LEABHARLANN CHOLÁISTE NA TRÍONÓIDE, BAILE ÁTHA CLIATH | TRINITY COLLEGE LIBRARY DUBLIN
Ollscoil Átha Cliath | The University of Dublin

Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

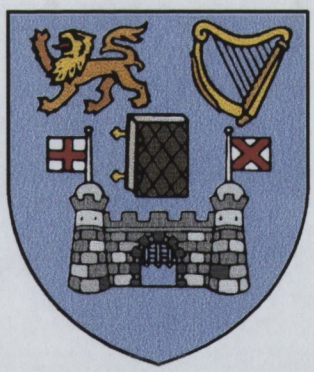
By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Localized Basis Approach to Many-Body Effects and Dielectric Response

Conor David Hogan

A Thesis submitted to
The University of Dublin
Trinity College
for the degree of
Doctor of Philosophy.



DEPARTMENT OF PHYSICS
TRINITY COLLEGE
UNIVERSITY OF DUBLIN

Oct 2000

© C. D. Hogan & University of Dublin

TRINITY COLLEGE
20 MAR 2001
LIBRARY DUBLIN

Thesis
X 5984+

Declaration

This thesis has not been submitted as an exercise for a degree at any other University. Except where otherwise stated, the work described herein has been carried out by the author alone.

This thesis may be borrowed or copied upon request with the permission of the Librarian, University of Dublin, Trinity College. The copyright belongs jointly to the University of Dublin and Conor David Hogan.

Signature of Author ..... Conor Hogan .....

Conor David Hogan

October 2000

Acknowledgements

I would like to thank everyone that has supported me during this work:

Foremost, I am indebted to my supervisor, Dr. Charles Patterson, for his continual guidance and support throughout the last four years; for taking an active interest in all aspects of this research, be it practical or academic, and for maintaining an impressive level of good humour and patience with me in lieu of frustration. Somehow, he always found the time to help me with any aspect of this work, and for that I am extremely grateful.

I have been very fortunate to have worked so closely with my colleague, Dr. Michael Kuzmin, who was a constant source of good ideas, and without whom this work would probably not have been so successful.

Many thanks go to Prof. Chris Wijers for several fruitful discussions on discrete dipole theory and nonlocality, and to Prof. Rex Godby for helpful advice with the self-energy code. Thanks to Dr. Mike Towler for useful advice on basis sets and useless advice on women. I would also like to acknowledge Dr. Simon Elliott and Dr. Simon Cox for discussions on various scientific issues, and the former for some proof-reading.

I owe many thanks to Geoff Bradley for much help and advice with computational matters, and for devoting so much time towards improving the computer systems. I would also like to thank Dr. Eamonn Cahill for being particularly helpful and for resolving several programming problems. In addition, I am grateful to Marc in het Panhuis for bringing his own brand of patient system administration to the world.

For financial support I would like to acknowledge the following Enterprise Ireland grants: SC/95/441 on 'Nonlinear optical properties of surfaces', SC/96/737 on 'Many-body effects for new physics in semiconductors', and SC/99/255 on 'Numerical study of excitons and plasmons in materials'. I would also like to acknowledge

Trinity Trust funds towards attending EPIOPTICS 5 in Erice, Sicily.

To the end of an era: thanks to all inhabitants of G30 over the years for livening up the place — Joe, Bosco, Hayes, Geoff, Michael, Marco, Svjetlana, Lola, Adam, Kate, Sadhbh, Simon amháin agus Simon eile, Tim, Finn, Judith, Brendan, Ferruccio, Stuart & Paul, Marc, Des, Stefan, Robert and of course... Sadikah.

I have had several great years in Dublin during this work; most of this is due to the great company of one inane Paddy and the usual miscreants: Fiachra, Malcolm, Beci, Treasa, Breedette and Kieran.

Luce, I couldn't have done it without you.

Finally: Paul, Barry and Anita, thanks for the support throughout; Mum and Dad, you've always been tremendous.

Summary

This thesis is a study of the optical and dielectric properties of semiconductors and insulators, and of the way in which these properties provide insight into the microscopic response of the interacting system. The central component of this work is the wavevector and frequency dependent dielectric function, $\varepsilon(\mathbf{q}, \omega)$, which provides the link between the observed experimental phenomena of absorption and reflection on one side, and the quantum mechanical electronic structure that determines the fundamental nature of the response mechanisms on the other.

A concept which is thoroughly explored in this thesis is so-called 'local field effect': its origins in the microscopic variation in the electric field that occur in an inhomogeneous system, and its influence on the optical and dielectric properties in the disparate systems of the surface and bulk. In the bulk, the local field effect is calculated by adopting a dielectric matrix approach within the framework of density functional theory, which in conjunction with the dielectric bandstructure technique is found to be essential for characterizing the distinct response features of several materials of variable ionicity, namely, Si, diamond, NaCl, MgO and GaN.

At the surface, the local fields are accounted for directly by considering a discretized form of the dielectric function, the discrete dipole model. The model is applied to the problem of calculating the reflection anisotropy of several vicinal surfaces of silicon, and it is shown that many features of the experimental spectra may be attributed to the effect of the crystal termination on the local field at the surface, which is classified as an intrinsic contribution to the reflectance anisotropy.

The dynamic role of the dielectric function for improving upon the Hartree-Fock mean-field approach to electronic structure calculations is examined by means of a many-body perturbation scheme within the independent quasiparticle approximation. This process requires a precise treatment of the screening and exchange mechanisms of the interacting system of particles via the calculation of the electron and hole self-energy within the GW-aproximation. It is found that inclusion of nonlocal and energy dependent interactions via the self-energy gives rise to a

marked improvement in the grossly exaggerated Hartree-Fock band gaps of Si, diamond and MgO, although quantitative agreement with experimental band gaps is not achieved. This finding suggests the use of a self-consistent iterative scheme starting from Hartree-Fock theory.

The quantum mechanical formalism that is adopted throughout is the linear combination of atomic orbitals (LCAO) scheme, in which the electronic Bloch wavefunctions are expressed in terms of a basis of localized Gaussian orbitals. The calculation of the dielectric function in the LCAO scheme represents a deviation from conventional plane-wave techniques and requires the development of an efficient and reliable computational scheme. Many features of this approach are presented in a detailed numerical appendix, which documents the necessary reciprocal space integration techniques, novel basis sets and explicit expressions for evaluating one-electron integrals.

Publications that have resulted from this thesis:

C.D. Hogan and C.H. Patterson.
"Reflectance Anistropy of Silicon Surfaces: Discrete Dipole Calculation."
Phys. Rev. B. **57**, 14843 (1998)

C.D. Hogan and C.H. Patterson.
"Reflectance Anisotropies of Silicon Surfaces: Analysis of Spectra in terms of Surface Excess Susceptibilities."
Surf. Sci. **402-404**, 586 (1998)

C.D. Hogan and C.H. Patterson.
"The Use of Gaussian Orbital Basis Sets in Dielectric Matrix and Self-Energy Calculations."
To be published.

To my parents

The presupposition of scientific work: belief in the unity and perpetuity of scientific work, so the individual may work at any part, however small, confident that his work will not be in vain — *Nietzsche, The Will to Power*

I hope I didn't brain my damage
— *Homer Simpson*

Contents

1 Introduction **1**

2 Quantum Theory of the Solid State **5**

2.1 The Many-Body Problem 6

2.2 The Hartree-Fock Approximation 7

2.3 Density Functional Theory 10

2.4 Green's Function Techniques 11

2.4.1 *GW*-Approximation 14

2.5 Choice of Basis 18

3 Theory of Optical Properties **20**

3.1 Linear Optical Response 21

3.1.1 Quantum Mechanical Approach 22

3.1.2 Longitudinal Dielectric Function 23

3.1.3 Calculations 28

3.2 The Influence of Inhomogeneity 31

3.2.1 Calculations 35

3.3 Beyond the Independent Electron Approach 36

3.3.1 Quasiparticle Corrections to Bandstructure 36

3.3.2 DFT-LDA Dielectric Function 37

3.3.3 Electron-Hole Interaction 38

3.4 Optical Properties of Surfaces 39

4 Static Dielectric Response **45**

4.1 Calculation of the Dielectric Matrix 46
4.1.1 Formalism 46
4.1.2 Basis Sets 48
4.1.3 Brillouin Zone Integration 55
4.1.4 **G**-vector Summation 56
4.2 Dielectric Bandstructure 57
4.2.1 Formalism 57
4.2.2 Results 59
4.2.3 Dielectric Eigenpotentials 61

5 Self-Energy Calculations **69**
5.1 Plane-Wave Representation 70
5.2 Plasmon-Pole Model 72
5.3 Self-Energy Matrix Element 75
5.4 Determination of Quasiparticle Energies 77
5.5 Results and Discussion 78

6 Surface Optical Properties **89**
6.1 Discrete Dipole Model: Formalism 90
6.2 Calculations 93
6.3 Results and Discussion 95

7 Conclusions **99**
7.1 Gaussian orbital framework 100
7.2 Dielectric properties 101
7.3 Quasiparticle calculations 102
7.4 Outlook 104

A Computer Code **106**
A.1 User Guide 106
A.1.1 Preliminary tasks 107
A.1.2 Input data structure: General parameters 108

A.1.3 Input data structure: Optional blocks 111
A.2 Programming Guidelines 120
A.2.1 General Structure of the Source Code 120
A.2.2 Memory Allocation 121
A.2.3 Global Strucures 123
A.2.4 **k**-Point Mesh Generation. 125

B Integral Evaluation **128**
B.1 Properties of Gaussian Type Functions 129
B.2 Integral Evaluation 131
B.2.1 Overlap integral 133
B.2.2 Optical matrix elements 134
B.2.3 Matrix elements of $e^{i\mathbf{q}\cdot\mathbf{r}}$ 135

C Reciprocal Space Integration **139**
C.1 Special-Point Technique 140
C.1.1 Theory 140
C.1.2 Special-Point Generation 141
C.2 Analytic Continuation Method 143
C.3 Auxiliary Function Technique 145

D Basis Sets **151**

List of Figures

2-1 Graphical representation of the self-energy in the GW-approximation. 17

3-1 Imaginary part of the dielectric function of silicon compared with the experimental data. 29

3-2 The effective number of electrons per atom as a function of cutoff frequency, for silicon. 31

3-3 Schematic diagram of the processes involved in photoemission and absorption. 38

3-4 Different models of linear response theory 41

4-1 Influence of basis set size and type on the total energy of silicon. . . . 49

4-2 Dependence of conduction band energy on the band index for silicon. 50

4-3 Dependence of conduction band energy on the band index for NaCl. . 50

4-4 Convergence of the dielectric matrix eigenvalues of silicon with respect to the number of conduction bands. 52

4-5 Dielectric bandstructures of silicon and diamond. 62

4-6 Dielectric bandstructures of NaCl and MgO 63

4-7 Dielectric bandstructure of GaN. 64

4-8 Dielectric eigenpotentials of silicon. 67

4-9 Dielectric eigenpotentials of NaCl. 68

5-1 Comparison between plasmon-pole model and fully calculated eigenvalues of the dielectric matrix, fitted using real energies. 76

5-2 Comparison between plasmon-pole model and fully calculated eigenvalues of the dielectric matrix, fitted using imaginary energies. 76

5-3 Schematic representation of the fitting procedure used to obtain the quasiparticle energies. 78
5-4 Quasiparticle and Hartree-Fock bandstructures of silicon. 85
5-5 Quasiparticle and Hartree-Fock bandstructures of diamond. 86
5-6 Quasiparticle and Hartree-Fock bandstructures of MgO. 87
5-7 Real part of selected matrix elements of the correlation self-energy of silicon. 88

6-1 Side view of the diamond structure along the $(1\bar{1}0)$ direction. 93
6-2 Real part of the dressed polarizability at the surface of a slab with a (001) surface and a (110) surface. 94
6-3 Calculations of reflectance anisotropy spectra for various surfaces of silicon, compared with experimental data. 96

A-1 Flow diagram of the computer code. 122

C-1 Imaginary part of the dielectric function of silicon calculated using the root-sampling and Gilat-Raubenheimer schemes. 146
C-2 Accuracy of the integration scheme for singular functions. 147

List of Tables

3.1 Influence of local fields on the calculated dielectric constant of silicon. 35

4.1 Selected elements of the dielectric matrix $\varepsilon_{\mathbf{GG'}}(\mathbf{q})$ of silicon, calculated using different special-point meshes in the full Brillouin zone. . 55

4.2 Convergence properties of the dielectric matrix eigenvalues of silicon with respect to the size of the matrix. 57

4.3 Smallest inverse eigenvalues ε_n^{-1} of the dielectric matrix at $\mathbf{q} \to 0$ for some typical covalent and ionic solids. 64

5.1 Hartree-Fock and *GW*-approximation band-gaps of Si, diamond and MgO. 80

5.2 Hartree-Fock and *GW*-approximation band-gaps of silicon and diamond obtained from the literature. 82

A.1 Number of **k**-points in the special point mesh at different values of shrinking factor. 127

B.1 Analytical expressions for integals of $e^{i\mathbf{Q}\cdot\mathbf{r}}$ over Gaussian primitive functions. 138

C.1 Special-points mesh in the full Brillouin zone for the FCC lattice. . . 143

C.2 Auxiliary functions used in reciprocal space integration, and their integrals over the Brillouin zone. 148

Chapter 1

Introduction

> For every problem, there is a solution that is neat, plausible and wrong — *H.L. Mencken*

The interaction of light with matter is one of the oldest and most challenging problems in science. A geometrical description of the reflection of light from the surface of a solid or liquid was deduced in ancient times by the Greeks, while modern progress in optics began with the experimental discovery of the law of refraction by Snell (1591–1626), as early as 1621. A more formal description is found in the work of Fermat (1605–1665), and other significant contributions to the field were given by Brewster (1781–1868), who considered the effects of optical polarization, and Fresnel (1788–1827), who discussed the influence of the optical density of the medium on the properties of transmission and refraction. To this day, many of these theories still provide a sufficent description of macroscopic optical phenomena.

Soon after light was shown to be an electromagnetic wave by Maxwell (1831–1879), the focus of scientific research was directed towards understanding the microscopic origins of optical properties, and attempts were made by Lorentz (1853–1928) and Drude (1863–1906) to apply the ideas of classical mechanics to the problems of absorption and dispersion. With the advent of quantum mechanics, the true nature of the microscopic world was revealed, and it was quickly realized that the interaction between light and matter (as manifested in optical spectra) could provide

important information about the physical structure of matter and the fundamental electronic processes therein. The quantity that defines the relationship between the measurable optical and dielectric properties and the electronic structure of solids is the dielectric function, $\varepsilon(\mathbf{q}, \omega)$; it is a very powerful concept, and contains much of the information concerning the physical properties of the medium.

The aim of this thesis is to contribute towards the calculation, interpretation and application of the dielectric function in the interacting many-body system, with the intention of elucidating its properties within a flexible and intuitive computational framework. The dielectric properties of the interacting system can be divided into two categories: those which involve details of the dielectric function, such as optical reflectivity, absorption, plasmon excitation, etc.; and those which involve an averaging over the electronic response, such as the self-energy of the electron. The first category requires a careful treatment of the microscopic interactions, and for the systems of interest in this thesis — namely, semiconductors and insulators — the calculation of the dielectric function presents a major computational task. With the aim of providing a formalism which is more suitable for investigating complex materials, the traditional plane-wave approach to the electronic structure is usurped in favour of a localized quantum chemical scheme, the Gaussian orbital representation. Within this formalism, it is shown how the static and dynamic properties of the dielectric function may be accurately investigated; of particular interest are the nonlocal microscopic processes that contribute to the macroscopic polarization, and the application of Green's function techniques to the evaluation of the interacting many-body system.

In considering phenomena such as reflection, the dielectric function may be combined with the Fresnel model, in which the response at the surface is identical to the response of a bulk-terminated continuous solid, so the fact that real solids have surfaces with microscopic structure is ignored. For optical wavelengths this is often a reasonable assumption, as light has a considerable penetration depth, and hence the contribution from the surface is minimal. To investigate surface phenomena an alternative to the dielectric function description is preferable, i.e., a scheme which is

based on a discrete kind of modelling. For bulk properties, the two descriptions are of course equivalent, and the only difference between a particular discrete framework and the dielectric function formalism lies in the approximations that are adopted when modelling the interaction with light and when describing the electronic structure. By adopting a simplified discrete model one can focus on specific aspects of the surface optical response, and in particular the local variations in the electric field that may occur near the surface. Effects due to electronic nonlocality are also considered in the continuum approach (albeit in a more convoluted manner), but at the surface they can become increasingly important. Their relevance to the surface optical spectra is not clear, which is a question that is also addressed in this thesis.

The work in this thesis is presented as follows: The next chapter will outline the fundamental concepts behind the *ab initio* formalisms to be used for calculating the dielectric function, namely, Hartree-Fock theory and density functional theory. An outline of the Green's function method for calculating the self-energy will be presented, and the chapter will conclude with a discussion of the Gaussian orbital scheme which has been adopted throughout. In Chapter 3 a detailed analysis of the longitudinal dielectric function will be given, and the connection with the optical spectra established. The matrix representation of the dielectric function will be discussed, which leads to the explicit treatment of the so-called 'local field effects' in the discrete dipole model, at the end of the chapter. In Chapter 4 the static dielectric matrix will be discussed in more detail: the first section will be mostly concerned with computational aspects associated with the use of a Gaussian basis set, whereas the second section will examine the information contained in the dielectric matrix through the concept of 'dielectric bandstructure'. Chapter 5 is concerned with the use of the dielectric matrix in the calculation of many-body effects in the self-energy at the level of the '*GW* approximation', and explicit calculations of the quasiparticle bandstructures of various materials will be presented. Discrete dipole calculations and investigation of the effects of nonlocality in the electric field on surface optical properties are presented in Chapter 6. Finally, in Chapter 7, some conclusions are drawn and an outlook for possible future work is given.

Algorithmic and computational details of the various calculations are presented in the appendices. These include explicit evaluations of optical matrix elements in the Gaussian orbital representation, a list of specially developed basis sets for the calculation of the dielectric matrix, as well as specific techniques for reciprocal space integration. In addition, the structure of the computer code that was developed as part of this work is outlined, and a detailed guide is presented for the benefit of future users and to aid further development of the code. A thorough knowledge of the information in the appendices is therefore not vital for an understanding of the basic conclusions of the thesis, although it should be appreciated that the appendices detail the more challenging aspects of the work that was accomplished.

Chapter 2

Quantum Theory of the Solid State

> A theory is charming not least because it is refutable: that is just what attracts the better minds to it — *Nietzsche, Beyond Good and Evil*

In order to elucidate the response of a material to an external perturbation, it is first necessary to describe adequately the fundamental properties of the unperturbed system. This in itself can pose a formidable task, and in general, the problem of calculating the electronic states and energy levels of a system consisting of $\sim 10^{23}$ of interacting electrons and nuclei is an intractable one. In practice, a number of approximations are adopted; usually, the complicated, many-body system is replaced by a self-consistent 'mean-field' description, in which particles are allowed to move independently of one another, but whose movements are determined by some average potential which is intended to account for the effects of exchange and correlation.

The focus of this chapter is to provide the foundations of the electronic structure formalisms which are to be used throughout this thesis. The chapter begins with an initial discussion of the many-body problem, after which two mean-field theories are outlined, namely, the Hartree-Fock (HF) and density functional theories (DFT).

The DFT bandstructure will be used in the calculation of the dielectric function in Chapters 3 and 4, whereas the Hartree-Fock framework is included primarily as a route to calculating the quasiparticle bandstructure. The self-energy formalism which describes the quasiparticle bandstructure is the subject of a separate section. Finally, the chapter closes on a practical issue: the linear combination of atomic orbitals (LCAO) approach, as applied to periodic systems, is discussed.

2.1 The Many-Body Problem

In calculating the electronic structure of a system of interacting particles, the basic problem is that of solving the stationary Schrödinger equation,

$$\hat{H}\Psi = E\Psi, \tag{2.1}$$

in which the general form of the total many-body Hamiltonian $\hat{H}$ is given by[1]

$$\begin{aligned} \hat{H} &= -\sum_I \frac{1}{2M_I}\nabla_I^2 + \sum_{I<J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \\ &- \sum_i \frac{1}{2}\nabla_i^2 + \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i,I} \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_i|} \end{aligned} \tag{2.2}$$

[1]Conventions used in this thesis:

1. **Units.**
   Atomic units are used throughout, unless noted otherwise: $\hbar = m_e = |e| = 1$. The units of energy and length are the Hartree and the Bohr, respectively.
   1 Hartree (Har.) = 2 Rydberg (Ry.) $\approx$ 27.2114 eV $\approx 4.35978 \times 10^{-18}$ J
   1 Bohr $\approx$ 0.529177 Å

2. **Fourier transforms.**

$$f(\mathbf{r}) = \sum_{\mathbf{q}} f(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{r}} \qquad f(\mathbf{q}) = \frac{1}{V}\int f(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}}\, \mathrm{d}\mathbf{r}$$

3. **Brillouin zone (BZ) integration.**

$$\frac{1}{V}\sum_{\mathbf{k}\in \mathrm{BZ}} f(\mathbf{k}) = \frac{1}{(2\pi)^3}\int_{\mathrm{BZ}} f(\mathbf{k})\mathrm{d}\mathbf{k}$$

   where sample volume $V = N\Omega$, for N cells of volume Ω. Sum over spin is implicitly included.

where I, J and i, j refer to the nuclei and electrons respectively, and Z denotes the nuclear charge. For the large number of interacting particles that comprise a solid, it is clear that finding a solution to equation (2.1) is impossible, simply due to the sheer computational difficulty involved. In order to obtain, therefore, an accurate description of the crystalline ground state, one must invariably apply several simplifying approximations; in fact, the above Hamiltonian already omits some of the smaller terms that account for relativistic effects, in addition to the weaker interactions that involve the electron spin and orbital angular momentum. Fortunately, these omissions are not expected to introduce significant errors to the calculations in this thesis. The standard approximation that is applied to equation (2.2) is the Born-Oppenheimer adiabatic approximation, in which the nuclear and kinetic energy term is decoupled, to yield a separate eigenvalue problem for the electronic system. The remaining Hamiltonian,

$$\hat{H} = -\sum_i \frac{1}{2}\nabla_i^2 + \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i,I} \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_i|} \tag{2.3}$$

still presents a very complicated many-body problem, however, due to the electron-electron interaction term. At this point a number of different *ab initio* schemes enter, although a common procedure is to approximate the electron-electron term by an effective one-electron potential. One of these schemes will now be considered in more detail.

2.2 The Hartree-Fock Approximation

An approximate solution to equation (2.1) is usually obtained through a variational procedure. As the true wavefunction Ψ is not normally known, a trial wavefunction $\tilde{\Psi}$ is constructed from a set of linearly independent basis functions, and having assumed some form for $\hat{H}$, the energy expectation value is obtained from

$$\langle \tilde{E} \rangle = \frac{\langle \tilde{\Psi} | \hat{H} | \tilde{\Psi} \rangle}{\langle \tilde{\Psi} | \tilde{\Psi} \rangle}. \tag{2.4}$$

According to the variational principle, the expectation value $\langle\tilde{E}\rangle \geq E_0$, where E_0 denotes the global minimum in energy for the system that is found only with the exact electronic many-body wavefunction Ψ. Therefore, an improved $\tilde{\Psi}$ may be found by minimising $\langle\tilde{E}\rangle$, which in practice is done by varying the linear expansion coefficients of the trial wavefunction.

The Hartree-Fock approach follows naturally from this procedure if a determinantal form is assumed for $\tilde{\Psi}$. For an N-electron system, the many-body wavefunction for an arbitrary state can be approximated by an $N \times N$ 'Slater determinant', which is simply given by an antisymmetrized[2] product of one-electon orbitals:

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N) \approx |\Phi_0\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_2(\mathbf{r}_1) & \cdots & \psi_N(\mathbf{r}_1) \\ \psi_1(\mathbf{r}_2) & \psi_2(\mathbf{r}_2) & \cdots & \psi_N(\mathbf{r}_2) \\ \psi_1(\mathbf{r}_3) & \psi_2(\mathbf{r}_3) & \cdots & \psi_N(\mathbf{r}_3) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\mathbf{r}_N) & \psi_2(\mathbf{r}_N) & \cdots & \psi_N(\mathbf{r}_N) \end{vmatrix}. \tag{2.5}$$

The Hartree-Fock approximation consists of replacing the true many-electron wavefunction by the Slater determinant with the lowest energy. This ground state determinant is obtained by varying the single-particle states, $\psi_i(\mathbf{r})$, in the manner described above, which leads to a set of eigenvalue equations:

$$[\hat{h}_0(\mathbf{r}) + V_{H,i}(\mathbf{r})]\psi_i(\mathbf{r}) + \int V_{\mathrm{ex},i}(\mathbf{r}, \mathbf{r}')\psi_i(\mathbf{r}')\mathrm{d}\mathbf{r}' = E_i\psi_i(\mathbf{r}), \tag{2.6}$$

in which V_H represents the 'Hartree' potential and V_{ex} is a nonlocal exchange potential:

$$V_{H,i}(\mathbf{r}) = \sum_{j\neq i}^{\mathrm{occ}} \int \frac{|\psi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|}\mathrm{d}\mathbf{r}', \tag{2.7}$$

$$V_{\mathrm{ex},i}(\mathbf{r}, \mathbf{r}') = -\sum_{j\neq i}^{\mathrm{occ}} \frac{\psi_j(\mathbf{r})\psi_j^*(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \tag{2.8}$$

[2]The requirement that the wavefunction be antisymmetric is derived from the Pauli exclusion principle, which forbids two electrons with the same spin to occupy the same orbital.

This set of coupled equations constitutes the basic Hartree-Fock problem; for a more precise treatment the reader is referred to the standard literature.[1, 2]

On examining equations 2.6–2.8, the physical principles behind the Hartree-Fock approximation become clear. The electron can be understood to move in an average potential which is the sum of three terms: a potential due to the ions; a potential V_H corresponding to the classical Coulomb interaction between the electron and the average electronic charge density; and a nonlocal exchange potential that arises as a consequence of the fermionic nature of the electron. However, there is one critical omission in this description of the electronic structure. Hartree-Fock theory neglects the polarizability of the surrounding electron gas that produces 'correlation' between the electrons, which is a considerable omission in solids, although for atomic systems the influence is smaller. The omission of electron correlation has serious repercussions for calculating the optical properties of solids, in that band gaps are consistently overestimated. However, the Hartree-Fock scheme has an important role in this thesis as a reference point beyond which the effects of correlation may be included, and as such it will be used in constructing the self-energy.

A natural progression beyond Hartree-Fock theory is to expand the ground and excited states of the system over a basis of Slater determinants, termed 'configurations'. The resulting multi-determinantal wavefunction is of the form:

$$|\Psi\rangle = C_0|\Phi_0\rangle + \sum_{a,r} C_a^r|\Phi_a^r\rangle + \sum_{\substack{a<b \\ r<s}} C_{ab}^{rs}|\Phi_{ab}^{rs}\rangle + \ldots, \tag{2.9}$$

where $|\Phi_0\rangle$ is the minimum energy single determinant — i.e., the Hartree-Fock ground state — and $|\Phi_a^r\rangle, |\Phi_{ab}^{rs}\rangle$, etc., represent singly and doubly, etc., excited determinants. This representation is indicative of the 'Configuration Interaction' [1] formalism, which in principle can account for all possible correlation effects in the solid; unfortunately, however, this scheme in its original form is thoroughly impractical for large systems. An alternative approach, which attempts to include both correlation and exchange effects into the Hamiltonian is density functional theory, which is discussed in the next section.

2.3 Density Functional Theory

The most successful technique for calculating the ground state properties of solids is density functional theory (DFT), usually applied within the framework of the local density approximation (LDA). The origins of DFT can be found in the work of Hohenberg and Kohn[3], and Kohn and Sham[4]; detailed accounts of the theoretical foundations are presented elsewhere [5], so only a brief outline of the method is given.

The fundamental tenet is that the energy of the interacting many-particle system may be expressed as a functional of the density, $\rho(\mathbf{r})$, alone, and while the existence of this functional is guaranteed by the Hohenberg and Kohn theorem, its exact form is not generally known. By constructing a ficticious system of one-electron Kohn-Sham orbitals, ϕ_i, the DFT-LDA problem can be reduced to one of solving the set of one-electron equations,

$$[\hat{h}_0(\mathbf{r}) + V_H(\mathbf{r}) + V_{\mathrm{xc}}(\mathbf{r})]\phi(\mathbf{r}) = \epsilon_i\phi(\mathbf{r}), \tag{2.10}$$

where the non-interacting Hartree and kinetic energy parts of the functional have been separated out to leave an effective 'exchange-correlation' functional, V_{xc}. An approximation for V_{xc} may be obtained by appealing to a physical model such as the homogeneous electron gas. The connection between this ficticious non-interacting system and the real system of strongly interacting electrons is rigorously given by the condition that the Kohn-Sham density matches the actual density exactly,

$$\rho^{KS}(\mathbf{r}) = \rho^{real}(\mathbf{r}) = 2\sum_{i}^{occ} |\phi_i(\mathbf{r})|^2, \tag{2.11}$$

which introduces a condition of self consistency into equation (2.10), as V_H and V_{xc} both depend on $\rho^{KS}(\mathbf{r})$. This condition is assured if all the electronic interaction terms of the real system are included in equation (2.10).

By comparison with the Hartree-Fock equations, [equation (2.6)], the Kohn-Sham orbitals may be interpreted as a set of single-particle states, whereby the (local) exchange-correlation potential V_{xc} replaces the (nonlocal) Hartree-Fock ex-

change potential V_{ex}. This is not strictly accurate, as DFT applies only to the ground-state. Nevertheless, it is extensively used to construct bandstructures, *i.e.*, including excited states, with the result that the DFT-LDA energy gap is underestimated by up to 50%[6, 7], particularly for materials with a strongly inhomogenous charge distribution such as semiconductors and insulators. The Kohn-Sham orbitals and energies are used thoughout this thesis for calculating the optical properties via the dielectric function, and in general it will be shown that a good agreement with the experimental spectra is achieved.

2.4 Green's Function Techniques

The correct formalism for describing single-particle excitation energies is the quasiparticle (QP) approach, which is rigorously described in many-body perturbation theory by the Green's function method. A brief outline of the relevant aspects for this thesis is presented here; a more detailed discussion of the method may be found in standard text-books on many-body theory (e.g. Inkson [2] and Mahan [8]) or in review articles by Hedin and Lundqvist [9] and Aryasetiawan and Gunnarson [10].

In the Green's function method, the effects of exchange and correlation on the propagation of a particle through an interacting system are fully described by the nonlocal and energy dependent self-energy $\Sigma(\mathbf{r}, \mathbf{r}', \omega)$. In principle, the single particle energies and wavefunctions may be obtained by solving the equation:

$$[\hat{h}_0(\mathbf{r}) + V_H(\mathbf{r})]\Psi(\mathbf{r}) + \int \Sigma(\mathbf{r}, \mathbf{r}'; E)\Psi(\mathbf{r}')\mathrm{d}\mathbf{r}' = E\Psi(\mathbf{r}). \tag{2.12}$$

The single particle states $\Psi(\mathbf{r})$ in this equation are known as quasiparticle states, in the sense that the interaction between the particle and the surrounding system is included in an effective single particle state. There is no guarantee that such states do exist, however, although the very success of the single particle theories discussed in the previous two sections suggests that well-defined quasiparticles do in fact exist. In the independent quasiparticle approximation, it is assumed that there is no interaction between these ficticious particles, which negates the possibility of

treating multiple-particle processes such as excitons. This thesis will only consider the independent quasiparticle picture, although the influence of excitonic states on the optical spectra will be briefly discussed in Section 3.3.

On comparing with the Hartree-Fock equations [2.6], equation (2.12) may be interpreted as an extension of Hartree-Fock theory whereby the static unscreened exchange potential is replaced by an energy-dependent operator, the self-energy. Alternatively, the self-energy may be regarded as the nonlocal and energy dependent equivalent of the exchange-correlation potential V_{xc} in the Kohn-Sham equations [2.10]. Despite these similarities, the calculation of the self-energy presents a fundamentally different problem than Hartree-Fock or DFT; it is a true many-body problem, and is far removed from a mean-field description of the interacting system. In particular, the variational principle cannot be used to find a solution.

