prone usage of a 24×24 Slater table instead of a 6×6 Slater table. The interplay between the advantage of the full power of group theory and practical considerations leads us to the following policy: we use such a subgroup of the group of the wave vector that no irreps with dimension higher than 2 occur. Connected with this is that we never use more than 12 symmetry operations.

4.3 Unfolding of symmetrised electron states

The block matrices that \underline{H}' consists of after rearranging rows and columns as explained in section 4.1 are diagonalized separately. The resulting eigenvalues constitute the band structure $E_n(\mathbf{k})$ at point \mathbf{k} in 1BZ, with n the band index. We find as many bands as the dimension N of the matrix \underline{H}' , but only the lowest few of these are usually of interest. It is also only this set of lowest energy levels that may be considered accurate, since for the highest levels the interactions with states $\psi_{m,\mathbf{k}}(\mathbf{r})$ ($m > N$) will be important and these were not taken into account because of the truncation of the expansion (4.1). The resulting eigenvectors for each block are the expansion coefficients of the wave function $\psi_{n,\mathbf{k}}(\mathbf{r})$ in SPW's that transform according to the same row of the same irrep. With the help of (part of) the matrix \underline{U} we can determine the expansion coefficients $C_{n,\mathbf{k}}(\mathbf{G})$ of the wave function in PW's $\exp(i(\mathbf{k}+\mathbf{G}) \cdot \mathbf{r})$. These $C_{n,\mathbf{k}}(\mathbf{G})$ for occupied states n are the ingredients for the calculation of the Fourier components $\rho(\mathbf{G})$ of the

valence-charge density (see chapter 3, equation (3.22)). Formally this is achieved as follows: the eigenvectors resulting from a $M \times M$ -block with associated SPW's F_j ($j=1, \dots, M$) are denoted by $C'_{n,k}(F_j)$. The $C_{n,k}(G_i)$ ($i=1, \dots, N$) are found by:

$$C_{n,k}(G_i) = \sum_{j=1}^M (\tilde{U}^T)_{ij} C'_{n,k}(F_j), \quad (4.15)$$

where \tilde{U}^T is the $N \times M$ -matrix formed by omitting all columns from the $N \times N$ -matrix U^T that do not correspond to the M functions F_j . For instance, in case of example 2(B) we have the following situation: both F_3 and F_5 belong to the same block, whereas PW's $\varphi_2, \varphi_3, \varphi_5$ and φ_6 are contained in both F_3 and F_5 . So by the construction (4.15) the expansion coefficients connected with $\varphi_2, \varphi_3, \varphi_5$ and φ_6 will have contributions from F_3 and F_5 . Since we have excluded 3- and higher-dimensional irreps, we can never have more than two contributions to one $C_{n,k}(G)$.

We further note that although a n_p -dimensional irrep results in n_p identical blocks, and thus in n_p sets of identical eigenvalues and eigenvectors, the electron states that belong to these n_p -fold degenerate energy levels will not be identical. This is true since, the identical expansion coefficients one finds relate to different SPW's. A way of explaining this is by observing that the matrices resulting by omitting columns from U^T are different for the two blocks giving the same eigenvalue. Of course both states should be different because of the Pauli exclusion principle.

We conclude this chapter by mentioning situations in which incorporating the exploitation of symmetry in computer codes may be less profitable. When very large numbers of special points N_{sp} are required (e.g., $N_{sp} > 60$), the fraction of general k points gets larger and starts dominating the computational effort of diagonalization. Another example is when the method of chapter 2 is used when the unit cell already is of low symmetry. In that case the symmetry of the special points is in general even lower, which may mean: no symmetry at all. This occurs, for instance, when slight distortions of a unit cell are invoked in order to determine phonon frequencies from the change in total energy upon distortion (frozen-phonon method [63]). But even in these situations the exploitation of crystal symmetry can

be of help. There will always be problems just past the limit of computational capacity available that can be handled with the use of symmetry but cannot be handled without its use.

CHAPTER 5

SECRETS DE CUISINE : CAPITA SELECTA

The present chapter is devoted to a discussion of some characteristic properties of the calculational scheme described in the previous chapters. Although these properties are of a somewhat technical nature, they certainly deserve attention, since they make possible a more convenient application of the method. A good understanding of these technical points is of great help, not only in reducing the computational work, but also in choosing reliable self-consistency criteria and cutoff parameters.

Here, we discuss three of such technical points:

- (i) In section 5.1 we show how the total energy is to be calculated in intermediate stages of the self-consistent (s.c.) calculation in order to obtain rapid convergence of E_{tot} .
- (ii) Section 5.2 is devoted to numerical noise on the total-energy-versus-volume curve, which is due to the finiteness of the cutoff parameter E_{pv} (see section 2.5) and the finite number of special points N_{sp} (see section 3.1).
- (iii) In section 5.3 we present support for the conjecture that special points are much more suited to integrate charge densities $\rho_{\mathbf{k}}(\mathbf{r})$ over 1BZ than they are to integrate energy bands $E_n(\mathbf{k})$ over 1BZ. This leads to an important computational reduction in the procedure to obtain the total energy.

The title of this chapter we owe to several of the lecturers at the Corsondonck Advanced Study Institute [1,16,67], who used the phrase in their lectures in connection with questions of the kind treated here.

5.1 Non-self-consistency correction

If eq. (2.58) is used to calculate the total energy per unit cell, E_{tot} , after each cycle in the self-consistency process -for instance, in order to judge to what accuracy E_{tot} has converged-, it should be realized that in the obtained E_{tot} the energy eigenvalues $E_n(\mathbf{k})$ are calculated by solving the set of equations (2.39) in which an input

screening potential is used, with Fourier components $V_{scr}^{in}(G)$ (see (2.69)). This input potential has been calculated by using a density n^{in} , which is different from the density n^{out} following from the solutions $C_{n,k}(G)$ of (2.39), with the aid of which the energy terms E_H^i and ΔE_{xc} in (2.58) are calculated. Only when self-consistency has been reached n^{in} and n^{out} are the same. In spite of this fact, it is possible to correct the "wrong" sum over band-structure energies $\sum_{(n,k)} E_n(k)$ -to be abbreviated by $\sum_{(i)} E_i$ in the following- for the fact that self-consistency has not been reached yet. After this correction, E_{tot} only depends on output quantities of one self-consistency cycle.

From eqs. (2.39) in chapter 2 the expression (2.57) was derived for $\sum_{(i)} E_i$. From this expression we infer that within each cycle of the self-consistent calculation the following identity holds (We use (2.54b) and (2.69)):

$$\sum_{(i)} E_i = E_{kin} + \Omega_c \sum_G V_{scr}^{in}(G) (n^{out}(G))^* + E_{ec}^i. \quad (5.1)$$

In (5.1) E_{kin} and E_{ec}^i are to be expressed, by means of (2.54a) and (2.54d), in terms of the solutions $C_{n,k}(G)$ of eqs. (2.39) (output quantities). These $C_{n,k}(G)$ are also used to calculate $n^{out}(G)$ using (3.22). So the RHS of (5.1) depends on n^{in} only through the Fourier components $V_{scr}^{in}(G)$. This dependence on n^{in} is explicitly removed by adding to $\sum_{(i)} E_i$ the *non-self-consistency correction* δ_{sc} :

$$\delta_{sc} = \Omega_c \sum_G \left[V_{scr}^{out}(G) - V_{scr}^{in}(G) \right] (n^{out}(G))^*, \quad (5.2)$$

where $V_{scr}^{out}(G)$ are Fourier components of the screening potential calculated from n^{out} . Since n^{out} may be expected to approximate the self-consistent density better than n^{in} , we expect $\sum_{(i)} E_i + \delta_{sc}$ to approximate the s.c. $\sum_{(i)} E_i$ better than expression (5.1).

From (5.1) and (5.2) we see that if V_{scr}^{out} by chance would be the fully converged s.c. screening potential, in the sense that all terms in the total energy have converged to within some prescribed accuracy, the correction results in the s.c. value for $\sum_{(i)} E_i$, whereas if V_{scr}^{in} is already s.c., the correction (by the definition of self-consistency) vanishes. These observations illustrate that δ_{sc} may be

regarded as a measure for the degree of self-consistency. Note that δ_{sc} in (5.2) includes the $G=0$ -term. We remind the reader that $V_H(G=0)$ was chosen to equal zero (see chapter 2), so only $V_{xc}(G=0)$ contributes to $V_{scr}(G=0)$. From eqs. (2.58), (2.59), and (5.1) one infers that the $G=0$ -term in (5.2) can be omitted if the energies $E_n(k)$ and the energy term ΔE_{xc} are calculated with $V_{xc}(G=0)$ put equal to zero. This may also be expressed by saying that E_{tot} given by (2.58) and corrected by adding (5.2) does not depend on the value of $V_{xc}(G=0)$ or by saying that self-consistency is only determined to within an additive constant in the potential.

As an illustration of the usefulness of δ_{sc} , we present in table 5.1 results of a s.c. calculation for cubic silicon carbide (with zincblende structure and denoted by 3C SiC) for the experimental value [64] of the lattice constant: $a_c = 4.3596 \text{ \AA}$. A kinetic-energy cutoff E_{pv} of 10.3 Ry and 2 special points ($q=4$ in section 3.2.1; Chadi-Cohen points) are used. Norm-conserving pseudopotentials from ref.[40] and the Kohn-Sham-Wigner (local) density functional for exchange and correlation are used (see chapter 2). The input screening potential for the first cycle is calculated from the valence-charge density obtained from an EPM-calculation using the form factors of ref.[65]. In table 5.1, for each of the presented eight cycles of the s.c. calculation the respective terms constituting the total energy per unit cell, E_{tot} , are given (cf. (2.58)), as well as δ_{sc} and a quantity δV defined as:

$$\delta V = \max_G |V_{scr}^{out}(G) - V_{scr}^{in}(G)|. \quad (5.3)$$

In (5.3) the maximum is taken over the set of all reciprocal-lattice vectors. In table 5.1, $E_{tot,corr}$ equals E_{tot} plus the correction δ_{sc} . The two last terms in the RHS of (2.58) are constants during the s.c. calculation and in this case are equal to -20.923537 Ry for the Ewald term and 4.416820 Ry for the constant split off from the ionic pseudopotentials.

The conclusion from table 5.1 is that the corrected total energy $E_{tot,corr}$ converges much faster than the uncorrected E_{tot} : after 3 cycles $E_{tot,corr}$ has converged to within an accuracy of 10^{-4} Ry, whereas E_{tot} takes 8 cycles to reach this accuracy.

	cycle 1	cycle 2	cycle 3	cycle 4
$\sum_{(i)} E_i$	- 1.581819	- 1.984517	- 1.995020	- 2.001523
ΔE_{xc}	1.698853	1.735050	1.734583	1.734244
E_H^0	1.286403	1.706164	1.703295	1.700233
E_{tot}	- 17.676086	- 18.462348	- 18.470449	- 18.474229
δ_{sc}	- 0.709153	- 0.020797	- 0.012979	- 0.009237
$E_{tot,corr}$	- 18.385239	- 18.483145	- 18.483428	- 18.483466
δV	5.3×10^{-2}	5.0×10^{-3}	2.7×10^{-3}	1.5×10^{-3}
	cycle 5	cycle 6	cycle 7	cycle 8
$\sum_{(i)} E_i$	- 2.008456	- 2.011289	- 2.012023	- 2.012206
ΔE_{xc}	1.734066	1.734064	1.734065	1.734065
E_H^0	1.698589	1.698554	1.698555	1.698557
E_{tot}	- 18.479696	- 18.482496	- 18.483230	- 18.483415
δ_{sc}	- 0.003778	- 0.000978	- 0.000244	- 0.000060
$E_{tot,corr}$	- 18.483474	- 18.483474	- 18.483474	- 18.483475
δV	4.3×10^{-4}	1.1×10^{-4}	2.6×10^{-5}	6.4×10^{-6}

Table 5.1. For each cycle of the self-consistency process, the quantities constituting the total energy per unit cell E_{tot} are given, as well as the non-self-consistency correction δ_{sc} (see eq. (5.2)), the corrected total energy per cell $E_{tot,corr}$, and the quantity δV , defined by eq. (5.3) (all entries are in Rydberg).

Table 5.1 is also helpful in clarifying a point of confusion in the literature (see footnote 12 in ref.[47]). In ref.[66] it was claimed that the total energy per atom was stable to within 10^{-5} Ry if δV was 10^{-4} Ry, whereas ref.[47] stated that with δV equal to 10^{-4} Ry the total energy per atom was only stable to within 10^{-3} Ry and that the remarks about stability in ref.[66] were not understood. This discrepancy of two orders of magnitude for the accuracy of E_{tot} is of course very relevant for the choice of a suitable self-consistency criterion for δV . Although the calculations in refs.[47] and [66] were performed for silicon and other details of the calculation are also different, table 5.1 strongly suggests that the calculations in ref.[66] were performed with the correction δ_{sc} , whereas in ref.[47]

this correction was ignored. We have found, also in other cases, that it is safe to assume that the total energy per atom calculated by (2.58) and corrected by adding $\delta_{s.c}$ differs from the s.c. result by less than 10^{-p-1} Ry if $\delta V = 10^{-p}$ Ry ($p=2,3,4,\dots$). This is true irrespective of the pseudopotentials and cutoffs used and we believe it to be characteristic for the calculational scheme.

We conclude by noting that the corrections discussed in refs.[67] and [68] are essentially the same as the correction introduced in this section.

5.2 Numerical noise on total-energy-versus-volume curves

The method presented in this work enables one to study the behaviour of the total energy as a function of the lattice parameters. In this way one may find for a given crystal structure and given types of atoms in the unit cell the set of equilibrium values of the lattice parameters, i.e., those values for which $E_{t.o.t}$ has a minimum. In case of one lattice parameter only, the unit-cell volume Ω_c is uniquely determined by this parameter, and by variation of this parameter a total-energy-versus-volume curve $E_{t.o.t}(\Omega_c)$ may be obtained. From this curve, various properties of the ground state of the crystal may be deduced (see chapter 6). If there are more lattice parameters, $E_{t.o.t}$ may be minimized with respect to variation of these parameters at some fixed value of Ω_c . An $E_{t.o.t}(\Omega_c)$ -curve is found by repeating this procedure for several Ω_c . If a unit cell contains more than one atom, we also have to consider additional *internal parameters*, that determine the positions of atoms in the unit cell. For instance, in the wurtzite structure described in section 3.4, c and a are lattice parameters and u is such an internal parameter. Concerning these internal parameters, a procedure similar to the one for lattice parameters can be followed to obtain $E_{t.o.t}(\Omega_c)$.

In the above procedure to obtain $E_{t.o.t}(\Omega_c)$, unfortunately, the finite cutoff $E_{p,v}$ (see section 2.5) and the finite (and usually very small) number of special points $N_{s,p}$ (see section 3.1) result in discontinuities in the $E_{t.o.t}(\Omega_c)$ -curves. These discontinuities cause difficulties in the interpretation of actually obtained values for $E_{t.o.t}$. This problem was already given some attention in refs.[69],

[70], and [71]. To simplify the discussion of these discontinuities, we restrict ourselves to the case of only one lattice parameter and no internal parameters. In this case discontinuities are most pronounced. An example is the zincblende structure with lattice parameter a_c (cell volume $\Omega_c = \frac{1}{4}a_c^3$; see section 3.4).

All values of the lattice parameter a for which discontinuities occur (denoted by a_d) are for given E_{PV} and given special-point set $\{\mathbf{k}_s\}$ ($s=1\dots N_{sp}$) found to be values for which the equality,

$$\frac{\hbar^2}{2m}|\mathbf{k}_s+\mathbf{G}|^2 = E_{PV} , \quad (5.4)$$

holds. In (5.4) \mathbf{G} is any reciprocal-lattice vector. Note in this respect that the expression in the left-hand side (LHS) of (5.4) is simply proportional to a^{-2} . Equation (5.4) follows from the criterion introduced in section 2.5, that selects PW's that are to be included in the expansion of the wave function. When increasing the lattice parameter, PW's with wave vector $\mathbf{k}_s+\mathbf{G}$ are abruptly *added* to the expansion of the wave function if the involved \mathbf{G} vectors are such that eq. (5.4) is satisfied. Such a discrete change in the basis set for special point \mathbf{k}_s results in a discontinuity in E_{tot} . One expects E_{tot} to drop at such values a_d due to the increased degree of freedom offered to the wave functions. It is furthermore expected that the discontinuity is larger if \mathbf{k}_s has a larger weighting factor w_s and also in cases in which the number of abruptly added PW's is larger. We define an effective change in the basis-set size ΔN :

$$\Delta N = \sum_{s=1}^{N_{sp}} w_s \Delta N(\mathbf{k}_s), \quad (5.5)$$

where $\Delta N(\mathbf{k}_s)$ is the change in basis-set size for special point \mathbf{k}_s . Of course the absolute magnitude of the discontinuity will also depend on whether ΔN is significant compared to the number of PW's already included in the basis set.