Nevertheless, equation (2.12) may be solved approximately if a simplified form for $\Sigma(\mathbf{r}, \mathbf{r}', E)$ is adopted. One of the most widely followed approaches, and the one used in this thesis, is the GW-approximation (GWA) due to Hedin [11]. To arrive at this approximation, we start from Hedin's set of coupled equations [11, 12]. These equations feature the usual Hartree Hamiltonian $\hat{H}_0 = \hat{h}_0 + V_H$ and bare Coulomb interaction v, as well as a number of additional quantities that arise as part of the Green's function method. Before examining this set of equations, it is necessary to briefly discuss what these quantities represent.

The first of these is the one-particle Green's function G, which describes the propagation of an electron or hole through the interacting system. The Green's function has many useful properties: it may be used to obtain the expectation value of single-particle operators in the ground state; it may be used to obtain the ground state energy of the system; and, of most relevance to this work, it can be used to obtain the one-electron excitation spectrum [10]. Mathematically, the excitation spectrum — or, equivalently, the quasiparticle bandstructure — is determined by the poles of the Green's function; this will become more apparent when an explicit form is assumed for G in Chapter 5.

The second quantity that appears in the equations to follow is the polarizability

function P: this describes the response of the system to the total potential, in a similiar manner to the classical polarizability. In the simplest approximation, P describes a dielectric response characterized by the creation of electron-hole pairs that do not interact; additional effects in the polarization appear through the so-called vertex correction Γ. The form that is assumed for P in the GW-approximation is the same as that used in constructing the dielectric function in the following chapters on optical response; the relation between the polarizabillty and the dielectric function is (schematically) given by $\varepsilon = 1 - vP$.

Of the remaining quantities, the physical meaning of the self-energy Σ has been discussed in the text above. The final undefined term is W, which is the screened Coulomb interaction. The use of a screened Coulomb interaction in place of the usual bare interaction, v, in the mathematical description of the interacting system of electrons is a very appropriate choice, as it appeals to the principle that a particle in an interacting system does not repond to the external potential but rather to the total potential, thus taking into account any induced potential that arises from the response of the system to the particle itself.

The set of equations may now be given:

$$\left[i\frac{\partial}{\partial t_1} - H_0(1)\right] G(1,2) - \int \mathrm{d}[3]\, \Sigma(1,3)\, G(3,2) = \delta(1,2) \tag{2.13}$$

$$\Sigma(1,2) = i \int \mathrm{d}[34]\, W(1^+,3)\, G(1,4)\, \Gamma(4,2;3) \tag{2.14}$$

$$W(1,2) = v(1,2) + \int \mathrm{d}[34]\, W(1,3)\, P(3,4)\, v(4,2) \tag{2.15}$$

$$P(1,2) = -i \int \mathrm{d}[34]\, G(2,3)\, G(4,2)\, \Gamma(3,4;1) \tag{2.16}$$

$$\Gamma(1,2;3) = \delta(1,2)\delta(1,3) + \int \mathrm{d}[4567]\, \frac{\delta\Sigma(1,2)}{\delta G(4,5)}\, G(4,6)\, G(7,5)\, \Gamma(6,7;3) \tag{2.17}$$

where the notation $1 \equiv \mathbf{r}_1, t_1$ is used for the purpose of clarity, and 1^+ indicates that the time variable t_1 is augmented by a positive infinitesimal. In addition, G, W and Γ satisfy Dyson-like equations:

$$G(1,2) = G_0(1,2) + \int \mathrm{d}[34]\, G_0(1,3)\, \Sigma(3,4)\, G(4,2) \tag{2.18}$$

Equation (2.13) is usually transformed to the eigenvalue problem given by equation (2.12) by using the spectral representation of the Green's function, which will be explicitly given in Section 5.1.

2.4.1 *GW* Approximation

Equations 2.13–2.17 constitute a self-consistent set of coupled equations which can in principle be solved to yield the self-energy. An iterative solution may be obtained by choosing an appropriate initial approximation for Σ before using the equations to generate higher order approximations. In mathematical terms, equations 2.13–2.17 comprise a summation over an infinite perturbation series of interaction diagrams, which may be divergent or conditionally convergent. In fact, a straightforward expansion of the self-energy in terms of the screened interaction W may even yield unphysical results [10]. The choice of which diagrams to sum is usually dictated by physical intuition as there is (so far) no systematic procedure for doing so.

The simplest approximation for the self-energy is given by Hedin's GW-approximation [11], which considers only a single term in the expansion of Σ in terms of the screened interaction. In this approach, an iterative solution to equations 2.13–2.17 is found by making the starting assumption that $\Sigma = 0$, which corresponds to neglecting all exchange and correlation effects in the self-energy, i.e., the Hartree approximation. Insertion of $\Sigma = 0$ into equation (2.17) gives a simple form for the vertex function:

$$\Gamma(1,2;3) = \delta(1,2)\delta(1,3). \tag{2.19}$$

From equation (2.16) the polarizability may be obtained:

$$P(1,2) = -iG(2,1)G(1,2), \tag{2.20}$$

which corresponds to the independent particle, or RPA, polarizability[3]. P may be used to construct W from equation (2.15), and Σ is finally obtained from equation

[3]The RPA ('random-phase approximation') polarizability is also known as the time-dependent Hartree polarizability, which is consistent with the original choice of $\Sigma = 0$.

(2.14) as

$$\Sigma(1,2) = iG(1,2)W(1^+,2), \tag{2.21}$$

which explains the origin of the term '*GW* approximation'.

The calculation of the self-energy Σ first requires a calculation of the Green's function G, which is constructed from a set of one-electron orbitals. To be consistent with the above scheme it is perhaps appropriate to use the Hartree G, which already includes some of the important characteristics of the system, such as a self-consistent treatment of the Coulomb forces via an effective field. However, the Hartree approximation includes large orbital self-interactions[4] and is well known to yield poor results for bandstructures and wavefunctions [14], and therefore is not an optimal starting point. Modern *GW* calculations tend to be based upon density-functional theory within the local-density approximation instead, even though there is no *a priori* reason for doing so. Nevertheless, the GWA-LDA scheme is widely used, and has been shown to produce results for the band gaps and photoemission data of many semiconductors and insulators that are in excellent agreement with experiment [10, 6, 15].

The computation of bandstructures and quasiparticle states within the framework of GWA-LDA is made significantly easier by the fact that the LDA and GWA orbitals are often very similar [16], so that it is sufficient to calculate the expectation value of the self-energy with respect to the LDA wavefunctions when calculating the QP energies. On combining equation (2.12) and equation (2.10), the quasiparticle eigenvalues may be obtained from:

$$E^{\mathrm{QP}}_{m\mathbf{k}} = E^{\mathrm{LDA}}_{m\mathbf{k}} - \langle \psi^{\mathrm{LDA}}_{m\mathbf{k}} | V_{\mathrm{xc}} | \psi^{\mathrm{LDA}}_{m\mathbf{k}} \rangle + \langle \psi^{\mathrm{LDA}}_{m\mathbf{k}} | \Sigma(E_{m\mathbf{k}}) | \psi^{\mathrm{LDA}}_{m\mathbf{k}} \rangle. \tag{2.22}$$

This expression supports the conjecture that the use of LDA wavefunctions is not a natural choice: as part of the GWA-LDA scheme, one must first solve the usual DFT problem, with the effects of exchange and correlation included at the LDA level via

[4]The self-interaction term (J_{ii}, in standard notation [1]) is completely cancelled by exchange terms (K_{ii}) in Hartree-Fock theory, or approximately in SIC (self-interaction correction) schemes [13].

V_{xc}. However, to compute the QP eigenvalues, one must subsequently subtract the $V_{\rm xc}$ contributions from the LDA eigenvalues, before adding the 'true' exchange and correlation effects as contained in the self-energy. Based on empirical evidence (*i.e.*, the good agreement with experiment), the scheme works quite well, although it is somewhat counter-intuitive.

An alternative approach, which is the one adopted in this thesis, is to calculate the self-energy in the Hartree-Fock basis. Like the LDA orbitals, Hartree-Fock orbitals are in many cases found to closely resemble the GWA orbitals ([16], and also concluded in Section 5.5 of this work), which similarly implies that only the diagonal matrix elements of the self-energy operator in the HF basis are required. This choice results in a more intuitive form for calculating the QP eigenvalues. Referring to equation (2.15), we define the screening contribution to W as $W_{\rm s}$ where $W = v + W_{\rm s}$, which allows us to divide Σ into a static and a dynamic part (schematically):

$$\Sigma = \Sigma_{\rm ex} + \Sigma_{\rm s}, \tag{2.23}$$

where $\Sigma_{\rm ex}$ is equivalent to the (static) Hartree-Fock exchange potential $V_{\rm ex}$, originating from the bare Coulomb interaction v. In the basis of Hartree-Fock orbitals, equation (2.22) thus becomes

$$E^{\rm QP}_{m\mathbf{k}} = E^{\rm HF}_{m\mathbf{k}} + \langle \psi^{\rm HF}_{m\mathbf{k}} | \Sigma_{\rm s}(E_{m\mathbf{k}}) | \psi^{\rm HF}_{m\mathbf{k}} \rangle. \tag{2.24}$$

It is clear, therefore, that the GWA corresponds to replacing[5] the static Coulomb interaction v in Hartree-Fock theory by a dynamically screened interaction W, such that the static, nonlocal HF exchange potential $V_{\rm ex}$ is replaced by a nonlocal energy-dependent self-energy $\Sigma(E)$. In other words, Hartree-Fock theory provides the reference for a proper investigation of the effects of correlation and screening. The dynamical screening mechanism is of major importance in solids, which explains

[5]Note that the term 'dynamically screened Hartree-Fock' is a misnomer: in addition to the screened HF potential there is also a Coulomb hole contribution; this will become clear when explicit expressions for $\Sigma(E)$ are presented in Chapter 5.

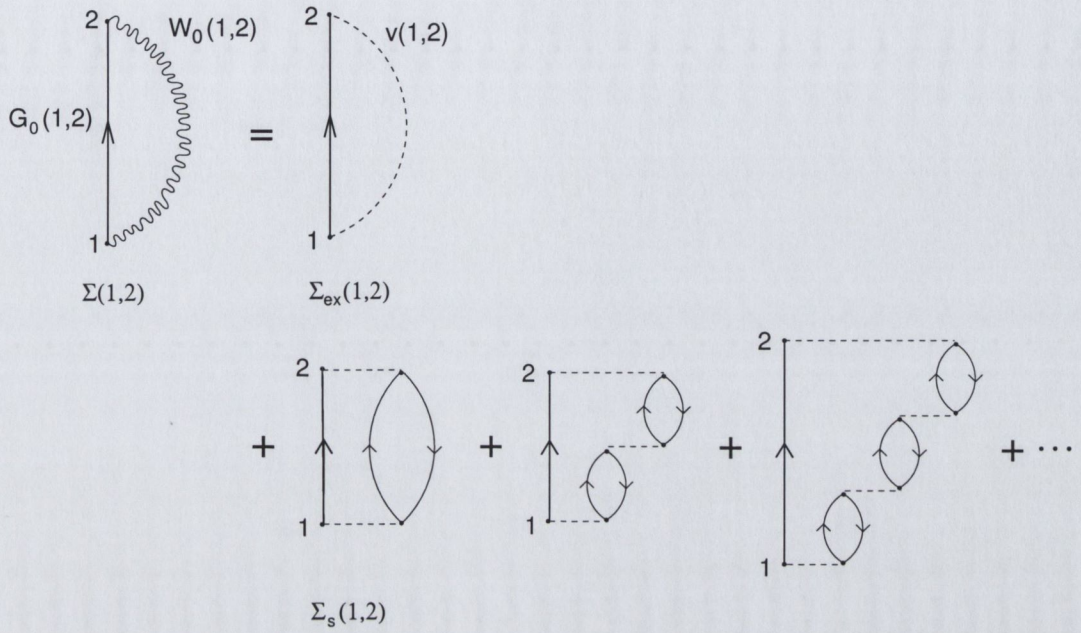


Figure 2-1: Graphical representation of the self-energy $\Sigma(1,2)$ in the GW-approximation. $\Sigma(1,2)$ is composed of the bare-exchange Hartree-Fock part $\Sigma_{\text{ex}}(1,2)$ and the polarization part $\Sigma_{\text{s}}(1,2)$. The latter is represented by an infinite series of non-interacting electron-hole pair bubble diagrams.

why HF theory usually gives unsatisfactory results for the band gaps of semiconductors and insulators. A graphical representation of the exchange and correlation effects within GWA is given in Figure 2-1.

In spite of the relative clarity offered by the GWA-HF scheme, it has not been widely adopted [17] — the preferred approach being the GWA-LDA scheme outlined above. One major reason for this is the popularity of DFT-LDA theory in the solid-state community; it has proven to be very successful (at least for predicting ground state properties) and a huge amount of experience has been gained in the discipline over the last few decades. Conversely, efficient and reliable Hartree-Fock calculations for periodic solids have only been realized in the last decade or so, with, for example, the CRYSTAL family of codes [18]. With regard to self-energy calculations, DFT-LDA provides a good starting point for the iteration of equations 2.14–2.17 in that the initial band energies are often relatively close to the quasiparticle energies, which reduces the need for a proper, self-consistent treatment. This may not be true for

the GWA-HF scheme, as band gaps of semiconductors and insulators are generally overestimated by a much larger amount in Hartree-Fock calculations due to the unscreened interactions. On the other hand, the Hartree-Fock scheme may offer a better initial bandstructure in the case of wide gap materials or transition metal oxides. These topics will be re-addressed in Chapter 5 where calculations using the GWA-HF scheme are presented.

2.5 Choice of Basis

An important aspect of the computational procedure is that of the basis or framework used for the *ab initio* calculations. Traditionally, the plane-wave representation is used because of its natural suitability for constructing Bloch functions in periodic structures and its mathematical simplicity in the evaluation of matrix elements. The plane-wave basis is usually then complemented with a pseudopotential to represent the core states. This approach has met with considerable success in DFT-LDA and GW calculations for bulk systems [7, 6]; ultimately, though, it is not suitable for studying optical properties in systems with large unit cells, systems with reduced symmetry (surfaces, clusters), or systems with high l quantum states ($3d$, $4f$, transition metals). An improvement on the basic plane-wave approach is the real-space/imaginary-time method [19, 20], which is found to scale favourably with the system size.

An alternative first-principles approach to DFT and GWA calculations is the localized orbital approach [21, 22]. This method requires only moderately sized basis sets and has been applied to the study of semi-core states in Si, Ge and CdS [23] and to the Si surface [24]. In addition, the method offers a more physically intuitive description of the wavefunction, being more closely linked to atomic wavefunctions; however, it does involve an increase in mathematical and computational complexity due to the loss of orthogonality of orbitals on neighbouring sites.

In this work we adopt the Linear Combination of Atomic Orbitals (LCAO) approach based on the *ab initio* package CRYSTAL95 [18], in which the atomic or-

bitals are expanded over a basis of Gaussian functions. The CRYSTAL95 code can be used to calculate the ground state energy and electronic wavefunction of periodic systems within the framework of Hartree-Fock or Density Functional theory, using all-electron or pseudopotential basis sets. The one-electron wavefunction or Crystalline Orbital $\psi_i(\mathbf{r};\mathbf{k})$ (equivalent to a molecular orbital for non-periodic systems) is defined over a basis of one-electron Bloch functions $\phi(\mathbf{r};\mathbf{k})$:

$$\psi_i(\mathbf{r};\mathbf{k}) = \sum_{\mu} a_{\mu,i}(\mathbf{k})\phi(\mathbf{r};\mathbf{k}) \tag{2.25}$$

Each Bloch function is expanded over the crystal lattice vectors $\mathbf{R}$ and defined in terms of local functions (atomic orbitals) $\varphi_\mu(\mathbf{r})$:

$$\phi_i(\mathbf{r};\mathbf{k}) = \sum_{R} \varphi_\mu(\mathbf{r} - \mathbf{A}_\mu - \mathbf{R})e^{i\mathbf{k}\cdot\mathbf{R}} \tag{2.26}$$

A single exponent Gaussian Type Function (GTF) may be used for each atomic orbital, but it is more efficient to use linear combinations of individually normalised GTFs, in what is termed a contraction:

$$\varphi_\mu(\mathbf{r} - \mathbf{A}_\mu - \mathbf{g}) = \sum_{j}^{n_G} d_j \ \mathrm{G}[a_j;\mathbf{r} - \mathbf{A}_\mu - \mathbf{g}] \tag{2.27}$$

Contractions of GTFs are traditionally grouped into 'shells' of the same type, i.e. $s, p,$ and d-type, to improve computational efficiency. Although the wavefunctions are ultimately expanded in terms of localised Gaussian orbitals, it should be stressed that the real basis function are the Bloch functions. A concise account of the definitions and properties of the Gaussian Type Functions used in this work is given in Appendix B, along with explicit derivations of the integrals involved in evaluating the optical matrix elements that are used throughout this work.

Chapter 3

Theory of Optical Properties

Light is the principal person in the picture

— *Edouard Manet*

Having outlined the *ab initio* formalism for calculating the electronic structure of solids, the link with the optical properties can now be established by introducing the dielectric function. It is possible to describe the response of the electronic system via the dielectric function in a number of ways, depending on the complexity of the system being examined and on the nature of the interactions which are understood to be important. This chapter is structured along a path of increasing complexity, in the sense that the initial treatment of the dielectric properties will adopt a mean-field approach for describing a simple homogeneous system, and afterwards consider more complex systems and more complicated processes.

Following some general definitions, the first section of this chapter will focus on the derivation of the longitudinal dielectric function, ε^L, in the limiting approximation of the homogeneous, nearly-free electron model. A comparision will be made between the calculated dielectric function in this limiting form and the corresponding experimental result for a real material. The conclusions drawn in this section motivate an improved treatment of the optical response by considering the influence of the charge inhomogeneity which is present in real systems, and this comprises the central topic of the second section. A brief discussion follows, in which further improvements to the calculation of the dielectric function are addressed, in particular

the role of many-body effects in the electronic structure. Finally, a limiting form of the dielectric function is considered in order to introduce the discrete dipole model of optical response, which is specifically designed to treat the description of optical reflection at surfaces.

3.1 Linear Optical Response

In its most general form, the dielectric function may be defined in terms of its inverse as

$$V^{\text{tot}}(\mathbf{r},t) = \int \mathrm{d}\mathbf{r}' \int_{-\infty}^{t} \mathrm{d}t'\, \varepsilon^{-1}(\mathbf{r},\mathbf{r}',t') V^{\text{ext}}(\mathbf{r}',t'), \tag{3.1}$$

in that it relates the total potential $V^{\text{tot}}(\mathbf{r},t)$ inside a material to the external potential $V^{\text{tot}}(\mathbf{r}',t')$, existing at all other positions and all earlier times. It is, in other words, a response function [25]; and being a property of the electronic structure of a solid, it can in principle be used to relate experimental observations of optical and dielectric phenomena to their microscopic origins. In this thesis, the form of the dielectric function that is mostly used is the wavevector and frequency dependent Fourier representation, $\varepsilon(\mathbf{q},\omega)$. In this form, the dielectric function is most useful in its application: in one limit, $\varepsilon(\mathbf{q}\to 0,\omega)$ describes both the absorption and refraction of light, as well as the collective excitations of the Fermi sea; in another limit, $\varepsilon(\mathbf{q},\omega\to 0)$ describes the electrostatic screening of the Coulomb interactions in the many-body system. In a system with infinitesimal translational invariance (e.g., the uniform electron gas), the fields at $\mathbf{r}$ and $\mathbf{r}'$ are related only by the difference $|\mathbf{r}-\mathbf{r}'|$, which allows the inverse dielectric function to be written simply as $\varepsilon^{-1}(|\mathbf{r}-\mathbf{r}'|,t')$. On Fourier transforming equation (3.1) with this form of ε^{-1} we obtain the relation

$$V^{\text{tot}}(\mathbf{q},\omega) = \varepsilon^{-1}(\mathbf{q},\omega) V^{\text{ext}}(\mathbf{q},\omega), \tag{3.2}$$

which defines $\varepsilon(\mathbf{q},\omega)$, as required. Removing the approximation of a homogeneous system requires a modified form of the dielectric function, which will be discussed in Section 3.2.

In general, the dielectric function is a complex[1] quantity, $\hat{\varepsilon} = \varepsilon_1 + i\varepsilon_2$; in which the real and imaginary parts are formally related by the Kramers-Kronig dispersion relation,

$$\varepsilon_1(\omega) = 1 + \frac{2}{\pi} P \int_0^\infty \frac{\omega' \varepsilon_2(\omega')}{(\omega')^2 - \omega^2} \mathrm{d}\omega', \tag{3.3}$$

and general formulae exist [25, 26] that connect the real and imaginary parts with directly observable quantities such as reflection and absorption. In the following treatment, the dielectric function is expressed in terms of the electronic structure by deriving an expression for the optical absorption directly, and then relating this expression back to the dielectric function.

3.1.1 Quantum Mechanical Approach

In the conventional quantum mechanical approach, due to Bassani and Parravicini [27], the connection between the electronic bandstructure of a material and its optical properties is made by neglecting the influence of system inhomogeneity, collective excitations and other many-body phenomena (i.e., plasmons, excitons, *etc.*), considering only the contributions of direct interband transitions. The basis of the approach is to treat the vector potential of the electromagnetic field as a small perturbation on the kinetic energy of the electrons. The Hamiltonian for a system of N electrons is then written as the sum of an unperturbed part arising from the kinetic $\hat{H}_{\text{K.E.}}$ and potential $\hat{H}_{\text{P.E.}}$ energy of the electrons, plus an extra contribution $\hat{H}_{\text{E.M.}}$ due to the presence of the electromagnetic field:

$$\hat{H} = \hat{H}_{\text{K.E.}} + \hat{H}_{\text{E.M.}} + \hat{H}_{\text{P.E.}} = \frac{1}{2} \sum_{i=1}^{N} [\hat{\mathbf{p}}_i + \mathbf{A}(\mathbf{r}_i, t)]^2 + \hat{H}_{\text{P.E.}}. \tag{3.4}$$

This assumes that the scalar potential ϕ is taken to be zero by applying the appropriate (Coulomb) gauge transformation. By disregarding nonlinear effects (terms in

[1] The accent denoting complex character will generally be omitted for clarity, i.e. $\hat{\varepsilon} \to \varepsilon$.

$\mathbf{A}^2$), the interaction Hamiltonian may be rewritten as

$$\hat{H}_{\text{E.M.}} = \sum_{i=1}^{N} \mathbf{A}(\mathbf{r}_i, t) \cdot \hat{\mathbf{p}}_i \tag{3.5}$$

The link with the bandstructure is achieved by treating $\hat{H}_{\text{E.M.}}$ as a time dependent perturbation on the electronic states of the system. This term will cause electrons to make transitions between occupied and unoccupied bands, and from the transition probability rate one can derive expressions for the optical constants. Since light is a transverse wave, the quantity that is obtained by following this procedure is called the transverse dielectric function, $\hat{\varepsilon}^{\text{T}}$. However, later in this thesis we will be investigating the effects of phenomena which originate from microscopic fluctuations in the dielectric response (the local-field effects), which are easier to consider when deriving the longitudinal dielectric function, $\hat{\varepsilon}^{\text{L}}$. At optical wavelengths (long wavelength limit), an isotropic medium cannot distinguish between electric fields parallel or perpendicular to the wavevector of the perturbation, which in fact means that the longitudinal and transverse dielectric functions become equal [25, 28]. Therefore, it is sufficient to restrict ourselves to considering the longitudinal dielectric function when describing the optical properties.

3.1.2 Longitudinal Dielectric Function

In their classic paper, Ehrenreich and Cohen [29] discussed the response of an electronic system to an external charge density, which is equivalent to considering the influence of a scalar potential. In this manner, the quantity that is obtained is the longitudinal dielectric function[2]. Their approach was to use a self-consistent field approximation (SCF) in the derivation. The basic tenet of the SCF approach is as follows: in the presence of an external perturbing potential V^{ext}, the charges within the system become redistributed, and this induced charge density has a potential which reduces, or screens, the original external potential. The total self-consistent potential V^{SCF} is composed of the external potential V^{ext} plus the induced potential

[2]In the Coulomb gauge, $\nabla \cdot \mathbf{A} = 0$, and the scalar potential is static and therefore longitudinal.

V^{ind},

$$V^{\text{SCF}} = V^{\text{ind}} + V^{\text{ext}}, \tag{3.6}$$

such that the final charge distribution is consistent with the total potential.

The principle ideas in the derivation can easily be illustrated in the context of the free-electron gas, and subsequently be extended to periodic or strongly inhomogeneous systems. The following treatment is based on the derivation by Wooten [25]. The unperturbed Hamiltonian satsifies the Schrödinger equation, $\hat{H}_0|\mathbf{k}\rangle = E_{\mathbf{k}}|\mathbf{k}\rangle$, where the free electron eigenstates are given by

$$|\mathbf{k}\rangle = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}}, \qquad V = \text{ normalization volume.} \tag{3.7}$$

The total single-particle Hamiltonian $\hat{H} = \hat{H}_0 + V^{\text{SCF}}(\mathbf{r}, t)$, is related to the total density operator $\hat{\rho} = \hat{\rho}_0 + \hat{\rho}'$ (where $\hat{\rho}_0$ is the density operator for the unperturbed system), via the Liouville equation of motion,

$$i\frac{\partial \hat{\rho}}{\partial t} = \left[\hat{H}, \hat{\rho}\right], \tag{3.8}$$

Using the Liouville equation for the unperturbed system, $i\partial\hat{\rho}_0/\partial t = [\hat{H}_0, \hat{\rho}_0]$, equation (3.8) may be linearised to give

$$i\frac{\partial \hat{\rho}^{\text{SCF}}}{\partial t} = \left[\hat{H}_0, \hat{\rho}'\right] + \left[\hat{V}^{\text{SCF}}(\mathbf{r}, t), \hat{\rho}_0\right]. \tag{3.9}$$

We take matrix elements between states $|\mathbf{k}\rangle$ and $|\mathbf{k}+\mathbf{q}\rangle$ to yield the equation of motion for the matrix elements:

$$i\frac{\partial}{\partial t}\langle \mathbf{k}+\mathbf{q}|\hat{\rho}'|\mathbf{k}\rangle = \langle \mathbf{k}+\mathbf{q}|\left[\hat{H}_0, \hat{\rho}'\right]|\mathbf{k}\rangle + \langle \mathbf{k}+\mathbf{q}|\left[\hat{V}^{\text{SCF}}(\mathbf{r}, t), \hat{\rho}_0\right]|\mathbf{k}\rangle. \tag{3.10}$$

The total perturbing SCF potential is Fourier decomposed as

$$V^{\text{SCF}}(\mathbf{r}, t) = \sum_{\mathbf{q}'} V^{\text{SCF}}(\mathbf{q}', t) e^{i\mathbf{q}'\cdot\mathbf{r}}, \tag{3.11}$$

and by making use of the property of the unperturbed density operator $\hat{\rho}_0 |\mathbf{k}\rangle = f(E_\mathbf{k})|\mathbf{k}\rangle$, where $f(E_\mathbf{k})$ is the Fermi-Dirac distribution function, we may evaluate the commutators to leave

$$i\frac{\partial}{\partial t}\langle \mathbf{k}+\mathbf{q}|\hat{\rho}'|\mathbf{k}\rangle = (E_{\mathbf{k}+\mathbf{q}} - E_\mathbf{k})\langle \mathbf{k}+\mathbf{q}|\rho'|\mathbf{k}\rangle + [f(E_\mathbf{k}) - f(E_{\mathbf{k}+\mathbf{q}})]\, V^{\mathrm{SCF}}(\mathbf{q},t). \quad (3.12)$$

At this point we have to make an adiabatic approximation for the time dependence of the perturbation in order to produce a dielectric function that is both physical and complex. The external potential is assumed to have the form $V^{\mathrm{ext}}(\mathbf{r},t) = V^{\mathrm{ext}}(\mathbf{r},\omega)e^{-i\omega t}e^{\eta t}$, where η is an infinitesimal, and the limit $(\eta \to 0^+)$ is to be taken at the end of the calculation. This form for the external potential shares a time dependence with the induced and total potentials, as well as the corresponding density fluctuations. Therefore, equation (3.12) reduces to [25]

$$\langle \mathbf{k}+\mathbf{q}|\hat{\rho}'|\mathbf{k}\rangle = \frac{f(E_{\mathbf{k}+\mathbf{q}}) - f(E_\mathbf{k})}{(E_{\mathbf{k}+\mathbf{q}} - E_\mathbf{k}) - \omega - i\eta} V^{\mathrm{scf}}(\mathbf{q},t) \quad (3.13)$$

We now use Poisson's equation, $\nabla^2 V^{\mathrm{ind}}(\mathbf{r},t) = -4\pi\delta n$, to relate the induced screening potential to the induced change in electron density, δn. Making use of the following expression of the charge density [25],

$$n = \mathrm{Tr}[\,\hat{\rho}\,\delta(\mathbf{r}_e - \mathbf{r})], \quad (3.14)$$

where $\delta(\mathbf{r}_e - \mathbf{r})$ is the electron position operator, the induced density δn may be shown to be given by

$$\delta n = \frac{1}{V}\sum_{\mathbf{k},\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}}\langle \mathbf{k}+\mathbf{q}|\hat{\rho}'|\mathbf{k}\rangle. \quad (3.15)$$

The induced potential is Fourier analysed in the same manner as equation (3.11), and when combined with Poisson's equation and equation (3.15) yields

$$V^{\mathrm{ind}}(\mathbf{q},t) = \frac{4\pi}{q^2 V}\sum_{\mathbf{k}}\langle \mathbf{k}+\mathbf{q}|\hat{\rho}'|\mathbf{k}\rangle. \quad (3.16)$$

An expression for the dielectric function may now be derived by recalling that

$V^{\mathrm{SCF}}(\mathbf{q},t) = V^{\mathrm{ext}}(\mathbf{q},t)/\varepsilon(\mathbf{q},\omega)$, and that the total, induced and external potentials are related by equation (3.6). On combining equation (3.13) for the total SCF potential with equation (3.16) for the induced potential we obtain the Lindhard dielectric function:

$$\varepsilon(\mathbf{q},\omega) = 1 - \frac{4\pi}{q^2 V}\sum_{\mathbf{k}} \frac{f(E_{\mathbf{k}+\mathbf{q}}) - f(E_{\mathbf{k}})}{E_{\mathbf{k}+\mathbf{q}} - E_{\mathbf{k}} - \omega - i\eta}. \tag{3.17}$$

This form of $\varepsilon(\mathbf{q},\omega)$ is also known as the random-phase approximation (RPA) dielectric function, which is how it will be generally referred to in this thesis.

This expression has been derived for the uniform electron gas. It is easily extendable to the periodic solid by considering Bloch functions instead of simple plane waves:

$$|n\mathbf{k}\rangle = \frac{1}{\sqrt{V}} u_{n\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{3.18}$$

where $u_{n\mathbf{k}}(\mathbf{r})$ is the spatially periodic part of the wavefunction. The derivation results in the following expression:

$$\varepsilon(\mathbf{q},\omega) = 1 - \frac{4\pi}{q^2 V}\sum_{n,n'}\sum_{\mathbf{k}} \frac{f(E_{n'\mathbf{k}+\mathbf{q}}) - f(E_{n\mathbf{k}})}{E_{n'\mathbf{k}+\mathbf{q}} - E_{n\mathbf{k}} - \omega - i\eta} |\langle n'\mathbf{k}+\mathbf{q}|e^{i\mathbf{q}\cdot\mathbf{r}}|n\mathbf{k}\rangle|^2, \tag{3.19}$$

which is the final form of the longitudinal dielectric function in the RPA.