In table 5.2, we give values a_d in the neighbourhood of 5.43 Å (the experimental lattice constant for silicon [64]) using $E_{PV}=11.2$ Ry and two special points ($q=4$ in section 3.2.1). The corresponding ΔN are also given.

a_d (in Å)	ΔN	a_d (in Å)	ΔN
4.7297	3.25	5.2250	3.00
4.7816	6.00	5.2719	4.50
4.8328	2.25	5.3185	6.00
4.8836	3.75	5.3646	5.25
4.9338	6.00	5.4104	3.75
4.9835	1.50	5.4558	3.00
5.0327	4.50	5.5008	6.75
5.0815	6.00	5.5454	2.25
5.1297	1.50	5.5897	3.25
5.1776	3.00	5.6336	4.50

Table 5.2. Values a_d of the lattice parameter of a fcc lattice for which discontinuities in the $E_{tot}(\Omega_c)$ -curve occur, when using $E_{PV} = 11.2$ Ry and 2 special points. (Only values between 4.70 Å and 5.65 Å are given). The corresponding effective change in basis-set size (see eq. (5.5)) is denoted by ΔN .

In fig. 5.1, we show part of the total-energy-versus-lattice-constant curve for silicon in the diamond structure, calculated with the cutoffs as given above, clearly exhibiting the discontinuities. The numbers between brackets are the numbers of PW's in the basis set for the two k points (weighting factors: $w_1=3/4$, $w_2=1/4$).

There are two independent ways of reducing the discontinuities:

(i) Increasing $N_{s,p}$ will increase the number of a_d -values in a certain interval, but will reduce ΔN for a single a_d -value, since with increasing $N_{s,p}$ the average w_s decreases.

(ii) Increasing E_{PV} will also put consecutive a_d -values closer together (since the LHS in (5.4) is proportional to a^{-2}), but increases N_{PV} and therefore makes the effect on E_{tot} of the change ΔN less significant.

A way to completely get rid of the discontinuities is to keep N_{PV} constant for every k point (instead of E_{PV}). This can be achieved by letting E_{PV} scale as a^{-2} (see eq. (5.4)). However, in such a procedure the wave functions are not described with the same resolution for each value of the lattice parameter (see section 2.5). In practice, the

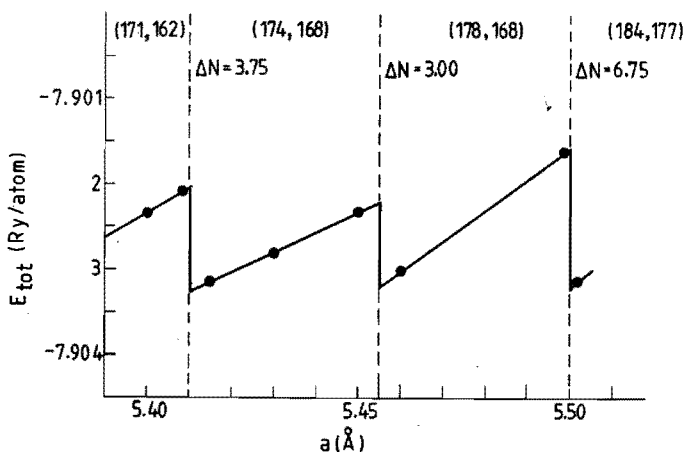


Fig. 5.1 Total energy E_{tot} (in Ry per atom) versus lattice constant a (in Å) for silicon in the diamond structure around the experimental value of a (5.43 Å) using $E_{PV} = 11.2$ Ry and 2 special points. Discontinuities occur at values of the lattice constant from table 5.2 because of an effective change in basis-set size ΔN . The numbers between brackets give the numbers of plane waves in the basis set for the two k points (weighting factors: $w_1=3/4$, $w_2=1/4$). Dots represent actually calculated values.

requirement of constant resolution appears to be more important to obtain converged values of total-energy differences [66,71]. We further remark that an exact procedure to keep N_{PV} constant in case of more than one lattice parameter generally cannot be given.

The magnitudes of the above described discontinuities not only depend on E_{PV} and N_{sp} , but also on the types of atoms in the unit cell, i.e., on the pseudopotentials. To illustrate this point, results are shown in fig. 5.2 of total-energy calculations for cubic silicon carbide (3C SiC) with $E_{PV}=20.6$ Ry and the same set of two special points as in the earlier example. The a_d -values and corresponding ΔN are given in table 5.3 (cf. fig. 5.2). The point to note is that although many more PW's are used in the expansion of the wave function (the numbers between brackets in fig. 5.2), the largest discontinuity is about a factor of 5 larger than in fig. 5.1 for Si. The reason for this is that the pseudopotential of the carbon atom is much deeper

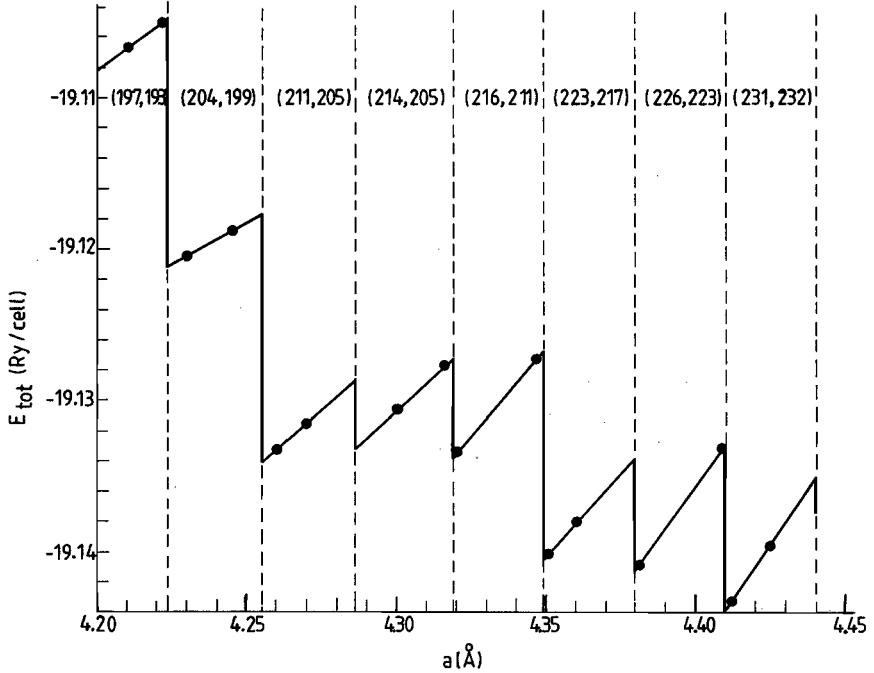


Fig. 5.2 Total energy E_{tot} (in Ry per cell) versus lattice constant a (in Å) for cubic silicon carbide around the experimental value of a (4.3596 Å) using $E_{PV} = 20.6$ Ry and 2 special points. Discontinuities occur at values of the lattice constant from table 5.3 because of an effective change in basis-set size ΔN . The numbers between brackets give the numbers of plane waves in the basis set for the two k points (weighting factors: $w_1=3/4$, $w_2=1/4$). Dots represent actually calculated values.

than the one for the silicon atom (see fig. 2.1). This asks in fact for the inclusion in the basis set of a still larger number of PW's. From fig. 5.2 one may also deduce an approximate proportionality of ΔN and the magnitude of the discontinuity (compare table 5.3 and fig. 5.2).

To reduce the effect of discontinuities on calculated total energies, we adopt the following strategy: we choose E_{PV} and N_{sp} such that the discontinuity for the largest ΔN occurring is smaller than

a_d (in Å)	ΔN	a_d (in Å)	ΔN
4.1594	4.50	4.3489	6.75
4.1915	3.75	4.3797	3.75
4.2235	6.75	4.4102	6.00
4.2552	6.75	4.4405	5.25
4.2866	2.25	4.4708	3.00
4.3179	3.00	4.5007	3.75

Table 5.3. Values a_d (between 4.15 Å and 4.52 Å) of the lattice parameter of a fcc lattice for which discontinuities in the $E_{t.o.t}(\Omega_c)$ -curve occur, when using $E_{PW} = 20.6$ Ry and 2 special points. The corresponding effective change in basis-set size (see eq. (5.5)) is denoted by ΔN .

some prescribed tolerance. It is furthermore advantageous to choose the lattice parameters for which $E_{t.o.t}$ is calculated not too close to each other nor to the a_d -values. In this way a sensible choice for E_{PW} and N_{sp} can be made, while circumventing the computationally much more demanding task of studying the convergence of properties derived from $E_{t.o.t}(\Omega_c)$ -curves as a function of E_{PW} and N_{sp} .

Finally, we remark that in case there are more lattice parameters the picture gets more complicated, since there will be points for which discontinuities occur distributed all over lattice-parameter space. On the $E_{t.o.t}(\Omega_c)$ -curve this will have the effect that the saw-tooth picture seen in figs. 5.1 and 5.2 turns into a curve with some noise on it.

5.3 Accuracy of energy-band integrations using special points

In this section we will show that in a total-energy calculation special points are much more suited to integrate the \mathbf{k} -dependent charge density $\rho_{\mathbf{k}}(\mathbf{r})$ over 1BZ than they are to integrate sums of energy bands $\sum_{(n)} E_n(\mathbf{k})$. In this connection we refer to section 3.3, where the convergence of the special-points approximation was studied when applied to a model energy band.

N_{sp}	$\sum_{(i)} E_i$	$\Delta(\sum_{(i)} E_i)$	N_{sp}	E_{tot}	ΔE_{tot}
2	- 2.0123	0	2	- 18.4835	0
6	- 2.0506	- 0.0383	6	- 18.5218	- 0.0383
10	- 2.0400	- 0.0277	10	- 18.5111	- 0.0276
19	- 2.0474	- 0.0351	19	- 18.5185	- 0.0350
28	- 2.0449	- 0.0326	28	- 18.5160	- 0.0325
44	- 2.0423	- 0.0300	44	- 18.5135	- 0.0300

(a) (b)

Table 5.4. Sum of one-electron energies and total energy per unit cell (in Rydberg) of cubic SiC with kinetic-energy cutoff $E_{PW} = 10.3$ Ry.

(a) Contribution $\sum_{(i)} E_i$ of one-electron energies to the total energy by summing over different numbers of special points N_{sp} . All one-electron energies are calculated with the potential that was self-consistently obtained with $N_{sp} = 2$. $\Delta(\sum_{(i)} E_i)$ is the difference with $\sum_{(i)} E_i$ for $N_{sp} = 2$.

(b) Total energy E_{tot} obtained self-consistently with various N_{sp} . ΔE_{tot} is the difference with E_{tot} for $N_{sp} = 2$.

We start from the s.c. calculation described in section 5.1, using two special points. The resulting s.c. values (after 8 cycles and including the correction δ_{sc} (see table 5.1)) are $E_{tot} = -18.4835$ Ry and $\sum_{(i)} E_i = -2.0123$ Ry. The s.c. potential resulting from this calculation with $N_{sp} = 2$ was now used to improve the approximation to $\sum_{(i)} E_i$ by calculating $\sum_{(i)} E_i$ using successively larger sets of 6, 10, 19, 28 and 44 special points ($q=6, 8, 10, 12, 14$ in section 3.2.1). From these improved values of $\sum_{(i)} E_i$ we calculate the difference $\Delta(\sum_{(i)} E_i)$ with the value calculated for $N_{sp} = 2$. The results are given in table 5.4(a). Next, we iterate to self-consistency (i.e., E_{tot} stable to within 10^{-4} Ry) using the same larger sets of special points also in the density calculation. The results are given in table 5.4(b) together with the difference ΔE_{tot} with the E_{tot} -value calculated using $N_{sp} = 2$. The fact to be noted is that ΔE_{tot} equals $\Delta(\sum_{(i)} E_i)$ to within an accuracy of 10^{-4} Ry. This shows that the improvement in the total-energy calculation by using larger sets of special points is entirely due to

a better determination of the band structure term and is not due to a better determination of the charge density ρ . Equivalently, regarding the calculation of E_{tot} , ρ is already sufficiently accurate using $N_{sp}=2$, whereas $\sum_{(i)} E_i$ is not. So in order to obtain the s.c. value of E_{tot} at a given E_{PV} and at a large value of N_{sp} , it is equally effective to iterate to self-consistency with a small value of N_{sp} and to supplement this by an improved calculation of $\sum_{(i)} E_i$ using the large value of N_{sp} . We have observed this in other calculations as well.

We have to emphasize that although the same value for E_{tot} is obtained in this way, this is not the case for the individual terms into which E_{tot} is decomposed. For instance, the fully s.c. value of $\sum_{(i)} E_i$ using $N_{sp}=6$ is -2.0605 Ry, which significantly differs from the value -2.0506 Ry in table 5.4(a).

The procedure suggested in this section may prove to be useful in calculations where, due to computational limitations, only a very small N_{sp} can be allowed in the s.c. calculation. As an example, we mention calculations for surfaces -these must be performed in a supercell geometry to establish periodicity-, where often only one special point is used. Our procedure now allows for an estimate of the accuracy of using one special point only, by calculating $E_n(\mathbf{k})$ for a larger set of special points \mathbf{k} only once, using the one-point self-consistent screening potential.

CHAPTER 6

APPLICATION TO SILICON, DIAMOND, AND SILICON CARBIDE

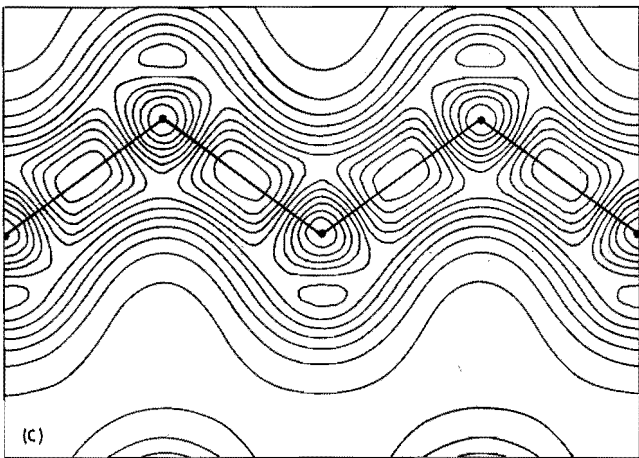
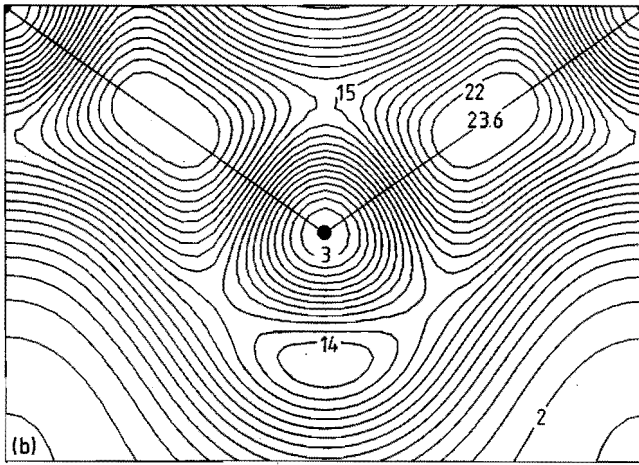
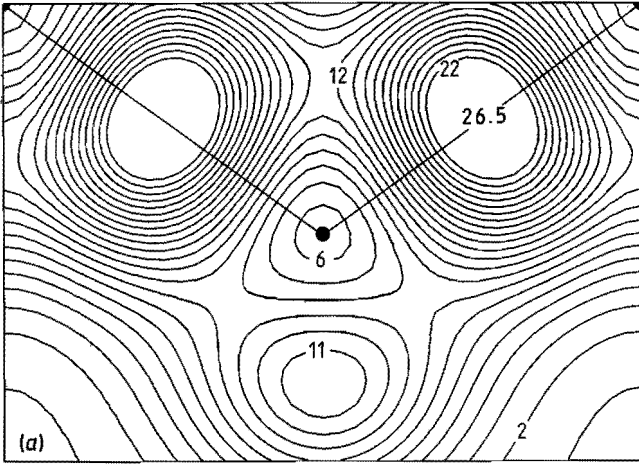
In the preceding chapters 2 to 5, the pseudopotential-density-functional method in momentum space has been described along with calculational details. These details involve the exploitation of the symmetry of lattice and crystal as well as characteristic properties of the calculational scheme. In the present chapter, we apply the method and techniques described to the prototypical semiconductor silicon (Si), the insulator diamond, and the cubic and wurtzite modifications of the semiconducting *polytypes* of silicon carbide (denoted by 3C SiC and 2H SiC, respectively).

Polytypism is the occurrence of different modifications of one compound, consisting of identical layers of structure whose stacking sequence differs [72,73]. The origin of polytypism is until now unexplained, although several theories exist, e.g., those based on screw-dislocations [74] or thermodynamic considerations [75]. Polytypism in the silicon carbides (SiC) -where it is studied most frequently- is interesting for a number of other reasons. The polytypes of SiC are semiconductors with a varying band gap, which may make them technologically important. Currently, they are most interesting, from a materials-science point of view, because of the strong fibres (so-called whiskers) that can be made from them [76]. From a fundamental point of view SiC is interesting -besides the above mentioned occurrence of polytypism- because its cubic modification, 3C SiC, is the only IV-IV compound with zincblende structure (e.g., GeC (germanium carbide) does not exist and GeSi (germanium silicide) occurs only in disordered structures). It therefore is an intermediate type of solid between III-V semiconductors (GaAs, AlAs, GaP, etc.) and crystals with the diamond structure (C, Si, Ge, α -Sn). Finally, the polytypes of SiC constitute a playground for the method described in this work, because they are natural superlattices, they pose problems similar to those found in artificially grown superstructures (e.g., large unit cells) and much can be learned from an application to SiC of how this method may be of use in connection with superlattices. Furthermore, the presence of the carbon atom requires large numbers of

plane waves in the basis set (see chapters 2 and 5) and makes application of the method computationally demanding. Therefore the limitations of the method can be explored as well as means to cope with these limitations.

Section 6.1 presents self-consistent valence-charge densities and screening potentials for silicon and diamond. Charge densities are given in the form of both contour plots, to get a view of the covalent bonding, and tables of Fourier components, to serve as reference and means of comparison with other calculations, as well as with experimental values. We also give the band structure at the high-symmetry points Γ , X, and L. Ground-state properties of Si and 3C SiC are obtained in section 6.2 by fitting results of total-energy calculations to equations of state for solids proposed in the literature. In section 6.3 the valence-charge density in 3C SiC is presented as well as the band structure resulting from the pseudopotential-density-functional method. We compare the band structure to other calculations and to experiment. The method of equivalent special-point sets -introduced in section 3.4- is used in section 6.4 to accurately determine the (small) energy differences between the structurally distinct crystals of (i) cubic-diamond and hexagonal-diamond Si and (ii) zincblende (3C) and wurtzite (2H) SiC. In section 6.5 we study the valence-charge density of 2H SiC, as well as the effect of variation of the lattice parameters in 2H SiC.

Fig. 6.1 The total valence-charge density of Si in the (110) plane. Units are numbers of electrons per unit-cell volume. The contourstep is 1 in (a) and (b), and 2 in (c). The black dots represent atomic positions and the straight lines connect bonded atoms. The bond-charge maxima are also given. (a) Result of EPM-calculation. (b) Self-consistent result of a calculation with norm-conserving pseudopotentials. (c) As (b), but showing a larger part of the same plane.



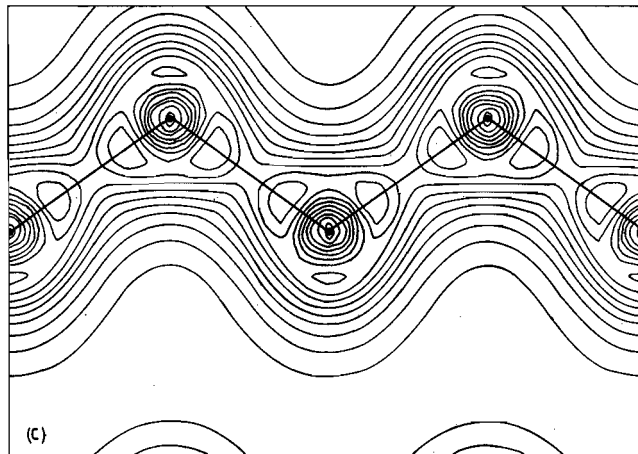
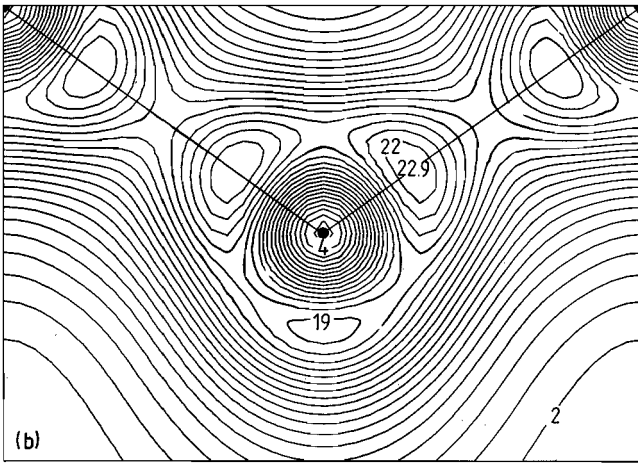
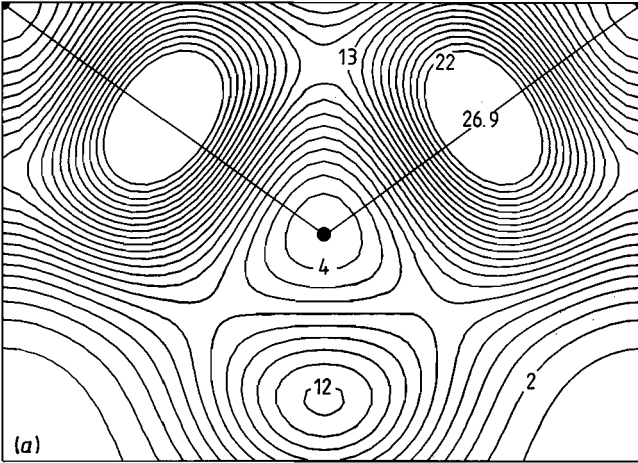
6.1 Self-consistent valence-charge density of silicon and diamond

We have calculated the self-consistent (s.c.) valence-charge density ρ of silicon in the diamond structure and carbon in the diamond structure (i.e., diamond). The diamond structure is the common structure for crystalline silicon, whereas carbon also crystallizes in the graphite structure (and is called graphite in that case). For both silicon and diamond we performed the calculations in this section for their experimental lattice constants: $a(\text{Si}) = 5.43 \text{ \AA}$ and $a(\text{C}) = 3.57 \text{ \AA}$ [64].

As start potential, we use the screening potential calculated from the valence-charge density obtained with an empirical pseudopotential (see section 2.3.3). In the latter calculation, we use the following form factors (in Rydberg): for Si: $\psi(3) = -0.2241$, $\psi(8) = +0.0551$, $\psi(11) = +0.0724$ [77], whereas for C: $\psi(3) = -0.696$, $\psi(8) = +0.337$, $\psi(11) = +0.132$ [65], where $\psi(q)$ is the form factor for a reciprocal-lattice vector with length q (in units $(2\pi/a)^2$; a is the lattice constant). With these form factors a band structure in agreement with experiment is found. The number of plane waves in the calculation with the empirical pseudopotentials (EPM) is about 90. We note that the s.c. ρ in our scheme should of course be independent of the start potential, although this cannot be proved rigorously. We did check, however, that in the calculation for silicon by starting with the s.c. screening potential obtained for diamond (scaled to the silicon crystal) the same s.c. ρ is obtained as by starting with the EPM-potential. As a matter of course, a good starting guess for the potential decreases the number of self-consistency cycles enormously.

Both in the EPM-calculation and in the s.c. calculation with norm-conserving pseudopotentials (SCNCP) we used two special points ($q=4$ in section 3.2.1) to integrate $\rho_{\mathbf{k}}(\mathbf{r})$ over 1BZ. As kinetic-energy cutoff E_{PV} (see section 2.5) we used 11.2 Ry for Si, resulting in about 170 plane waves in the basis set for each \mathbf{k} point, and 30.5 Ry for diamond, resulting in about 220 plane waves for each \mathbf{k} point. More computational details are given in ref.[44].

Fig. 6.2 Analogous to fig. 6.1, but for diamond.



In figs. 6.1 and 6.2 we show equi-density contour plots of the valence-charge density in the (110) plane. This plane contains a zigzag chain of bonded atoms and therefore a contour plot of this plane gives a qualitative idea of the bonding. Figure 6.1 gives results for silicon: fig. 6.1(a) shows the result from the EPM-calculation, fig. 6.1(b) from the SCNCP-calculation, whereas figure 6.1(c) represents the SCNCP-density in a larger part of the (110) plane. In all figures we see the feature characteristic of covalent bonding; the electrons are captured in bonds between the atoms. In fig. 6.1(c) we see that there are also regions with negligible density, which is the reason why crystals in the diamond structure are sometimes called *open structures*. The main difference between the results of the EPM- and SCNCP-calculation is the value of the *bond anisotropy ratio*, which is defined in ref.[78] as the ratio L_1/L_2 of the lengths of the outermost density contour surrounding the bond parallel (L_1) and perpendicular (L_2) to the bond axis, respectively. For figs. 6.1(a) and (b), L_1/L_2 equals 0.82 and 1.16, respectively. The experimentally observed ratio is 1.4, but the interpretation of the X-ray measurements is not completely rigorous [78,79]. Experimental results are always for the total electronic charge density and therefore include the core-electron charge density, which is not included in the (valence-)charge density calculated by us. Usually the core-electron density is subtracted from the experimental electron density by making some choice for the core contribution.

For diamond, plots are given in figs. 6.2(a)-(c). The most remarkable feature is that the EPM-density (fig. 6.2(a)) does not have the double hump in the bond charge found in the SCNCP-calculation. This double hump is a well established feature in the covalent carbon bond [24,80,81]; other groups [30,82] also find a very small double-hump structure in silicon by using a larger cutoff E_{PW} and more special points than we did. The fact that it is not found in the EPM-calculation for diamond illustrates the inadequacy of a local potential such as used in the EPM to describe the carbon core pseudopotential. Aside from the different bonding topology, fig. 6.2(c) for diamond is much like fig. 6.1(c) for silicon.

In tables 6.1 and 6.2 we compare the Fourier components of $n(\mathbf{r})$ as found in various calculations and those obtained from experiment. The experimental values are taken from tables in ref.[83], where they were

G	EPM	SCNCP ($N_{sp}=2$)	SCNCP ($N_{sp}=28$)	exp.
000	8.000	8.000	8.000	8.000
111	-1.658	-1.762	-1.747	-1.780
211	0.313	0.076	0.071	0.084
221	0.459	0.346	0.340	0.338
222	0.509	0.346	0.335	0.380
220	0.238	0.373	0.370	0.364
322	0.025	-0.087	-0.089	-0.052
332	-0.002	-0.138	-0.138	-0.126
333	-0.004	-0.112	-0.110	-0.097
320	-0.002	-0.076	-0.075	-0.076
422	0.021	-0.034	-0.034	-0.080

Table 6.1. Fourier components $n(\mathbf{G})$ (in numbers of electrons per cell volume) of the valence-charge density for silicon in the diamond structure, calculated with an empirical pseudopotential (EPM), self-consistently using a norm-conserving pseudopotential (SCNCP; with 2 and 28 special points), and "experimental" results from ref.[83] (see text). \mathbf{G} vectors are given as sets of components with respect to basis vectors $\{b_1, b_2, b_3\}$ (see section 3.2.1).

obtained from X-ray scattering data from which a core-electron density was subtracted (Therefore they cannot be taken too literally (see the discussion above)). From table 6.1 for Si, we see that n is a smooth function of r , since the Fourier components $n(\mathbf{G})$ drop rapidly with increasing $|\mathbf{G}|$. Furthermore, $n_{\mathbf{k}}(\mathbf{r})$ is a smooth function of \mathbf{k} , since the results obtained by using 28 special points ($q=12$ in section 3.2.1) do not differ by more than a few percent from those obtained with 2 special points. The agreement of the SCNCP-calculation with experiment is much better than that of the EPM-calculation. To illustrate the independence of the present results of the choice of norm-conserving pseudopotential, the choice of (local) XC-potential, and of kinetic-energy cutoff, we quote results from ref.[47], where values of $n(111)$ are given using the pseudopotential from ref.[14], the Ceperley-Alder functional for correlation, (2.66), and twice the number of plane waves we used. These values are: -1.7691 (electrons

G	EPM	SCNCP	exp.
000	8.000	8.000	8.000
111	-1.648	-1.911	-1.946
211	0.518	-0.322	-0.295
221	0.533	0.072	0.124
222	0.553	0.227	0.298
220	0.178	0.172	0.310
322	0.005	-0.047	0.009
332	0.020	-0.142	-0.060
333	-0.031	-0.129	-0.002
320	-0.025	-0.100	-0.002
422	0.027	-0.112	-0.097

Table 6.2. Analogous to table 6.1, but for diamond. SCNCP is calculated with 2 special points.

per cell) using $N_{sp}=2$ and -1.7539 (electrons per cell) using $N_{sp}=28$.

Similar general remarks can be made concerning table 6.2 for diamond.

Table 6.3 contains Fourier components of the exchange-correlation potential V_{xc} and of the Hartree potential V_H calculated from the $n(\mathbf{G})$ of tables 6.1 and 6.2 according to the description given in section 2.3.3. The Fourier components $V_{xc}(\mathbf{G})$ and $V_H(\mathbf{G})$ are generally larger (absolute valued) in diamond than in silicon. This can almost completely be attributed to the fact that these components scale with the reciprocal lattice parameter a^{-1} ; the components for diamond are generally a factor of about 1.5 larger than those for silicon, while $a(\text{Si})/a(\text{C}) = 1.52$. Put differently, the screening potential, which is the sum of V_{xc} and V_H , is -apart from the above scaling factor 1.52- nearly the same in diamond and silicon. So the main difference between silicon and diamond, concerning the effective potential the valence electrons experience, is constituted by the difference in ion-core pseudopotentials. One difference is that nonlocal contributions to the effective potential in diamond are much more important than in the case of silicon.

G	SILICON				DIAMOND			
	EPM		SCNCP		EPM		SCNCP	
	V_{xc}	V_H	V_{xc}	V_H	V_{xc}	V_H	V_{xc}	V_H
000	-0.6664	0	-0.6673	0	-0.9572	0	-0.9715	0
111	0.0501	-0.1372	0.0512	-0.1458	0.0799	-0.2073	0.0794	-0.2404
211	-0.0314	0.0097	-0.0075	0.0024	-0.0320	0.0244	0.0014	-0.0152
221	-0.0084	0.0104	-0.0068	0.0078	-0.0134	0.0183	-0.0029	0.0025
222	-0.0046	0.0105	-0.0021	0.0072	-0.0077	0.0174	-0.0014	0.0071
220	-0.0013	0.0037	-0.0062	0.0058	-0.0005	0.0042	-0.0042	0.0041
322	-0.0007	0.0003	0.0023	-0.0011	-0.0011	0.0001	0.0018	-0.0009
332	-0.0002	-0.0000	0.0026	-0.0014	-0.0002	0.0003	0.0036	-0.0022
333	0.0001	0.0000	0.0019	-0.0010	0.0002	0.0004	0.0030	-0.0018
320	0.0002	0.0000	0.0013	-0.0007	0.0004	0.0004	0.0025	-0.0014
422	0.0004	0.0002	0.0011	-0.0003	0.0010	0.0003	0.0031	-0.0013

Table 6.3. Fourier components $V_{xc}(G)$ and $V_H(G)$ of the screening potentials for silicon and diamond (in Rydberg), calculated with an empirical pseudopotential (EPM) and self-consistently with a norm-conserving pseudopotential (SCNCP). G vectors are given as sets of components with respect to basis vectors $\{b_1, b_2, b_3\}$ (see section 3.2.1).