Since we are interested in the macroscopic dielectric function, i.e., the quantity that is directly observable in experiments, we evaluate equation (3.19) in the limit of vanishing photon wave-vector, $\mathbf{q}$, which to a good approximation corresponds to the wavelength of light inside the solid.[3] This may be done by using the relation [30]

$$\langle c\mathbf{k}+\mathbf{q}|e^{i\mathbf{q}\cdot\mathbf{r}}|v\mathbf{k}\rangle = \frac{\langle c\mathbf{k}+\mathbf{q}|[e^{i\mathbf{q}\cdot\mathbf{r}}, \hat{H}]|v\mathbf{k}\rangle}{E_{v\mathbf{k}} - E_{c\mathbf{k}+\mathbf{q}}}, \tag{3.20}$$

which in the $\mathbf{q} \to 0$ limit introduces matrix elements of the velocity operator $\hat{\mathbf{v}}$, defined by

$$\hat{\mathbf{v}} = \frac{1}{i}\left[\hat{\mathbf{r}}, \hat{H}\right]. \tag{3.21}$$

[3] Wavelength of light ≈ 5000Å $\gg$ Cell dimensions ~ 1Å

Equations (3.20) and (3.21) are valid expressions for any single-particle Hamiltonian satisfying $\hat{H}|n\mathbf{k}\rangle = E_{n\mathbf{k}}|n\mathbf{k}\rangle$. However, the transformation between the velocity matrix elements and the more useful momentum matrix elements is affected by the presence of a nonlocal potential:

$$\hat{\mathbf{v}} = \frac{1}{i}[\hat{\mathbf{r}}, \hat{H}] = \hat{\mathbf{p}} + \frac{1}{i}[\hat{\mathbf{r}}, V_{\mathrm{nl}}], \tag{3.22}$$

where we have written the Hamiltonian as the sum of local (l) and nonlocal (nl) potentials, $\hat{H} = \hat{\mathbf{p}}^2/2 + V_{\mathrm{l}} + V_{\mathrm{nl}}$ [30]. This has two major implications. Firstly, a reliable calculation of the optical properties derived from a DFT-LDA bandstructure must either account explicitly for the second term of equation (3.22) if a nonlocal pseudopotential is used to represent the core states. This is a common scheme used in conjunction with a plane-wave basis set, as it considerably reduces the size of the basis set in that case (core states are very demanding in plane-wave calculations; the sharply-peaked, localized structure of the wavefunction is not well described by the delocalized plane-waves.) An alternative approach is to utilize an all-electron basis set, which does not introduce any element of nonlocality into the Hamiltonian. The latter approach is quite attainable within the Gaussian orbital framework, and is the one which has been adopted in this work. Secondly, it also effects the calculation of the self-energy in Chapter 5, which requires the computation of the RPA dielectric function at long wavelength. However, the Hartree-Fock bandstructure will be used initially in the self-energy calculation, and this, of course, is inherently nonlocal in character. This means that some parameters have to be estimated at $\mathbf{q} \to 0$, a point which is re-addressed in the later chapters.

Neglecting these nonlocal corrections, and incorporating the momentum matrix element form, we arrive at the final[4] expression for the long-wavelength dielectric

[4]The dielectric function is strictly speaking a second-rank tensor, $\varepsilon_{\alpha\beta}(\omega)$, which for the crystals of cubic symmetry considered in this thesis, is simply diagonal, $\varepsilon_{\alpha\beta}(\omega) = \delta_{\alpha\beta}\varepsilon(\omega)$, with the elements given as in the text.

function:

$$\varepsilon_M(\omega) = 1 + \frac{16\pi}{V} \sum_{c,v} \sum_{\mathbf{k}} \frac{1}{[E_{c\mathbf{k}} - E_{v\mathbf{k}}]} \frac{|\langle c\mathbf{k}|\mathbf{p}|v\mathbf{k}\rangle|^2}{[(E_{c\mathbf{k}} - E_{v\mathbf{k}})^2 - (\omega + i\eta)^2]}. \tag{3.23}$$

In the above, The Fermi functions $f(E)$ have been taken to define occupancies of 0 for conduction bands or 1 for valence bands, and the spin variable has been explicitly integrated out.

3.1.3 Calculations

The calculation of the RPA dielectric function given in equation (3.23) requires the generation of one-electron wavefunctions, eigenvalues and momentum matrix elements, for a range of valence and conduction band indices throughout the irreducible Brillouin zone. This may be quite a demanding task, especially considering that the **k**-point integration must be done extremely accurately to yield a smooth function at values of ω close to transition resonances. The integration scheme that has been adopted to achieve this is the 'analytic continuation' technique of Gilat and Raubenheimer, the details of which are presented in Appendix C.2. To reduce this computational demand, it is common to consider in isolation the imaginary part of equation (3.23), which may be transformed into the following form:

$$\varepsilon_2(\omega) = \frac{8\pi^2}{V} \sum_{c,v} \sum_{\mathbf{k}} \frac{|\langle c\mathbf{k}|\mathbf{p}|v\mathbf{k}\rangle|^2}{[E_{c\mathbf{k}} - E_{v\mathbf{k}}]^2} \delta(E_{c\mathbf{k}} - E_{v\mathbf{k}} - \omega) \tag{3.24}$$

where use has been made of the standard relation (P denotes principal value)

$$\lim_{\eta \to 0^+} \frac{1}{X - i\eta} = P\frac{1}{X} + i\pi\delta(X). \tag{3.25}$$

The calculation of $\varepsilon_2(\omega)$ has now been made significantly easier because the integrand need only be evaluated at particular values of ω. The real part can then be obtained from the imaginary part by making use of the Kramers-Kronig relation, equation (3.3). However, in this work we have chosen to compute the real and imaginary parts directly from equation (3.23).

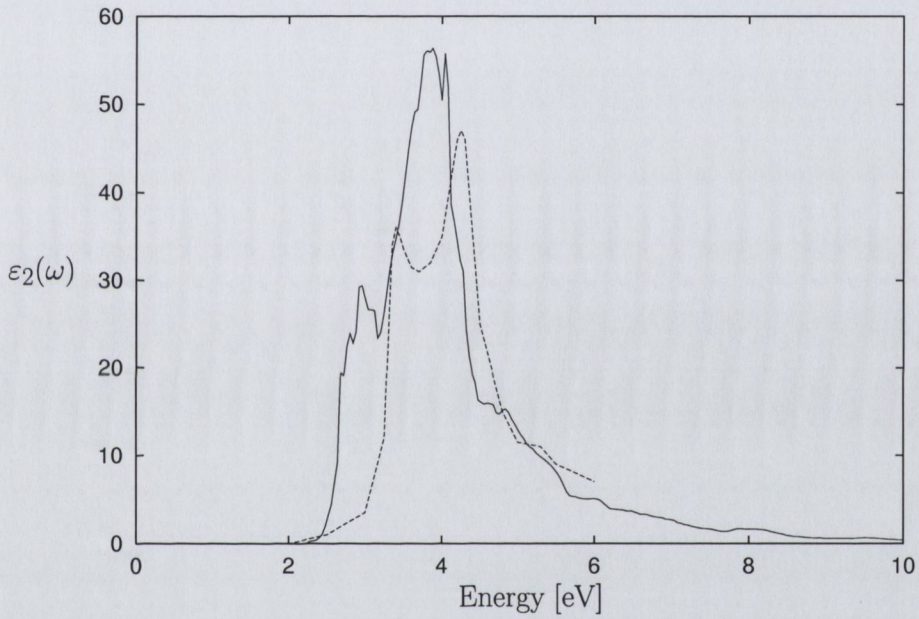


Figure 3-1: Imaginary part of the dielectric function of silicon (solid line), compared with the experimental value of Aspnes *et al.* [31] (dotted line), as a function of energy. The amplitude of the main peak in the calculated spectrum slightly overestimates the experimental amplitude and appears at a lower energy by 0.5 eV. The calculations were performed using an all-electron 8-1111G(d) Gaussian basis set with 30 conduction bands, and the Brillouin zone integration was carried out using the analytic continuation technique of Gilat and Raubenheimer [32], with 195 special-points in the irreducible zone.

A sample calculation of $\varepsilon_2(\omega)$ obtained in this manner is presented in Figure 3.1.3 for silicon. Although the lineshape and amplitude are generally in qualitative agreement with the experiment, there are several discrepancies if one examines the results more closely. Firstly, the dominant peaks in the spectrum occur at a lower frequency than expected by about 0.5 eV. This may be attributed to the underestimation of the band gap in the LDA for silicon. Secondly, the lineshape is poorly reproduced with respect to the relative magnitude of the E_1 and E_2 peaks, as well as with the overall intensities. The reasons for this are not so obvious. What is certain though, is that they must derive from the various assumptions made that comprise the one-electron description, such as effects due to exchange and correlation in the electronic system or effects deriving from spatial inhomogeneity.

The overall reliability of the calculated spectrum may be tested via the Thomas-Reiche-Kuhn f-sum rule, which is quite valid within the RPA [30, 33]. It is given by:

$$\int_0^\infty \omega\varepsilon_2(\omega)\mathrm{d}\omega = \frac{1}{2}\pi\omega_p^2, \tag{3.26}$$

where the plasma frequency ω_p is given by $\omega_p = \sqrt{4\pi n}$ and n denotes the density of electrons contributing to the optical properties. This of course depends very much on the frequency interval being considered, and therefore it is common to define the sum rules in terms of an effective number of electrons per atom, η_{eff}, taking part in transitions over a finite frequency range,

$$\int_0^{\omega_c} \omega\varepsilon_2(\omega)\mathrm{d}\omega = \frac{1}{2}\pi(4\pi N_a)\eta_{\text{eff}}(\omega_c) \tag{3.27}$$

where N_a is the density of atoms.

Figure 3-2 demonstrates the success of the optical sum rules for silicon. There are in total four valence electrons per Si atom that primarily contribute to the optical transitions at low frequencies, as confirmed by the figure; however, the sum rule (which is effectively a valence electron sum rule in this case) is violated above an energy of 90 eV, which marks the onset of optical transitions involving deep core states. In theory, the calculation could be extended to extremely high frequencies

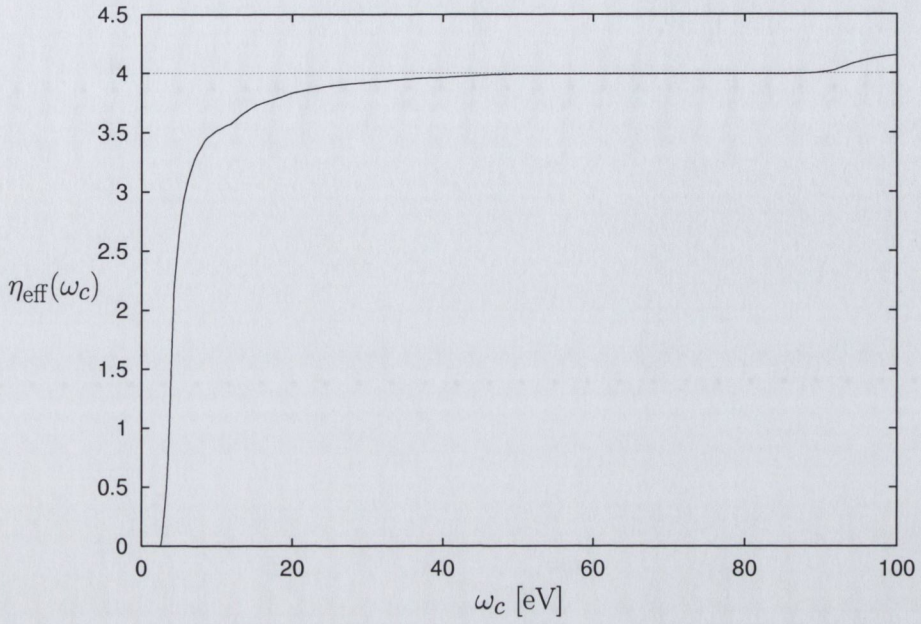


Figure 3-2: The effective number of electrons per atom as a function of cutoff frequency for silicon, obtained from a numerical integration of the calculated $\varepsilon_2(\omega)$ in Figure 3-1. $\eta_{\text{eff}}(\omega_c)$ converges to the value 4.0 electrons per Si atom, in agreement with the theoretical prediction for the valence electrons. For cutoff frequencies $\omega_c > 90$ eV, optical transitions involving core states contribute to $\eta_{\text{eff}}(\omega_c)$. The same computational conditions were used as in Figure 3-1.

such that even the core excitations are exhausted, although in practice this would require a prohibitively large set of states and a very accurate integration in this region. Nevertheless, it is apparent, based on the comparison with the experimental spectra, that several phenomena which influence the optical properties must be coming into play, and the following section attempts to deal with one of these.

3.2 The Influence of Inhomogeneity

The treatment of the dielectric function in the previous section is mainly suited towards systems of a homogeneous character (such as metals), since the derivation is based upon the model system of the electron gas. In materials such as semiconductors, insulators and transition metals, however, the periodicity of the lattice potentials has the effect of localizing the electronic wavefunctions relative to the plane-wave states, which in general forces the electronic response to be spatially inhomogeneous. This has implications on the type of approximation one may make

when deriving the dielectric function, and it is easier to investigate the type of phenomena which must be accounted for by re-examining the problem in a real-space representation.

The most general form of the longitudinal dielectric function in a real-space representation was expressed by equation (3.1). Since we wish to discuss the spatial properties only, the form of equation (3.1) may be simplified by performing a Fourier transform on the time variables only, to yield:

$$V^{\text{tot}}(\mathbf{r},\omega) = \int \mathrm{d}\mathbf{r}' \varepsilon^{-1}(\mathbf{r},\mathbf{r}',\omega) V^{\text{ext}}(\mathbf{r}',\omega). \tag{3.28}$$

Although written in a real space representation, this relation between total and external fields is quite distinct from the expression that was presented earlier [i.e., equation (3.2)], in that it is a true nonlocal relation, in the sense that the dielectric function explicitly depends on the both variables $\mathbf{r}$ and $\mathbf{r}'$ and not just the difference $|\mathbf{r} - \mathbf{r}'|$. This is necessary because for inhomogeneous systems the induced density at a point $\mathbf{r}$ does not simply depend on the distance from the source at $\mathbf{r}'$ but on its local environment; the local polarizability will be different in a bond than in an interstitial site, for instance. Note that the dielectric function may be written in terms of the polarizability as [2]:

$$\varepsilon(\mathbf{r},\mathbf{r}',\omega) = \delta(\mathbf{r} - \mathbf{r}') - \int \mathrm{d}\mathbf{r}'' v(\mathbf{r},\mathbf{r}'') P(\mathbf{r}'',\mathbf{r}',\omega), \tag{3.29}$$

where $v(\mathbf{r},\mathbf{r}')$ is the Coulomb potential given by $v(\mathbf{r},\mathbf{r}') = 1/|\mathbf{r} - \mathbf{r}'|$.

In the continuum approach which we have been following, it is appropriate to make use of the translational symmetry of the crystal, which requires that

$$\varepsilon(\mathbf{r} + \mathbf{R}, \mathbf{r}' + \mathbf{R}, \omega) = \varepsilon(\mathbf{r},\mathbf{r}',\omega), \tag{3.30}$$

$\mathbf{R}$ denoting a lattice vector. This enables the dielectric function to be written in a

Fourier representation, whereby equation (3.28) is transformed to

$$V^{\text{tot}}(\mathbf{q}+\mathbf{G},\omega) = \sum_{\mathbf{G}'} \varepsilon^{-1}(\mathbf{q}+\mathbf{G},\mathbf{q}+\mathbf{G}',\omega) V^{\text{ext}}(\mathbf{q}+\mathbf{G}',\omega). \tag{3.31}$$

Here the wavevector $\mathbf{q}$ is restricted to lie inside the first Brillouin zone, while the sum on $\mathbf{G}'$ spans the (infinite) reciprocal lattice.

The above expression may be interpreted as follows: a perturbing electric field of frequency ω and wavevector $\mathbf{q}+\mathbf{G}$ gives rise to many microscopic fields of frequency ω and wavevector $\mathbf{q}+\mathbf{G}'$, which in real-space corresponds to a rapidly oscillating (in space) microscopic field within the unit cell. This phenomenon has been termed a 'local field effect' on the basis that the microscopic polarizing electric field is different from the macroscopic electric field as a result of these short range fluctuating responses.

We can see clearly where the local field effects arise by deriving an explicit form for the dielectric function. Although $\varepsilon(\mathbf{q}+\mathbf{G},\mathbf{q}+\mathbf{G}',\omega)$ in theory spans an infinite number of reciprocal lattice vectors, in practice we may truncate the sum in equation (3.31) and write it in the form of a matrix in the $\mathbf{G}$ and $\mathbf{G}'$ vectors: $\varepsilon_{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega)$. This is the form we shall assume from now on, to indicate a finite representation in reciprocal space. The extension of the RPA longitudinal dielectric function [equation (3.19)] to include the local field effects was carried out by Adler [28] and Wiser [34], and essentially corresponds to altering equation (3.11) so as to include a summation over $\mathbf{G}$-vectors in the expansion of the potential:

$$V^{\text{AW}}(\mathbf{r},t) = \sum_{\mathbf{q},\mathbf{G}} V^{\text{AW}}(\mathbf{q},\mathbf{G},t) e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}. \tag{3.32}$$

This inclusion of Umklapp processes is necessary to describe the rapid variations in potential within the cell that we require. On carrying out the derivation we obtain the extended RPA dielectric matrix:

$$\varepsilon^{\text{AW}}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) = \delta_{\mathbf{G},\mathbf{G}'} - v_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) P_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega), \tag{3.33}$$

where $v_{\mathbf{GG'}}(\mathbf{q})$ is the Fourier transform of the bare Coulomb interaction, i.e. $v_{\mathbf{GG'}}(\mathbf{q}) = 4\pi e^2 \delta_{\mathbf{GG'}}/|\mathbf{q}+\mathbf{G'}|^2$, and the polarizability $P_{\mathbf{GG'}}(\mathbf{q},\omega)$ is given by ($\eta \to 0^+$):

$$\begin{aligned} P_{\mathbf{GG'}}(\mathbf{q},\omega) &= \frac{2}{V}\sum_{n,n'}\sum_{\mathbf{k}} \frac{f(E_{n'\mathbf{k}+\mathbf{q}}) - f(E_{n\mathbf{k}})}{E_{n'\mathbf{k}+\mathbf{q}} - E_{n\mathbf{k}} - \omega - i\eta} \\ &\quad \times \langle n'\mathbf{k}+\mathbf{q}|e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}|n\mathbf{k}\rangle\langle n\mathbf{k}|e^{-i(\mathbf{q}+\mathbf{G'})\cdot\mathbf{r}}|n'\mathbf{k}+\mathbf{q}\rangle \end{aligned} \tag{3.34}$$

where the sum over $\mathbf{k}$ is performed inside the first Brillouin zone, and n, n' are the band indices labelling the Bloch eigenstates $|n\mathbf{k}\rangle$, $|n'\mathbf{k}\rangle$ of energy $E_{n\mathbf{k}}$, $E_{n'\mathbf{k}}$, in the solid. The factor of 2 arises from an explicit integration over the spin variable.

In this representation, the macroscopic dielectric function $\varepsilon_M(\omega)$ is obtained by inverting numerically the dielectric matrix in the long-wavelength ($\mathbf{q} \to 0$) limit and taking the $\mathbf{G} = 0, \mathbf{G'} = 0$ elements:[28, 34]

$$\varepsilon_M(\omega) = \frac{1}{\lim\limits_{\mathbf{q}\to 0} \varepsilon^{-1}_{0,0}(\mathbf{q},\omega)}. \tag{3.35}$$

Equation (3.35) contrasts sharply with the RPA result given by equation (3.23), which is equivalent to the following:

$$\varepsilon_M(\omega) = \lim_{\mathbf{q}\to 0} \varepsilon_{0,0}(\mathbf{q},\omega). \tag{3.36}$$

This confirms the fact that the common approximation of equation (3.36) is equivalent to neglecting the local field effects altogether. Finally, the relationship between the microscopic potentials $V(\mathbf{q}+\mathbf{G},\omega)$ and the macroscopic potential $V(\mathbf{q},\omega)$ is given by [35]

$$V(\mathbf{q}+\mathbf{G},\omega) = \frac{\varepsilon^{-1}_{0,\mathbf{G}}}{\varepsilon^{-1}_{0,0}} V(\mathbf{q},\omega). \tag{3.37}$$

These expressions demonstrate that the local field effects are contained in the off-diagonal elements of the dielectric matrix, and for homogeneous systems, the dielectric matrix becomes diagonal. As mentioned previously, it should be stressed that these effects strictly arise due to the inhomogeneity of the crystal and are intrinsically linked to the nonlocal nature of the polarizability, $P(\mathbf{r},\mathbf{r}',\omega)$. However,

$N_{\mathbf{G}}$	ε_∞	$N_{\mathbf{G}}$	ε_∞
1	11.729	65	10.545
15	11.012	89	10.541
27	10.857	113	10.540
59	10.549	169	10.540

Table 3.1: Influence of local fields on the calculated dielectric constant of silicon, ε_∞, through inversion of the dielectric matrix of size $N_{\mathbf{G}}$. The dielectric matrix is calculated at wavevector $\mathbf{q} = \frac{2\pi}{a}(\frac{1}{16}\frac{1}{16}\frac{1}{16})$, which is not sufficiently small to yield an accurate comparison with the experimental value of 11.4 (0 K) [39]. However, results of ε_∞ published elsewhere [36, 37, 38] for $\mathbf{q} \to 0$ are found to systematically overestimate the experimental value, in some cases by up to 20% [38].

when expressed in the form of equation (3.33), the dielectric matrix does not immediately lend itself to an intuitive understanding of the physical nature of the screening processes and local field effects described within. In order examine this further, in Chapter 4 we utilize a technique which enables such information to be visualized through a set of real-space equipotential surfaces. Firstly though, the magnitude of the off-diagonal response will be briefly examined by means of a few calculations.

3.2.1 Calculations

To illustrate the importance of the local fields on the dielectric function, Table 3.1 presents a calculation of the dielectric constant $\varepsilon_\infty = \varepsilon_M(\omega \to 0)$, with and without the inclusion of local fields. The first element of the table, $N_{\mathbf{G}} = 1$, corresponds to the RPA dielectric constant, as given by equation (3.36); the remaining elements correspond to the Adler-Wiser dielectric constant (i.e., with local-fields included), calculated for increasing size of the dielectric matrix. The value of $\varepsilon_\infty = 1/\varepsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\omega \to 0)$ is well converged at 169 **G**-vectors, and it can be easily deduced from the data that the inclusion of the local-field effects causes a 10.1% reduction in the macroscopic dielectric constant. Similar reductions have been observed by several other authors for silicon: 10.4% [36], 10.2% [37], 9.5% [38].

3.3 Beyond the Independent Electron Approach

It is apparent both from Figure 3-1 and from the results in the previous section that the macroscopic dielectric function $\varepsilon_M(\omega)$ calculated within the RPA, even with the effects of the local fields included, produces only a reasonable quantitative agreement with experiment. This discrepancy may be attributed to a number of factors of varying complexity, and although this thesis will not directly consider the treatment of such effects it is worth examining them briefly because of their relevance to the calculation of the self-energy in Chapter 5 and the use of the DFT-LDA bandstructure in calculating the dielectric function.

3.3.1 Quasiparticle Corrections to Bandstructure

The calculation of the longitudinal dielectric function depends critically upon the quality of both the set of band eigenvalues which specify the optical transition energies as well as the Bloch eigenvectors which determine the optical matrix elements. The *ab initio* techniques that are at our disposal for calculating this initial bandstructure are LDA-DFT and Hartree-Fock, techniques which are well known to underestimate and overestimate, respectively, the magnitude of the optical gap in a range of materials. To correct this flaw one needs to treat effects that lie beyond the independent electron model and use instead the many-body quasiparticle bandstructure.

One formalism that has been developed which attempts to include the effects of the corrected bandstructure in the optical response at the level of a mean-field theory is the scissors operator (SO) scheme [40]. In this approach, the many-body correction to the DFT-LDA bandstructure via the self-energy is introduced into the dielectric function by using the LDA eigenvalues with a **k**-dependent energy shift in the conduction bands, while keeping the same LDA wavefunctions (and hence, the same matrix elements). A further simplification of this scheme is to apply a rigid 'shift' to the conduction bands, which corresponds to a **k**-*independent* scissors operator [41, 42]. This results in a new expression for the dielectric function of the

form ($\eta \to 0^+$):

$$\begin{aligned}
\varepsilon^{\mathrm{SO}}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) &= \delta_{\mathbf{G},\mathbf{G}'} - v_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) P^{\mathrm{SO}}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega), \\
P^{\mathrm{SO}}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) &= \frac{2}{V} \sum_{n,n'} \sum_{\mathbf{k}} \frac{f(E_{n'\mathbf{k}+\mathbf{q}}) - f(E_{n\mathbf{k}})}{E_{n'\mathbf{k}+\mathbf{q}} - E_{n\mathbf{k}} + \Delta(\delta_{n'c} - \delta_{nc}) - \omega - i\eta} \\
&\quad \times \langle n'\mathbf{k}+\mathbf{q}| e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} |n\mathbf{k}\rangle \langle n\mathbf{k}| e^{-i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}} |n'\mathbf{k}+\mathbf{q}\rangle, \qquad (3.38)
\end{aligned}$$

where only the Δ term is related to the scissors operator, and c denotes a conduction band.

In Chapter 5 the self-energy shift is calculated as a correction to the Hartree-Fock bandstructure, rather than to the more traditional LDA bandstructure. However, it was not attempted to apply the scissors operator scheme to the HF formalism for reasons that will become apparent later; in particular, the corrected bandstructures do not approach those derived from an LDA starting point.

3.3.2 DFT-LDA Dielectric Function

Since the Bloch states that have been used to compute $\varepsilon(\mathbf{q},\omega)$ contain the effects of exchange and correlation at the level of LDA, it is more consistent to also include these effects when describing the polarization. This is achieved in the framework of time-dependent density functional theory, which leads to the following expression for the longitudinal dielectric function:

$$\varepsilon^{\mathrm{DFT}}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) = \delta_{\mathbf{G}\mathbf{G}'} - v_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) P_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) [\delta_{\mathbf{G}\mathbf{G}'} - f^{\mathrm{xc}}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) P_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega)]^{-1} \qquad (3.39)$$

where $P_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega)$ is the polarization function of independent particles (RPA), and the response kernel $f^{\mathrm{xc}}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega)$ describing the exchange-correlation effects is taken to be the functional derivative of the exchange-correlation potential V_{XC}, with respect to the electron density. The inclusion of these effects is found to bring about only moderate changes to the RPA result for the dielectric constant ε_∞, of the order of 4% [43] or less [44]. However, it has also been recently argued [45] that the theorems governing DFT do not, in their original form, apply to the problem of

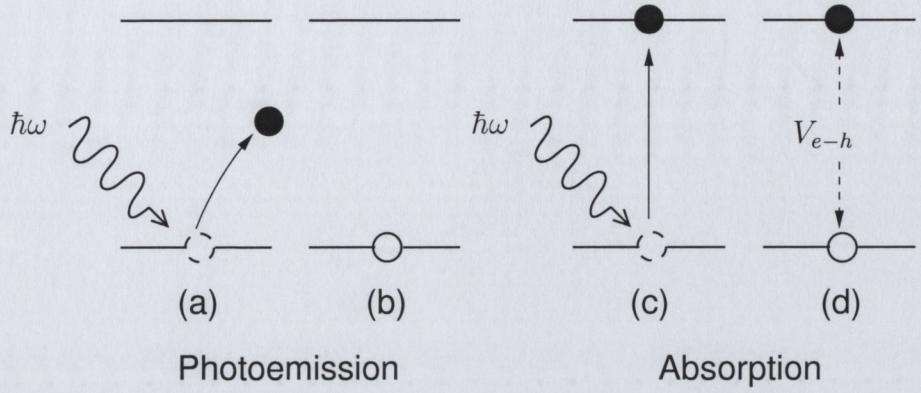


Figure 3-3: Schematic diagram of the processes involved in photoemission and absorption. In the former, a photon ejects an electron from a valence state (a), leaving a hole (b); the electron is assumed to leave the crystal. In the absorption process, an electron is excited to the conduction band (c) where it is free to interact with the remaining hole in the valence band (d).

an infinite, periodic and nonmetallic system in the presence of a long-wavelength perturbation [44], which may explain the failure of DFT-LDA to reproduce the correct dielectric constant.

3.3.3 Electron-Hole Interaction

The final point which we mention here regards the ultimate inability of a single-particle description based on the band approximation to properly describe the optical processes which are probed by experiment. Optical absorption involves the creation of electron-hole pairs that can interact to form bound states such as excitons, and these have been shown to have a marked effect on the calculation of the dielectric function, often bringing the calculated prediction in excellent quantitative agreement with the experimental results. Bandstructures, on the other hand, are more appropriately probed by photoemisson and inverse photoemission experiments, which are essentially single-particle or single-hole processes (see Figure 3-3).

The electron-hole interaction can be included by dressing the vertex function beyond the bare form adopted in Section 2.4 which led to the derivation of the GW approximation [12]. P must now be constructed from a true iteration of the Hedin

equations by setting $\Sigma = iGW$ in equation (2.17), to give the integral equation for Γ

$$\Gamma(1,2;3) = \delta(1,2)\delta(1,3) + i\int \mathrm{d}[67]W_0(1^+,2)G(1,6)G(7,2)\Gamma(6,7;3) \tag{3.40}$$

where W_0 is the dynamically screened Coulomb interaction based on the RPA polarizability. This can be [12] transformed into an equation for a generalized P and combined with equation (2.15) to eventually give the Bethe-Salpeter equation:

$$S(1,1';2,2') = S_0(1,1';2,2') + \int \mathrm{d}[33'44']S_0(1,1';3,3')I(3,3';4,4')S(4,4';2,2'), \tag{3.41}$$

where the interaction term I contains two contributions:

$$I(1,1';2,2') = -i\delta(1,1')\delta(2,2')V_{\mathrm{ex}}(1,2) + i\delta(1,2)\delta(1',2')W_0(1,1'). \tag{3.42}$$

In the above, the term S_0 represents the polarization function for the noninteracting electron-hole pair, $S_0(1,1';2,2') = G(1',2')G(2,1)$, which is equivalent to the RPA expression, and the interaction term is comprised of the unscreened electron-hole exchange $V_{\mathrm{ex}}(1,2)$ and the screened electron-hole interaction W_0.

As stated earlier, the treatment of these excitonic effects lies beyond the scope of this thesis; however, the Bethe-Salpeter equation may be solved [12] by expressing it in terms of the wavefunctions and energies that are obtained from the *ab initio* quasiparticle calculations of Chapter 5. The procedure outlined above is therefore a natural extension of the formalism developed in that chapter.

3.4 Optical Properties of Surfaces

Up to now the treatment of optical properties has centred around two concepts: the use of a dielectric function to describe the interaction between a crystalline solid and a perturbing electric field, and the transformation of this dielectric function to a reciprocal space representation. This approach, which we will term the 'continuum' model, is generally suited to the treatment of bulk properties of materials, as it

is closely linked to the idea of the periodic Bloch wavefunction and the success of bandstructure theory. As discussed in Section 3.2, the model features a quantum mechanical treatment of the nonlocal polarizability and the local field effects inside the cell, which are necessary for a proper description of the optical and dielectric response of the bulk.

Alternative models have been designed to account for optical phenomena that involve variations in the polarization across many cells, such as reflection at surfaces, but where an approximate description of the polarizability is often sufficient. Two of these techniques are schematically presented in Figure 3.4 along with the continuum model; arising from a discrete analysis of the optical response, they are referred to as 'discrete' models. The simpler of these two schemes, and the one of most relevance to this thesis, is the discrete dipole model (DDM) [46, 47, 48]. The model offers a physically appealing and logically consistent interpretation of surface phenomena, in particular those observed using experimental spectroscopies, by adopting a semiclassical microscopic picture of the optical response.

The basic formalism is as follows: the solid is approximated as a regular assembly of polarizable dipoles, situated at, for example, atomic sites or bond centres. In the presence of an external field, each of these dipoles may be polarized independently of the surrounding dipoles by the local field, which is composed, as before, of the external and induced fields at each site. The induction of a dipole moment $\mathbf{p}_i$ at a site i is assumed to occur according to the expression:

$$\mathbf{p}_i = \alpha_i \mathbf{E}_i^{\text{loc}}, \tag{3.43}$$

where α_i represents the local polarizability tensor for that site. The key element of the model lies in the determination of the local field $\mathbf{E}_i^{\text{loc}}$, which is given by

$$\begin{aligned} \mathbf{E}_i^{\text{loc}} &= \mathbf{E}_i^{\text{ext}} + \mathbf{E}_i^{\text{ind}} \\ &= \mathbf{E}_i^{\text{ext}} + \sum_{j \neq i} \mathbf{T}_{ij} \mathbf{p}_j. \end{aligned} \tag{3.44}$$

The quantity $\mathbf{T}_{ij}$ is the so-called 'transfer-tensor' for the dipole-dipole interaction.