The lowest few energy eigenvalues calculated with our s.c. potential at the high-symmetry points Γ , X , and L are given in table 6.4 and compared to previous calculations using the same method and to experimental values. The same E_{PW} as in the s.c. calculations are used. There is a good agreement between both calculations using the pseudopotential-density-functional method, although computational details differ slightly. The agreement with experiment is good for the valence bands, but large differences occur for the conduction bands. The latter discrepancy also holds for the indirect band gap, which for silicon we found to be 0.47 eV (with a conduction-band minimum at 0.85X), whereas the experimental value is 1.17 eV (conduction-band minimum at 0.82X). The bad agreement with experiments is a common feature of all calculations using the pseudopotential-density-functional method, but is not surprising in view of the fact that DFT

	SILICON			DIAMOND		
	This work	Yin/Cohen	exp.	This work	Yin/Cohen	exp. ^f
Γ_1	-11.74	-11.95	-12.5 ^b	-21.55	-21.45	-21.0
Γ_{25}'	0	0	0	0	0	0
Γ_{15}'	2.53	2.54	3.4 ^a	5.30	5.40	6.0
Γ_2'	3.50	3.39	4.2 ^a	13.15	13.38	15.3
X_1	-7.65	-7.80	—	-12.85	-12.65	—
X_4	-2.84	-2.92	-2.5 ^b	-6.36	-6.22	—
X_1	0.54	0.62	1.3 ^c	4.34	4.63	—
X_4	10.08	9.99	—	16.61	16.73	—
L_2'	-9.42	-9.57	-9.3 ^b	-15.73	-15.57	-15.2
L_1	-6.92	-7.01	-6.7 ^a	-13.63	-13.35	-12.8
L_3'	-1.18	-1.23	-1.2 ^a	-2.87	-2.81	—
L_1	1.50	1.52	2.1 ^d	8.60	8.92	—
L_3	3.29	3.37	4.2 ^e	8.16	8.31	—

Table 6.4. Energy eigenvalues (in eV) at symmetry points Γ , X , and L , calculated with the self-consistently determined potential. The Yin/Cohen values are taken from ref.[40]. The experimental values originate from: (a) ref.[64], (b) ref.[66], (c) estimated from conduction-band minimum and longitudinal effective mass, (d) ref.[86], (e) ref.[87], (f) ref.[88].

is a theory for the ground state. The discrepancy is caused by an inadequate inclusion of many-body effects [84]. Only recently the accurate calculation of excitation energies from first-principles has become possible [85].

6.2 Ground-state properties of silicon and cubic SiC

Some ground-state properties of a solid may be derived by calculating the total energy for different values of the lattice parameter(s). By varying the lattice constant of the diamond or zincblende structure, we simulate the effect of applying isotropic pressure to the crystal. In our method we are, contrary to experiments, not restricted to

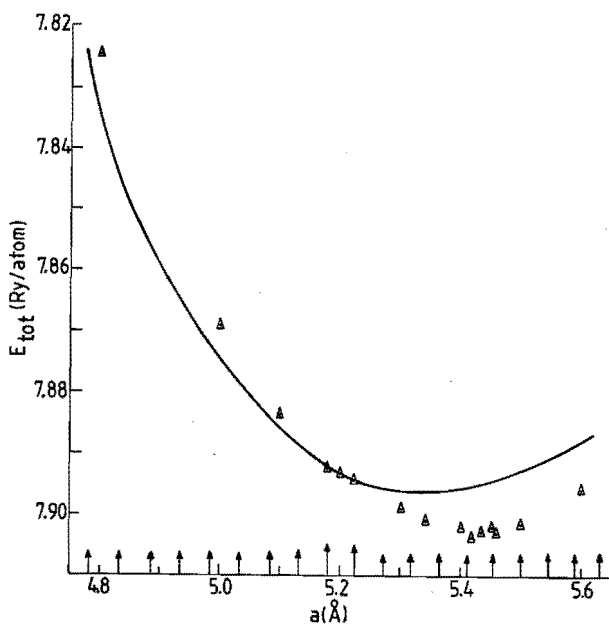


Fig. 6.3 Total energy E_{tot} of silicon (in Ry per atom) as a function of lattice constant a (in Å), keeping the number of plane waves constant (solid curve) and keeping the kinetic-energy cutoff E_{PV} constant (marked points). Arrows denote values of the lattice constant for which the curve through the marked points would exhibit discontinuities.

positive values of the pressure. The position of the minimum of the total-energy-versus-volume curve, $E_{tot}(\Omega_c)$, determines the equilibrium lattice constant a_{eq} , the value of the minimum is related to the cohesive energy of the crystal, whereas second and higher order derivatives of E_{tot} at a_{eq} determine the compressibility and its pressure dependence (this is made more explicit below). A preliminary report of the results in this section was given in ref.[70].

In fig. 6.3 we collect total-energy calculations for Si. In all calculations of this section, the same two special points as before, k_1 and k_2 , were used. The marked points indicate values obtained by keeping E_{PV} constant at 11.2 Ry. The curve through these points (not drawn) has discontinuities for values of the lattice constant denoted

by arrows on the horizontal axis (see also section 5.2 and table 5.2). The full curve in fig. 6.3 was calculated using fixed numbers of plane waves; 149 for k_1 and 150 for k_2 . This curve does not have discontinuities and coincides with the former curve for lattice constants in the interval between the two larger arrows in fig. 6.3. To the left of this interval the marked points are calculated using less plane waves than in the calculation of the full curve, whereas to the right more plane waves are used for the marked points. Because of the smaller and larger variational freedom offered to the wave functions, respectively, the marked points are situated above the full curve for $a < 5.18 \text{ \AA}$ and below the full curve for $a > 5.23 \text{ \AA}$.

We combine the results for the total energy with Murnaghan's equation of state for solids [89], which for the volume-dependent energy $E(V)$ takes the form:

$$E(V) = \frac{B_0 V}{B'_0} \left[\frac{(V_0/V)^{B'_0}}{B'_0 - 1} + 1 \right] + E_0, \quad (6.1a)$$

where E_0 is a constant equal to:

$$E_0 = E(V_0) - \frac{B_0 V_0}{B'_0 - 1}. \quad (6.1b)$$

In (6.1) V_0 is the equilibrium volume, B_0 the (equilibrium) bulk modulus defined by:

$$B_0 = V \left. \frac{d^2 E}{dV^2} \right|_{V=V_0}, \quad (6.2)$$

and B'_0 the derivative at zero pressure of the pressure-dependent bulk modulus $B(P)$. The pressure at constant temperature is defined as:

$$P(V) = - \frac{d}{dV}(E(V)), \quad (6.3)$$

and the pressure-dependent bulk modulus as:

$$B(P) = V \left. \frac{d^2}{dV^2}(E(V)) \right|_{V=V(P)}. \quad (6.4)$$

	a_{eq} (Å)	B_0 (Mbar)	B'_0	$E_{tot,min}$ (Ry/atom)
(1)	5.334	1.20	3.3	-7.896
(2)	5.452	1.00	2.6	-7.902
(3)	5.449	0.95	3.3	-7.903
exp.	5.429	0.99	4.2	-7.925

Table 6.5. Ground-state properties of Si, calculated in three different ways (see text) and compared to experiment. a_{eq} is the equilibrium lattice constant, B_0 the equilibrium bulk modulus, B'_0 the pressure-derivative of the bulk modulus, and $E_{tot,min}$ the minimum total energy.

Equation (6.1) is easily derived (for constant temperature) from the assumption that $B(P)$ varies linearly with pressure:

$$B(P) = B_0 + B'_0 P. \quad (6.5)$$

This assumption is certainly valid in the neighbourhood of $P=0$. The bulk modulus B_0 is the reciprocal of the isothermal compressibility, a quantity more commonly dealt with in thermodynamics.

We have made least squares fits of three sets of calculated points to Murnaghan's equation of state (6.1):

(1) Points on the solid curve of fig. 6.3, which were calculated with constant N_{PV} .

(2) Six points calculated with constant $E_{PV}=11.2$ Ry for lattice constants lying closest to the equilibrium value of the lattice constant, but not too close to values for which discontinuities occur (see the discussion in section 5.2).

(3) Eight points calculated with constant $E_{PV}=11.2$ Ry, selected as under (2).

The results of these fits are compared to experimental values [64] in table 6.5.

The quantity $E_{tot,min}$ is not directly accessible to experiment, since in the total energy E_{tot} calculated by us, the interaction energy within the cores is not accounted for, neither are zero-

temperature vibrations of the atoms. An expression for E_{tot} in terms of quantities that are accessible to experiment is obtained as follows: if we call $E_{core,j}$ the energy of an isolated core of atom j , and E_{vib} the energy due to zero-temperature vibrations of the atoms, we have for the actual total energy per unit cell of the crystal, $E_{tot,crystal}$ (in the frozen-core approximation (see section 2.2)):

$$E_{tot,crystal} = E_{tot} + \sum_j E_{core,j} + E_{vib}, \quad (6.6)$$

where the sum over j is over atoms in the unit cell. The energy of atom j , $E_{atom,j}$, may then be written as:

$$E_{atom,j} = - \sum_{i=1}^{Z_j} \phi_i(j) + E_{core,j}, \quad (6.7)$$

where $\phi_i(j)$ is the ionization potential of the i^{th} out of Z_j valence electrons of atom j . Since the cohesive energy per unit cell, E_{coh} , is the difference between the sum of $E_{atom,j}$ over atoms in the unit cell and $E_{tot,crystal}$, we obtain:

$$E_{tot} = - \left[\sum_j \sum_{i=1}^{Z_j} \phi_i(j) + E_{coh} + E_{vib} \right]. \quad (6.8)$$

The experimental values of $\phi_i(j)$ are found in ref.[90] for all atoms, while E_{coh} is also accessible to experiment. An experimental value for E_{vib} may be obtained from the formula [91]:

$$E_{vib} = \frac{9}{8} k_B \theta_0, \quad (6.9)$$

where k_B is Boltzmann's constant and θ_0 the Debye temperature, which can be obtained from experiment. Usually E_{vib} is small, but not negligible. This can be seen in table 6.6; we note in this respect that E_{tot} is of the order of 10 Ry per atom, but E_{coh} is of the order 0.5 Ry per atom.

In ref.[30] more than three times the number of plane waves (N_{PW}) were used than we did, resulting in $a_{eq} = 5.40 \text{ \AA}$ and $B_0 = 0.93$. From a comparison of the latter results with table 6.5, we conclude that

	θ_D (in K)	E_{vib} (in Ry/atom)
Si	645	0.0046
C	2220	0.0158
3C SiC	1430	0.0102

Table 6.6. Debye temperature and associated zero-temperature vibration energy (see text) for silicon, diamond, and cubic silicon carbide. Values for θ_D are extrapolated to $T = 0$ K [92].

keeping E_{PV} constant (results denoted by (2) and (3) in table 6.5) gives better convergence of ground-state properties than keeping N_{PV} constant (denoted by (1) in table 6.5). We also conclude that with a cutoff of $E_{PV}=11.2$ Ry a_{eq} and B_0 have converged to within 1% and 7%, respectively. Taking into account the numerical uncertainty due to discontinuities, the value of B_0' cannot be determined very well, since energies resulting from the equation of state (6.1) are not very sensitive to the value of B_0' .

From the $E_{tot,min}$ -value in table 6.5, we obtain a theoretical (calculated) value for the cohesive energy by subtracting the sum of $E_{tot,min}$ and E_{vib} ($=-7.898$ Ry per atom) from the energy of an isolated pseudo-atom $E_{ps,atom}$ ($=-7.497$ Ry); to $E_{ps,atom}$ we first add in a rather ad hoc fashion a correction for the spin-polarisation energy E_{sp} ($=-0.058$ Ry per atom [66,93]). A pseudo-atom is an atom without interaction within the core; for the calculation of $E_{ps,atom}$ the same pseudopotential is used as in the calculation for the solid [94]. We have:

$$E_{coh} = (E_{ps,atom} + E_{sp}) - (E_{tot,min} + E_{vib}). \quad (6.10)$$

This results in a theoretical value of $E_{coh} = 0.343$ Ry/atom = 4.67 eV/atom, whereas the experimental value is 4.63 eV/atom [95]. The excellent agreement is in fact slightly misleading, since $E_{tot,min}$ drops by an amount of about 0.3 eV/atom when E_{PV} approaches infinity [30,66].

Similar calculations have been performed for 3C SiC. As for Si, 2 special points are used and E_{PV} is chosen such that the largest

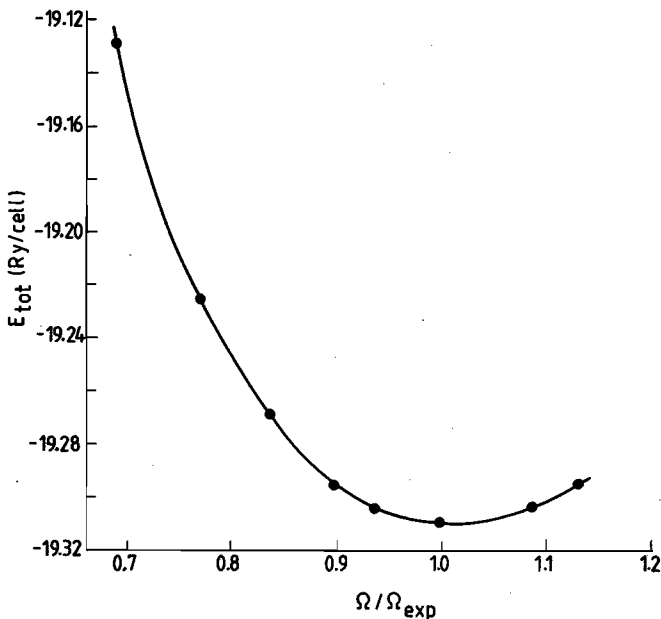


Fig. 6.4 Total energy E_{tot} of 3C SiC (in Rydberg per cell) as a function of the ratio of volumes $\Omega/\Omega_{\text{exp}}$ (Ω_{exp} denotes the experimental volume). The solid curve is the fit to Murnaghan's equation of state (see text).

discontinuities in the $E_{\text{tot}}(\Omega_c)$ -curve are about as large as these are for Si with $E_{\text{PV}}=11.2$ Ry and 2 special points. We find $E_{\text{PV}}=29.7$ Ry, leading to about 380 plane waves in the basis set if the lattice constant is around the experimental value. In that case the discontinuities are always smaller than 2 mRy/atom. By choosing lattice constants for which E_{tot} is calculated midway between a_d -values (see section 5.2), we estimate the uncertainty of calculated total-energy differences due to finite E_{PV} and N_{sp} to be 1 mRy/atom. The $E_{\text{tot}}(\Omega_c)$ -curve is given in figure 6.4, together with the fit to Murnaghan's equation of state.

In table 6.7 we compare results of various total-energy calculations and various fits with each other and with experiment:

- (1) Fit to Murnaghan's equation of state of total energies calculated using the lower cutoff $E_{\text{PV}}=20.6$ Ry for values of the lattice constant selected as in the case of silicon (see also fig. 5.2).

	a_{eq} (Å)	B_0 (Mbar)	B'_0	$E_{tot.min}$ (Ry/cell)
(1)	4.411	3.08	2.1	-19.140
(2)	4.351	2.00	7.3	-19.308
(2')	4.351	2.09	7.4	-19.308
(3)	4.365	2.20	3.4	-19.309
(3')	4.364	2.17	3.8	-19.309
exp.	4.360	2.24*	—	-19.410

Table 6.7. Ground-state properties (as in table 6.5) of 3C SiC, calculated in 5 different ways (see text) and compared to experiment. The star denotes that the experimental B_0 is only an estimate [97].

(2) Fit of the 6 points from fig. 6.4 ($E_{PV}=29.7$ Ry) closest to a_{eq} to Murnaghan's equation of state.

(2') Same points as (2), but fitted to the energy-volume relation which follows from Birch's equation of state for solids [96]:

$$E(V) = \frac{9B_0V_0}{16} \left[(B'_0-4) \left[\frac{V_0}{V} \right]^2 + (14-3B'_0) \left[\frac{V_0}{V} \right]^{4/3} + (3B'_0-16) \left[\frac{V_0}{V} \right]^{2/3} \right] + E_0, \quad (6.11a)$$

where

$$E_0 = E(V_0) - \frac{9B_0V_0}{16} (B'_0 - 6), \quad (6.11b)$$

and B_0 , V_0 and B'_0 have the same meaning as in (6.1).

(3) Fit of all 8 calculated points from fig. 6.4 to eq. (6.1).

(3') Same points as (3), but fitted to (6.11).

Concerning table 6.7 we make the following remarks:

(1) Fits to (6.1) and to (6.11) give practically the same results (compare (2) with (2') and (3) with (3')).