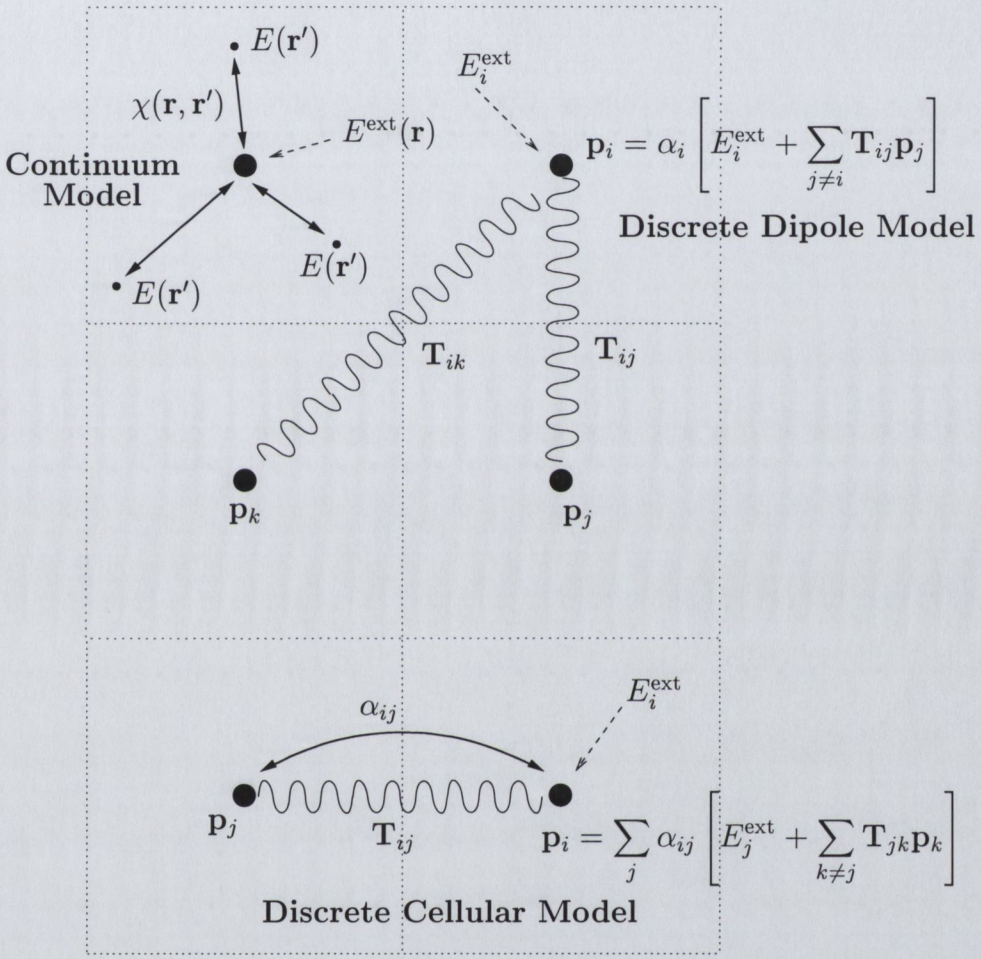


Figure 3-4: Different models of linear response theory: in the continuum discription, the induced polarization $\mathbf{P}(\mathbf{r})$ is related to the external field by the nonlocal susceptibility $\chi(\mathbf{r},\mathbf{r}')$. In the discrete dipole and discrete cellular descriptions, the induced polarization is discretized to give the dipole moment $\mathbf{p}_i$ at each site i, where the local polarizing electric field at i depends on the induced moments $\mathbf{p}_j$ via the 'transfer-tensor' $\mathbf{T}_{ij}$, as defined in the text. The discrete cellular approach also attempts to account for nonlocal effects in the induction process through an intercellular polarizability α_{ij}, which is a local quantity in the discrete dipole model.

It is a nonlocal quantity, in the sense that the electromagnetic fields coupling the dipoles emanate from sources at many different sites, and is responsible for a large part of the 'local field effect' described by the continuum model, the remainder originating in the nonlocality of the polarizability $\mathbf{P}(\mathbf{r}, \mathbf{r}', \omega)$. This part of the local field effect is most important in strongly inhomogeneous systems (i.e. surfaces), where the local field at a site may deviate significantly from its value in the bulk.

Of course, the approximation that the polarizability is local[5] within the cell is quite crude, and strictly the model should only be applied to non-interacting systems, like for instance, solid argon (in the metallic limit, the model breaks down completely). In spite of this, however, the discrete dipole model has been moderately successful in investigations of the optical properties of semiconductor surfaces, based on the fact that a proper description of the local field is of vital importance in this case.

The model can be improved upon by considering nonlocalities in the polarizability itself. In the context of a discrete model, this must be done at the intercellular level. In the 'discrete cellular' model (Figure 3.4) of Wijers and de Boeij[49], the induction of equation (3.43) is modified to

$$\mathbf{p}_i = \sum_k \alpha_{i,k} \mathbf{E}_k^{\mathrm{loc}}, \tag{3.45}$$

which improves on the assumption of independently polarizable cells by taking into account the overlap of the wavefunctions of adjoining cells. This 'quantum induction' effect is quite distinct from the electromagnetic nonlocality discussed above that arises in the transfer tensor. Unfortunately, however, the calculation of the nonlocal polarizabilities is very computationally demanding, and therefore in the investigation of local field effects at surfaces presented in Chapter 6 we will limit ourselves to the a local polarizability and the discrete dipole model.

Before continuing, it is worth mentioning the connection between the discrete

[5]In the DDM, the polarizability is not strictly local, as it is obtained from an averaging of the charge density response within the cell or bond, but is local in the sense that it is represented by the quantity α_i.

dipole theory and the dielectric matrix formalism. In the limit of independently polarizable entities, the connection between the discrete polarizability and the dielectric constant (i.e., $\mathbf{q} \to 0$) of continuum theory is given by the Clausius-Mossotti (CM) relation

$$\left(\frac{\varepsilon_\infty - 1}{\varepsilon_\infty + 2}\right) = \frac{4\pi\alpha}{3\Omega} \tag{3.46}$$

where α denotes the microscopic polarizability. This is derived based on the assumption of a bulk cubic system consisting of one dipole per cell, although Poppe [50] and others [51] have demonstrated that the expression also holds for lattices of other symmetry. The CM relation may be successfully applied to materials where the atomic or ionic polarizability remains largely unaffected by the electronic environment, or in other words, is a local quantity. In Section 3.2 it was shown how an accurate description of the dielectric response of a strongly inhomogeneous solid needed a matrix representation in reciprocal space to account for the rapid short range variations in potential or field within the unit cell. This is facilitated by a nonlocal cell-periodic polarizability. To obtain the CM relation from the continuum model, one therefore needs to consider the quantum-mechanical limit of independent wavefunctions, that is, the extreme tight-binding model (ETB). This was carried out by Nicastro and Patterson [52], who showed that the ETB dielectric matrix can be reduced to the form $\varepsilon = 1 - \alpha\mathbf{T}$, where $\mathbf{T}$ represents the same dipole sums (transfer tensor) of equation (3.44). The connection with the Clausius-Mossotti relation may then be established. Similar conclusions were reached by Adler [28] and Sinha [53] by considering the limiting cases of model dielectric functions.

The discrete dipole and discrete cellular methods offer a very clear theoretical description of the optical response, in stark contrast to the continuum approach. To a large extent, this is because the distinction between sources and fields, and the treatment of the corresponding induction, is kept very clear throughout. Equation (3.43) may be viewed [49] as the discretization of the macroscopic equation

$$\mathbf{P}(\mathbf{r}) = \int \mathrm{d}\mathbf{r}' \chi(\mathbf{r}, \mathbf{r}')\mathbf{E}(\mathbf{r}'), \tag{3.47}$$

which is the conventional way to obtain a source term [the polarization $\mathbf{P}(\mathbf{r})$] from the field term [$\mathbf{E}(\mathbf{r}')$], via the susceptibility $\chi(\mathbf{r}, \mathbf{r}')$, which is itself closely related to the dielectric function. Note, however, that the determination of $\chi(\mathbf{r}, \mathbf{r}')$ should be strictly a quantum mechanical issue; the separate procedure of obtaining the field $\mathbf{E}(\mathbf{r}')$ from the source is correctly supplied in the discrete formalism by Maxwell's equations, i.e., equation (3.44). One may therefore treat the nonlocality due to the polarizability (as contained in the discrete cellular model) separately to the electromagnetic nonlocality that arises in the definition of the local field. These distinct concepts are mixed together in the dielectric matrix of the continuum approach.

This concludes the discussion of the basic concepts of the discrete approach to optical properties and its relation to the dielectric matrix formalism. In Chapter 6 the discrete dipole model is applied to the problem of calculating the reflectance anisotropy spectra of several surfaces of silicon. The majority of that chapter derives from Ref. [54].

Chapter 4

Static Dielectric Response

> The tragedy of beautiful theories is that they are often destroyed by ugly facts
> — *Thomas Huxley*

The static dielectric matrix, $\varepsilon_{\mathbf{GG'}}(\mathbf{q}, 0)$, is a central component of this thesis. It performs a dual purpose: firstly, it is used in the construction of a model for the dynamic dielectric matrix, which is used for the efficient calculation of the self-energy in the GW approximation, as introduced in Chapter 2. The details of this 'plasmon-pole' model of dielectric response is dealt with in more detail in Chapter 5. Secondly, in the limit of zero frequency, the static dielectric matrix provides detailed information regarding the electrostatic screening processes in a crystal which arise from the electron-electron interactions, explicitly taking into account the effects of inhomogeneity inside the cell. Elucidating the matrix representation of $\varepsilon_{\mathbf{GG'}}(\mathbf{q}, 0)$, and determining what information it contains about the characteristic response of particular classes of materials, comprise the central issues of this chapter. The theoretical tool that is used for this analysis is the concept of 'dielectric bandstructure' [55], which allows the complicated matrix structure $\varepsilon_{\mathbf{GG'}}(\mathbf{q}, 0)$ to be represented in a graphical way, and in particular emphasizes the role of local-field effects in the dielectric response. The formalism behind the dielectric bandstructure is presented in Section 4.2, and is followed by explicit calculations and analysis for a range of materials.

Before this topic is discussed in more detail, it is worthwhile to discuss the computational scheme involved in generating the matrix. The calculation of the dielectric matrix is the most computationally intensive part of the work presented in this thesis, and therefore merits a careful examination. The Gaussian orbital framework has not been traditionally used for calculating $\varepsilon_{\mathbf{GG'}}(\mathbf{q},0)$ to a good precision[1], and therefore this point will be addressed in detail. In addition, we will briefly examine the usual convergence criteria which enter the expression for the dielectric matrix: Sections 4.1.3 and 4.1.4 consider the dependence on the number of **k**-points and the number of **G**-vectors (i.e. the actual size of the matrix), respectively.

4.1 Calculation of the Dielectric Matrix

4.1.1 Formalism

In Chapter 3 the frequency and wavevector dependent longitudinal dielectric matrix $\varepsilon_{\mathbf{GG'}}(\mathbf{q},\omega)$ was derived, with local field effects included, in the self-consistent field approximation. The result is given by equation (3.19). In the limit of zero frequency, equation (3.19) reduces to the form:

$$\varepsilon_{\mathbf{GG'}}(\mathbf{q},0) = \delta_{\mathbf{G},\mathbf{G'}} + \frac{16\pi}{V|\mathbf{q}+\mathbf{G}|^2} \sum_{v,c} \sum_{\mathbf{k}} \frac{\langle c\mathbf{k}+\mathbf{q}|e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}|v\mathbf{k}\rangle\langle v\mathbf{k}|e^{-i(\mathbf{q}+\mathbf{G'})\cdot\mathbf{r'}}|c\mathbf{k}+\mathbf{q}\rangle}{E_{c\mathbf{k}+\mathbf{q}} - E_{v\mathbf{k}}}. \tag{4.1}$$

In order to calculate this expression, it is first necessary to compute matrix elements of $e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}$ for all $\mathbf{q}+\mathbf{G}$ over a set of valence and conduction Bloch states spanning the full Brillouin zone. In this chapter, we will use a DFT-LDA Hamiltonian for the purpose of generating the eigenstates $|n\mathbf{k}\rangle$ and eigenvalues $E_{n\mathbf{k}}$. The use of the LDA bandstructure for calculating the RPA dielectric matrix has been shown previously by many authors [37, 36] to yield results that agree quite well with the experimental values, such as for the dielectric constants of several semiconductors. We therefore expect that this scheme will produce realistic information about general dielectric

[1] As distinct from calculations that involve an averaging over the dielectric matrix, e.g. [21, 24].

properties, and in particular illustrate the nature of the static screening processes.

Some insight into the computational aspects can be gained by focussing on the long wavelength limit of the matrix, which allows us to rewrite the matrix elements of $e^{i\mathbf{q}\cdot\mathbf{r}}$ in terms of matrix elements of the momentum operator, $\hat{\mathbf{p}}$ [56, 30]. This transformation was previously given by equation (3.20) and confirms the gauge invariance of the optical properties. Taking the $\mathbf{q} \to 0$ limit of this equation explicitly yields the relation

$$\lim_{\mathbf{q}\to 0}\langle c\mathbf{k}+\mathbf{q}|e^{i\mathbf{q}\cdot\mathbf{r}}|v\mathbf{k}\rangle = \frac{\mathbf{q}\cdot\langle c\mathbf{k}|\mathbf{p}|v\mathbf{k}\rangle}{(E_{c\mathbf{k}}-E_{v\mathbf{k}})}, \tag{4.2}$$

which, as stated earlier, strictly holds only in the absence of nonlocal potentials in the Kohn-Sham Hamiltonian, a condition we have avoided by considering only all-electron basis sets. Each occurrence of $\langle c\mathbf{k}+\mathbf{q}|e^{i\mathbf{q}\cdot\mathbf{r}}|v\mathbf{k}\rangle$ in equation (4.1) can therefore be substituted with equation (4.2), which introduces an extra factor of $(E_{c\mathbf{k}}-E_{v\mathbf{k}})$ into the denominator each time; for example, the 'head' $(\mathbf{G}=\mathbf{G}'=0)$ element can be expressed as

$$\varepsilon_{0,0}(\mathbf{q}\to 0) = 1+\frac{16\pi}{V}\sum_{v,c}\sum_{\mathbf{k}}\frac{|\langle c\mathbf{k}|\mathbf{p}|v\mathbf{k}\rangle|^2}{(E_{c\mathbf{k}}-E_{v\mathbf{k}})^3}, \tag{4.3}$$

with similar expressions for the 'wing' $(\mathbf{G}=0,\mathbf{G}'\neq 0$ or $\mathbf{G}\neq 0,\mathbf{G}'=0)$ elements. It then becomes trivial to understand how the convergence properties of the matrix are largely determined by the form of the denominator for the different elements. For a particular valence band v, the presence of an energy cubed term in the denominator would cause the summation over conduction bands c to die off rapidly, a behaviour unique to the 'head' element. Similarly, the single $(E_{c\mathbf{k}}-E_{v\mathbf{k}})$ term that characterises the 'body' elements implies that the summation converges very slowly with the number of conduction bands, whereas the 'wing' elements display intermediate convergence behaviour. This has consequences for the choice of basis set used when calculating the electronic states, as discussed in the next section.

4.1.2 Basis Sets

The construction and use of an appropriate basis set constitutes a critical factor in most *ab initio* calculations, but is particularly important within the Gaussian orbital framework. We will now examine the origins and implications of this statement with regard to the calculation of the dielectric matrix.

On re-examining the derivation of the dielectric matrix within the self-consistent field approximation, we may infer from equation (3.14) (evaluated in some basis) that the appearance of the energy difference term $(E_{c\mathbf{k}+\mathbf{q}} - E_{v\mathbf{k}})$ in equation (4.1) originates from a *mathematical* completeness relation. This is particularly true if Kohn-Sham eigenvalues are interpreted as the single-particle energies, as they are unphysical quantities in the context of dielectric properties. The same conclusion may be reached by considering the fact that within DFT the static dielectric matrix must be a property of the ground state of the system [43].

In many electronic structure calculations, the completeness relation

$$\sum_{n\mathbf{k}} |n\mathbf{k}\rangle\langle n\mathbf{k}| \approx 1 \tag{4.4}$$

can be fulfilled to a desired degree of accuracy by specifying an appropriate energy cut-off parameter [38]. Unfortunately, there is no single criterion that may be used to construct a complete basis set in the Gaussian orbital framework. For most physical applications (i.e., calculations of ground state properties), a consideration of the total energy proves to be sufficient, as implied by the variational theorem. However, this is a weak criterion for constructing virtual states. As we shall see, this poses problems for the calculation of the dielectric matrix, which involves a summation over both occupied and unoccupied states, as shown by equation (4.1). Without any other guiding criterion for now, we must assume that the accurate calculation of the dielectric matrix requires at least a reasonable value for the total energy. Within the Gaussian basis scheme, the total energy itself depends on a number of criteria: it depends on the values of the exponents and contraction coefficients of the Gaussian primitives, the type of function used (s, p, d etc.) as well as the actual size of the

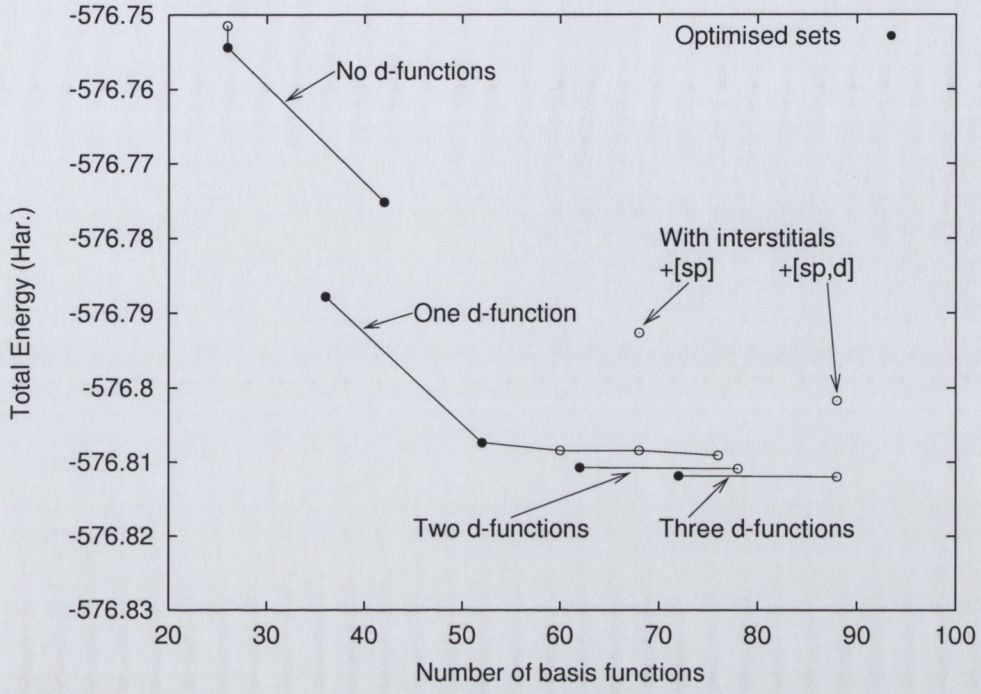


Figure 4-1: Influence of basis set size and type on the total energy of silicon. The solid lines link sets with the same number of d-orbitals but with increasing number of s and p orbitals, all of which are situated on the atom. In general, the total energy decreases as the basis set gets larger, or as polarization functions (i.e., d-orbitals in silicon) are added. However, the addition of further atomic orbitals of a particular character has only a minor influence on the total energy after a minimum threshold in size of about 50 basis functions. The importance of optimization is emphasized by the difference in energy between the (unoptimized) basis sets with interstitial functions, and the fully and partially optimized sets of comparable size. Methods of constructing and optimizating these sets are described in Appendix D.

basis set.

Figure 4-1 demonstrates the influence of the Gaussian basis set on the total energy of silicon. The sets that are labelled 'Optimized' have been specially constructed using the procedure outlined in Appendix D: they correspond to a global minimum in the exponent *vs.* total energy curve for a particular basis set type, and are highly accurate for the purpose of calculating ground state properties. The figure suggests that while polarization functions in the basis set (i.e. d-shells in silicon) may have a large impact on the total energy curve, both the addition of subsequent functions of the same type (albeit after a certain size) and the optimization procedure have only a minimal effect on the total energy.

A closer look at the eigenenergies, however, reveals that the use of the total energy as the sole criterion for choosing a basis set may indeed be misguided.

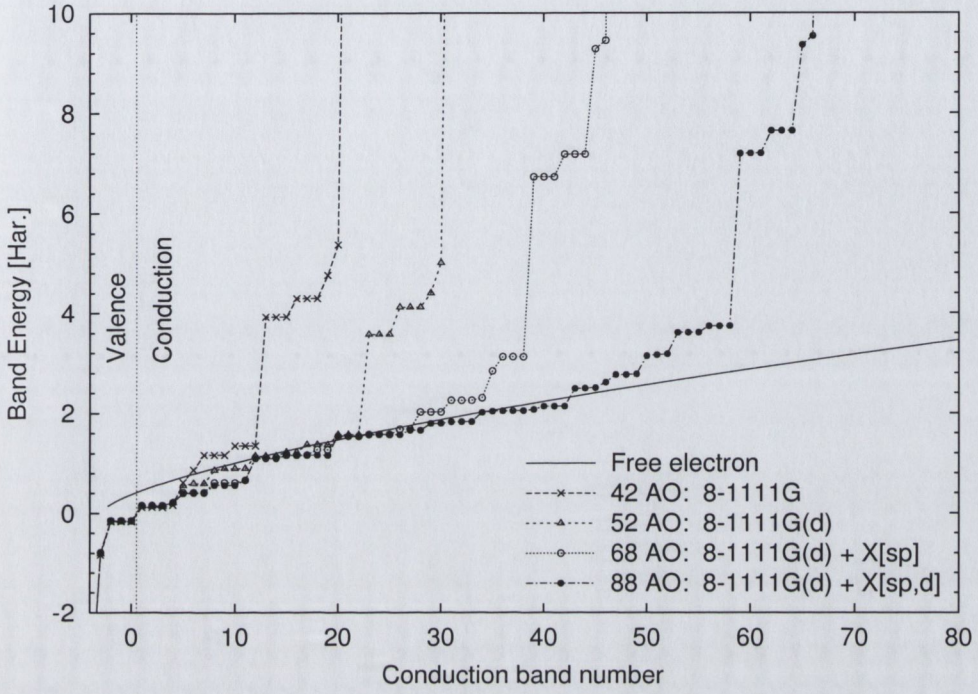


Figure 4-2: Dependence of conduction band energy (in Hartree) on the band index for silicon. The zero level of energy is aligned with the bottom of the valence band for the purpose of comparison with the free electron prediction. Energies shown are the eigenvalues at the Γ point; a more accurate comparison with the solid curve would require an average over the Brillouin zone for each band. See Appendix D for explanation of notation.

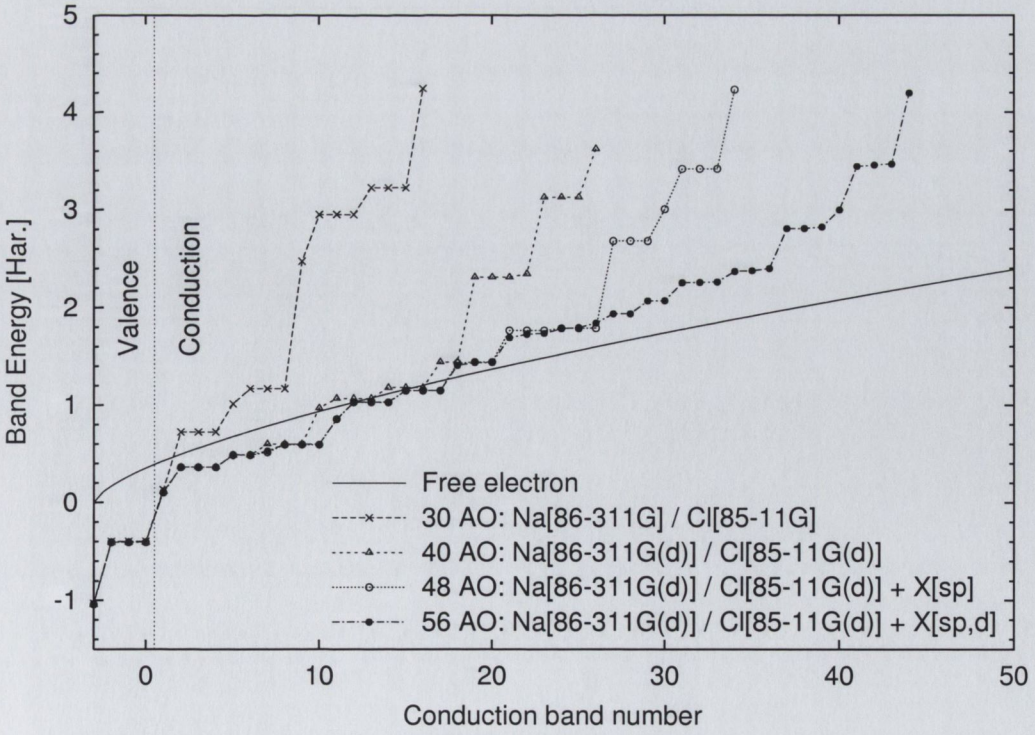


Figure 4-3: Dependence of conduction band energy on the band index for NaCl.

Figures 4-2 and 4-3 illustrates further the problem of incompleteness that arises with the use of Gaussian basis sets, which makes them rather unsuitable for calculating accurate optical spectra. As the index n of the conduction band increases, the energy $E(n)$ of that state would be expected to approach the theoretical free electron result given by

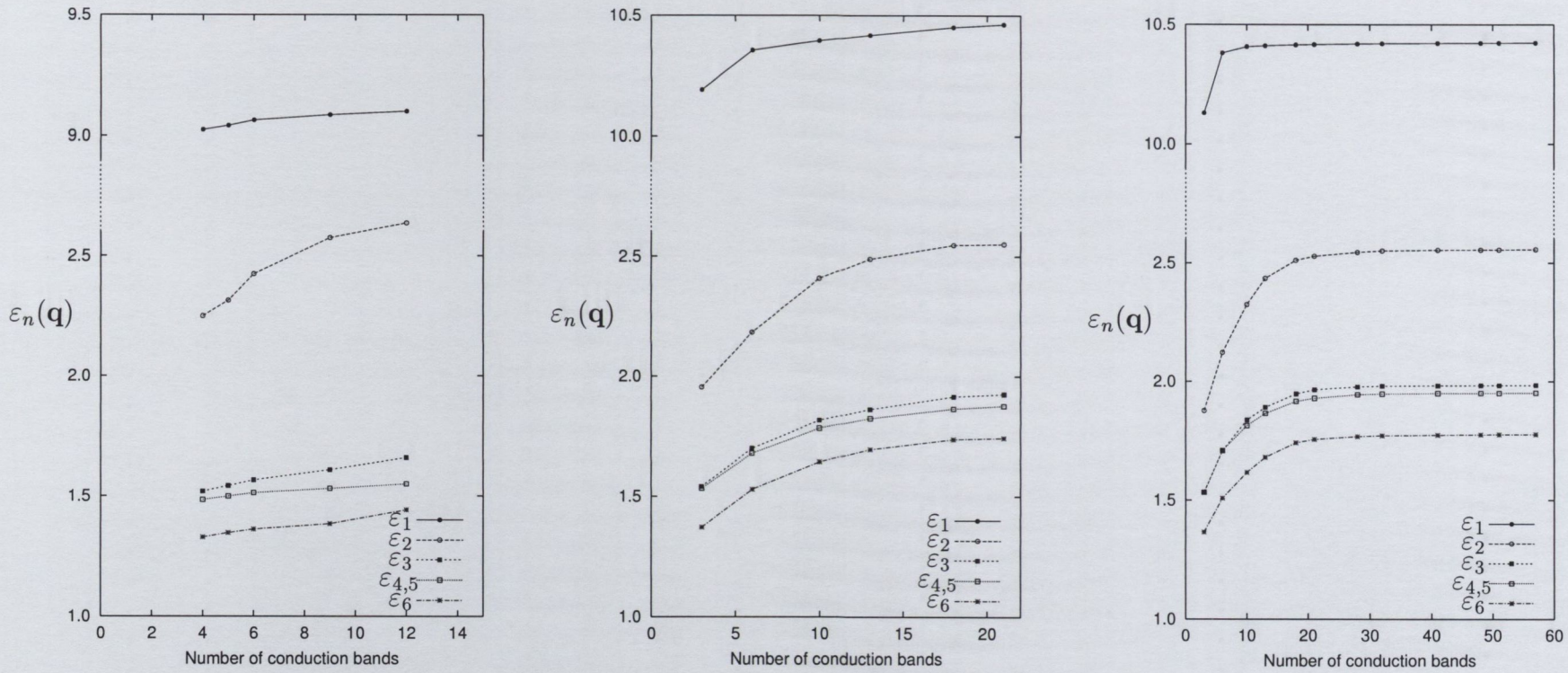
$$E(n) = \frac{1}{2}\left(\frac{6\pi^2}{\Omega}\right)^{\frac{2}{3}} n^{\frac{2}{3}}, \tag{4.5}$$

as electrons occupying energy states far above the band gap tend to become delocalised and their wavefunctions become free electron-like. However, because the Bloch states are constructed from localized Gaussian orbitals, the correct multinodal character of the higher bands (i.e., $3d, \ldots, 4d, \ldots, 5f, \ldots$, etc.) is not necessarily created within the limited set (s, p, d-type only) of functions available, which effectively reduces the useful range of states available during summation. The effect is more pronounced in all-electron basis sets than in pseudopotential sets since the energy range encompassed by the electronic states is more limited in the latter case.

The figures highlight a serious possible failure of the Gaussian orbital framework in accurately calculating properties of excited states or properties that require an almost complete set of electronic states. It is likely that previously published calculations of dielectric constants [21] and quasiparticle energies [21, 24] using a Gaussian scheme may not be fully converged, as has been suggested elsewhere [20]. For instance, in the case of an uncontracted and fully optimised set such as[2] 8-1111(G), only about 10 conduction band states lie at realistic energies compared to the free-electron curve – even though the set would be more than sufficient for many calculations of ground state physical properties. The plane-wave representation may be superior in this respect; reports in the literature regularly cite the use of hundreds of bands in the calculation of the dielectric function: e.g. 210 bands [37], 285–580 bands [38] and 196 bands [36] (all for silicon) — as compared with an upper limit of ~60 bands in the present calculations.

To illustrate how the basis set problem affects the calculation of the dielectric

[2]See Appendix D for explanation of notation.



(a) 8-1111G (b) 8-1111G(d) (c) 8-1111G(d)+X[spd]

Figure 4-4: Convergence of the dielectric matrix eigenvalues (at small $\mathbf{q}$) of silicon with respect to the number of conduction bands used in the summation over states, for (a) an uncontracted and fully optimized basis set, (b) same as (a) plus a polarization function, and (c) a specially constructed basis set with interstitial functions. The highest eigenvalue is denoted by ε_1, the next highest by ε_2, and so on. Note the difference between the scales used in (a–c). Although each set has a higher number of conduction bands in total, it would be misleading to plot the contribution to ε_n from those bands at high energies whose eigenvalues do not lie close to the free-electron eigenvalues (c.f. Figure 4-2) as this would visually suggest a higher degree of convergence than that actually obtained.

matrix, Figure 4-4 shows the convergence of the dielectric matrix eigenvalues with respect to the number of conduction bands involved in the summation of equation (4.1). As will be explained later in this chapter, the eigenvalues of $\varepsilon_{\mathbf{GG'}}(\mathbf{q})$ effectively summarise the information contained within the matrix and constitute an alternative to considering the convergence behaviour of explicit matrix elements. Figure 4.4(a) and Figure 4.4(b) demonstrate the level of convergence for two fully optimized basis sets, 8-1111G and 8-1111G(d). Both these sets are constructed from a standard 8-31G set by uncontracting the valence orbitals and re-optimizing (c.f. Figure 4-1), and in the second set a polarization function is added. It should be noted that the sum over states is truncated at the upper limit of the range of *continuous* conduction bands, which was illustrated in Figure 4-2. States lying above this cut-off index contribute little to the summation, but are omitted from the diagram so as not to give a false impression of convergence. It is clear that the 8-1111G set is poorly converged as expected from the small effective set of unoccupied states. The 8-1111G(d) offers a definite improvement on this, but does not seem to be fully converged either. Incidentally, the importance of *type* of atomic orbital for calculating the dielectric matrix is emphasized by the difference between the highest eigenvalues in Figure 4.4(a) and Figure 4.4(b).