(ii) The uncertainty due to the fitting and the choice of equation of state, which does not follow from first principles, is estimated by comparing (2) and (3). For this type of uncertainty we have: 0.3% for a_{eq} , 10% for B_0 , less than 10^{-3} Ry/cell for $E_{tot.min}$, whereas B'_0 can vary by a factor of 2.

(iii) The uncertainty due to the cutoff $E_{PV}=29.7$ Ry is estimated to be the same as in silicon with $E_{PV}=11.2$ Ry: less than 1% for a_{eq} and about 10% for B_0 . This estimate follows from the fact that energy

differences are all that is needed to calculate a_{eq} , B_0 , and B'_0 , while these differences have the same uncertainty in SiC for $E_{PV}=29.7$ Ry as in Si for $E_{PV}=11.2$ Ry.

(iv) From the results for $E_{PV}=20.6$ Ry, we infer that a_{eq} converges faster than the other quantities.

(v) In view of (ii) and (iii), the agreement with experiment of the results for $E_{PV}=29.7$ Ry is excellent for a_{eq} , although we expect a_{eq} to drop when more plane waves are used, but not by more than 1% (just as in silicon). The experimental value of B_0 is only an estimate [97]; B_0 is hard to measure, because the relatively large crystals of 3C SiC that are needed are hard to fabricate. Both the calculated and experimental value of B_0 are intermediate between the bulk moduli of silicon and diamond (0.99 Mbar and 4.42 Mbar, respectively). They furthermore agree with the result of a semi-empirical calculation [98] of the elastic constants C_{11} and C_{12} . These are related to B_0 by the equality, $B_0 = (C_{11} + 2C_{12})/3$ (cubic crystals [99]). From ref. [98] we then infer: $B_0 = 2.11$ Mbar. An experimental value of B'_0 has never been reported, but since silicon and diamond both have B'_0 close to 4, the same is expected to be true for 3C SiC. The experimental value of $E_{tot,min}$ is calculated as explained above (eq. (6.8)); we obtain the experimental cohesive energy of 3C SiC needed in this calculation from the cohesive energies of Si and C [95] and the heat of formation of 3C SiC [92], following ref.[100]: $E_{coh}(3C) = 6.33$ eV/atom.

Recently, other calculations applying the pseudopotential-density-functional method to 3C SiC have been reported; these calculations use a larger E_{PV} . The results are in general agreement with table 6.7 and the above remarks about the accuracy of the calculated properties: $a_{eq} = 4.323$ Å, $B_0 = 2.50$ Mbar, $B'_0 = 3.2$ in ref.[101] and $a_{eq} = 4.318$ Å, $B_0 = 2.34$ Mbar in ref.[102].

To complete the discussion of ground-state properties of 3C SiC obtainable from our calculations, we show in table 6.8 the dependence of E_{tot} on E_{PV} and N_{PV} . All these calculations use 2 special points and the experimental lattice constant. The experimental value of E_{tot} is calculated as explained with table 6.7. From fig. 1 in ref.[102] we estimate E_{tot} to be -19.40 Ry/cell for $E_{PV}=50$ Ry. We conclude that by increasing E_{PV} E_{tot} converges to a value very close to our experimental estimate. From the lowest E_{tot} -value in table 6.8 (for $E_{PV}=36$ Ry), we calculate the cohesive energy for 3C SiC in precisely

N_{PV}	E_{PV} (Ry)	E_{tot} (Ry)
(77,74)	10.3	-18.48
(143,150)	15.2	-18.93
(223,217)	20.6	-19.14
(262,274)	23.3	-19.21
(382,374)	29.7	-19.31
(508,516)	36.0	-19.35
	exp.	-19.41

Table 6.8. Calculated total energy E_{tot} (in Ry per cell) of 3C SiC as a function of kinetic-energy cutoff E_{PV} for the experimental value of the lattice constant $a_{exp} = 4.3596 \text{ \AA}$ [64]. Under N_{PV} the numbers of plane waves for the two k points used are listed. The experimental value of E_{tot} was deduced from the experimental values of the cohesive energy E_{coh} , the Debye temperature Θ_D , and the ionization potentials ϕ_i of the atoms Si and C (see eq. (6.8)).

the same way as we did for silicon. Besides quantities for the silicon pseudo-atom given before, we need: $E_{p.s. atom}(C)$ ($=-10.676$ Ry/atom [94]) and $E_{sp}(C)$ ($=-0.103$ Ry/atom [24]). We thus find a cohesive energy for 3C SiC of 6.75 eV/atom; in reasonable agreement with the experimental value of 6.33 eV/atom. The agreement becomes less if the more converged value for E_{tot} of ref.[102] is used: $E_{coh} = 7.09$ eV/atom. This general feature of overestimating the cohesive energy is mainly due to the inadequacy of the use of pseudopotentials (and/or the local-density approximation) to calculate the energy of the pseudo-atom. This can be inferred from the fact that the calculated E_{tot} converges, with increasing E_{PV} , to a value close to the experimental E_{tot} , whereas the converged, calculated E_{coh} overestimates the experimental E_{coh} . The additional ingredient to the calculation of E_{coh} is basically the calculated pseudo-atom energy.

Since all results in this section are obtained using only 2 special points to perform integrations over 1BZ, we may conclude that energy differences can be determined quite accurately with this small number of special points. This is due to the fact that technical approximations are made in an equivalent way (see sections 2.5 and 3.4).

6.3 Valence-charge density and band structure of cubic SiC

In fig. 6.5 we compare valence-charge densities in 3C SiC, calculated for the experimental lattice constant. Figure 6.5(a) is the result of an EPM-calculation using the form factors of ref.[65] and about 90 plane waves (PW's) for the expansion of the wave functions. Figure 6.5(b) results from a self-consistent calculation with norm-conserving pseudopotentials (NCPSP) using about 380 PW's. Apparently the empirical pseudopotential for Si is too weak compared to the one used for C. This is less an effect of the locality of the atomic carbon pseudopotential -as was the case for the discrepancy between EPM- and SCNCP-densities for diamond (section 6.1)- than of the fact that the empirical pseudopotentials were not submitted to transferability criteria (see section 2.2).

In fig. 6.5(b) we see that ρ displays features typical of both ionic bonding (superposed spherical charge distributions centred on the atomic positions) and covalent bonding (charge accumulated in bonds between the atoms), illustrating that 3C SiC is partially ionic. The charge density resembles the one for III-V compounds (e.g., GaAs [100]). However, in SiC the ionicity of the bond is not the result of a difference in charge of the two cores, but of a difference in size of the two cores.

The band structure of 3C SiC is shown in fig. 6.6, calculated with the EPM-potential (dashed lines) as well as with a self-consistent screening potential and norm-conserving ionic potentials (SCNCP; full lines). In the latter calculation $E_{PV}=23.3$ Ry is used; employing $E_{PV}=29.7$ Ry changes individual energy levels by less than 0.2 eV. The three highest valence-band levels from both calculations agree very well, but the valence-band width differs significantly: 15.4 eV for SCNCP and 19.4 eV for EPM. The form factors of the EPM-calculation were adjusted to reproduce -among other optical transitions- the experimental indirect Γ to X band gap of 2.40 eV. The SCNCP-result for the gap is almost 60% off (just as for silicon; see the discussion of table 6.4).

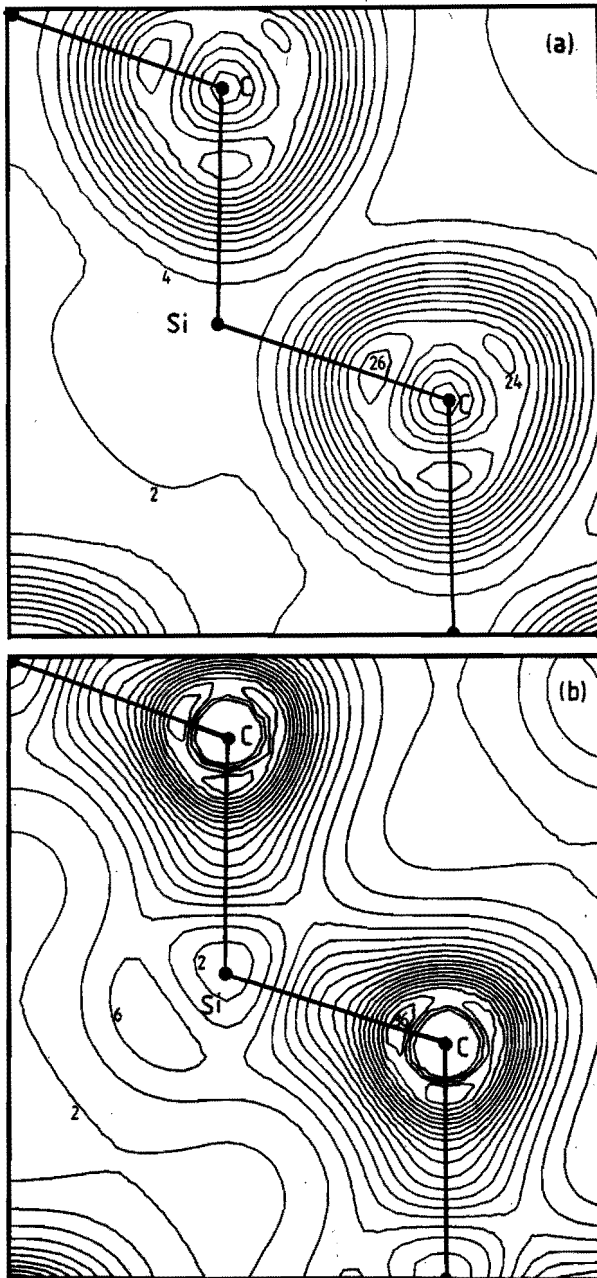


Fig. 6.5 Total valence-charge density of 3C SiC in the (110) plane. Units are numbers of electrons per unit-cell volume. The contour step is 2. Black dots represent atomic positions and straight lines connect bonded atoms. (a) Result of EPM-calculation. (b) Self-consistent result using norm-conserving pseudopotentials.

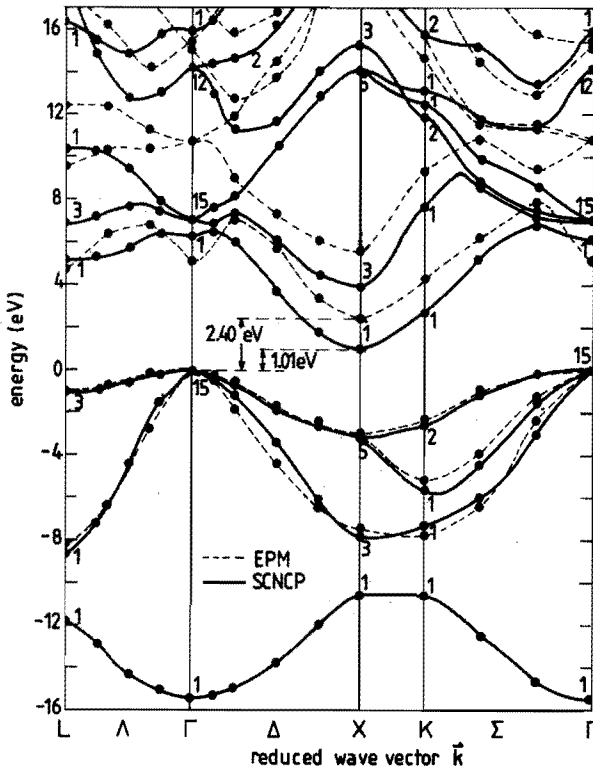


Fig. 6.6 Band structure of 3C SiC, calculated with an empirical pseudopotential (EPM) and self-consistently with norm-conserving pseudopotentials (SCNCP). Labels at high-symmetry points denote the irreducible representation of the band concerning.

In table 6.9 we compare the SCNCP-band-structure at high-symmetry points for $E_{PV}=29.7$ Ry with an all-electron density-functional calculation using the Augmented-Spherical-Wave (ASW) method modified to treat open structures [103,104]. We also compare with experimental results, which are deduced from tables in refs.[64,65,105]. There is reasonable agreement between pseudopotential and all-electron calculations. SCNCP and ASW agree with experiment for the valence bands, but disagree for the conduction bands. The dispersion of the conduction bands also disagrees between SCNCP and experiment, i.e., the discrepancy varies with the reduced wave vector k .

	SCNCP	ASW	exp.
Γ_1	-15.2	-15.2	—
Γ_{15}	0	0	0
Γ_1	6.4	5.8	6.0 ^{b, c}
Γ_{15}	7.1	7.1	7.8 ^a
X_1	-10.3	-10.4	—
X_3	-7.8	-7.8	-7.3 ^a
X_5	-3.1	-3.2	-2.8 ^{a, b}
X_1	1.1	1.4	2.4 ^{a, b, c}
X_3	4.0	4.2	5.5 ^b
L_1	-11.7	-11.8	—
L'_1	-8.5	-8.5	—
L_3	-1.0	-1.1	-1.2 ^c
L_1	5.3	5.1	4.2 ^c
L_3	7.0	7.2	8.5 ^{a, c}

Table 6.9. Energy eigenvalues for 3C SiC at high-symmetry points Γ , X , and L . SCNCP denotes results from the present method and ASW denotes results from all-electron calculations with the Augmented-Spherical-Wave method [104]. The experimental values originate from tables in: (^a) ref.[105], (^b) ref.[65], (^c) ref.[64].

6.4 Accurate energy differences and equivalent special-point sets

In section 6.2 we showed that accurate energy differences between crystals with the same crystal structure, but with different lattice constants can be obtained if technical approximations are made in an equivalent way, meaning constant E_{PV} and special points chosen as explained in section 3.4. In the latter section we already argued that it is more difficult to obtain such accuracy if the crystals are structurally different. We developed a systematic way to proceed in the latter case and introduced *equivalent special-point sets* (ESPS). In this section we give an application of this approach by calculating total-energy differences δE_{tot} between cubic-diamond (CD) silicon and hexagonal-diamond (HD) silicon as well as between zincblende SiC (3C) and wurtzite SiC (2H). These structures are described in section 3.4. The energy differences between the structures mentioned are expected to be very small, since each atom has 4 nearest and 12 second-nearest neighbours at the same distances in both structures.

Some of the results in this section were published in ref.[106].

In table 6.10 we present results of total-energy calculations for CD and HD silicon using the ESPS of table 3.4. For CD the experimental lattice constant $a_c=5.43 \text{ \AA}$ is used, whereas for HD the corresponding lattice parameters $a = \frac{1}{2}a_c\sqrt{2}$ and $c = (2a/3)\sqrt{6}$ are used (see section 3.4). We remark that HD silicon is a hypothetical solid, since it is not found in nature. This, however, does not prevent us from a calculation with our method. It is gratifying to see in table 6.10 that by employing successively larger special-point sets, the change in total energy has the same sign for both structures, when applying ESPS. However, the fact that δE_{tot} converges not faster than the individual total energies is disappointing. The latter fact merely indicates that the special-point sets for both structures are not equivalent enough.

We remark that the difference in Ewald energy (see section 2.3) of 0.0124 Ry/atom [107] is apparently canceled almost completely by the other terms in the total energy.

$(N_{sp}(fcc), N_{sp}(hex))$	(2,3)	(6,8)	(10,14)	(19,20)	(28,30)
$E_{tot}(CD)$	-7.9029	-7.9079	-7.9085	-7.9087	-7.9086
$E_{tot}(HD)$	-7.8944	-7.9040	-7.9072	-7.9077	-7.9077
δE_{tot}	0.0085	0.0039	0.0013	0.0010	0.0009

Table 6.10. Total energies (in Ry per atom) for cubic-diamond (CD) and hexagonal-diamond (HD) silicon using equivalent special-point sets. $\delta E_{tot} = E_{tot}(HD) - E_{tot}(CD)$ (in Ry per atom). $N_{sp}(fcc)$ and $N_{sp}(hex)$ are the numbers of special points used for CD and HD, respectively. A kinetic-energy cutoff $E_{PV} = 11.2$ Ry was used.

In table 6.11(a) and (b) we show results of total-energy calculations (in Ry per pair Si-C) for 3C SiC and 2H SiC, respectively, using the ESPSPs of table 3.4 and various values of the cutoff E_{PV} . Table 6.11(c) gives the difference in total energies in mRy per pair Si-C calculated with the same E_{PV} and equivalent special-point sets. For 3C the experimental lattice constant is used, whereas for 2H the corresponding c and a are used (see above). In the table crosses denote values that could not be obtained in less than 6 hours of processing time on a Burroughs 7900 computer; values marked by a star have been calculated on a CYBER 205 vector computer.