In an effort to ameliorate this problem somewhat, extra sets of orbitals are added to the basis which are located at interstitial sites. This specifically aims to improve the flexibility of the basis set throughout the cell and attempts to reproduce the highly nodal structure of the free-electron plane-wave eigenstates. Ideally this would be achieved by adding higher order functions (f, g, h etc.) to the atomic sites only, but these have not, as yet, been implemented in the CRYSTAL95 code. The inclusion of the extra functions brings a marked improvement to the convergence properties, as can be seen for the 8-1111G(d) + X[spd] basis set in Figure 4.4(c). A true test of convergence would require an arbitrarily large basis for comparison, which is not obtainable; the figures do however suggest that the eigenvalues have become well converged with the addition of the auxiliary functions. As a comparison with the literature, the use of 60 conduction bands in the calculation of the 'head'

(i.e., $\varepsilon^{\mathrm{RPA}} \equiv \varepsilon_{00}^{\mathrm{AW}}$) corresponds to a convergence of 99.99% [37], whereas convergence levels of 99.5% [$\mathbf{G} = \mathbf{G}' = (\bar{1}\bar{1}\bar{1})$], 63.5% [$\mathbf{G} = \mathbf{G}' = (\bar{2}\bar{2}\bar{2})$] and 29.8% [$\mathbf{G} = \mathbf{G}' = (004)$] [36] are reported for other elements in the matrix. Corresponding values may not be quite as severe for the largest eigenvalues of the matrix, due to the mixing of elements, although exact values are not available for comparison. It is well known [38] that the local field effects in silicon are manifested in the large numbers of (small) off-diagonal elements of the matrix, rather than in the (relatively large) diagonal elements. Nevertheless, the calculated shift in ε_∞ with the inclusion of local fields is comparable with the general literature values, as shown in Section 3.2.1, which suggests that the treatment of the local field effects is at least *relatively* accurate.

Using interstitial basis functions as a means of increasing the effective range of conduction band states is not, however, an ideal solution. An increase in the number of orbital sites inside the cell can give rise to a large overlap between the basis functions, and a quasi-linear dependence can occur due to numerical limitations in the CRYSTAL95 code. Such behaviour is particularly common during Hartree-Fock calculations, in which case the Gaussian exponents of orbitals situated at both interstitial sites and atomic sites have to be carefully adjusted (usually, made less diffuse) in order to allow the total energy calculation to converge. This is an arduous and impractical task; a thorough optimization of extended basis sets utilizing several interstitial functions is almost certainly impossible (c.f. Figure 4-1). A better solution, perhaps, would be to complement the Gaussian orbital basis with a plane wave basis in order to improve closure. Nevertheless, the extended basis sets were used throughout this thesis in the calculation of the dielectric bandstructures [Section 4.2] and the self-energy correction [Section 5.5], and are presented in Appendix D.

We will now examine the remaining factors that influence the calculation of the static dielectric matrix.

$\varepsilon_{\mathbf{G},\mathbf{G}'}$	G	G′	$N_\mathbf{k} = 32$	$N_\mathbf{k} = 256$	$N_\mathbf{k} = 2048$
head	(0 0 0)	(0 0 0)	15.7339	**11.7399**	11.9293
wing	(0 0 0)	(1 0 0)	0.6290	**0.5808**	0.5801
	(0 0 0)	(2 2 2)	0.1220	**0.0945**	0.0944
body	(1 0 0)	(1 0 0)	**1.5718**	1.5531	1.5531
	(2 2 2)	(2 2 2)	**1.0541**	1.0535	1.0535
	(1 0 0)	(2 2 2)	**0.0278**	0.0270	0.0271

Table 4.1: Selected elements of the dielectric matrix $\varepsilon_{\mathbf{G}\mathbf{G}'}(\mathbf{q})$ of silicon, calculated using different special-point meshes in the full Brillouin zone. Results are shown for $\mathbf{q} = \frac{2\pi}{a}(\frac{1}{8}\frac{1}{8}\frac{1}{8})$. Assuming that the densest mesh is fully converged, use of the settings in bold results in errors of at most 2% for the head and wings, or 3% for the body elements. A fairly small basis set (30 AO's) had to be used due to the limitations in computer memory associated with the densest **k**-point mesh.

4.1.3 Brillouin Zone Integration

After the basis set, the Brillouin zone integration is the most important factor that influences the convergence of the dielectric matrix elements. In practice, the integration is carried out using the Monkhorst-Pack special-points scheme, in which the integral is approximated by a weighted sum of the integrand over a carefully chosen set of $N_\mathbf{k}$ equally spaced **k**-points, $\{\mathbf{k}_i\}$:

$$\frac{\Omega}{(2\pi)^3}\int_{\mathrm{BZ}} \mathrm{d}\mathbf{k}\, f(\mathbf{k}) \approx \frac{1}{N}\sum_{\mathbf{k}\in BZ} f(\mathbf{k}) \approx \sum_{i}^{N_\mathbf{k}} w_i f(\mathbf{k}_i), \tag{4.6}$$

where there are N unit cells of volume Ω and w_i are weighting factors such that $\sum_i w_i = 1$. A full description of the Monkhorst-Pack integration scheme is presented in Appendix C.1.

The special-points technique does not need to be as accurate as the Gilat-Raubenheimer scheme that was adopted for the calculation of the frequency-dependent dielectric function in Section 4.1.1, since the denominator of equation (4.1) always remains finite at zero frequency. The density of special points required for convergence depends on the element of the matrix under consideration, in a similar manner to the dependence on the number of conduction bands; the presence of the energy cubed term in the 'head' [c.f. equation (4.3)] means that the denominator rapidly oscillates during the band summation, thus requiring a much denser mesh than that needed for the 'body' and 'wing' elements. This is verified in Table 4.1 for various

matrix elements of $\varepsilon_{\mathbf{G},\mathbf{G}'}$. Based on this convergence behaviour, the option of using a fairly sparse mesh when calculating the 'body' elements and a denser mesh for calculating both the 'head' and the 'wing' elements has been built into the computer code. In most of the calculations in this thesis a 256-point mesh has been used for calculating head and wing elements of the dielectric matrix and a 32-point mesh otherwise. According to highly accurate calculations reported in the literature [37, 36], the use of 256 **k**-points for all elements results in an error (overestimation) of 7% in the calculation of ε_∞ and ε_{00}, which imposes an upper bound on the error in the present results.

4.1.4 G-vector Summation

In addition to fulfilling the convergence criteria for separate matrix elements, the dielectric matrix as a whole must be sufficiently large to ensure that the *inverse* matrix elements are properly converged. A large number of **G**-vectors is necessary to obtain a well converged macroscopic dielectric function including local field effects, as defined by equation (3.35), or a converged set of matrix eigenvalues. Table 4.2 shows the convergence of the dielectric matrix eigenvalues of silicon with respect to the number of **G**-vectors. Well converged results are achieved after 59 **G**-vectors, at least for the highest eigenvalues presented, a conclusion which is also reached for the diamond calculation. This choice is also supported in the literature (the use of 59 **G**-vectors is sufficient to converge ε_∞ to less than 1% for silicon [37, 38]). However, the degree of convergence is not quite as good for the ionic compounds, NaCl and MgO, a fact which can be attributed to the limited number of conduction bands that are available for summation in both these materials (see Figure 4-3).

Finally, we just state that the convergence properties with respect to the number of conduction bands, the number of **k**-points and the number of **G**-vectors is the same for the RPA dielectric matrix using Hartree-Fock eigenvalues, so that the conclusions reached in this chapter are also applicable to the calculations of the self-energy in Chapter 5. It should be re-iterated at this point that the convergence requirements are less severe for the self-energy calculations, which feature an

N_G	ε_1	ε_2	ε_3	ε_4	ε_5
1	10.233	-	-	-	
9	10.377	2.086	1.862	1.862	1.857
15	10.379	2.553	1.875	1.875	1.875
27	10.389	2.554	1.968	1.940	1.940
59	**10.423**	**2.556**	**1.984**	**1.951**	**1.951**
89	10.424	2.559	1.984	1.951	1.951
113	10.425	2.561	1.984	1.951	1.951
169	10.425	2.561	1.986	1.953	1.953

Table 4.2: Convergence properties of the dielectric matrix eigenvalues of silicon with respect to the size of the matrix, for the extended basis set 8-1111(d) + X[spd], where X denotes interstitial functions. Good convergence is achieved for 59 **G**-vectors, as indicated in bold. Inconsistencies at 169 **G**-vectors stem from the poor calculation of the high **G** elements, which need a much larger set of basis functions.

averaging over the dielectric matrix.

4.2 Dielectric Bandstructure

4.2.1 Formalism

The concept of dielectric bandstructure (DBS) was introduced by Baldereschi and Tosatti [55] as a means of elucidating the dielectric matrix through a graphical representation. In the same manner that a Hamiltonian matrix may be diagonalised to give a characteristic energy bandstructure and set of wavefunctions, the dielectric matrix may be diagonalised to yield a dielectric bandstructure and set of eigenpotentials. As stated previously, the static dielectric matrix $\varepsilon_{\mathbf{G},\mathbf{G}'}(\mathbf{q})$ describes the response of a system of charged particles to an external perturbation within an inhomogeneous environment. The application of an external potential induces screening charges in the system that are largely determined by the degree of ionicity or covalency of the material, which means that the dielectric bandstructures can be similarly classified. The eigenpotentials in turn graphically depict the screening modes which may be excited, and reflect the nature of the local field effects.

The theory is quite straightforward [55]: we omit the frequency label for clarity

and rewrite equation (3.31) as

$$V_{\mathbf{G}}^{\text{tot}}(\mathbf{q}) = \sum_{\mathbf{G}'} \varepsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q}) V_{\mathbf{G}'}^{\text{ext}}(\mathbf{q}). \tag{4.7}$$

Defining the symmetrized dielectric matrix as

$$\tilde{\varepsilon}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) = \frac{|\mathbf{q}+\mathbf{G}|}{|\mathbf{q}+\mathbf{G}'|} \varepsilon_{\mathbf{G}\mathbf{G}'}(\mathbf{q}), \tag{4.8}$$

we note that $\tilde{\varepsilon}_{\mathbf{G}\mathbf{G}'}(\mathbf{q})$ is a Hermitian operator (real symmetric if the crystal has an inversion centre) and therefore can be diagonalized as

$$\sum_{\mathbf{G}'} \tilde{\varepsilon}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) \tilde{V}_{n\mathbf{G}'}(\mathbf{q}) = \varepsilon_n(\mathbf{q}) \tilde{V}_{n\mathbf{G}}(\mathbf{q}). \tag{4.9}$$

This introduces a set of eigenvalues $\varepsilon_n(\mathbf{q})$ and eigenvectors $\tilde{V}_{n\mathbf{G}}(\mathbf{q})$, and it is then easy to show that $V_{n\mathbf{G}}(\mathbf{q}) = |\mathbf{q}+\mathbf{G}|^{-1} \tilde{V}_{n\mathbf{G}}(\mathbf{q})$ are the eigenvectors of $\varepsilon_{\mathbf{G}\mathbf{G}'}(\mathbf{q})$, with the same eigenvalues $\varepsilon_n(\mathbf{q})$.

The central idea is that the eigenvalue spectrum defined by $\varepsilon_n(\mathbf{q})$, as $\mathbf{q}$ spans the Brillouin zone, represents (along with the corresponding eigenpotentials) the essential information contained in the static dielectric matrix. The real space dielectric eigenpotential is related to the eigenvectors by the expression:

$$V_n(\mathbf{q}, \mathbf{r}) = e^{i\mathbf{q}\cdot\mathbf{r}} \sum_{\mathbf{G}} V_{n\mathbf{G}}(\mathbf{q}) e^{i\mathbf{G}\cdot\mathbf{r}}. \tag{4.10}$$

The physical meaning of the eigenpotential may be deduced by examining equation (4.7) in the context of an external perturbing potential $V_{\mathbf{G}}^{\text{ext}}(\mathbf{q}) = V_{n\mathbf{G}}(\mathbf{q})$. The screened potential is then given by

$$V_{\mathbf{G}}^{\text{tot}}(\mathbf{q}) = \sum_{\mathbf{G}} \varepsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q}) V_{n\mathbf{G}}(\mathbf{q}) = \varepsilon_n^{-1}(\mathbf{q}) V_{n\mathbf{G}}(\mathbf{q}), \tag{4.11}$$

which has the same $\mathbf{r}$-dependence as the perturbation, but is scaled by a constant factor $\varepsilon_n^{-1}(\mathbf{q})$. For the purposes of understanding screening, it is therefore more intuitive to consider the DBS formed from the *inverse* eigenvalues $\varepsilon_n^{-1}(\mathbf{q})$. The

real-space eigenpotentials are discussed in more detail in Section 4.2.3.

4.2.2 Results

Calculations of full dielectric bandstructures are presented in Figures 4-5 to 4-7 for two covalent materials (Si, diamond), two ionic[3] materials (NaCl, MgO) and the intermediate compound GaN. Also shown are the corresponding diagonal bandstructures (so-called 'empty lattice' bandstructures), obtained from the diagonal elements of $\varepsilon_{\mathbf{GG'}}(\mathbf{q})$ (i.e., neglecting local field effects). As mentioned in Chapter 3, omitting the off-diagonal elements of the dielectric matrix when calculating the inverse (to obtain, for example, the macroscopic dielectric function, $\varepsilon_M(\omega)$) is equivalent to omitting the local field effects, or equivalently, the microscopic fluctuations in potential that arise due to the inhomogeneity of the system. The same can also be said for the inverse eigenvalues ε_n^{-1}, such that the difference between the diagonal and fully calculated DBS for a particular material directly reflects the size of these effects. However, it is apparent from the figures that the full DBS are not strikingly different from the empty lattice results, which explains why approximating the full matrix by the diagonal $\varepsilon(\mathbf{q})$ is often adopted. This observation is most relevant to silicon, whereas the largest difference between the two bandstructures occurs for the ionic materials. Several trends can be observed in the eigenvalues of the five materials under consideration, and these can be examined more clearly in Table 4.3, which presents the dielectric eigenvalues corresponding to the most screened eigenpotentials of each material at $\mathbf{q} \to 0$. The more obvious trend is that the inverse eigenvalues generally get larger with increasing ionic character, which is to be expected, since the magnitude of the electronic polarizability, and hence the electronic screening, should correlate with the magnitude of the band gap in each material [57, 58]. This overall trend can be easily observed by considering the smallest inverse eigenvalue: it is typically dominated by the contribution of the large 'head' element of the direct matrix, which tends to conceal the effects of the smaller off-

[3]Note that only the *electronic* polarizability is considered in the calculation of the dielectric matrices.

diagonal elements. Silicon, being a narrow gap semiconductor, has the most diffuse electronic wavefunction and displays the largest screening $[(\varepsilon_n^{-1})_{\min} \equiv \varepsilon_1^{-1} = 0.080]$, whereas the insulators MgO and NaCl have their valence electon density highly localized at the anion, have a smaller electronic polarizability and therefore have the highest eigenvalues [$\varepsilon_1^{-1} = 0.305$ (MgO), 0.328 (NaCl)]. Diamond [$\varepsilon_1^{-1} = 0.175$], and the polar semiconductor GaN [$\varepsilon_1^{-1} = 0.174$], have an intermediate character with respect to the screening properties. Note that this trend across the materials implies that only the *relative* importance of the DBS eigenvalues to their diagonal counterpart must be considered when discussing the magnitude of the local fields.

Strong local field effects are most prevalent in those materials which deviate furthest from the ideal model of a homogeneous electron density and a uniform screening mechanism. In the homogeneous limit, the screening eigenpotential is a simple plane-wave [see equation (4.10)]; for strongly ionic or covalent materials it is generally found that the inclusion of non-zero off-diagonal elements has the effect of lowering the DBS eigenvalues from the diagonal values, which reflects the increased screening of the dielectric eigenpotentials over plane-waves [55]. The trend in increasing local field correction therefore follows the trend associated with the overall screening magnitude. In MgO, for example, the lowest (eight-fold) degenerate eigenvalue in the diagonal bandstructure at $\varepsilon_2^{-1} = 0.728$ is screened by the local fields to such an extent that one component is shifted to the value $\varepsilon_1^{-1} = 0.305$, which is below the $\mathbf{G} = 0$ eigenvalue in the DBS at $\varepsilon_2^{-1} = 0.354$, and represents an increase in screening of nearly 60%.

In the fully calculated bandstructure, eigenvalues may be classified corresponding to the symmetry properties of the corresponding eigenpotential, as has been done by several authors [55, 36]. Eigenvalues belonging to particular irreducible representations may be screened more than others depending on the electronic environment or the nature of the chemical bond. Such classification has not, however, been rigorously done in this work. Nevertheless, we find (based on degeneracy) that the ordering of bands in the DBS is sometimes different from that obtained in pure-plane wave calculations (i.e., Bloch functions expanded over a plane-wave

basis as distinct from Gaussian orbitals). In particular, we are unable to reproduce the magnitude of the 'longitudinal-transverse' (L-T) splittings noted elsewhere from pure plane-wave calculations that are associated with the difference between the $\mathbf{q} = 0$ and $\mathbf{q} \to 0$ eigenvalues of Γ_{15} symmetry. In diamond, the ordering of the L-T bands is even found to be reversed.

However, almost identical results to those presented here have appeared in calculations that expand the dielectric matrix, Coulomb potential and polarizability over an auxiliary basis of Gaussian functions [59]. In this form, the static dielectric matrix is written as $\varepsilon_{\beta\beta'}(\mathbf{q})$, where β, β' denote the non-orthogonal auxiliary basis. Of course, the DBS eigenvalues should be independent of the basis chosen, and excellent agreement between the two schemes has been found for diamond and MgO.

The concept of dielectric bandstructure allows us to graphically compare the dielectric matrices of different materials, or even to comment upon the validity of different models of the dielectric response, as has been done by Car *et al.* [60] for silicon. However, the information in the dielectric matrix is only partially conveyed by the eigenvalue dispersion; the remaining information relating to the real-space screening mechanisms is contained in the eigenpotentials.

4.2.3 Dielectric Eigenpotentials

A physical picture of the (static) screened potentials and charge transfer processes inside the cell is given by the set of dielectric eigenpotentials, which in a real-space representation is defined by equation (4.10). We may recall from equation (4.11) that the application of a perturbing eigenpotential to the system causes a response with the same $\mathbf{r}$-dependence, albeit screened by the appropriate inverse eigenvalue. A typical perturbing potential is of course not in the form of an eigenpotential (e.g. $V^{\text{ext}}(\mathbf{q}, \mathbf{r}) \sim e^{i\mathbf{q}\cdot\mathbf{r}}$); however, the set of eigenpotentials comprises a complete orthonormal set such that any external potential can be decomposed as

$$V^{\text{ext}}(\mathbf{q}, \mathbf{r}) = \sum_n C_n(\mathbf{q}) V_n(\mathbf{q}, \mathbf{r}) = \sum_n C_n(\mathbf{q}) \sum_{\mathbf{G}} V_{n\mathbf{G}} e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}, \qquad (4.12)$$

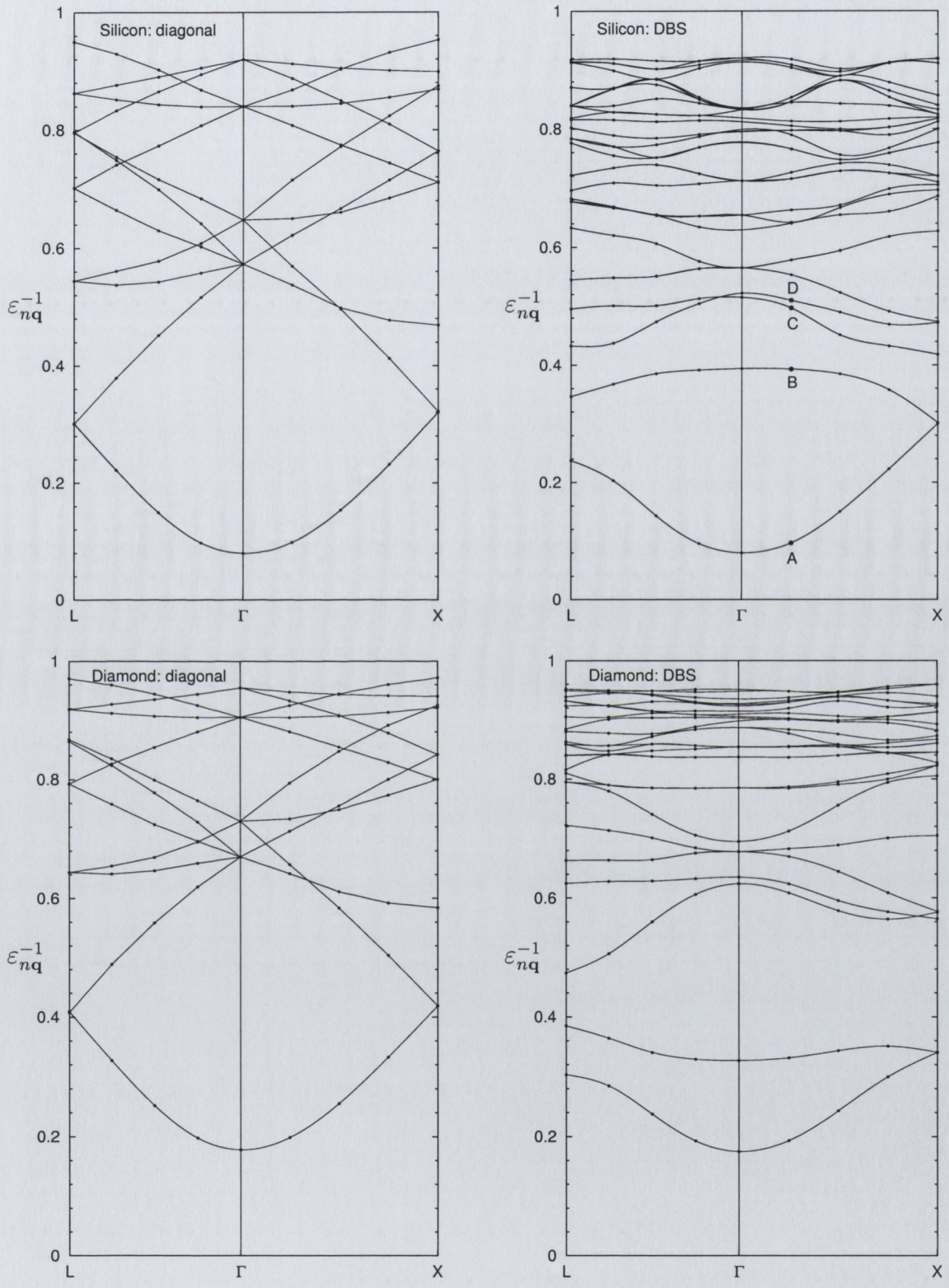


Figure 4-5: Dielectric bandstructures of silicon and diamond: full calculation, and diagonal approximation.

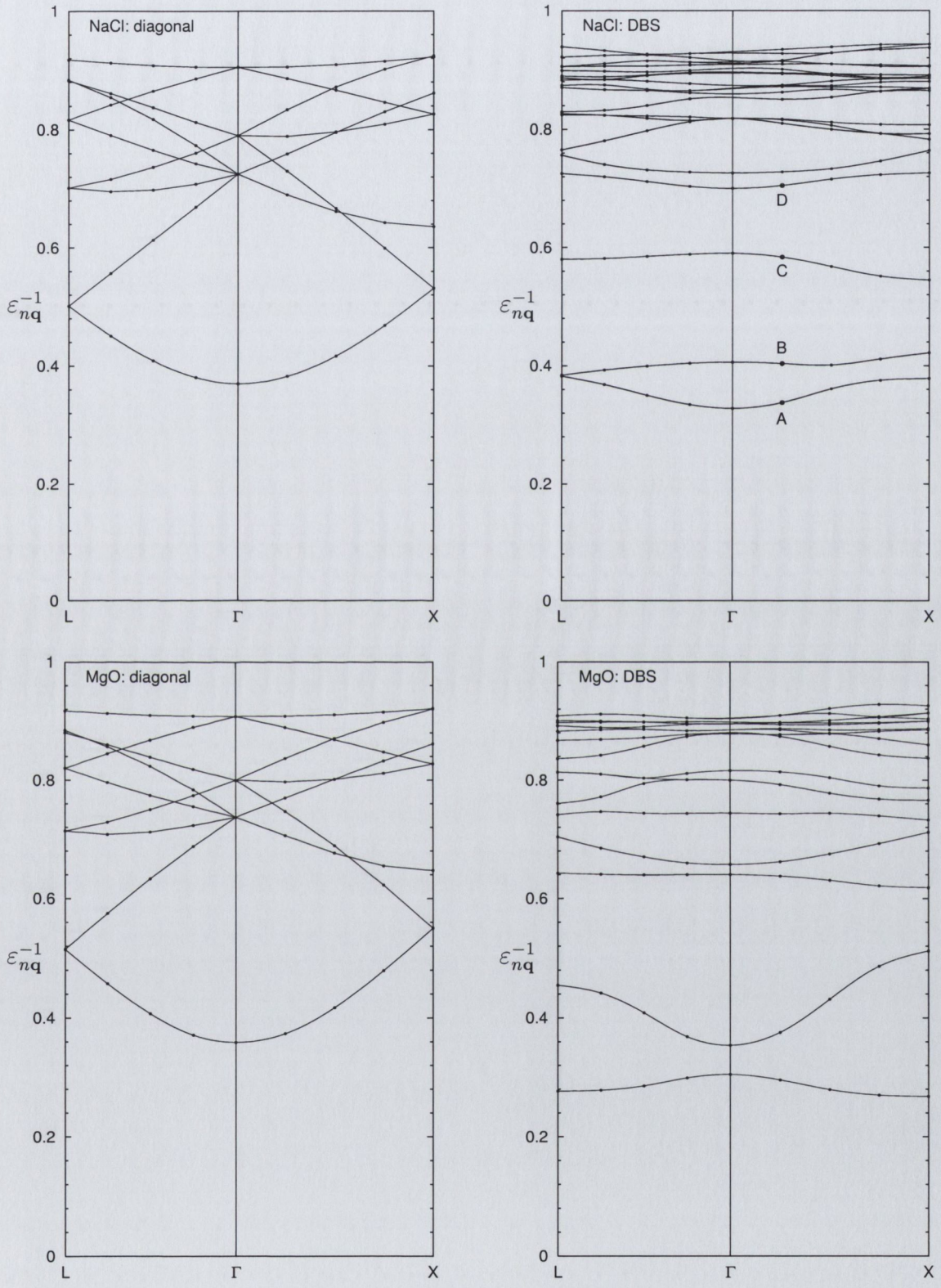


Figure 4-6: Dielectric bandstructures of NaCl and MgO: full calculation, and diagonal approximation.

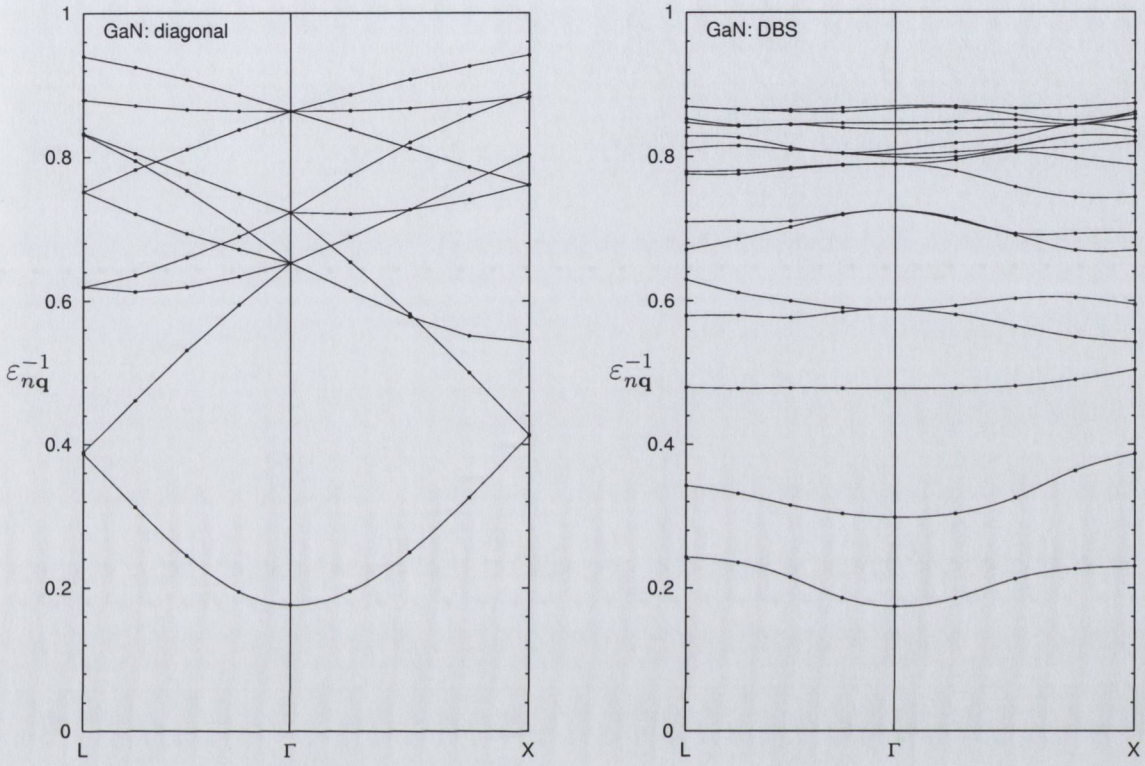


Figure 4-7: Dielectric bandstructure of GaN: full calculation, and diagonal approximation.

	ε_1^{-1}	ε_2^{-1}	ε_3^{-1}	ε_4^{-1}	ε_5^{-1}	ε_6^{-1}	ε_7^{-1}
Si	0.080	0.394	0.515(2)	0.522	0.565	0.632	0.654 (2)
C	0.175	0.327	0.624	0.635(2)	0.678(3)	0.695	0.785 (2)
GaN	0.174	0.298	0.478	0.588(2)	0.589	0.725(3)	0.789 (3)
MgO	0.305	0.354	0.640(2)	0.657	0.744	0.800(2)	0.877 (3)
NaCl	0.328	0.406	0.590(2)	0.700	0.726	0.816(2)	0.818 (2)

Table 4.3: Smallest inverse eigenvalues ε_n^{-1} of the dielectric matrix at $\mathbf{q} \to 0$ for some typical covalent and ionic solids. The number in brackets indicates the degeneracy.

where $C_n(\mathbf{q})$ denotes the expansion coefficient. In addition, the summation is simplified by the fact that an arbitrary potential of a given symmetry can decomposed into eigenpotentials of the same symmetry [55], although again we mention that the symmetry classification was not carried out in the thesis. The eigenpotentials may therefore be regarded as independent modes of screening which may be induced by an external perturbation. Since they originate in the off-diagonal response, the eigenpotentials are intimately linked to the nonlocality of the dielectric matrix and the short range local fields.

Calculations of the four most screened eigenpotentials corresponding to the lowest inverse eigenvalues ε_n^{-1} in silicon and sodium chloride are displayed in Figures 4-8 and 4-9. These are presented in the form of isosurfaces (equipotential surfaces), where red and blue denote positive and negative isosurfaces respectively. Because of the form of equation (4.10), there is an additional degeneracy with regard to real and imaginary parts of $V_n(\mathbf{q}, \mathbf{r})$, and we have chosen to plot the component with the largest magnitude in each case. Also, the $e^{i\mathbf{q}\cdot\mathbf{r}}$ term in equation (4.10) imposes a slow modulation at finite $\mathbf{q}$ on the eigenpotential across several cells, and as this tends to obscure the interesting short range information we have chosen to plot the cell-periodic part only at $\mathbf{q} = \frac{2\pi}{a}(\frac{1}{8}, 0, 0)$, as indicated by the points $A \ldots D$ in Figures 4-5 and 4-6. The most screened eigenpotential reflects the main electronic polarization centre in both materials: silicon, being covalent, exhibits a longitudinal dipole centred on the bond [Figure 4-8(A)], whereas in the ionic material all the eigenpotentials that are shown are located at the anion. The lesser screened modes for silicon tend to be localized on the atomic centres. Figure 4-8(B) corresponds to a monopolar charge transfer mode, while Figure 4-8(C) and (D) are the most screened transverse and longitudinal modes that are involved in L-T splitting. Practically all of the screening eigenpotentials of interest in sodium chloride are located on the anion, which is consistent with classical ionic polarizabilities which allocate only about 6% of the cell polarizability to the cation [57]. It is dangerous, however, to try to impart too much physical information onto specific eigenpotentials, as noted by Engel and Farid [61] and as implicitly suggested by equation (4.12); for realistic

external potentials, the eigenpotentials are best viewed as a mathematical construct, i.e., an orthogonal basis set.

The results obtained above from the real space decomposition of the dielectric matrix suggests the use of an efficient set of basis functions for schematizing the polarizability. Simple models of the dielectric matrix have been found to reproduce the screening mechanism due to the diagonal elements very well [60, 21]; off diagonal screening, however, depends more on the details of the electronic bandstructure, and is therefore more difficult to describe. The traditional use of plane waves for representing the dielectric matrix is motivated here by the ease of comparison with other calculations, but having deconstructed the matrix into the real space eigenpotentials it is evident that most of the screening, with the effects of local fields included, could be easily reproduced with the use of a few carefully chosen functions. The shape of the most screened eigenpotential in silicon lends credence to the *ansatz* of the discrete dipole model which locates the polarizable entities at the bond sites in the semiconductor. However, it is clear that the effects of nonlocality inside the cell cannot be included in such a description, and would require further polarization centres on the atoms.

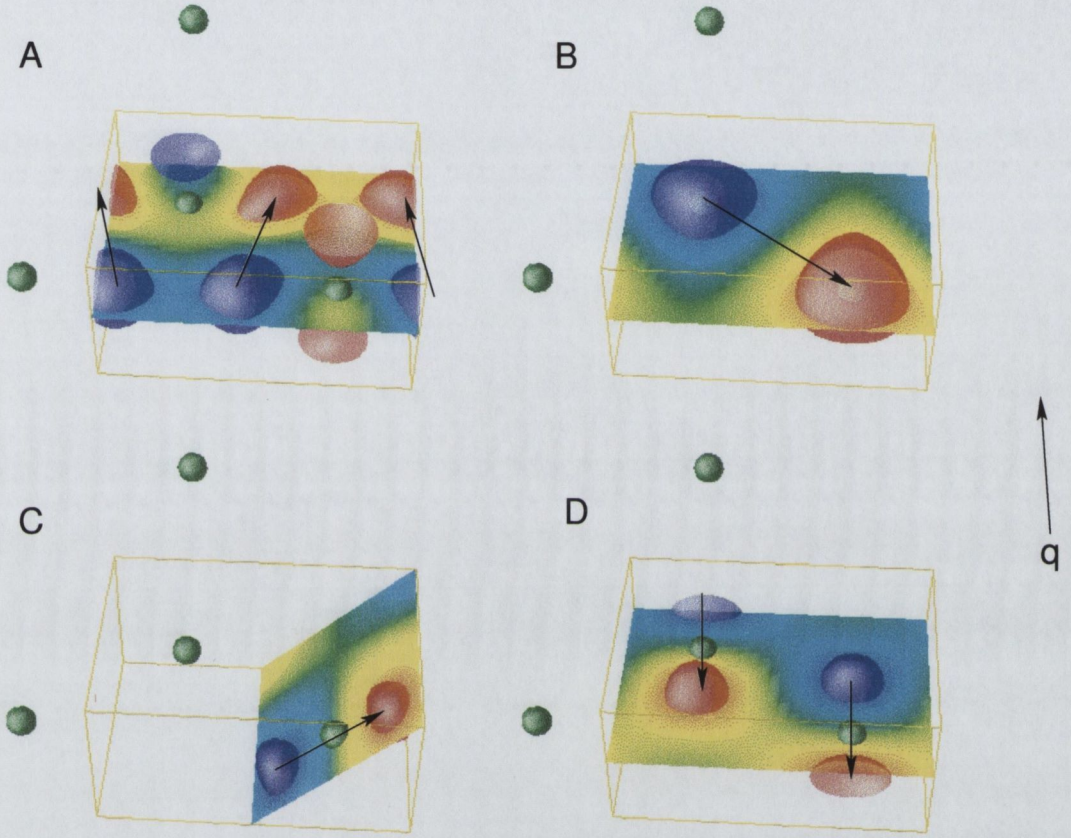


Figure 4-8: Cell periodic part of selected dielectric eigenpotentials of silicon, corresponding to the four largest eigenvalues. The labels $A \ldots D$ refer to the points marked on the dielectric bandstructure of Figure 4-5, and the direction of the wavevector $\mathbf{q}$ is indicated. Where real and imaginary components of the eigenpotentials are available, the component containing the isosurface of greatest magnitude is depicted.

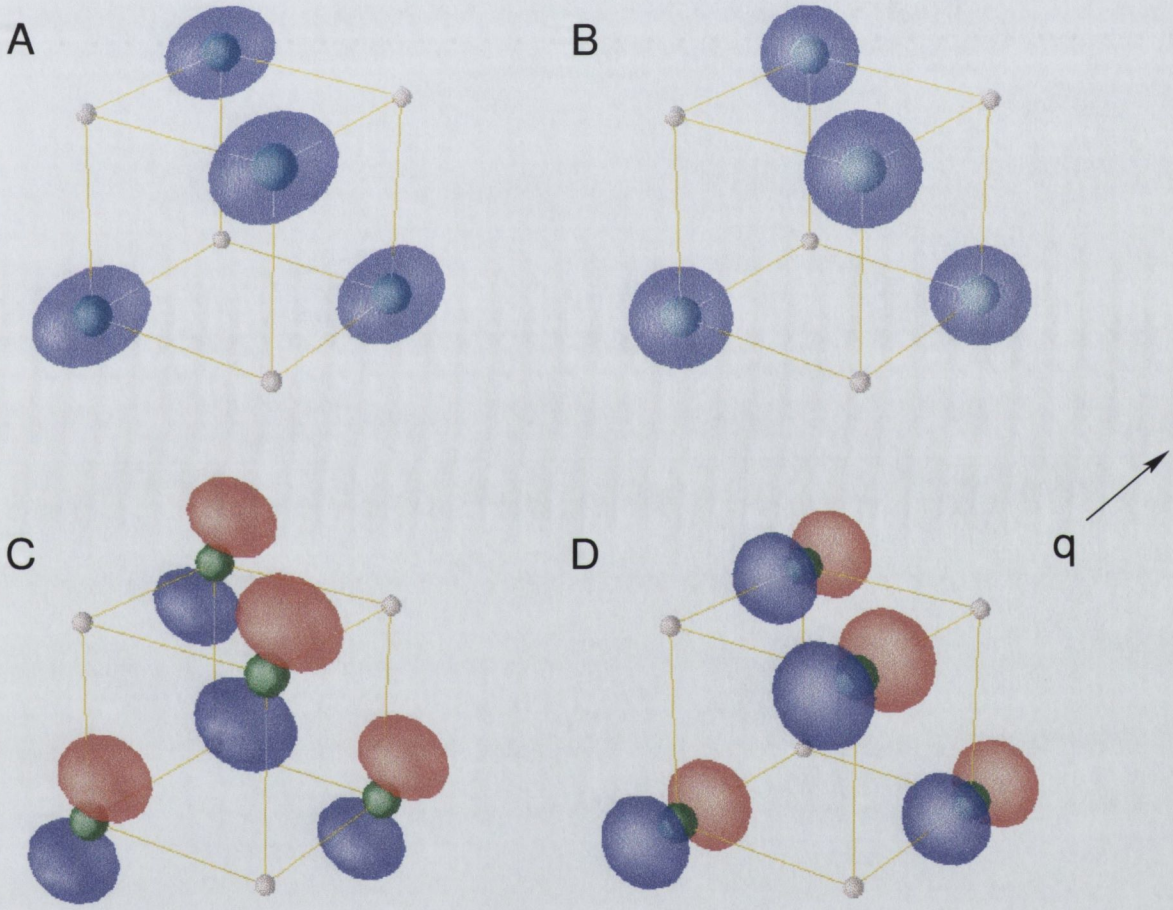


Figure 4-9: Cell periodic part of selected dielectric eigenpotentials of NaCl, corresponding to the four largest eigenvalues. The labels $A \ldots D$ refer to the points marked on the dielectric bandstructure of Figure 4-6, and the direction of the wavevector $\mathbf{q}$ is indicated. Where real and imaginary components of the eigenpotentials are available, the component containing the isosurface of greatest magnitude is depicted.

Chapter 5

Self-Energy Calculations

αγεωμέτρητος μηδεὶς εισίτω

Let no man enter who does not know mathematics

— *Sign at entrance to Plato's academy*

The main focus of this chapter is the application of the dielectric function to a more complex problem, namely, the calculation of the self-energy and the quasiparticle bandstructure. The most commonly adopted approach to computing the self-energy, and the scheme used in this thesis, is the *GW* approximation due to Hedin. The first full *ab initio* computation of the self-energy was performed by Hybertson and Louie in 1985 [62] on a range of semiconductors and insulators, and close agreement between quasiparticle states and experimental photoemission and inverse photoemission results was obtained. Several different approaches have followed since then [19, 63], although two common elements are prevalent: the use of the DFT-LDA bandstructure in the construction of the Green's function and the dielectric matrix, and the expansion of the latter in a plane-wave basis.

In the first section of this chapter the plane-wave representation of the self-energy in the *GW* approximation is presented, although there is an important distinction with regard to the conventional approach, in that the crystalline Bloch functions are expressed in terms of localized Gaussian orbitals. This mixed-basis scheme is not an optimal approach; however, it does facilitate the use of the dielectric matrices

that have been examined in detail in the previous chapters. In the following section, a plasmon-pole approximation for the dynamical screened Coulomb interaction is explained in some detail, and some comments are made about the modifications that must be applied to the reciprocal space integration scheme. The procedure for calculating the quasiparticle energies from the frequency dependent self-energy is then outlined, and the chapter ends with several calculations and a discussion of the merits and problems associated with the Hartree-Fock implementation of the *GW* scheme that has been adopted in this work.

5.1 Plane-Wave Representation

In Section 2.4 the expectation value of the self-energy operator was expressed in a basis of Hartree-Fock orbitals, which made it possible to decompose the self-energy into a static part, corresponding to the Hartree-Fock exchange potential, and a dynamic part, corresponding to the polarization of the medium. The Hartree-Fock exchange part is self-consistently included in the ground state *ab initio* calculation, while the polarization part has to be explicitly evaluated. Within the *GW* approximation, computation of the self-energy involves the calculation of the zeroth-order Green's function G_0, which is constructed from the Hartree-Fock orbitals, as well as the screened interaction W, which is closely related to the RPA polarizability examined in Chapter 4.

Up to this point, however, the analysis has been in a real-space representation, whereas the RPA polarizability has been expressed in reciprocal space. Although it is possible to calculate the self-energy in real-space (notably the real-space/imaginary time code of Rojas *et al.* [19]), it is easiest to continue the dielectric matrix analysis of Chapter 4 and work in the plane-wave (PW) basis. This approach takes advantage of the translational lattice symmetry and allows direct comparisons to be made with the literature [6, 64, 63]. The most efficient and consistent scheme is to represent the screened interaction W in a Gaussian orbital (GO) basis, distinct from the one used to expand the Bloch functions [24, 59]. Nevertheless, in this thesis we adopt

a mixed (HF-GO, GWA-PW) approach, primarily for the purpose of evaluating the Hartree-Fock based scheme as a basis for future work, although it should be noted that it is preferable to calculate certain elements of the GWA-GO scheme [24] (notably the Coulomb interaction matrix) in a plane-wave basis and transform this into the GO basis.

The formal expression in a real-space and energy basis for the self-energy within the *GW* approximation is given by

$$\Sigma(\mathbf{r},\mathbf{r}',E) = \frac{i}{2\pi}\int e^{-i\omega 0^+} G(\mathbf{r},\mathbf{r}',E-\omega)W(\mathbf{r},\mathbf{r}',\omega)\mathrm{d}\omega. \tag{5.1}$$

As discussed in Chapter 2, the single particle Green's function G is approximated by the zero-order Hartree-Fock Green's function:

$$G(\mathbf{r},\mathbf{r}',E) = \sum_{n\mathbf{k}} \frac{\psi^{\mathrm{HF}}_{n\mathbf{k}}(\mathbf{r})\psi^{\mathrm{HF}*}_{n\mathbf{k}}(\mathbf{r}')}{E - E^{\mathrm{HF}}_{n\mathbf{k}} + i0^+\mathrm{sgn}(E^{\mathrm{HF}}_{n\mathbf{k}} - \mu)}, \tag{5.2}$$

where μ denotes the chemical potential, and the eigenfunctions $\psi^{\mathrm{HF}}_{n\mathbf{k}}$ and energies $E^{\mathrm{HF}}_{n\mathbf{k}}$ are obtained from a full self-consistent Hartree-Fock calculation using the CRYSTAL95 code. The screened interaction W is given by a convolution of the inverse time-ordered dielectric function within the RPA and the Coulomb potential:

$$W(\mathbf{r},\mathbf{r}',\omega) = \int \mathrm{d}\mathbf{r}''\varepsilon^{-1}(\mathbf{r},\mathbf{r}'',\omega)v(\mathbf{r}'',\mathbf{r}'), \tag{5.3}$$

and is related to its Fourier transform by wavevector basis via the expression

$$W(\mathbf{r},\mathbf{r}',\omega) = \sum_{\mathbf{q},\mathbf{G},\mathbf{G}'} e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} W_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) e^{-i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'}, \tag{5.4}$$

where $\mathbf{G}$ is a reciprocal lattice vector and $\mathbf{q}$ is a wavevector in the first Brillouin zone. The Fourier coefficients in equation (5.4) are given by

$$\begin{aligned} W_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) &= \varepsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega)v_{\mathbf{G}\mathbf{G}'}(\mathbf{q}) \\ &= \varepsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega)\frac{4\pi e^2}{V}\frac{1}{|\mathbf{q}+\mathbf{G}|}\frac{1}{|\mathbf{q}+\mathbf{G}'|}. \end{aligned} \tag{5.5}$$

The time-ordered dielectric matrix differs from the linear-response (causal) dielectric matrix in Chapter 4 by the sign of the infinitesimal in the denominator (see equation (3.34)). The Hermitian form of the matrix has also been adopted.

On combining the Fourier transformed expression for W with the spectral representation of G, we obtain for the diagonal matrix element of the self-energy operator [65]:

$$\begin{aligned}\langle m\mathbf{k}|\Sigma(E)|m\mathbf{k}\rangle &= \int \mathrm{d}\mathbf{r} \int \mathrm{d}\mathbf{r}' \Psi^*_{m\mathbf{k}}(\mathbf{r})\Sigma(\mathbf{r},\mathbf{r}',E_{m\mathbf{k}})\Psi_{m\mathbf{k}}(\mathbf{r}') \\ &= \frac{4\pi}{V} \sum_{\mathbf{q},\mathbf{G},\mathbf{G}'} \sum_n \frac{\langle m\mathbf{k}|e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}|n\mathbf{k}+\mathbf{q}\rangle}{|\mathbf{q}+\mathbf{G}|} \frac{\langle n\mathbf{k}+\mathbf{q}|e^{i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}}|m\mathbf{k}\rangle}{|\mathbf{q}+\mathbf{G}'|} \\ &\quad \times \frac{i}{2\pi} \int_{\omega\in\mathbf{R}} \mathrm{d}\omega\, e^{-i\omega 0^+} \frac{\varepsilon^{-1}_{-\mathbf{G},-\mathbf{G}'}(-\mathbf{q},\omega)}{E-\omega-E_{n,\mathbf{k}+\mathbf{q}}+i0^+\mathrm{sgn}(E_{n,\mathbf{k}+\mathbf{q}}-\mu)} \end{aligned} \tag{5.6}$$

The HF label has been dropped from the eigenstates and eigenvalues for clarity throughout.

5.2 Plasmon-Pole Model

It is apparent from a brief inspection of equation (5.6) that a direct computational scheme would involve the calculation and inversion of the dielectric matrix for many values of ω. Although such a scheme is possible[15, 66], it is extremely time consuming in the present plane-wave framework. Therefore, we utilize a plasmon-pole model to approximate the frequency dependence of the dielectric matrix: within this scheme, only the static matrix $\varepsilon_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega\to 0)$ has to be fully calculated, and a model form is assumed that allows equation (5.6) to be integrated analytically. For many systems with s,p-type character,[1] the imaginary part of the screened interaction (Im W) is identified by a strong peak at a $\mathbf{q}$-dependent frequency $\omega_{\mathbf{q}}$ that corresponds to the excitation of a plasmon, which suggests that Im W could be realistically modelled by a pole at this frequency. The plasmon is a well-defined excitation at $\mathbf{q}\to 0$; at non-zero $\mathbf{q}$ it merges with the particle-hole excitations and is

[1] *i.e.*, simple semiconductors and insulators whose chemical properties are mostly determined by the outermost valence s and p orbitals.

not quite so well represented by a single peaked function. This is even more relevant for more complex systems of d, f-character, where the plasmon-pole model may not be applicable[10]. In addition, use of the plasmon-pole model prevents the quasi-particle life-times from being determined, and we are thus restricted to calculating the energy spectrum of the QP states.

Building on the concept of the dielectric bandstructure discussed in Chapter 4, we adopt a plasmon-pole model based on the generalized model of von der Linden and Horsch[64]. In this variation, the *ansatz* is made that the eigenvectors of the dielectric matrix are independent of frequency, which implies that all the energy dependence is projected onto the eigenvalues. The inverse Hermitian dielectric matrix is approximated by:

$$\varepsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega) = \sum_l V_{l,\mathbf{G}}(\mathbf{q})\, \varepsilon^{-1}_{\mathbf{q}l}(\omega)[V_{l,\mathbf{G}'}(\mathbf{q})]^*, \tag{5.7}$$

where the model form for the eigenvalues is given by

$$\varepsilon^{-1}_{\mathbf{q}l}(\omega) = 1 + \frac{z_{\mathbf{q}l}\omega_{\mathbf{q}l}}{2}\left(\frac{1}{\omega - \omega_{\mathbf{q}l} + i0^+} - \frac{1}{\omega - \omega_{\mathbf{q}l} - i0^+}\right), \tag{5.8}$$

where $z_{\mathbf{q}l}$ are weighting parameters, and $\omega_{\mathbf{q}l}$ are the plasmon-pole frequencies.

Other proponents of the plasmon-pole model generally rely on the use of the Johnson f-sum rule[35] to determine the fitting parameters $z_{\mathbf{q}l}$ and $\omega_{\mathbf{q}l}$. However, this approach is not valid in the presence of a nonlocal potential in the single-particle Hamiltonian[30], which is of course relevant to the calculations in this thesis (the Hartree-Fock formalism explicitly contains a nonlocal exchange potential). A calculation similiar to that presented in Figure 3.1.3 was carried out to test the validity of the f-sum rule applied to the RPA dielectric function obtained from a Hartree-Fock bandstructure. As expected, it was observed that the sum-rule was violated: η_{eff} for silicon was calculated at 1.9 valence electrons, which disagrees badly with the correct value of 4. Evidently, a quite distinct approach is necessary. However, since the main focus of this thesis concerns optical properties, and hence, the self-energy correction to the band gap, it is sufficient to accurately determine the

states around the Fermi level, which occur at a much lower energy that that of the plasmon. This is reflected in the differing forms of plasmon-pole model appearing in the literature (some model the frequency dependence of each dielectric matrix element separately[6], some attribute a separate pole to each eigenvalue[64, 21] — with varying degrees of theoretical inconsistency). Nonetheless, all these models agree on the use of the static dielectric matrix as a fitting parameter.

Based on these observations, a simple fitting procedure has been adopted to determine a pole frequency $\omega_{\mathbf{q}l}$ and pole strength $z_{\mathbf{q}l}$ for each eigenvalue. The dielectric matrix is calculated at two frequencies — at zero frequency and at some non-zero frequency ω_f — and it is then trivial to show that the two parameters are given by

$$\omega_{\mathbf{q}l}^2 = \frac{\omega_f^2[1 - \varepsilon_{\mathbf{q}l}^{-1}(\omega_f)]}{\varepsilon_{\mathbf{q}l}^{-1}(\omega \to 0) - \varepsilon_{\mathbf{q}l}^{-1}(\omega_f)} \tag{5.9}$$

$$z_{\mathbf{q}l} = 1 - \varepsilon_{\mathbf{q}l}^{-1}(\omega \to 0). \tag{5.10}$$

However, this procedure is quite unstable with respect to the choice of the w_f parameter. An improved technique is to calculate the non-zero frequency dielectric matrix at an imaginary frequency instead[67]. In this case the dielectric matrix obeys a Lorentzian-type frequency behaviour and is less sensitive to the rapid fluctuations which occur in $\varepsilon_{\mathbf{q}l}$ for energies around and above the gap. In this case the parameters are given by the slightly modified expressions

$$\omega_{\mathbf{q}l}^2 = \frac{-\omega_f^2[1 - \varepsilon_{\mathbf{q}l}^{-1}(i\omega_f)]}{\varepsilon_{\mathbf{q}l}^{-1}(\omega \to i0) - \varepsilon_{\mathbf{q}l}^{-1}(i\omega_f)} \tag{5.11}$$

$$z_{\mathbf{q}l} = 1 - \varepsilon_{\mathbf{q}l}^{-1}(\omega \to i0). \tag{5.12}$$

Figure 5-2 demonstrates how the true calculated eigenvalues compare with the values predicted by the plasmon-pole model along the imaginary energy axis. The agreement with the calculated values is quite good over the full energy range displayed in the figure, even at values of w_f corresponding to the problematic gap frequency. The eigenvalue fitting for real values of w_f is shown in Figure 5-1 and

does not exhibit a satisfactory agreement over the full energy range, although again the agreement with the fully calculated points is good within the low energy range. The only restriction in this case is that ω_f must be chosen to lie below the gap; a poorly chosen ω_f may even lead to unphysical plasmon-pole frequencies. The calculated values of the quasiparticle band-gap in silicon, obtained using the two fitting procedures, were found to differ by less than 0.02 eV. It should be stressed, however, that an imaginary w_f is used simply to determine the two plasmon-pole parameters, and that the frequency integration in equation (5.6) is still performed with respect to real energies.

5.3 Self-Energy Matrix Element

The plasmon-pole form of $\varepsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega)$ is inserted into equation (5.6) for the self-energy, which allows the frequency integral to be analytically evaluated. This yields the following expression for the expectation value of Σ [21]:

$$\langle m\mathbf{k}|\Sigma(E)|m\mathbf{k}\rangle = \frac{4\pi e^2}{V}\sum_{\mathbf{q},\mathbf{G},\mathbf{G}'}\sum_{n}\frac{\langle m\mathbf{k}|e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}|n\mathbf{k}+\mathbf{q}\rangle}{|\mathbf{q}+\mathbf{G}|}\frac{\langle n\mathbf{k}+\mathbf{q}|e^{i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}}|m\mathbf{k}\rangle}{|\mathbf{q}+\mathbf{G}'|}$$
$$\times\begin{cases}\displaystyle\sum_{l}V_{l,-\mathbf{G}}(-\mathbf{q})[V_{l,-\mathbf{G}'}(-\mathbf{q})]^*\left[-1+\frac{z_{-\mathbf{q}l}\omega_{-\mathbf{q}l}}{2}\frac{1}{E-E^{\mathrm{HF}}_{n\mathbf{k}+\mathbf{q}}+\omega_{-\mathbf{q}l}}\right] & \text{for } n\in\text{VB}\\ \displaystyle\sum_{l}V_{l,-\mathbf{G}}(-\mathbf{q})[V_{l,-\mathbf{G}'}(-\mathbf{q})]^*\left[\frac{z_{-\mathbf{q}l}\omega_{-\mathbf{q}l}}{2}\frac{1}{E-E^{\mathrm{HF}}_{n\mathbf{k}+\mathbf{q}}-\omega_{-\mathbf{q}l}}\right] & \text{for } n\in\text{CB}.\end{cases} \tag{5.13}$$

It is clear from this equation how the self-energy matrix element splits naturally into a bare exchange part (from the '-1' for n $\in$VB) and a part containing the contributions from screening and correlation:

$$\langle m\mathbf{k}|\Sigma(E)|m\mathbf{k}\rangle = \langle m\mathbf{k}|\Sigma_{ex}|m\mathbf{k}\rangle + \langle m\mathbf{k}|\Sigma_s(E)|m\mathbf{k}\rangle. \tag{5.14}$$

The exchange term is of course explicitly contained in the initial self-consistent Hartree-Fock calculation.

Equation (5.13) gives a clear indication of how computationally intensive the calculation of the self-energy is, even with a plasmon-pole approximation for $\varepsilon_{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega)$.

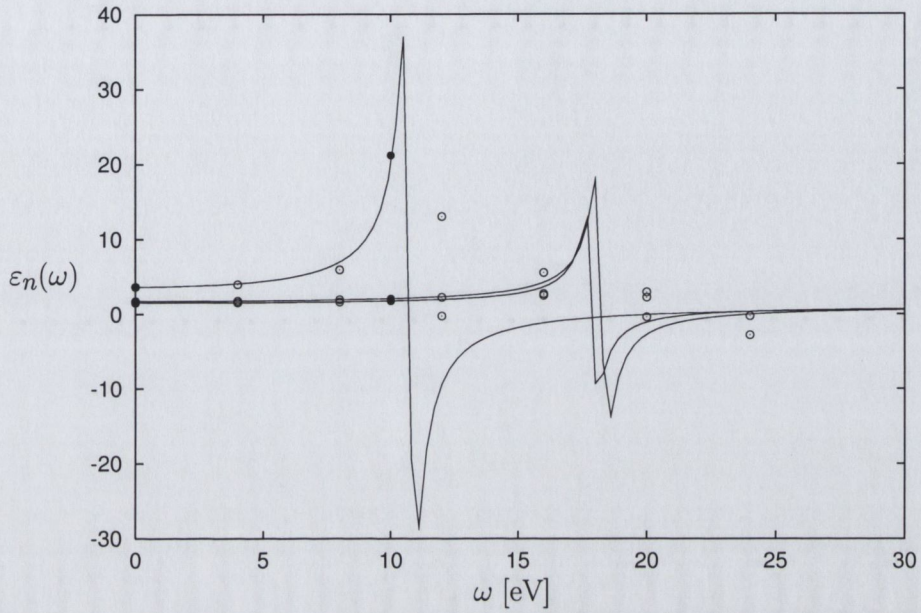


Figure 5-1: Comparison between plasmon-pole model (lines) and fully calculated eigenvalues of the dielectric matrix (circles) at several energies, for the three largest eigenvalues. The plasmon-pole spectra are fitted to the eigenvalues indicated by the solid circles at 0 eV and 10 eV. A small broadening parameter has been used for plotting the model spectra.

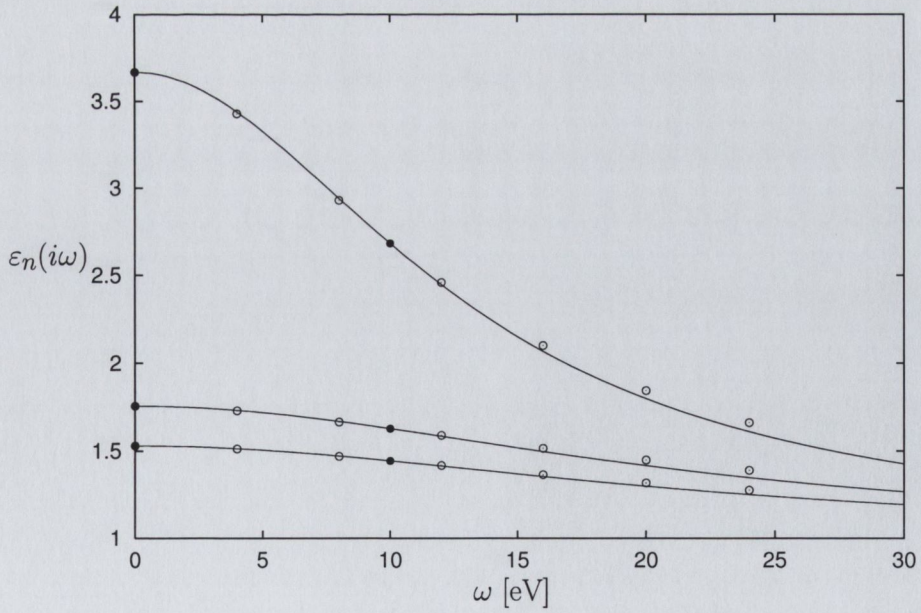


Figure 5-2: Comparison between plasmon-pole model (lines) and fully calculated eigenvalues of the dielectric matrix (circles) at several *imaginary* energies, for the three largest eigenvalues.

The full expression for $\Sigma(E)$ involves a double sum over reciprocal lattive vectors $\mathbf{G}, \mathbf{G}'$, as well as a sum over $\mathbf{q}$ within the full Brillouin zone, which supports the use of the plasmon-pole approximation, since the fine details of the summand are evidently not critical to the final result. Although demanding, the computation itself is quite straightforward once the optical matrix elements and dielectric bandstructure have been calculated. One important difficulty remains, however, in the summation over $\mathbf{q}$; the occurrence of singularities in equation (5.13) when $\mathbf{G} = \mathbf{G}' = 0$ or $\mathbf{G} \neq \mathbf{G}' = 0$ has the effect of rendering the special-point theorem invalid in this case. As noted by several authors[65, 68], this summation has to be treated with care. The auxiliary function technique of Gygi *et al.* [68] has been adopted for this reason and is discussed in detail in Appendix C.3.

5.4 Determination of Quasiparticle Energies

Once the matrix elements $\langle m\mathbf{k}|\Sigma_s(E)|m\mathbf{k}\rangle$ are evaluated, the problem still remains of calculating the self-energy at the actual quasiparticle energy $E^{\mathrm{QP}}_{m\mathbf{k}}$ to obtain the corrections to the bandstructure. As noted in Section 2.4, this is achieved by solving the equation

$$E^{\mathrm{QP}}_{m\mathbf{k}} = E^{\mathrm{HF}}_{m\mathbf{k}} + \langle m\mathbf{k}|\Sigma_s(E^{\mathrm{QP}}_{m\mathbf{k}})|m\mathbf{k}\rangle. \tag{5.15}$$

A simple quadratic interpolation scheme has been adopted whereby the matrix elements of Σ_s are calculated at the Hartree-Fock energy $E^{\mathrm{HF}}_{m\mathbf{k}}$ and also at two bounding energies $E^{\mathrm{HF}}_{m\mathbf{k}} \pm \Delta$, where Δ is an arbitrary fitting parameter, typically of the order 1–2 eV. Equation (5.15) is trivially solved by finding the intersection between the line $f(E) = E - E^{\mathrm{HF}}_{m\mathbf{k}}$ and the quadratic fit to the three fully calculated points, $f(E) = \Sigma^{\mathrm{Fit}}_{m\mathbf{k}}(E)$. A schematic representation of the procedure is given in Figure 5-3. Although the self-energy does not of course exhibit a true quadratic behaviour with respect to energy, it is sufficiently smooth around the quasiparticle energy to merit this approach so long as the Δ parameter is chosen carefully. The conventional technique for determining the quasiparticle energies utilizes a 'renormalization factor'

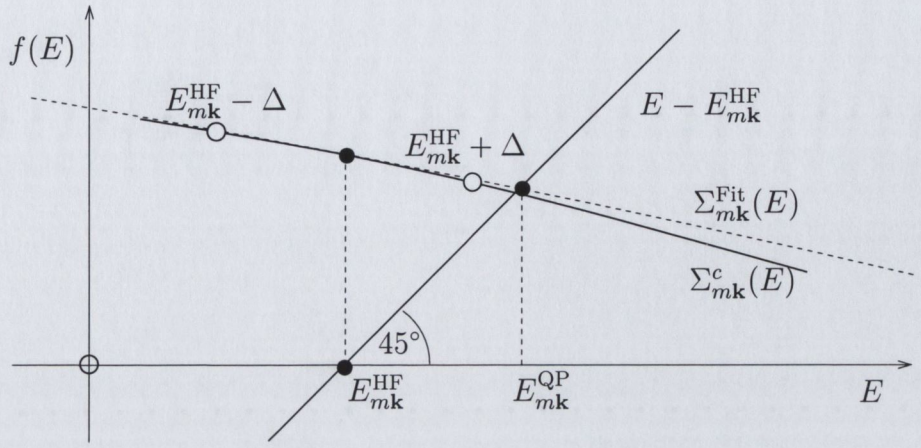


Figure 5-3: Schematic representation of the fitting procedure used to obtain the quasiparticle energies.