It is clear from table 6.11(a) and (b) that just as in section 6.2 energy differences are to be calculated by subtracting total energies calculated at the same E_{PV} , implying the same resolution for wave functions in r -space. Since the unit-cell volume of 2H is exactly twice the unit-cell volume of 3C, according to (2.82), the number of plane waves N_{PV} then is approximately twice as large in the 2H-calculation. Since computing time goes up with a power of N_{PV} between 2 and 3, calculations are more time consuming for 2H by a factor between 4 and 8 (Therefore table 6.11(b) contains more crosses than table 6.11(a)). We also infer from table 6.11(a) and 6.11(b) that the convergence with respect to increasing N_{sp} is slightly better the higher E_{PV} is. However, certainly for $E_{PV} = 10.3$ Ry and $E_{PV} = 15.2$ Ry the convergence is considerably less than for silicon with $E_{PV} = 11.2$ Ry (cf. table 6.10). We speculate that in order to obtain δE_{tot} with an

accuracy of 10^{-4} Ry/pair Si-C a cutoff $E_{PV}=29.7$ Ry and the (28,30) ESPS are required (cf. section 6.2).

The bad convergence with respect to increasing N_{sp} , combined with the fact that our ESPS are probably not equivalent enough, results in the situation that we are not able to accurately determine the energy difference between 3C and 2H (see table 6.11(c)). The feature of "non-equivalence" is also reflected in the fact that the sign of δE_{tot} is consistently negative for the (10,14) ESPS. Computational limitations restrict us to a best estimate from the $E_{PV}=15.2$ Ry results: $\delta E_{tot} = 1.1 \pm 0.9$ mRy/pair Si-C. This energy difference is of the same order of magnitude as δE_{tot} of CD and HD silicon. Very recently δE_{tot} of 3C and 2H was calculated using the supercell approach described in section 3.4 [57], in which better equivalence of the two special-point sets is achieved at the cost of larger computing times. In the latter calculation, plane waves with kinetic energy below 10 Ry are treated exactly as we did, whereas those with kinetic energy between 10 and 20 Ry are included by means of a perturbation technique. This results in $\delta E_{tot} = 0.73 \pm 0.02$ mRy/pair Si-C, in agreement with our result above. In the latter error estimate only the error of the special-points approximation is included and not the error due to the finiteness of E_{PV} .

Just as for CD and HD silicon, the difference in Ewald energy between 3C and 2H, 30.7 mRy/pair Si-C, is canceled almost completely by the other terms in the total energy.

Table 6.11. Total energies (in Ry per pair Si-C) for (a) 3C SiC and (b) 2H SiC, using various kinetic-energy cutoffs E_{PV} and various numbers of special points N_{sp} . (c) gives energy differences $\delta_{SiC} = E_{tot}(2H) - E_{tot}(3C)$ (in mRy per pair Si-C) found from total energies at the same E_{PV} and with (equivalent) special-point sets as indicated. Crosses denote values not obtainable within 6 hours of processing time on a Burroughs 7900 computer and stars denote values obtained on a CYBER 205 vector computer.

$N_{sp}(\text{fcc})$	2	6	10	19	28	44
E_{PV} (Ry)						
10.3	-18.4835	-18.5218	-18.5111	-18.5185	-18.5160	-18.5135
15.2	-18.9298	-18.9257	-18.9249	-18.9270	-18.9264	-18.9258
20.6	-19.1379	—	—	—	—	—
23.3	-19.2118	-19.2226	-19.2221	—	—	X
29.7	-19.3065	—	—	X	X	X
36.0	-19.3473	—	X	X	X	X

(a)

$N_{sp}(\text{hex})$	3	8	14	20	30	42
E_{PV} (Ry)						
10.3	-18.4727	-18.5038	-18.5131	-18.5137	-18.5127	-18.5151*
15.2	-18.8988	-18.9173	-18.9250	-18.9229	-18.9244	-18.9247*
20.6	-19.1224	—	—	X	X	X
23.3	-19.1986	-19.2175	-19.2229*	X	X	X
29.7	-19.2951	X	X	X	X	X
36.0	X	X	X	X	X	X

(b)

$(N_{sp}(\text{fcc}), N_{sp}(\text{hex}))$	(2,3)	(6,8)	(10,14)	(19,20)	(28,30)	(44,42)
E_{PV} (Ry)						
10.3		10.8	18.0	-2.0	4.8	3.3
15.2		31.0	8.4	-0.1	4.1	2.0
20.6		15.5	—	—	X	X
23.3		13.2	5.0	-0.8*	X	X
29.7		11.4	X	X	X	X
36.0		X	X	X	X	X

(c)

6.5 Wurtzite SiC: mapping and relaxation

In this last section of this chapter we compare the valence-charge density ρ in wurtzite SiC (2H SiC) with ρ in 3C SiC. Furthermore, we briefly discuss the relaxation of the lattice parameters c and a , i.e., we search for the minimum of E_{tot} when these parameters are allowed to vary.

In fig. 6.7 we show the valence-charge density in the equivalent planes of 2H SiC and 3C SiC that were described with fig. 3.3. Both were calculated using $E_{PV}=15.2$ Ry, and the usual 2 special points for 3C and 3 special points for 2H ($q_a=3$, $q_c=2$ in section 3.2.2). The similarity of figs. 6.7(a) and (b) is in accordance with the fact that polytypes only differ in the way of stacking layers identical in structure. A suitable choice for such a layer is indicated in fig. 6.7(b); the height of a layer is $\frac{1}{2}c$, half the height of the unit cell of 2H, and it contains one plane in which silicon atoms are arranged in equilateral triangles and one plane in which carbon atoms are arranged in equilateral triangles (see also fig. 3.4). A fairly good picture of the electron density in 2H SiC can be obtained from the density in 3C SiC by keeping the central layer (between dashed lines) fixed and rotating the layers just above and below over 180° around the bond axes protruding into these layers.

More precisely, a transformation equivalent to the transformation described above is formulated as follows (*density mapping*): suppose we have calculated the density in 3C, ρ_{3C} , at positions $r = \xi t_1 + \eta t_2 + \zeta t_3$, where t_1, t_2 , and t_3 are basis vector of the hexagonal Bravais lattice of 2H SiC. We choose the origin in a C-type position, i.e., the type of position not occupied by atoms in fig. 6.7(a) (see also fig. 3.4). The density for 3C is transformed into a density $\tilde{\rho}_{2H}$ for 2H by:

$$\tilde{\rho}_{2H}(\xi, \eta, \zeta) = \begin{cases} \rho_{3C}(\xi, \eta, \zeta) & 0 \leq \xi, \eta < 1, 0 \leq \zeta < \frac{1}{2}, \\ \rho_{3C}(\xi - \eta, \xi, \zeta - \frac{1}{2}) & 0 \leq \xi, \eta < 1, \frac{1}{2} \leq \zeta < 1. \end{cases} \quad (6.12)$$

Equation (6.12) shows that the density in the complete unit cell of 2H SiC is obtained from the density in one layer of 3C SiC. The transformation defined by (6.12) rotates the density in the "building-block" layer over 60° around the z -axis through a C-type position and

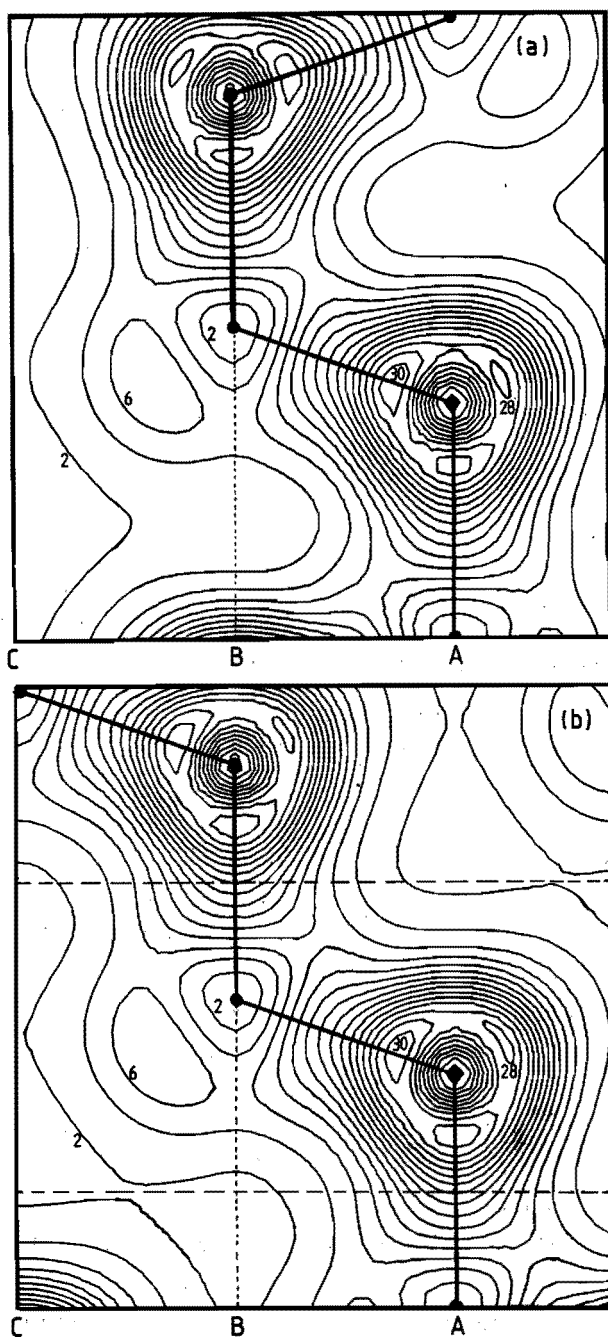


Fig. 6.7 Total valence-charge density (a) of 2H SiC in the (110) plane, compared to (b) the one of 3C SiC in the (110) plane. Both densities are calculated with $E_{FV} = 15.2$ Ry. The region between dashed lines in fig. 6.7(b) belongs to the "building-block" layer (see text).

G	$ \tilde{\rho}_{2H} $	$ \rho_{2H} $
000	8.000	8.000
100	1.400	1.420
002	1.997	1.990
101	1.226	1.234
102	0.421	0.399
110	0.193	0.170
103	0.119	0.110
200	0.174	0.189
112	0.201	0.201
201	0.161	0.149

Table 6.12. Fourier components of the valence-charge density (in numbers of electronic charges per unit-cell volume of 3C SiC), obtained by mapping a self-consistent density of 3C SiC according to (6.12) ($\tilde{\rho}_{2H}$) and by a self-consistent calculation for 2H SiC (ρ_{2H}). G vectors are given with respect to the basis vectors b'_1, b'_2 and b'_3 of the reciprocal lattice of 2H SiC (see section 3.2.2).

places the resulting layer on top of the "building-block" layer. There are 5 other possibilities to define this density mapping, viz., using all other operations of the space group of wurtzite (C_{6v}^4) that are accompanied by a nonprimitive translation. In fig. 6.8 we show an original density of 3C SiC and its mapped equivalent, which may be compared with the density for 2H SiC in fig. 6.7(a). In table 6.12 we make the comparison more quantitative by giving Fourier components of the mapped valence-charge density (fig. 6.8(b)) and of the actual valence-charge density calculated directly (fig. 6.7(a)). The agreement is reasonable.

By transformations similar to the one described above, the valence-charge density in all polytypes may be approximated from knowledge of the density in one layer of the simplest polytype, 3C SiC, only. Unfortunately, this does not give us a means of comparing the polytypes energetically, since the total energy as functional of the density is not known. For instance, in (2.58) the term $\sum_{(n,k)} E_n(k)$ is only determined by ρ through the solution of the

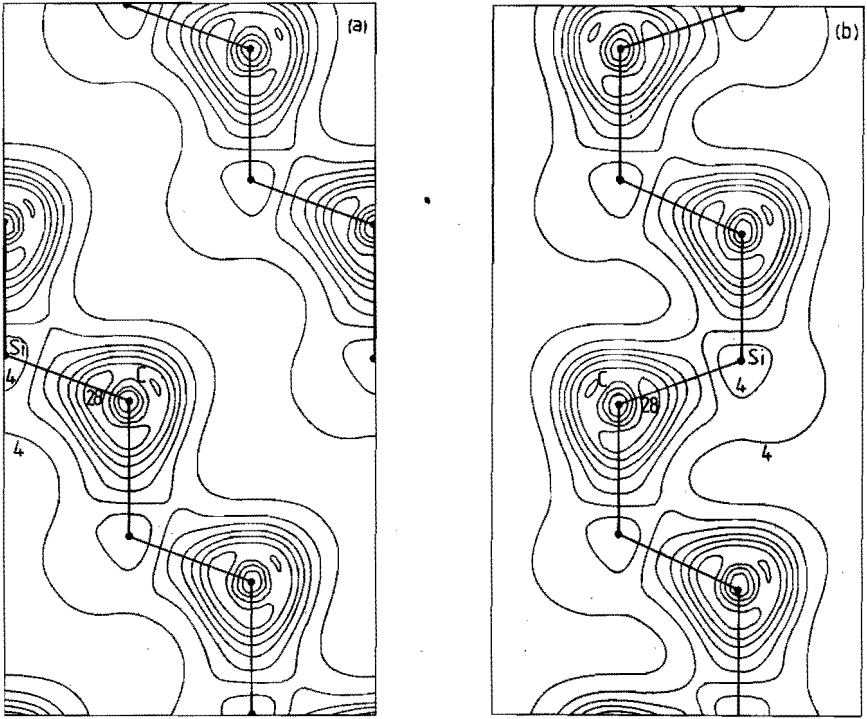


Fig. 6.8 Total valence-charge density of 3C SiC in the (110) plane, (a) before and (b) after the density-mapping transformation, eq. (6.12). The contour step is 4 electrons per unit-cell volume of 3C SiC. The density after mapping approximates the self-consistent density of 2H SiC (Fig. 6.7(a)) very well.

set of equations (2.39), in which the effective potential V_{eff} is determined by ρ . For polytypes with large unit cells, however, this solution is computationally prohibitive.

In fig. 6.9 results are shown of total-energy calculations for 2H SiC with varying lattice parameters c and a . The internal parameter u is eliminated as independent parameter by the restriction that all Si-C bond lengths are equal or, equivalently, that all atoms (Si and C) are situated exactly in the centre of regular tetrahedra with atoms of the other type (C and Si) at the corners. In the theoretical calculation this restriction is merely for our convenience -one independent variable is eliminated-, in experiments this condition

seems to be satisfied [73]. With this restriction u depends on c and a as follows:

$$u = \frac{1}{3} \left[\frac{a}{c} \right]^2 + \frac{1}{4}. \quad (6.13)$$

In fig. 6.9(a) the c/a -ratio is kept constant at the ideal value $(2/3)\sqrt{6}$ ($=1.633$). In fig. 6.9(b) we keep the lattice parameter a constant at $\frac{1}{2}a_c\sqrt{2}$, with a_c the experimental lattice constant of 3C SiC, implying $a = 3.0827 \text{ \AA}$. Finally, in fig. 6.9(c) the parameter c is kept constant at $(2a_c/3)\sqrt{3} = 5.0340 \text{ \AA}$. We note that these ideal (3C SiC derived) values of a and c differ only very slightly from the experimental 2H SiC values: $a = 3.0763 \text{ \AA}$, $c = 5.0480 \text{ \AA}$ and $c/a = 1.641$ [64].

The error bars in fig. 6.9 have been determined as follows; the calculations have been performed with $E_{PV}=23.3 \text{ Ry}$ and 3 special points. With these cutoffs the maximal effective change in basis-set size, ΔN_{\max} (see section 5.2), is 1.33. From calculations for 3C SiC we know that at $E_{PV}=23.3 \text{ Ry}$ discontinuities can be as large as 10 mRy per cell of 3C when $\Delta N = 7.5$. Because the unit cell of 2H SiC is twice as large, discontinuities can be as large as 20 mRy per cell when $\Delta N = 7.5$. Since $\Delta N_{\max} = 1.33$, we obtain an error estimate of: 3.5 mRy per cell of 2H SiC.

Because of the error bars in fig. 6.9, the minima of the 3 curves are only roughly determined. From the results in fig. 6.9, we deduce the equilibrium lattice parameters as follows; each of the figures 6.9(a), (b), and (c) yields a set of equilibrium lattice parameters c , a , and c/a (the minimum is determined from the parabola through the data points in figs. 6.9(a) and (c) and the third order polynomial through the data points in fig. 6.9(b)). The three results are averaged and the error is estimated to be the largest deviation from the average. We then find: $a = 3.11 \pm 0.03 \text{ \AA}$, $c = 5.10 \pm 0.08 \text{ \AA}$, $c/a = 1.64 \pm 0.04$. All calculated values are within 1% of the experimental values.