$Z_{m\mathbf{k}}$[10] defined by

$$Z_{m\mathbf{k}} = \left[1 - \frac{\partial \mathrm{Re}\Delta\Sigma_s(E_{m\mathbf{k}}^{\mathrm{QP}})}{\partial E}\right]^{-1} \tag{5.16}$$

such that the QP energies may be determined by

$$E_{m\mathbf{k}}^{\mathrm{QP}} = E_{m\mathbf{k}}^{\mathrm{HF}} + Z_{m\mathbf{k}}\,\mathrm{Re}\Sigma_s(E_{m\mathbf{k}}^{\mathrm{HF}}); \tag{5.17}$$

this technique is equivalent to a linear interpolation, accurate to first order in $E_{m\mathbf{k}}^{\mathrm{QP}} - E_{m\mathbf{k}}^{\mathrm{HF}}$. The results seem to be quite insensitive, however, to the choice of Δ or the interpolation scheme used: the size of Δ was varied to test the effect on the calculated quasiparticle band gaps, and it was observed that a change of 50% resulted in a change in the band gap of less than 0.01 eV. In the following section it will be shown that the slope of the self-energy is negative in the energy range of the quasiparticle states which implies that the general effect of renormalization is to reduce the magnitude of the self-energy correction.

5.5 Results and Discussion

Self-energy corrections to the Hartree-Fock bandstructure were calculated for the three materials Si, diamond and Mgo, using the method described above. In each case, highly accurate basis sets were developed to provide the large number of con-

duction band states that are necessary for a sufficiently accurate calculation of the inverse dielectric matrix (basis sets are documented in Appendix D). For the main self-energy calculation, various run-time parameters had to be determined to ensure well converged results; specifically, these include the density of **k**- and **q**-point meshes, the number of **G**-vectors in the dielectric matrix and the real-space cut-off for the Bloch sums that determine wavefunction normalization. 256 **k**-point Monkhorst-Pack meshes were used when computing the head and wings of the dielectric matrices for each material, whereas a 32 **k**-point mesh was used for the body elements as well as for the **q**-point integration scheme. These parameters yielded results for the quasiparticle band gaps at Γ within 0.2 eV of the converged values. In addition, Hartree-Fock dielectric bandstructures were generated in order to determine the two plasmon-pole model parameters $V_{l,\mathbf{G}}(\mathbf{q})$ and $z_{\mathbf{q},l}$ in the limit $\mathbf{q} \to 0$; these parameters are necessary for the **q**-integration scheme, as described in detail in Appendix C.3.

Figures 5-4,5-5 and 5-6 show the GWA-corrected bandstructures superimposed upon the Hartree-Fock bandstructures calculated self-consistently from CRYSTAL95, for Si, diamond and MgO, respectively. It is generally observed that the self-energy correction substantially reduces the band-gaps from the exaggerated HF values, by both raising the valence bands and lowering the conduction bands. As observed in GWA-LDA calculations, the HF (LDA) conduction bands are shifted uniformly (although, of course, in the opposite direction to GWA-LDA) whereas the valence bands are shifted asymmetrically, resulting in a narrowing of the valence band. However, although the inclusion of the self-energy corrections succeeds in reducing the size of the Hartree-Fock band-gaps in each of the materials studied, it does not succeed in producing an accurate agreement with experiment, as shown in Table 5.1. The direct and indirect band-gaps remain overestimated, while the valence band widths are almost unchanged from their Hartree-Fock values.

One possible reason for the origin of these discrepancies lies in the use of the Gaussian orbital formalism itself. In Section 4.1.2 it was demonstrated that the restricted size of the basis set in the LCAO scheme leads to a lack of convergence

	Si			Diamond			MgO		
	HF	GWA	Expt.	HF	GWA	Expt.	HF	GWA	Expt.
E_g^{dir}	8.58	4.40	3.40	14.37	9.11	7.30	15.59	11.94	7.83
E_g^{ind}	5.96	1.76	1.17	12.08	6.51	5.48	-	-	
E_{VB}	17.17	15.02	12.5±0.6, 11.2	29.66	26.68	24.2±1, 21±1	23.41	22.37	17.0

Table 5.1: Hartree-Fock (HF) and *GW*-approximation results for the direct band-gap E_g^{dir}, indirect band-gap E_g^{ind} and valence band width from the present calculations for Si, diamond and MgO, respectively, compared with the experimental values. Energies are given in eV. Experimental values taken from Ref. [69]

during the construction of the dielectric matrix, and that interstitial functions were deemed necessary to produce a basis set with more than ~20 low-lying conduction band states, for all the materials studied. However, it is found that the self-energy correction to the Hartree-Fock gap is not very sensitive to the size of the basis set. A test calculation reveals that the use of ~20 conduction bands in the summations of Equation (5.13) results in a minimum band-gap of 1.78 eV for silicon, compared with a value of 1.76 eV for ~60 bands. Similar behaviour is also reported in the GWA-DFT work of Rohlfing *et al.* [21]: the corrected band-gaps for calculations using 40 and 60 Gaussian orbitals respectively are found to differ by (at most) 0.06 eV. A lack of convergence due to basis set incompleteness may therefore be discounted. A second possible reason for the discrepancy, numerical inaccuracies associated with the number of special points and **G**-vectors, can also be disgarded as care has been taken to ensure a sufficiently large number of each in all parts of the GWA calculations. The lack of agreement with experiment is thus more likely to arise from some aspect of the GWA-HF formalism.

One possibility regards the implicit assumption of equation (2.24), i.e., that the GWA wavefunctions and the Hartree-Fock wavefunctions are very similar, which permits the expectation value for the self-energy to be expressed between the Hartree-Fock eigenstates. This assumption is supported by the literature [16] and by present calculations. In Table 5.2 a comparison is presented between the direct and indirect Hartree-Fock band-gaps of silicon and diamond calculated in a full self-consistent manner, both from this work and from several published sources, and the Hartree-

Fock gaps calculated as a by-product of a GWA-DFT calculation. The latter calculations are carried out via a 'perturbed' scheme, i.e.,

$$E_g^{\mathrm{HF}} = E_g^{\mathrm{LDA}} - E_{xc}^{\mathrm{LDA}} + \Sigma_{\mathrm{ex}}, \tag{5.18}$$

which is equivalent to omitting the correlation part from the self-energy, and only considering the bare exchange part. It is interesting to note that although the 'perturbed' HF results, as obtained from GWA-DFT calculations, are scattered over a range of about 1.3 eV (for instance, $E_g^{\mathrm{ind}}(\mathrm{HF})$ for silicon ranges from 5.05 eV [Ref.(a)] to 6.35 eV [Ref.(c)]), they agree fairly well with the HF band-gaps calculated fully self-consistently. In particular, Ref.(d) also used a Gaussian-orbital based scheme and predicts values for the Hartree-Fock gaps in close agreement with the present calculations. Remarkably, the GWA-DFT corrected gaps (where available) are found to agree with experiment within 0.2eV for both materials [with the notable exception of Ref.(f)] despite these large variations in the Hartree-Fock (bare-exchange self-energy) results. Therefore, if we recall that the bare-exchange self-energy does *not* depend on the HF eigenvalues [as explicitly shown by equation (C.22)], we may conclude that the use of Hartree-Fock wavefunctions in place of the true GWA wavefunctions, or indeed, the DFT wavefunctions, does not introduce a considerable error in the calculation. The poor agreement with experiment in the GWA-HF scheme must be attributed to the use of Hartree-Fock eigenvalues as a starting point in constructing both the Green's function and the screened interaction W.

In order to examine this point in more detail, the screening part of the self-energy, $\Sigma_s(E)$, as calculated in the GWA-HF framework, is presented for Si in Figure 5-7 along with a calculation from an accurate GWA-LDA calculation [67, 20]. Although the lineshapes and intensities are similiar for both types of calculation, the quantitative agreement is quite poor; this is particularly evident for the plasmon-pole frequencies, which occur on either side of the Fermi energy, around which only single-particle excitations are possible. The two sets of data are also quite disparate

			'Perturbed' HF				Self-consistent HF			
		Ref:	(a)	(b)	(c)	(d)	(e)	(f)	Present	Expt.
Si	E_g^{dir}	(HF)	7.94	8.5	8.92	8.56	9.27	-	8.58	
		(GW)	3.30	3.68	3.30	3.39	-	-	4.40	3.40
	E_g^{ind}	(HF)	5.05	5.2	6.35	~6.0	6.43	-	5.96	
		(GW)	1.24	1.21	1.01	1.25	-	-	1.76	1.17
C	E_g^{dir}	(HF)	-	15.0	-	-	-	15.0	14.37	
		(GW)	7.26	8.84	-	7.63	-	8.25	9.11	7.12
	E_g^{ind}	(HF)	12.0	12.8	-	-	-	~13.0	12.08	
		(GW)	5.33	7.16	-	5.67	-	~6.2	6.51	5.48

(a) Refs.[70] & [15]
(b) Estimated from Fig. 2 of Ref.[65].
(c) Ref.[63]
(d) Ref.[21]
(e) Ref.[68]
(f) Estimated from Fig. 5 of [17].
Experimental values taken from Ref. [69].

Table 5.2: Hartree-Fock (HF) and GW-approximation values in eV for the direct (E_g^{dir}) and indirect (E_g^{ind}) band-gaps, respectively, of silicon and diamond obtained from the literature, compared with the values calculated in this thesis and the experimental data. 'Perturbed' here refers to the method of estimating the Hartree-Fock band gap as a by-product of an LDA-GWA self-energy calculation, where the bare exchange self-energy Σ_{ex} replaces the full $\Sigma(E)$, as described by Equation (5.18). 'Self-consistent' refers to a full *ab initio* Hartree-Fock calculation, or semi-empirical in the case of (f). The relatively poor agreement with experiment shown by (b) may be due to an improper treatment of renormalization [equation (5.17)].

in the region around the Fermi level containing the positions of the quasiparticle energies. These differences occur as a direct consequence of using the exaggerated Hartree-Fock eigenvalues. For instance, the Hartree-Fock RPA dielectric matrix is known [71] to significantly underestimate (by ~50%) the value of the experimental dielectric constant, a trend which is also found in the present results. Similarly, it is observed that the Hartree-Fock RPA dielectric matrix consistently overestimates the plasmon-pole frequencies by up to 45%, taking the bulk plasma frequency as a guideline.

As discussed in Section 2.4, there is no *a priori* reason for using the RPA dielectric matrix in the calculation of the self-energy, except that it presents the most sophisticated choice as part of a 'best G, best W' philosophy[72]. However, this refers to the use of an RPA-LDA dielectric matrix, which is known to give good qualitative agreement with experiment. Although it provides a clear theoretical reference level for evaluating the effects of correlation and screening, the GWA-HF scheme used in this thesis does not therefore appear to provide a suitable starting point for self-energy calculations that rely on a purely perturbative scheme, at least for simple s, p-character semiconductors and insulators. As discussed in Section 2.4, conventional GWA calculations rely on a self-consistent DFT-LDA calculation to provide the initial set of states, which approximately includes the effects of exchange and correlation from the onset. Since the starting point is not greatly different from the experimental result, the perturbative GWA-DFT scheme succeeds in producing results in good agreement with experiment. The current GWA-HF scheme, however, only includes the unscreened exchange term during the full self-consistent calculation, and attempts to account for the (considerably large) effects of screening and correlation via a perturbative scheme, which is not possible.

An improvement on the current results may perhaps be achieved by means of a self-consistent treatment. Thus, instead of assuming W is taken within the RPA, and G_0 is constructed from the Hartree-Fock eigenfunctions, (schematically)

$$\Sigma = G_0 W_0 = G_0 v(1 - v G_0 G_0)^{-1}, \tag{5.19}$$

both G and W may be updated as part of an iterative scheme, i.e.

$$\Sigma = GW \tag{5.20}$$

$$\text{where:} \quad G^{-1} = G_0^{-1} - \Sigma$$

$$W = v(1 - vGG)^{-1}.$$

However, many unresolved issues remain concerning the propriety of self-consistently calculating the self-energy without simultaneously accounting for correlations beyond the GW approximation, i.e. vertex corrections where $\Gamma \neq 0$. Schöne and Eguiluz [73] have performed a fully self-consistent GW calculation on silicon, and discovered that the converged band-gap has in fact a higher value than the $\Sigma = G_0W_0$ result (and hence, the experimental value) by at least 0.6 eV ($E_g^{\text{dir}} = 4.02$ eV,$E_g^{\text{ind}} = 1.91$ eV) Therefore, it is possible that a self-consistent treatment of the calculations in this chapter would also converge to values above the experimental data, which would imply that the GWA-HF scheme could not be used (in this form) to achieve good agreement with photoemission data for the materials that give correct values in GWA-DFT. It may be more worthwhile to apply the GWA-HF method to systems that are not well described by LDA, such as manganites and transition-metal oxides.

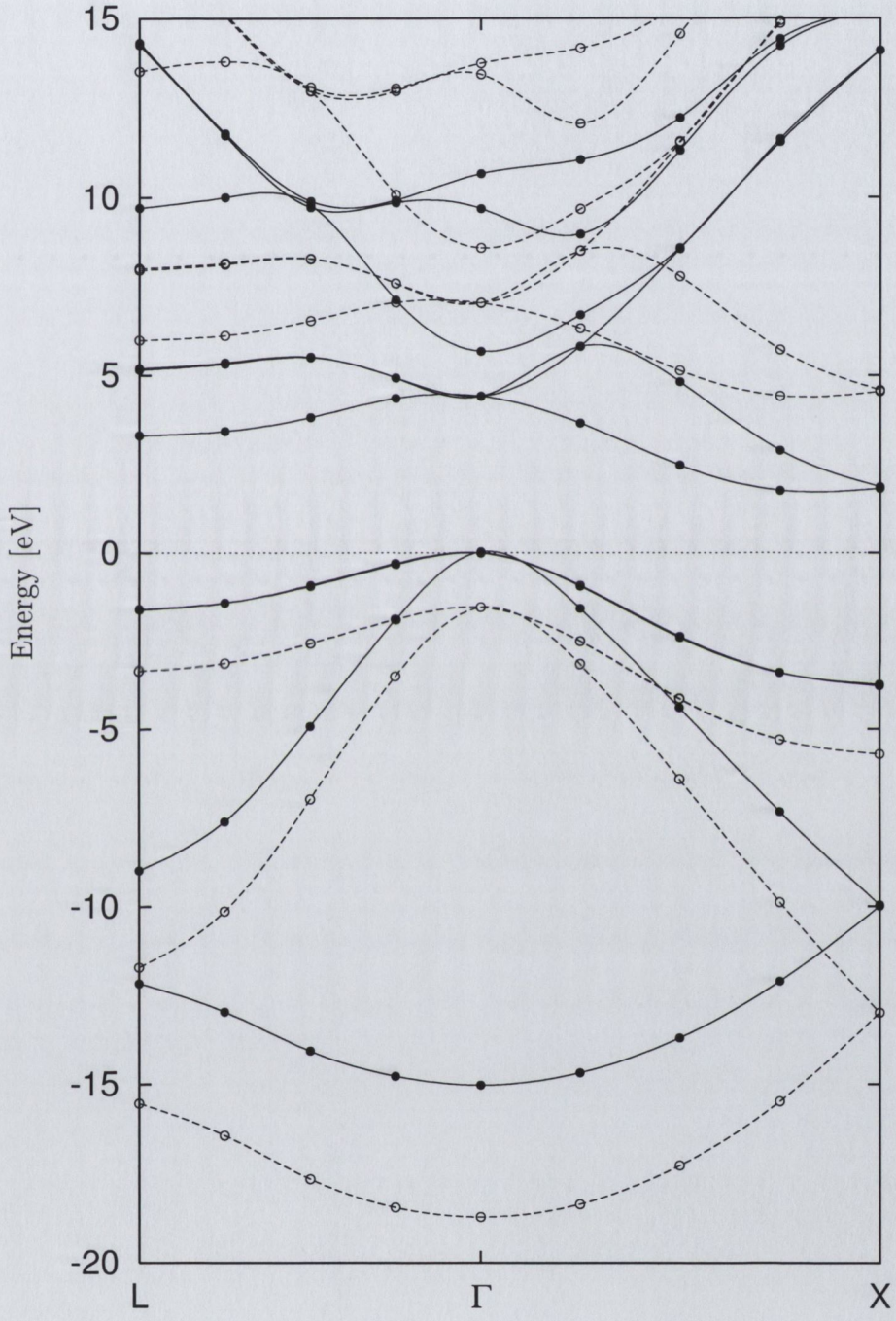


Figure 5-4: Quasiparticle bandstructure of Si (solid lines) and the initial Hartree-Fock bandstructure (dashed lines); all energies are relative to the top of the quasiparticle valence bands.

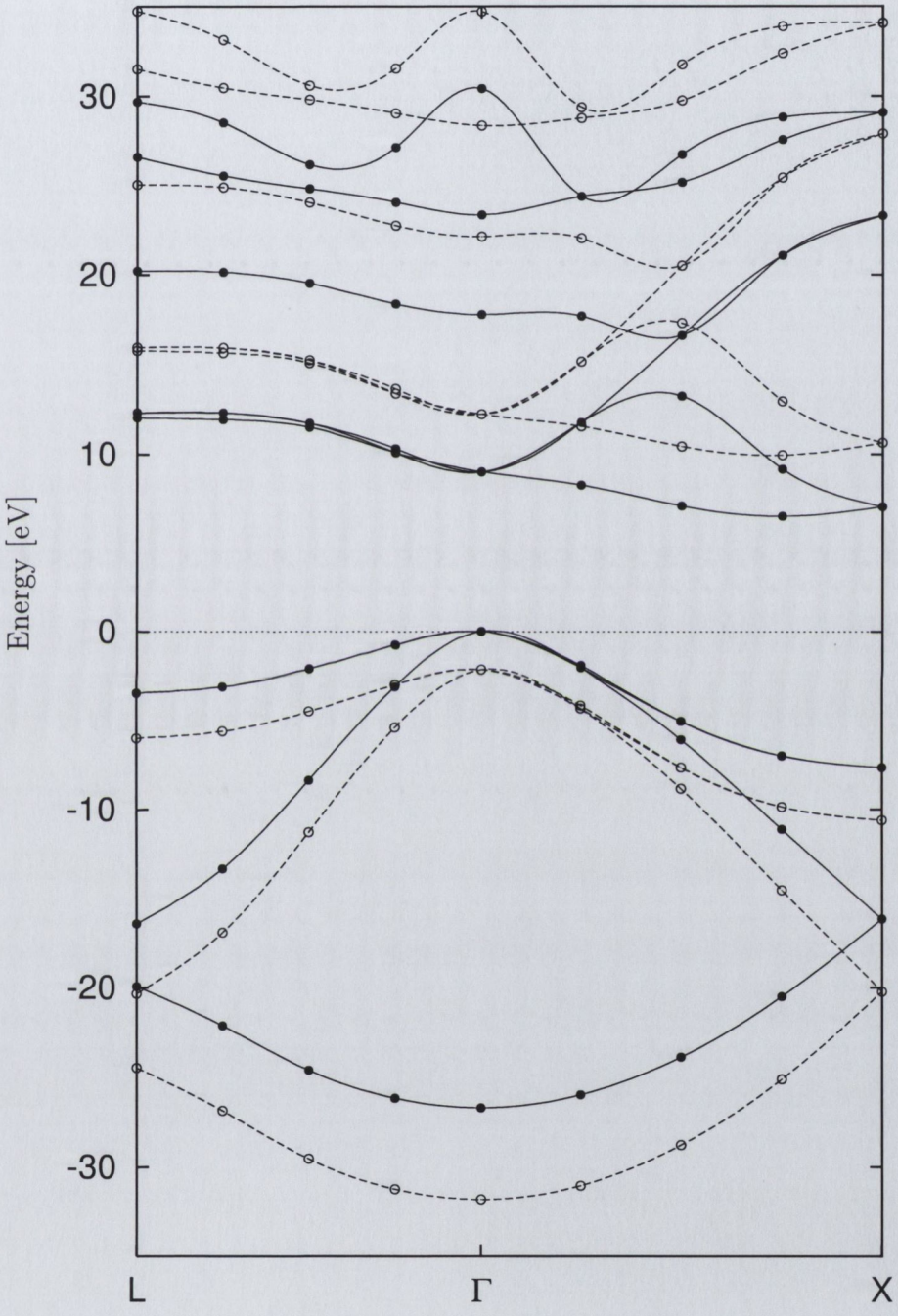


Figure 5-5: Quasiparticle bandstructure of diamond (solid lines) and the initial Hartree-Fock bandstructure (dashed lines); all energies are relative to the top of the quasiparticle valence bands.

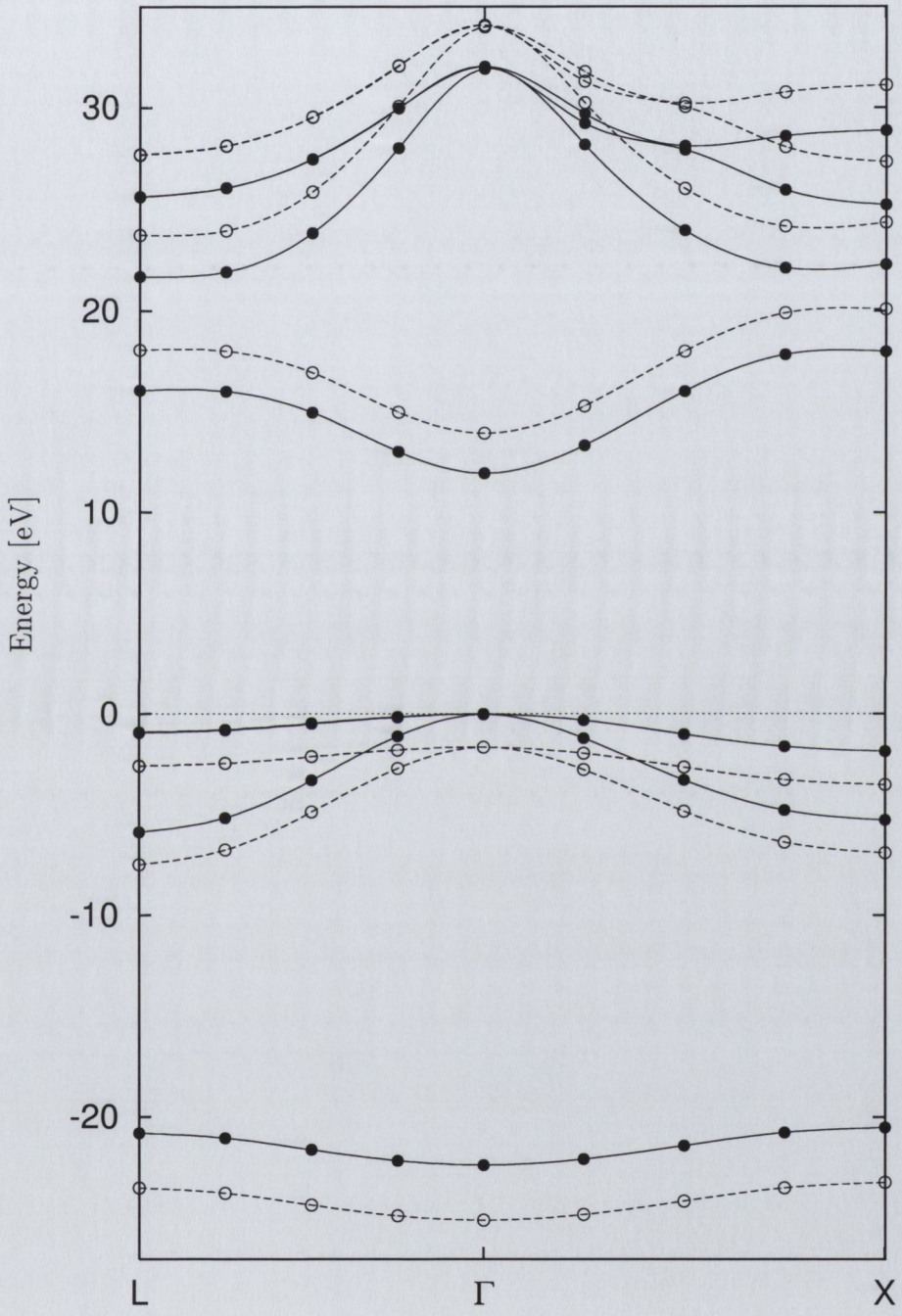


Figure 5-6: Quasiparticle bandstructure of MgO (solid lines) and the initial Hartree-Fock bandstructure (dashed lines); all energies are relative to the top of the quasiparticle valence bands.

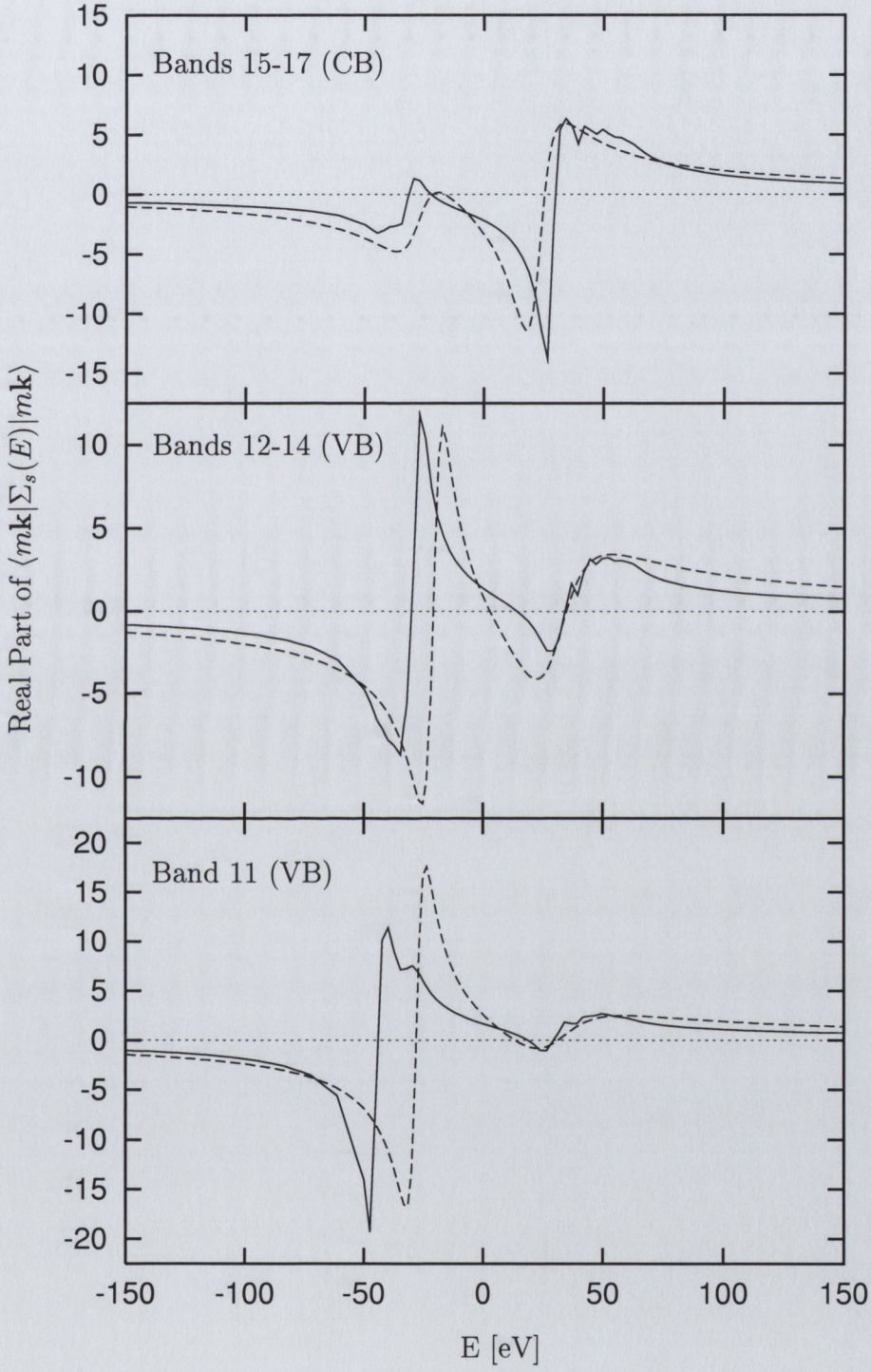


Figure 5-7: Real part of selected matrix elements of the correlation self-energy $\langle m\mathbf{k}|\Sigma_s(E)|m\mathbf{k}\rangle$ of silicon, for the bands indicated, at Γ. Full lines are the calculations of this thesis; dashed lines are the data of [67, 20]. A broadening parameter of 2 eV has been used in the calculated spectra.

Chapter 6

Surface Optical Properties

> Science may be described as the art of
> systematic oversimplification
> — *Karl Popper*

In Section 3.4, the discrete dipole model (DDM) was introduced as a suitable technique for describing the optical properties of surfaces, and a brief outline of the theoretical basis was given. Owing to its simplicity, the DDM presents a tractable framework that allows thick slabs of semiconductor ($\sim$80 Å) to be treated at a relatively low computational cost, such that surfaces with large unit cells may be examined, and due to the explicit microscopic modelling of the response, the influence of complicated surface reconstructions may be treated in a realistic manner. The central features of the model are the direct inclusion of local field effects and a discretized inductive process, where polarizable units are localized at discrete sites throughout the system. In the symmetric bulk system, the physics is simply described by the Clausius-Mossotti relationship [equation (3.46)]; at the surface, the local fields are responsible for some interesting phenomena.

The main focus of this chapter is the relationship between the local field effects and the surface optical properties, in particular its relevance to the experimental technique of reflectance anisotropy spectroscopy (RAS). The discrete dipole model has been used previously to obtain optical spectra in good qualitative agreement with experimental data [48, 74, 47, 75], and has been used to investigate the mag-

nitude of the local field effect on various linear and nonlinear calculations [76, 77]. Mochán and Barrera [47], for example, found a good comparison with the reflectance anisotropy spectra of Si and Ge(110) surfaces. However, they used the Clausius-Mossotti relation [equation (3.46)] to obtain a polarizability for each site from the *experimental* dielectric function, thus effectively including the self-interaction of each polarizable entity, the effects of exchange and correlation, and any non-dipolar interactions that may be contributing to the dielectric response. It is therefore unsurprising that such a good agreement is obtained without the need for scaling or shifts in energy in the calculated spectra. Nevertheless, it prevents the model from being used to relate features of an RA spectrum to surface reconstructions, dimers, *etc.* The approach that has been adopted in this thesis has this ability, but crucially, it properly accounts for the local field effects. Therefore, the approach is expected to yield at least qualitative information about the influence of the crystal termination on the RA spectra.

The outline of this chapter is as follows. In the initial section, the basic theory behind the model, which was briefly described in Section 3.4, is reconsidered in more detail. Following this discussion are some calculations of the RAS for several vicinal surfaces of silicon, along with an examination of the spatial variation in the optical response of the slab. Finally, these results are discussed in relation to their significance to calculations of the RAS and to the discrete dipole model in general.

6.1 Discrete Dipole Model: Formalism

The theoretical basis of the discrete dipole model (DDM) was previously formulated by equations (3.43) and (3.44). The dipole moment $\mathbf{p}$ at a site i is assumed to be induced via a local polarizability α_i and the local field $\mathbf{E}_i^{\text{loc}}$, such that

$$\mathbf{p}_i = \alpha_i \mathbf{E}_i^{\text{loc}} \tag{6.1}$$

$$= \alpha_i [\mathbf{E}_i^{\text{ext}} + \mathbf{E}_i^{\text{dip}}], \tag{6.2}$$

where $\mathbf{E}_i^{\text{dip}}$ represents the electric field radiated by the remaining dipoles in the system. The explicit form for the dipolar field may be derived using the Hertz vector formalism [46, 78]; for a harmonic driving field, this is given by[1]

$$\begin{aligned} \mathbf{E}_i^{\text{dip}} &= \sum_{j \neq i} [\nabla \nabla^{\mathrm{T}} + k^2] \frac{e^{ik|\mathbf{r}_i - \mathbf{r}_j|}}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j|} \mathbf{p}_j \\ &= \sum_{j \neq i} \mathbf{T}_{ij} \mathbf{p}_j, \end{aligned} \tag{6.3}$$

where a common frequency dependent term has been omitted. This explains the origin of the 'transfer tensor' introduced in Section 3.4. In the particular scheme adopted here, we assume that the speed of light $c \to \infty$, which implies $k \to 0$. Equation (6.3) may be evaluated to yield the expressions known as dipole sums, which constitute the central calculations of the discrete dipole model. The techniques for carrying out the dipole sums were originally devised by Ewald for a bulk system of dipoles and extended to treat slabs by Nijboer and de Wette [79] via a planewise approach. The actual means of solving the coupled set of equations equation (6.3) for a slab is discussed in more detail in Ref. [46]; briefly, the solutions are obtained by a normal mode analysis, and these modes (termed 'dipole waves') are allowed to couple to an external field if the frequency matches the eigenfrequencies of the modes.