It is clear that a discrimination between experimental 2H SiC lattice parameters ($a = 3.0763 \text{ \AA}$, $c = 5.0480 \text{ \AA}$) and ideal (3C SiC derived) lattice parameters ($a = 3.0827 \text{ \AA}$, $c = 5.0340 \text{ \AA}$) is out of the reach of the present calculation and probably even out of the reach of the whole method.

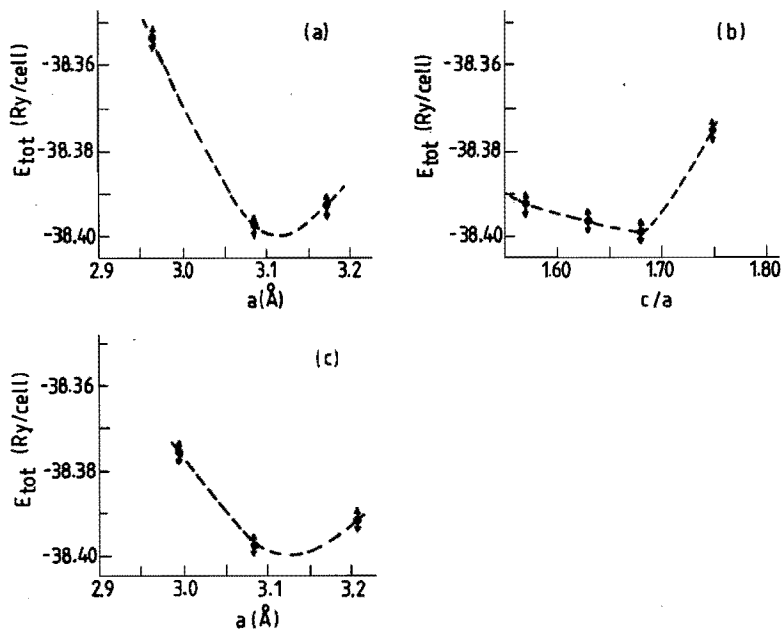


Fig. 6.9 Total energy of 2H SiC as a function of the lattice parameters c and a ; (a) variation of a with constant c/a -ratio, (b) variation of c/a with constant a , and (c) variation of a with constant c . The arrows indicate estimated error bars (see text).

CHAPTER 7

OUTLOOK: TOWARDS A FUNDAMENTAL DESCRIPTION OF CRYSTALS WITH LIMITED PERIODICITY

So far we have focused on *bulk crystals*, i.e., crystals that are periodic in three linearly independent directions with a period of the order of the lattice constant. We have found that an accurate description of the atomic and electronic structure of such crystals is possible by starting from properties of the constituting atoms and without the need to include parameters obtained from experiment. It should be noted that within the same theoretical framework also lattice dynamical properties such as phonon frequencies and elastic properties may be, and actually have been, calculated in agreement with and in prediction of experiments [30]. Furthermore, transition pressures for solid-solid phase transformations under the application of pressure have been calculated and predicted (see, e.g., ref.[108]).

In this chapter we briefly review the achievements so far and the possibilities in the future of the pseudopotential-density-functional method when applied to crystals with *limited periodicity*. Such crystals we will understand to be crystals for which the periodicity in one of the three directions is non-existent or greatly modified (enlarged) with respect to the bulk crystal. We here restrict crystals with limited periodicity to the following 4 categories:

- (i) *interfaces*: two different crystals grown on top of each other. A technologically important class of interfaces are the semiconductor-semiconductor interfaces called heterojunctions.
- (ii) *surfaces*: these can be seen as interfaces in which one of the "crystals" is the vacuum.
- (iii) *crystals with stacking faults*: bulk crystals with structural irregularities in the stacking of layers.
- (iv) *superlattices*: periodically repeated interfaces, thereby restoring the periodicity in the third direction, with periodicity lengths, however, which may be significantly larger than those in the directions parallel to the interfaces. Since around 1980 techniques such as Molecular Beam Epitaxy (MBE) and

Metal-Organic Chemical Vapour Deposition (MOCVD) have come available to artificially grow superlattices, which are becoming increasingly important in technology. Polytypes (see chapter 6) can be seen as a natural class of superlattices. The number of layers in a unit cell and their stacking sequence differs between different polytypes.

To apply the pseudopotential-density-functional (PDF) method using plane waves, it is absolutely necessary to restore the periodicity in the third direction if it is non-existent (categories (i), (ii), and (iii)). This is done by periodically repeating the interface, surface or stacking fault using such a large unit cell (called supercell) that neighbouring interfaces, surfaces or stacking faults do not influence each other. Such an influence can be investigated and controlled by calculations with supercells of different size.

What has been achieved until now in the mentioned categories using the PDF-method with norm-conserving pseudopotentials and plane waves? In answering this question we do not intend to be exhaustive, but we want to give a general idea of the applicability in practice of the method.

- (i) Van de Walle and Martin [109] calculated band discontinuities of semiconductor heterojunctions, which are of the order of 0.5 eV, to within an accuracy of 0.05-0.1 eV. In their calculations it has appeared to be possible to use low cutoffs - $E_{PV} = 6$ Ry (leading to about 35 plane waves per atom) and typically 4 special points - combined with small supercells containing typically 8 atoms. In this way an actual heterojunction could be described, i.e., charge densities and potentials were found to be bulk-like already one layer away from the interface. Heterojunctions of semiconductors in which both bulk solids have equal crystal structure and approximately equal lattice constants (e.g., GaAs and AlAs), as well as lattice strained heterojunctions (e.g., Si/Ge) were treated.
- (ii) Surfaces are more difficult to handle since the effect of the vacuum penetrates deeper into the crystal, leading to the necessity of large supercells. Pandey [110], Northrup [111] and the MIT-group of Joannopoulos [112,113] have performed surface-reconstruction calculations for various surfaces of Si and GaAs.

In the calculations of ref.[113] about 1700 plane waves are treated exactly -determining the size of the matrices to be diagonalized-, whereas an additional 3500 plane waves are included via perturbation theory. The energy gain of surface reconstruction is typically of the order of 0.5 eV per surface atom.

- (iii) Chou et al. [114] calculated stacking-fault energies in silicon, which are of the order of 0.02 eV, with an accuracy of 20%, using cutoffs, $E_{PV} = 10$ Ry (~ 70 plane waves per atom), $N_{s,p} = 16$, and supercells of 16 and 14 atoms.
- (iv) Polytypic energy differences were discussed in section 6.4. So far calculations for polytypes of SiC with up to 8 atoms were performed [57]. Energy differences here are of the order of 0.005 eV per atom. Very recently, Martin reported preliminary calculations of structural phase-transformations under pressure in superlattices of GaAs and AlAs [115].

The calculations mentioned above are generally performed at the limits of computational power (locally) available. If it appears to be necessary to include more atoms in the supercell, or if one wishes to lay a smaller claim on computing facilities, one has to replace the plane waves in the basis set by functions of which a smaller number per atom are needed. An example, employed in connection with norm-conserving pseudopotentials, are LCAO's [24], where the atomic orbitals are combinations of Gaussians. This approach was applied to the surface of diamond by Vanderbilt and Louie [116]. Gaussians were also used in the description of twin boundaries in silicon by DiVincenzo et al. [117] (supercells of 36-40 atoms). The calculational scheme connected with LCAO's is, however, significantly more complicated than the one associated with plane waves.

We now briefly discuss an approach that is sometimes followed to circumvent the need for calculations with large unit cells. In this so-called *parameter approach*, calculations for a few geometries with small unit cells provide the parameters with which the systems with large unit cells are described. We mention two examples:

- (1) Polytypes differ in the way layers of equal structure are stacked (see section 6.5 for the definition of such a layer in SiC). Inspired by a resemblance between the phase diagram of the Axial

Next Nearest Neighbour Ising (ANNNI) model and polytypism [118], the following expression for the energy of an arbitrary polytype was proposed:

$$E = J_0 + \alpha J_1 + \beta J_2. \quad (7.1)$$

In (7.1) α and β are constants (with absolute value between 0 and 1) which are simply related to the specific stacking sequence. The parameters J_0 , J_1 , and J_2 are interaction energies in the layer, between nearest-neighbour layers, and between next-nearest neighbour layers, respectively. Total-energy calculations for 3 polytypes suffice to determine the 3 parameters $\{J_0, J_1, J_2\}$ [57]. For these 3 polytypes one is free to choose those with the smallest unit cells, i.e., 3C, 2H, and 4H SiC (2,4, and 8 atoms per cell, respectively). From the calculated parameters the total energy for an arbitrary polytype, also one with a very large unit cell, is now easily calculated. However, we must emphasize that it remains to be investigated whether the expression (7.1) is adequate, i.e., whether interaction energies between layers further apart (J_3, J_4, \dots) may indeed be neglected. This can be done by calculations for additional polytypes.

- (2) The (7x7) reconstruction of the Si (111) surface has a much too large unit cell to be directly accessible to the type of calculation described. By means of the recently proposed Takayanagi-model for this reconstruction [119], the energy gain of this reconstruction may be expressed in terms of 3 parameters. These parameters were determined by Northrup [120] from surface calculations with smaller unit cells ($(\sqrt{3} \times \sqrt{3})$ and (2×2) reconstructions).

The approach, as sketched in the two examples, does not fully replace the need for calculations with large unit cells. It is, however, an acceptable alternative within the present computational potentialities.

From the above sketch of the state of the art and with the ever-increasing computing power in mind, we may conclude that the method described in this work will be used extensively in the near future to study ground-state properties of crystals with limited

periodicity. Experimentalists, however, are usually more interested in *excited-state properties* of surfaces and interfaces. The *theoretical* study from first principles of such properties has not come to its full growth yet. Only recently the first first-principles calculations for bulk solids of excited-state properties in agreement with experiment have been reported [85]. These calculations turn out to be much more demanding computationally than those for the ground state, which makes it uncertain in how far the method is of practical value to the determination of excited-state properties of surfaces and interfaces.

References

- [1] *Electronic Structure, Dynamics and Quantum Structural Properties of Condensed Matter*, J.T. Devreese, P.E. van Camp eds. (Plenum, New York, 1985).
- [2] V. Heine, in ref.[1].
- [3] *Science* **234**, Dec. '86, p. 1074 (1986).
- [4] P. Hohenberg, W. Kohn, *Phys. Rev.* **136**, B864 (1964).
- [5] W. Kohn, L.J. Sham, *Phys. Rev.* **140**, A1133 (1965).
- [6] L.H. Thomas, *Proc. Cambr. Philos. Soc.* **23**, 542 (1926).
E. Fermi, *Z. Phys.* **48**, 73 (1928).
- [7] J. Callaway, N.H. March, *Solid State Phys.* **38**, 135 (1984).
- [8] *Theory of the inhomogeneous electron gas*, S. Lundqvist, N.H. March eds. (Plenum, New York, 1983).
- [9] M.T. Yin, M.L. Cohen, *Phys. Rev. Lett.* **45**, 1004 (1980).
- [10] J.C. Phillips, L. Kleinman, *Phys. Rev.* **116**, 287 (1959).
- [11] W.C. Herring, *Phys. Rev.* **57**, 1169 (1940).
- [12] D.R. Hamann, M. Schlüter, C. Chiang, *Phys. Rev. Lett.* **43**, 1494 (1979).
- [13] V. Heine, *Solid State Phys.* **24**, 1 (1970).
- [14] G.B. Bachelet, D.R. Hamann, M. Schlüter, *Phys. Rev. B* **26**, 4199 (1982).
- [15] M.T. Yin, M.L. Cohen, *Phys. Rev. B* **25**, 7403 (1982).
- [16] J.T. Devreese, F. Brosens, in ref.[1], chapter 3 and 4.
- [17] L.I. Schiff, *Quantum Mechanics*, 2nd ed., chapter V (McGraw-Hill, New York, 1955).
- [18] G.P. Kerker, *J. Phys. C* **13**, L189 (1980).
- [19] S.G. Louie, S. Froyen, M.L. Cohen, *Phys. Rev. B* **26**, 1738 (1982).
- [20] U. von Barth, C.D. Gelatt, *Phys. Rev. B* **21**, 2222 (1980).
- [21] T.L. Gilbert, *Phys. Rev. B* **12**, 2111 (1975).
- [22] B.N. Taylor, W.H. Parker, D.N. Langenberg, *Rev. Mod. Phys.* **41**, 375 (1969).
- [23] J. Ihm, A. Zunger, M.L. Cohen, *J. Phys. C* **12**, 4409 (1979).
- [24] J.R. Chelikowsky, S.G. Louie, *Phys. Rev. B* **29**, 3470 (1984).
- [25] H.J.F. Jansen, A.J. Freeman, *Phys. Rev. B* **30**, 561 (1984).
- [26] O.K. Andersen, *Phys. Rev. B* **12**, 3060 (1975),
H.L. Skriver, *The LMTO Method* (Springer, Berlin, 1984).
- [27] M.T. Yin, M.L. Cohen, *Phys. Rev. B* **26**, 3259 (1982).

- [28] J.D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975).
- [29] P.P. Ewald, *Ann. Physik* **64**, 253 (1921).
- [30] O.H. Nielsen, R.M. Martin, *Phys. Rev. B* **32**, 3792 (1985).
- [31] M. Schlüter, J.R. Chelikowsky, S.G. Louie, M.L. Cohen, *Phys. Rev. B* **12**, 4200 (1975).
- [32] J.R. Chelikowsky, M.L. Cohen, *Phys. Rev. B* **13**, 826 (1976).
- [33] M. Gell-Mann, K.A. Brueckner, *Phys. Rev.* **106**, 364 (1957).
- [34] E.P. Wigner, *Phys. Rev.* **46**, 1002 (1934).
- [35] J. Perdew, A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).
- [36] D.M. Ceperley, B.J. Alder, *Phys. Rev. Lett.* **45**, 566 (1980).
- [37] M.L. Cohen, V. Heine, *Solid State Phys.* **24**, 38 (1970).
- [38] G.P. Srivastava, *J. Phys. C* **15**, 707 (1982).
- [39] P. Bendt, A. Zunger, *Phys. Rev. B* **26**, 3114 (1982).
- [40] G.B. Bachelet, H.S. Greenside, G.A. Baraff, M. Schlüter, *Phys. Rev. B* **24**, 4745 (1981).
- [41] M. Abramowitz, I.A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972).
- [42] E. Merzbacher, *Quantum Mechanics* (Wiley, New York, 1970).
- [43] I.S. Gradshteyn, I.M. Ryzhik, *Tables of Integrals, Series, and Products*, p. 710, form. (6.615) (Academic, New York, 1980).
- [44] P.J.H. Denteneer, W. van Haeringen, *J. Phys. C* **18**, 4127 (1985).
- [45] P.J.H. Denteneer, W. van Haeringen, F. Brosens, J.T. Devreese, O.H. Nielsen, P.E. van Camp, V.E. van Doren, unpublished.
- [46] IMSL Library: *FORTRAN Subroutines for Mathematics and Statistics* (IMSL, Inc., 1984).
- [47] E. Holzschuh, *Phys. Rev. B* **28**, 7346 (1983).
- [48] A. Baldereschi, *Phys. Rev. B* **7**, 5212 (1973).
- [49] D.J. Chadi, M.L. Cohen, *Phys. Rev. B* **8**, 5747 (1973).
- [50] H.J. Monkhorst, J.D. Pack, *Phys. Rev. B* **13**, 5188 (1976).
- [51] A.H. MacDonald, *Phys. Rev. B* **18**, 5897 (1978).
- [52] R.A. Evarestov, V.P. Smirnov, *Phys. Stat. Sol.(b)* **119**, 9 (1983).
- [53] A. Schönflies, *Theorie der Kristallstruktur* (Bornträger, Berlin, 1923).
- [54] J.D. Pack, H.J. Monkhorst, *Phys. Rev. B* **16**, 1748 (1977).
- [55] L.P. Bouckaert, R. Smoluchowski, E.P. Wigner, *Phys. Rev.* **50**, 58 (1936).
- [56] W.C. Herring, *J. Franklin Inst.* **233**, 525 (1942).

- [57] C. Cheng, R.J. Needs, V. Heine, N. Churcher, *Europhys. Lett.* **3**, 475 (1987).
- [58] J.C. Slater, *Quantum Theory of Molecules and Solids*, Vol. 2 (McGraw-Hill, New York, 1965).
- [59] G.F. Koster, *Solid State Phys.* **5**, 173 (1957).
- [60] J.F. Cornwell, *Group Theory and Electronic Energy Bands in Solids* (North-Holland, Amsterdam, 1969).
- [61] J.C. Slater, *Quantum Theory of Molecules and Solids*, Vol. 1, Appendix 12 (McGraw-Hill, New York, 1965).
- [62] H. Schlosser, *J. Phys. Chem. Solids* **23**, 963 (1962).
- [63] H. Wendel, R.M. Martin, *Phys. Rev. B* **19**, 5251 (1979).
- [64] Landolt-Börnstein: *Numerical Data and Functional Relationships in Science and Technology*, O. Madelung ed., Group 3, Vol. 17, Part a (Springer, Berlin, 1982).
- [65] H.-G. Junginger, W. van Haeringen, *Phys. Stat. Sol.(b)* **37**, 709 (1970).
- [66] M.T. Yin, M.L. Cohen, *Phys. Rev. B* **26**, 5668 (1982).
- [67] O.H. Nielsen, R.M. Martin, in ref.[1].
- [68] S.G. Louie, in ref.[1].
- [69] M.T. Yin, in *Proceedings of the 17th International Conference on the Physics of Semiconductors* (San Francisco 1984), D.J. Chadi, W.A. Harrison eds. (Springer, New York, 1985).
- [70] P.J.H. Denteneer, W. van Haeringen, *Phys. Rev. B* **33**, 2831 (1986).
- [71] P. Gomes Dacosta, O.H. Nielsen, K. Kunc, *J. Phys. C* **19**, 3163 (1986).
- [72] A.R. Verma, P. Krishna, *Polymorphism and Polytypism in Crystals* (Wiley, New York, 1966).
- [73] W.F. Knippenberg, *Philips Res. Rep.* **18**, 161 (1963).
- [74] F.C. Frank, *Philos. Mag.* **42**, 104 (1951).
- [75] H. Jagodzinski, *Neues Jahrb. Mineral. Monatsh.* **3**, 49 (1954).
- [76] *Scientific American*, August '85, p. 50 (1985).
- [77] J.R. Chelikowsky, M.L. Cohen, *Phys. Rev. B* **10**, 5095 (1974).
- [78] A. Zunger, M.L. Cohen, *Phys. Rev. B* **20**, 4082 (1979).
- [79] Y.W. Yang, P. Coppens, *Solid State Comm.* **15**, 1555 (1974).
- [80] M.T. Yin, M.L. Cohen, *Phys. Rev. B* **24**, 6121 (1981).
- [81] N.A.W. Holzwarth, S.G. Louie, S. Rabi, *Phys. Rev. B* **26**, 5382 (1982).

- [82] P.E. van Camp, V.E. van Doren, J.T. Devreese, Phys. Rev. B 34, 1314 (1986).
- [83] V.K. Bashenov, V.E. Gorbachev, D.I. Marvakov, Phys. Stat. Sol.(b) 133, 285 (1985).
- [84] L.J. Sham, M. Schlüter, Phys. Rev. Lett. 51, 1888 (1983).
- [85] M.S. Hybertsen, S.G. Louie, Phys. Rev. B 34, 5390 (1986).
- [86] R. Hulthen, N.G. Nilsson, Solid State Comm. 18, 1341 (1976).
- [87] D. Straub, L. Ley, F.J. Himpsel, Phys. Rev. Lett. 54, 142 (1985).
- [88] F.J. Himpsel, J.F. van der Veen, D.E. Eastmann, Phys. Rev. B 22, 1967 (1980).
- [89] F.D. Murnaghan, Proc. Nat. Acad. Sci. U.S.A. 30, 244 (1944).
- [90] C.E. Moore, *Ionization Potentials and Ionization Limits Derived from the Analysis of Optical Spectra*, U.S. Nat. Bureau of Standards, Ref. Data Ser. 34 (US GPO, Washington D.C., 1970).
- [91] M. Born, K. Huang, *Dynamical Theory of Crystals* (Oxford, London, 1954).
- [92] *Handbook of Chemistry and Physics*, 63rd ed., R.C. Weast ed. (CRC, Boca Raton, 1982).
- [93] O. Gunnarson, B.I. Lundqvist, J.W. Wilkins, Phys. Rev. B 10, 1319 (1974).
- [94] O.H. Nielsen, private communication.
- [95] C. Kittel, *Introduction to Solid State Physics*, 5th ed., (Wiley, New York, 1976).
- [96] F. Birch, J. Geophys. Res. 57, 227 (1952).
- [97] D.H. Yean, J.R. Riter Jr., J. Phys. Chem. Solids 32, 653 (1971).
- [98] K.B. Tolpygo, Fiz. Tverd. Tela (Leningrad) 2, 2655 (1960) [Sov. Phys. Solid State 2, 2367 (1961)].
- [99] J.F. Nye, *Physical Properties of Crystals* (Oxford, London, 1957).
- [100] J. Ihm, J.D. Joannopoulos, Phys. Rev. B 24, 4191 (1981).
- [101] N.I. Churchev, K. Kunc, V. Heine, Solid State Comm. 56, 177 (1985).
- [102] J.L. Martins, A. Zunger, Phys. Rev. Lett. 56, 1400 (1986).
- [103] A.R. Williams, J. Kübler, C.D. Gelatt, Phys. Rev. B 19, 6094 (1979).
- [104] H.W.A.M. Rompa, private communication.

- [105] F. Herman, J.P. Van Dyke, R.L. Kortum, *Mat. Res. Bull.* **4**, S167 (1969).
- [106] P.J.H. Denteneer, W. van Haeringen, *Solid State Comm.* **59**, 829 (1986).
- [107] D. Lenstra, A.G. Roosenbrand, P.J.H. Denteneer, W. van Haeringen, *Physica* **138B**, 83 (1986).
- [108] K.J. Chang, M.L. Cohen, *Phys. Rev. B* **31**, 7819 (1985).
- [109] C.G. Van de Walle, R.M. Martin, *J. Vac. Sci. Technol. B* **3**, 1256 (1985), *ibid.* **B 4**, 1055 (1986), C.G. Van de Walle, Ph. D. thesis (Stanford University, 1986).
- [110] K.C. Pandey, *Phys. Rev. Lett.* **49**, 223 (1982).
- [111] J.E. Northrup, M.L. Cohen, *Phys. Rev. B* **29**, 5944 (1984).
- [112] J. Ihm, J.D. Joannopoulos, *Phys. Rev. B* **26**, 4429 (1982).
- [113] E. Kaxiras, Y. Bar-Yam, J.D. Joannopoulos, K.C. Pandey, *Phys. Rev. B* **33**, 4406 (1986).
- [114] M.Y. Chou, M.L. Cohen, S.G. Louie, *Phys. Rev. B* **32**, 7979 (1985).
- [115] R.M. Martin, in *Proceedings of the 18th International Conference on the Physics of Semiconductors* (Stockholm 1986), O. Engström ed. (World Scientific, Singapore, 1987).
- [116] D. Vanderbilt, S.G. Louie, *Phys. Rev. B* **29**, 7099 (1982).
- [117] D.P. DiVincenzo, O.L. Alerhand, M. Schlüter, J.W. Wilkins, *Phys. Rev. Lett.* **56**, 1925 (1986).
- [118] J. Smith, J. Yeomans, V. Heine, in *Modulated Structure Materials*, T. Tsakalakos ed. (Martinus Nijhoff Publishers, Dordrecht, 1984).
- [119] K. Takayanagi, Y. Tanishiro, M. Takahashi, S. Takahashi, *J. Vac. Sci. Technol. A* **3**, 1502 (1985).
- [120] J.E. Northrup, in *Proceedings of the 18th International Conference on the Physics of Semiconductors* (Stockholm 1986), O. Engström ed. (World Scientific, Singapore, 1987).

DE PSEUDOPOTENTIALAAL-DICHTHEIDSFUNCTIONAAL METHODE
TOEGEPAST OP HALFGELEIDER-KRISTALLEN

Samenvatting

In dit proefschrift wordt een methode beschreven ter berekening van de eigenschappen van vaste stoffen die voortvloeien uit het gedrag van de elektronen in die stoffen. Deze methode leidt tot kwantitatieve resultaten zonder de noodzaak om gegevens uit experimenten te kennen en kan dientengevolge, na succesvolle vergelijking met experimentele resultaten, voorspellende waarde hebben. De methode combineert de pseudopotentialaal-theorie, die het mogelijk maakt alleen de valentie-elektronen in de berekening op te nemen, en de dichtheidsfunctionaal-theorie, die een beschrijving geeft van de grondtoestand van veel-elektron-systemen in een externe potentiaal. Dit leidt tot de noodzaak van het zelf-consistent oplossen van de Schrödinger-vergelijking voor elektronen, waarin de potentiaal bepaald wordt door de elektronendichtheid, die op zijn beurt bepaald wordt door de golf functies die oplossing zijn van de Schrödinger-vergelijking.

Pseudopotentialen, die gegenereerd worden vanuit de eigenschappen van geïsoleerde atomen plus zekere overdraagbaarheidscriteria, maken het mogelijk de functies van belang (dichtheden, potentialen,...) te ontwikkelen in hanteerbare aantallen vlakke golven die de translatiesymmetrie van het kristal weerspiegelen. Aangezien geen aannames worden gedaan over de vorm van de van belang zijnde functies, is de methode in principe geschikt voor elk kristal en meer dan andere methoden geschikt voor kristallen met sterk inhomogene elektronendichtheden. Voorbeelden hiervan zijn praktisch alle halfgeleiders en ook isolatoren zoals diamant.

Deze laatste klassen van vaste stoffen hebben voor de beschreven methode het bijkomende voordeel, dat de integraties over de Brillouin zone bijzonder efficiënt kunnen worden uitgevoerd met de techniek der "speciale punten in de eerste Brillouin zone", waaraan een hoofdstuk is gewijd.

Verder wordt ingegaan op het benutten van de kristal-symmetrie bij het oplossen van de Schrödinger-vergelijking en op karakteristieke eigenschappen van het rekenschema, die benut kunnen worden om de hoeveelheid rekenwerk in te perken.

Er is een hoofdstuk gewijd aan de uitvoerige illustratie van de methode aan de hand van rekenresultaten die bereikt zijn met behulp van een computer-programma dat bovenbeschreven theorieën en technieken incorporeert. Er worden resultaten gepresenteerd voor silicium, diamant en de twee structureel extreme polytypen van silicium carbide.

Tot slot wordt een indruk gegeven van de toepasbaarheid van de methode op kristallen waarin de periodiciteit in één richting duidelijk afwijkt van die in andere richtingen of zelfs geheel ontbreekt. Dit verschijnsel treedt op bij oppervlakken, grensvlakken, superroosters en polytypen.

CURRICULUM VITAE

- feb. '59 : geboren te Brunssum
- sept. '71 - juni '77 : Gymnasium β
Bernardinuscollege te Heerlen
- sept. '77 - april '83 : Natuurkunde-studie
Rijksuniversiteit Utrecht
Afstudeerrichting: Theoretische Natuurkunde
Afstudeerbegeleider: Prof.dr. M.H.J.J. Ernst
Titel afstudeerscriptie: Diffusie in
wanordelijke systemen
Onderwijsbevoegdheden: Wiskunde, Natuurkunde
en Sterrenkunde
- juni '83 - mei '87 : Promotie-onderzoek
bij Prof.dr. W. van Haeringen,
vakgroep Theoretische Natuurkunde,
Faculteit der Technische Natuurkunde,
Technische Universiteit Eindhoven
Titel proefschrift: The pseudopotential-
density-functional method applied to
semiconducting crystals

STELLINGEN

behorende bij het proefschrift van

P.J.H. Denteneer

1. Het storingsrekening-schema genoemd naar Feenberg is ongeschikt voor toepassing op het Anderson-model voor sterk wanordelijke ketens, waarin de interatomaire overlap-energie klein is ten opzichte van de breedte van de uniforme verdeling van lokale atomaire energie niveaus.

P.M. Morse, H. Feshbach, Methods of Theoretical Physics, hoofdstuk 9 (McGraw-Hill, New York, 1953).

J. Ziman, Models of Disorder, hoofdstuk 9 (Cambridge University Press, Cambridge, 1979).

2. In de literatuur over diffusie in wanordelijke ketens is ten onrechte geconcludeerd, dat de effectieve medium benadering een uitdrukking voor de frequentie-afhankelijke diffusie-coëfficiënt $D(z)$ oplevert die, in de systematische ontwikkeling voor kleine z in machten van \sqrt{z} , exact is tot en met orde z . Een gevolg hiervan is, dat de algemene schaalhypothese voor dit type problemen beperktere geldigheid heeft dan mede op grond van numerieke berekeningen wordt aangenomen.

I. Webman, J. Klafter, Phys. Rev. B 26, 5950 (1982).

J.W. Haus, K. Kehr, K. Kitahara, Phys. Rev. B 25, 4918 (1982).

P.J.H. Denteneer, M.H. Ernst, J. Phys. C 16, L961 (1983).

3. Zijn δ_1 en δ_2 de verschillen in bindingsenergie (per eenheidscel in een laag) van de polytypen met stapeling ABAB.. en ABCBABC... met het polytype met stapeling ABCABC..., respectievelijk, dan is in zeer goede benadering de energie van de intrinsieke stapelfout $\delta_1 + 2\delta_2$ en de energie van de extrinsieke stapelfout $4\delta_2$. De door Chang en Cohen gegeven schatting van de stapelfout-energie voor germanium mist dan ook elke grond.

K.J. Chang, M.L. Cohen, Phys. Rev. B 34, 8581 (1986).

4. Bestudering van de systematische uitdovingen in het diffractiepatroon van het organische kristal triterpane E reduceert het aantal mogelijke ruimtgroepen voor dit kristal tot drie. De argumentatie van Smith om tot de uiteindelijke vaststelling van de ruimtgroep te komen kan aanmerkelijk versneld worden door op te merken, dat het samenstellende molecule een chirale verbinding van natuurlijke oorsprong is.

G.W. Smith, *Acta Cryst. B* **26**, 1746 (1970).

5. De voorfactoren in de uitdrukking voor de n° cumulant van een waarschijnlijkheidsverdeling in termen van produkten van de eerste n centrale momenten van die verdeling kunnen met behulp van een eenvoudig combinatorisch argument worden bepaald.

6. De methode van Vanderbilt om "gladdere" norm-behoudende pseudopotentialen te genereren garandeert niet, dat feitelijk met minder vlakke golven in de verzameling van basisfuncties kan worden volstaan. Hierdoor kan over het nut van deze methode geen uitspraak worden gedaan.

D. Vanderbilt, *Phys. Rev. B* **32**, 8412 (1985).

7. De huidige publicatie-woede in de fysica maakt het uitgeven van de handelingen van een conferentie tot een nutteloze activiteit.

8. Een instelling die wetenschappelijk onderzoek tot een van haar doelstellingen rekent dient de faciliteiten die zij hiertoe bezit 168 uur per week aan haar wetenschappelijke medewerkers ter beschikking te stellen.

9. De uitspraak: "Het boek was beter dan de film", is in de regel gebaseerd op een vergelijking van de literaire kwaliteiten van boek en film-scenario en is dan nietszeggend met betrekking tot het geheel van cinematografische kwaliteiten van de film.
10. In de gangbare definitie van de efficiëntie van een verzameling speciale punten ter integratie over de reciproke ruimte wordt ten onrechte geen rekening gehouden met de symmetrie van individuele punten. Een betere definitie wordt verkregen door in de gangbare definitie het aantal speciale punten te vervangen door het aantal ermee geassocieerde punten in de eerste Brillouin zone.

R.A. Evarestov, V.P. Smirnov, *Phys. Status Sol.(b)* 119, 9 (1983).
Dit proefschrift, hoofdstuk 3 en 4.

11. Aan de in de literatuur aanvaarde experimentele waarde voor de compressibiliteit van silicium carbide moet op grond van *ab initio* berekeningen getwijfeld worden.

Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology, Group 3, Vol.17, Part a, O. Madelung ed. (Springer, Berlin, 1982).
Dit proefschrift, hoofdstuk 6.

12. De door Chelikowsky en Cohen gegeven tabel van Fourier-componenten van de valentie-ladingsdichtheid in silicium is met de door hen gebruikte methode en de door hen gegeven parameters niet te reproduceren.

J.R. Chelikowsky, M.L. Cohen, *Phys. Rev. B* 10, 5095 (1974).
Dit proefschrift, hoofdstuk 6, tabel 6.1.