An important question which has been neglected up to this point concerns the choice of the polarizable unit. In principle, the unit should be chosen to be as small as possible in order to properly account for the spatial variations in the local fields. However, in their calculation of the reflectance anisotropy spectrum of the Ge(110) surface, Mochán and Barrera [47] found that better agreement with experiment could be reached if the polarizable unit was ascribed to each set of four tetrahedrally arranged bonds, rather than one unit to each atom. Clearly, the results obtained from the discrete dipole model are sensitive to this choice, particularly since the possibility of nonlocality in the polarizability itself has been neglected, as discussed in Section 3.4. In this thesis, the treatment of Herrendörfer [46] is followed, by which

[1] S.I. units are used for the expressions in this chapter, for consistency with Ref. [46]

the covalent bond is used as the polarizable unit for the system being studied, namely silicon. Since the valence electrons are concentrated in the bond region, and hence also the fluctuations in the electron density that are important for optical properties, it is believed that the polarizable bond model is a reasonable *ansatz*.

In order to have a widely applicable model, the site polarizabilities were derived from quantum mechanical calculations on clusters [46], in which the response of the electronic charge density of a silicon cluster under the influence of a static electric field is investigated. This procedure allows surface and bulk polarizabilities to be determined separately. It should be noted, however, that the same bare polarizabilities (i.e. with local field effects excluded) have been used throughout the slab. Spatial variations in the polarization may therefore only arise from the truncation of lattice planes at the surface. However, the anisotropy of the bond is taken into account through an axial polarizability tensor, and quantum mechanical corrections to the nearest neighbour dipole-dipole interaction have been explicitly included, again closely following the work of Herrendörfer [46].

The discrete dipole model is particularly instructive in the study of surface optical properties when it is used in conjunction with the experimental technique of reflectance anisotropy (RA) spectroscopy [80, 81]. RA spectroscopy is a non-destructive technique that is very sensitive to the spatial inhomogeneity of the surface and thus presents a useful tool for investigating the local surface structure, or for identifying, for example, particular reconstructions that are favoured during a chemical process. It is well known [47, 48] that local-field effects have a significant influence on surface spectra, and are understood to appear in the energy region above the optical gap where no optically active surface states are expected (such as those associated with adsorbates or reconstructions). However, it is not clear whether it is possible to assign peaks in the RA spectrum to differences between surface and bulk polarizabilities or to local field effects arising from the lattice truncation. The latter phenomenon is termed 'intrinsic' anisotropy, as it is exclusively due to this reduction of symmetry at the surface.

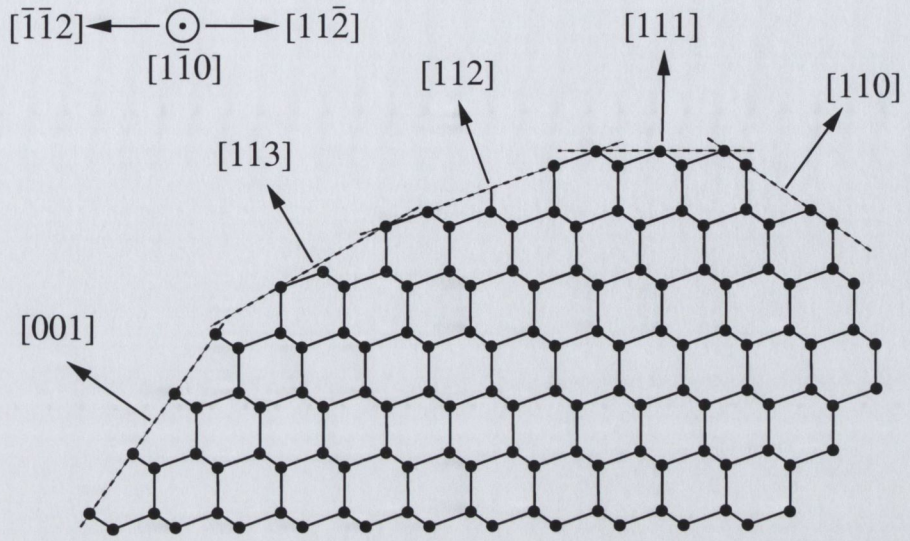


Figure 6-1: Side view of the diamond structure along the $[1\bar{1}0]$ direction. The surfaces being studied are indicated with arrows perpendicular to those surfaces.

6.2 Calculations

In order to investigate the importance of the intrinsic anisotropy to the RA spectra, the above discrete dipole formalism was applied to the calculation of the RA spectra of five bulk-terminated silicon surfaces: (001), (113), (112), (111) and (110), as shown in Figure 6-1. Reflectance anisotropy is based on the idea that the isotropic bulk contribution to the reflectivity can be excluded by measuring the difference between the reflectivities obtained for perpendicular directions of polarization, thus exposing that part of the response due to surface anisotropy. For the calculation we use the McIntyre-Aspnes model [82] to obtain the difference in reflectivity of a 3-layer system (vacuum-surface-bulk) from the Fresnel 2-layer system (bulk and vacuum only) in terms of the surface excess susceptibility χ_{xs} [83], (where $\chi_{xs} = \chi_s - \chi_b$):

$$\frac{\Delta R}{R} = 4k_z d \ \mathrm{Im}\left(\frac{\chi^i_{xs}\chi^r_b - \chi^r_{xs}\chi^i_b}{|\chi_b|^2}\right). \tag{6.4}$$

Surface and bulk contributions are denoted by s and b respectively. The susceptibilities χ_s and χ_b are found by summing over the dipole amplitudes from the solution

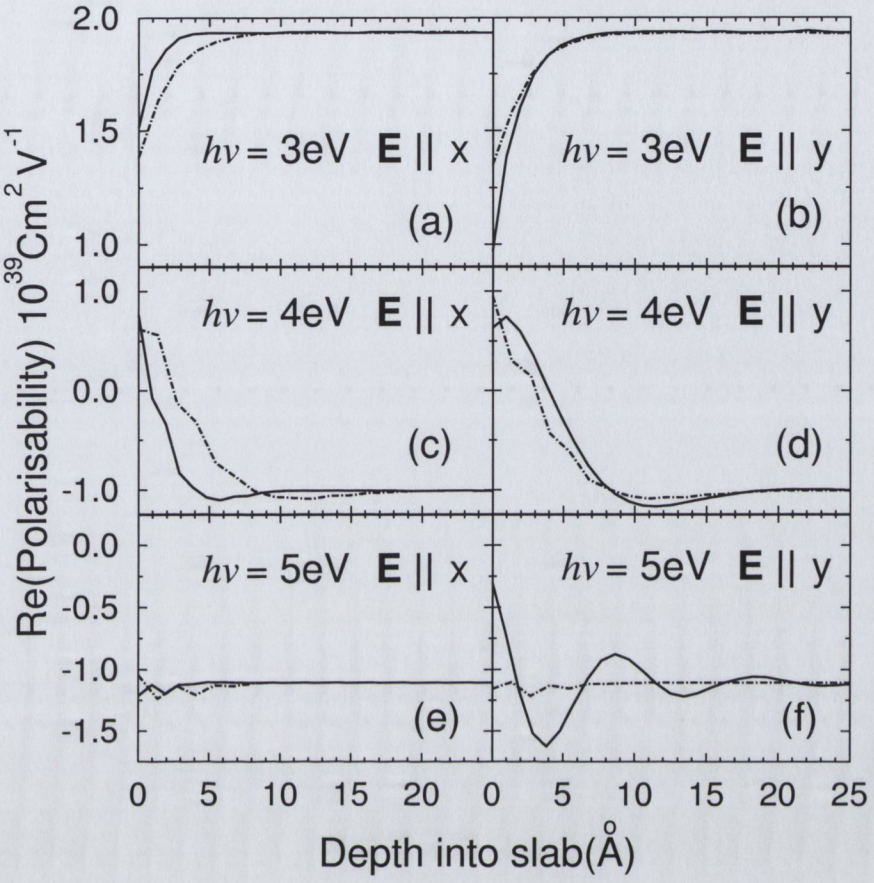


Figure 6-2: Real part of the dressed polarisability at the surface of a slab with a (001) surface (dot-dashed lines) and a (110) surface (solid lines). The energy and direction of the incident electric field with respect to the crystal axes are given in each panel.

of equation (6.1) in the appropriate spatial regions, i.e.:

$$\chi_b(\omega) = \frac{1}{\epsilon_0 \Omega_b} \sum_{i=1}^{4} \frac{\mathbf{p}_i(\omega)}{|\mathbf{E}^{ext}(\omega)|} \tag{6.5}$$

$$\chi_s(\omega) = \frac{1}{\epsilon_0 \Omega_s} \sum_{i=1}^{N_s} \frac{\mathbf{p}_i(\omega)}{|\mathbf{E}^{ext}(\omega)|}, \tag{6.6}$$

where Ω_b and Ω_s denote the volume of the bulk and surface unit cells, respectively. Finally, the reflectance anisotropy is calculated from

$$\mathrm{RA}(\omega) = \frac{\Delta R_x}{R} - \frac{\Delta R_y}{R}. \tag{6.7}$$

Note that in the present study we assign x to the $[1\bar{1}0]$ direction which is common to each surface, while y is directed along the dotted lines of Figure 6-1, in a right handed frame.

6.3 Results and Discussion

Results of our calculations [54] on the five surfaces are now presented. In Figure 6-2 we plot the 'dressed' polarizability, which is the site polarizability with the effects of the local fields included, for the (110) and (001) surfaces, as a function of the depth into the slab. In reality, the orientation of the anisotropic bond polarizability alternates from one sublayer to the next, as does the dressed polarizability, and therefore an average value over sets of four bonds attached to an atom in each layer is presented in order to provide a clear description of the variation with position. Note the oscillations in polarizability that persist for energies in the 4.5-6.0 eV range; this phenomenon is found to exist in the (110), (112) and (113) surfaces whereas the (001) and (111) calculations demonstrate more bulk-like behaviour for depths greater than 10 Å. It is interesting to note that the surface polarization penetrates quite deeply into the solid ($\sim$20Å), which suggests that a prohibitively large surface unit cell may be needed for calculations that aim to include local field effects at some surfaces.

The calculated $\Delta R/R$ 'spectra' are shown in Figures 6-3 [I(a)] and [I(b)], for each of the five surfaces. It is immediately evident that there are strong similarities between the spectra for the different surfaces, but that the RA spectra themselves (Figure 6-3 [I(c)]) are markedly different. Several features may be noted; for instance, the (001) and (110) surfaces have minima in $\Delta R/R$ at the same energy for both x and y directions, while the (112) and (113) surfaces have minima in $\Delta R/R$ that depend on the crystal azimuth. The derivative-like lineshape in the RA spectra for the (112) surface can be attributed to the small shift in energy of the main $\Delta R_y/R$ with respect to the peak in $\Delta R_x/R$ at 3.95eV.

Figure 6-3 [II] shows a comparison between experimental and calculated RA

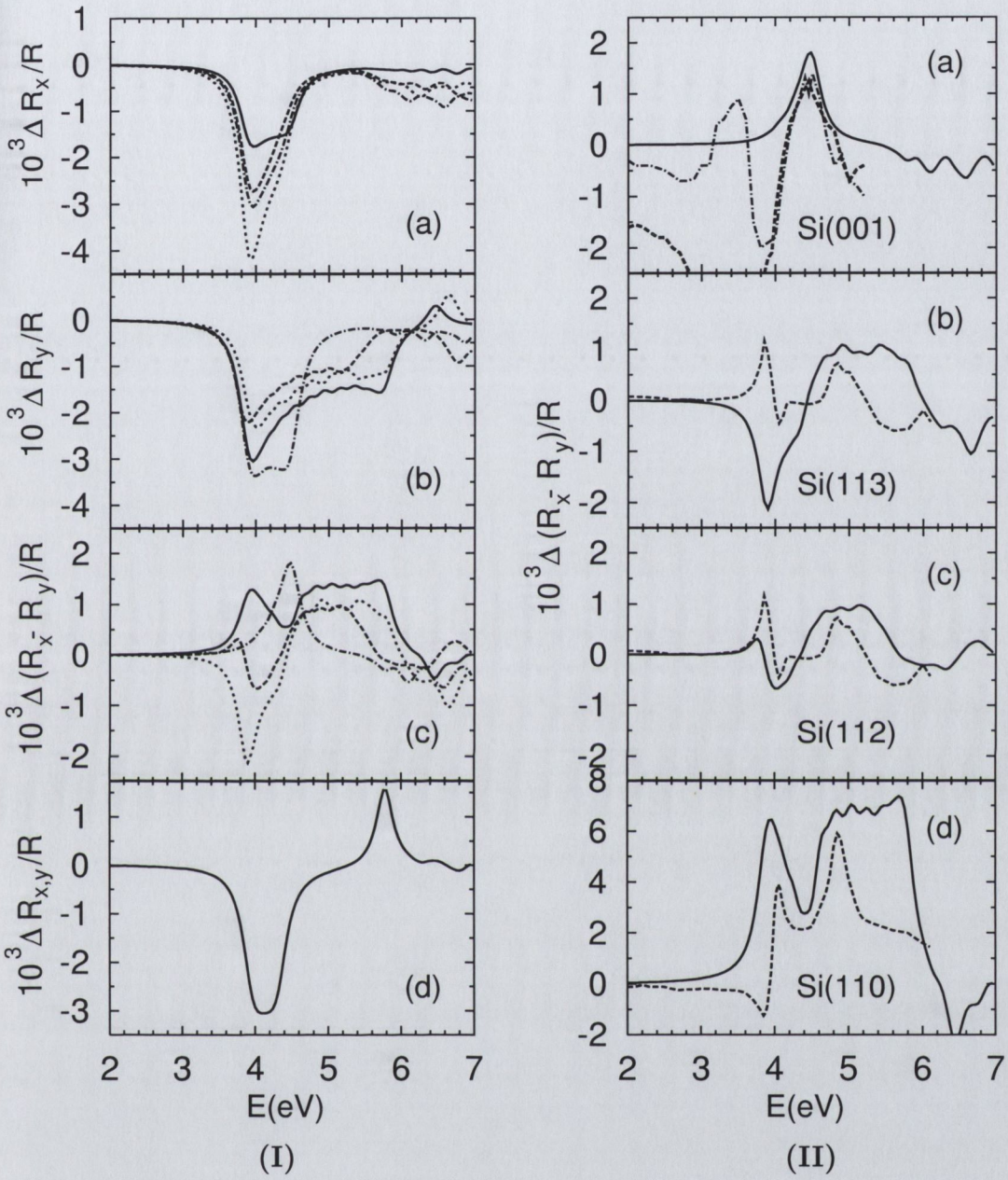


Figure 6-3: (I) (a) $\Delta R_x/R$, (b) $\Delta R_y/R$ and (c) RA spectra calculated using the discrete dipole model (DDM) for the surfaces (001) dot-dashed, (113) dotted, (112) dashed and (110) solid lines, (d) $\Delta R_x/R$ for the (111) surface.
Note $\Delta R_x/R = \Delta R_y/R$ from symmetry considerations for (111) surface.
(II) Calculated RA spectra using the DDM (solid lines) compared with experimental spectra (dashed lines). Note that experimental spectra are *shifted* in energy [+0.6 eV for (110) and (113); -0.3 eV for (001)], and that calculated spectra have been *reduced* by the factors shown in brackets below:
(a) Si(001) (DDM$\times\frac{1}{10}$), Si(001)-(1×2) (dashed line) [81] and Si(001)-(1×2):As (dot-dashed line) [84]
(b) Si(113) (DDM$\times\frac{1}{10}$) and Si(113)-H [85]
(c) Si(112) (DDM$\times\frac{1}{10}$) and Si(113)-H [85]
(d) Si(110) (DDM$\times\frac{1}{2}$) and Si(110)-H [85].

spectra. There is a reasonable agreement between theory and experiment for the (110) and (001) surfaces; it is found however that the best agreement to the (113) data results from a (112) surface calculation. This unusual finding may be attributed either to a possible facetting of the (113) surface to a (112) surface with narrow (111) terraces, or of course, to the inherent limitations and approximations of the discrete dipole model. It should be noted that the calculated intensities have been scaled by the factors given in the figure caption, and that the experimental spectra have been slightly shifted in energy (for (110) and (113) +0.6 eV; for (100) −0.3 eV) Similiar shifts in energy have been used in discrete dipole calculations by other authors [49] to account for exchange and correlation effects in the polarizability, which is consistent with the observation that a small change in polarizability (∼3%) is sufficient to shift the calculated spectra by ∼0.6 eV.

The response of the slab to an external, uniform field polarized parallel to the surface is characterized by the fact that the surface resonates at a slightly higher frequency than that of the underlying bulk. This is simply because the local field at the surface is weaker in magnitude than in the bulk. Local fields tend to reinforce an external field polarized parallel to the surface, but due to the lack of neighbouring planes above the surface, the surface local field is reduced. Therefore, minor differences occur between the surface and bulk dressed polarizabilities around the bulk resonance frequency which extend down to the first 5–10 atomic layers of the solid. Uwai and Kobayashi [86] have also noted long range surface polarization effects in the RA spectra of GaAs, which they attribute to the effect of crystal termination on the electronic wavefunction. Del Sole and Onida [87] have subsequently dismissed these findings, and based on realistic one-electron calculations of the optical properties of GaAs(110) and GaAs(100), have stated that surface effects are localized in a few monolayers below the surface. Their scheme does not, however, feature a description of the surface local field effect (as described by the discrete dipole model), and they suggest that such effects may be responsible for *bulk-like* lineshapes in the optical spectra.

Striking similarities have been reported [85] between the experimental bulk di-

electric function $\varepsilon_2(\omega)$ and the RA spectrum for the (110) H-terminated surface of silicon, and similiar comparisons have been made between the RAS of the (113) surface and the energy derivative of the reflectivity.[2] As we have shown in Figure 6-3 [II](d), the discrete dipole model succeeds in predicting a double-peaked structure in the RAS for the (110) surface, in agreement with the results of other authors for the Si(110) and Ge(110) surfaces [47]. In addition, the model also predicts the existence of derivative lineshapes in the (113) surface, purely from a consideration of the surface local field effect. Of course, the description of the electronic structure of the surface and bulk is greatly oversimplified in the DDM — in particular by the assumption that the bare polarizability is a local quantity — and it is therefore not surprising that a quantitative agreement with experiment is not achieved. Nevertheless, it is evident from the simplicity of the mathematical scheme presented by equations (6.1–6.3) that the discrete dipole model offers clear intuitive guidance for interpreting features in the RA spectra, and may perhaps be used most successfully in conjunction with more formal quantum-mechanical models. For example, the influence of the 'intrinsic' anisotropy on the optical spectra may be evaluated quite easily within the DDM scheme, as shown in the present study. Recent indications [49] have shown that by accounting for the nonlocality of the polarizability (as part of the discrete cellular scheme from Section 3.4), even excellent *quantitative* results of RA spectra may be derived from a discrete model of optical properties.

[2]Note that the optical reflectivity is also, for the most part, a property of the bulk.

Chapter 7

Conclusions

FELIX QUI POSUIT RERUM COGNOSCERE CAUSAS
Happy is he who has been able to understand
the causes of things — *Virgil, Georgics, II*

The thesis presented here deals with the optical and dielectric properties of solids, and looks at the way in which an investigation into these properties provides insight into the microscopic processes that occur within the interacting system. In carrying out this study, one central component has been focussed on throughout, namely, $\varepsilon(\mathbf{q},\omega)$, the wavevector and frequency dependent longitudinal dielectric function. Several aspects of the calculation, application and understanding of $\varepsilon(\mathbf{q},\omega)$ have been discussed in the thesis: the accurate computation of $\varepsilon(\mathbf{q},\omega)$ within the localized Gaussian orbital scheme; the analysis of the static and nonlocal dielectric response, via the matrix representation, $\varepsilon_{\mathbf{GG'}}(\mathbf{q},\omega \to 0)$, and the concept of dielectric bandstructure; the application of the dielectric function to the calculation of bulk and surface optical properties; the use of the dielectric function in the calculation of the self-energy and quasiparticle bandstructure as part of a GWA+Hartree-Fock scheme. The important conclusions derived from these calculations will now be briefly re-examined.

7.1 Gaussian orbital framework

The possibility of elucidating the optical and dielectric properties of complex systems — namely, surfaces and *d* type materials — motivated a detailed look at the practical merits of the Gaussian orbital scheme. The reduced size of the basis set involved in the LCAO scheme, as distinct from the more conventional plane-wave approaches, is well known to offer distinct advantages for examining the *physical* properties of such systems. However, their suitability for calculating accurate optical and dielectric properties has not been widely examined. As described in Section 3.1, it was found that the standard available basis sets, which consist entirely of atom-centred orbitals, are sufficently flexible for calculating the basic dielectric properties of the system, as given by the RPA dielectric function of Ehrenreich and Cohen (c.f. Figure 3.1.3). Fundamental deficiencies in the RPA approximation were cited as the source of quantitative differences between the experimentally obtained dielectric function and the calculated result, and methods for improving the agreement with experiment were outlined — most notably, the influence of the electron-hole interaction on the optical spectra was discussed. However, it was found that the limited effective number of conduction band states obtained using standard basis sets (as required in the calculation of the dielectric response function) cannot adequately describe the influence of spatial inhomogeneity on the system response, as manifested by a local variation in the polarizing electric field inside the cell.

An important conclusion drawn from this section, therefore, is the necessity of using specially adapted basis sets which include orbitals situated at interstitial sites in the cell. Using the free-electron description of the eigenvalue *vs.* band index relation as a guideline, the number of effective conduction band states can be increased by adding more functions throughout the cell, although specific procedures must be followed when choosing suitable Gaussian exponents and radial forms. The use of these extended basis sets in calculating the optical spectra and dielectric constants allow the influence of local field effects to be precisely and consistently accounted for. However, it was also found that filling the cell with interstitial orbitals increases the danger of quasi-linear overlap of Bloch functions, which has the undesirable ef-

fect of introducing numerical problems into the Hartree-Fock or DFT calculation. Ultimately, therefore, the Gaussian orbital scheme (at least, in its original form) tends to be impractical for highly accurate dielectric matrix calculations.

With regard to self-energy calculations, the Gaussian orbital scheme was found to be quite satisfactory in determining the quasiparticle band-gaps, as any lack of convergence in the dielectric matrix due to limitations of the basis set has only a relatively small influence on the self-energy correction. Similar observations are found in the work of Rohlfing [21], although again, there is a limit on the accuracy that can be obtained in comparision with large plane-wave calculations [20].

7.2 Dielectric properties

As stated above, incorporation of interstitial functions into the Gaussian basis set enables the dielectric matrix to be accurately calculated, and hence the influence of local field effects on dielectric and optical properties can be examined in detail. In this thesis, local field effects are investigated via two distinct, albeit related, schemes: the conventional "continuum" description of bulk response, as represented by the dielectric matrix, and the "discrete" picture offered by the discrete dipole model, which is used to examine the optical properties of surfaces. In order to examine the magnitude and spatial dependence of these local fields in bulk calculations, the method of dielectric bandstructure is applied to several covalent and ionic materials, namely, Si, diamond, GaN, NaCl and MgO. By offering a graphical representation of the dielectric matrix, the dielectric bandstructure technique illustrates both the absolute magnitude of the dielectric screening (found to be largest in the covalent semiconductors), as well as the influence of crystal inhomogeneity that gives rise to the local fields. With regard to the spatial distribution of the microscopic response, it is instructive to decompose the static dielectric matrix into a set of screening modes (eigenpotentials), which closely relate to the degree of ionicity of the material. The primary screening modes of covalent materials are found to be dipolar in nature, while those of ionic materials are predominantly monopolar. To some extent, this

correlates with the *ansatz* made in the discrete dipole model — that the principal polarizable unit in a semiconductor is a dipole — but it also emphasizes the fact that other screening mechanisms are not being described in the model.

In spite of this connection, the two methods are quite exclusive in their practical applicability. This arises as a result of different treatments of nonlocal phenomena in the two models of dielectric response. The dielectric matrix scheme incorporates two kinds of nonlocality, quantum mechanical (inter- and intra-site polarizability) and electromagnetic (local field effects); the discrete dipole model, focusses solely on electromagnetic nonlocality, which is expected to be the more influential phenomenon involved in long-range (inter-cellular) dielectric response. For this reason the latter method is suitable for modelling experimental techniques such as reflectance anisotropy spectroscopy. The RA spectra for several vicinal surfaces of silicon were calculated using the model, and it was found that some of the important features in the experimental spectra could be attributed to the 'intrinsic' anisotropy, i.e., the anisotropy of the optical response that solely arises from the termination of the periodic lattice at the surface, as distinct from features related to surface reconstructions or relaxations. However, spectra calculated using the model (in its present form) require significant scaling and shifting to achieve reasonable comparison with the experimental results. This feature may be attributed to the poor description of the (nonlocal) cell polarizability, which, as illustrated by the dielectric matrix formalism, is considerably more complex than the simple bond-centred dipole model.

7.3 Quasiparticle calculations

A challenging section of this thesis was the computation of the quasiparticle bandstructure, which is obtained from the calculation of the electron and hole self-energy within the *GW* approximation. Instead of adopting the conventional approach which utilizes a DFT-LDA bandstructure as a first approximation in calculating the self-energy, an alternative approach is presented, whereby the self-energy is calculated starting from a Hartree-Fock bandstructure. As explained in Section 2.4, the GWA

may be regarded as a Hartree-Fock approximation but with the effects of dynamical (RPA) screening included. This is achieved with the use of a dielectric matrix formalism; the physical description of the microscopic and nonlocal properties of the screening in self-energy calculations is therefore the same as that provided by the dielectric bandstructure analysis of Chapter 4.

The main result obtained in the calculations is a large reduction in the (grossly exaggerated) Hartree-Fock band-gap for all the materials studied. However, a good quantitative agreement with the experimental data — and hence, the published GWA-DFT data — is not found. In accounting for this discrepancy, numerical problems such a possible lack of convergence with respect to basis set size may be discounted, as may errors resulting from the assumption that the HF eigenvectors are quite similar to their GWA counterparts. It was inferred, therefore, that the use of the Hartree-Fock eigenvalues in constructing the initial Green's function G_0 and screened Coulomb interaction W_0 was the principal reason for this discrepancy as they do, in fact, differ considerably from the DFT and experimental eigenvalues. In particular, the dynamical screening is poorly described when Hartree-Fock eigenvalues are used; calculated dielectric constants and plasmon-pole frequencies are found to be greatly underestimated and overestimated, respectively. Since the basic GWA-HF and GWA-DFT methods are perturbative (i.e., not self-consistent) schemes, the initial choice of electronic structure is extremely important. For the simple *sp*-type semiconductors and insulators examined in this work, the GWA-DFT framework therefore presents a simpler and better solution; the GWA-HF scheme may, however, prove to be useful in the study of more complex systems, such as transition metal oxides. Both GWA-HF and GWA-DFT schemes may, of course, be extended to include a form of self-consistency although such calculations become considerably more difficult. Such a scheme would be expected to produce results in agreement with published self-consistent GWA-DFT results, which (for silicon) are found to overestimate the experimental band-gaps and also, therefore, the perturbative GWA-DFT results.

The work presented here demonstrates that the dielectric function is a powerful

tool for understanding the microscopic processes involved in optical and dielectric response, applicable to the study of many-body interactions. In the framework of a localized Gaussian orbital basis, it is hoped that the work in this thesis will be extended towards the examination of these phenomena at surfaces and in $3d$-type systems.

7.4 Outlook

As with any large-scale computational project, there needs to be a balance between the development of new codes and the improvement of existing ones. Taking the results of this thesis into account, much of the groundwork for accurately calculating the dielectric function has been accomplished. One aspect of the calculations which posed several problems was the use of the Gaussian orbital basis itself. It would, for example, be useful to develop new methods for solving the problems associated with the size of the basis set in a more efficient manner to that presented.

With regard to the existing computer code, many improvements may be made; in particular, the restrictions imposed by memory limitations may be removed by utilizing dynamic memory allocation (see Appendix A.2.2). Considerable improvements in computational time may be achieved by making use of symmetry properties during the calculation of the dielectric matrix elements and the Gaussian orbital integrals (Appendix B). Work also needs to be done in improving the treatment of the limiting form of the dielectric matrix $\varepsilon_{\mathbf{GG'}}(\mathbf{q} \to 0, \omega)$.

Some interesting possibilities for future work have already been discussed in the text. The importance of pursuing a self-consistent scheme for evaluating self-energies was mentioned in Section 5.5, and some extensions to the RPA dielectric function were described in Section 3.3. Of particular interest here is the Bethe-Salpeter equation (equations 3.41–3.42), which could be used to incorporate excitonic effects in the optical spectra, and follows naturally from the quasiparticle calculations.

Improvements to the discrete dipole model have been considered in the PhD thesis of Herrendörfer [46], and the influence of nonlocality in the polarizability has

been investigated by Boeij and Wijers [88, 49]. The connection between the discrete dipole model and the continuum model of dielectric response has been addressed by Nicastro and Patterson [52], and further work along this line of research would be interesting. In addition, a direct study of the response function relating the change in charge density to the external field (i.e., not the total field) is ongoing.

Appendix A

Computer Code

This section of the thesis focusses on the design and implementation of the FORTRAN77 computer code that was developed for the calculation of the various quantities which have been presented in the main text, such as the dielectric bandstructures of Chapter 4 and the self-energy of Chapter 5. The appendix is divided into two sections: Section A.1 gives an overview of how to run the program in conjunction with the CRYSTAL95 code, and documents the format of the input and output; while Section A.2 describes the actual structure of the code and details some of the algorithms that comprise it. There exists great potential for further development, e.g. self-consistent GW, implementation of the Bethe-Salpeter equation, etc., and the latter sections is intended as a guideline for such future work.

A.1 User Guide

The main calculations are carried out by the `COMBO` executable, which is controlled via *standard input*. The input file contains the information requesting the particular tasks as well as specifying the run-time parameters. Output is directed to the *standard output* stream and to several task-dependent files that are created locally, while error messages are written to *standard error*. The command line takes the form:

```
% ./COMBO < [input-file] > [output-file]
```

Auxiliary input is provided in the form of the two files `Xtal.dat` and `fort.9`[1] which must be generated prior to the main calculation by using the CRYSTAL95 code. All crystal structure and basis set information is introduced at this preliminary stage, as discussed in the following section.

A.1.1 Preliminary tasks

This section begins with a caveat: the quality of the final results ultimately depends on the quality of the basis set. Many of the problems which arise during a calculation can very often be traced back to inadequacies in the Gaussian orbital basis. It is vital to carefully test the capabilities and characteristics of a particular set; for instance, as discussed in Section 4.1.2, accurate optical properties often require an extended set of conduction bands to ensure a reasonably converged sum over states. This means that many of the standard basis sets developed by the quantum chemistry community are unsuitable for this type of calculation and therefore require some modification. A brief overview of this optimization process is presented in Appendix D along with the documentation of several highly accurate basis sets.

Assuming that a sufficiently accurate Gaussian orbital basis has been developed, the first task is to generate the portable properties file `Xtal.dat`. This is achieved by running the `integrals` and `scf` (or alternatively `scfdir`) parts of CRYSTAL95 using the input file `dat.inp`, which should contain the usual CRYSTAL95 basis set and lattice information. Standard output must be directed into a new file, entitled `dat.out`. Following this, the CRYSTAL95 `properties` code should be run as normal using the the following keywords in the input file (called `dat.pro` below): `NEWK KNETOUT EXTPRT COORPRT`. The new **k**-point mesh generated by the `NEWK` option should have the same Monkhorst-Pack shrinking factors as specified in the `scf` part. Output from `properties` is then appended to `dat.out`.

[1]The default names for FORTRAN units are given by `fort.X`, where `X` indicates the unit number. On some systems (HP) this is replaced by `ftn.X`. In addition, the `fort.9` properties file may be renamed `fort.99` on some Linux systems.

The command line should look as follows:

```
% integrals < dat.inp > dat.out
% scf >> dat.out
% properties < dat.pro >> dat.out
```

This process results in the creation of several new FORTRAN files. To summarise, the files in existence should include

Filename	Content	Created by	Format
`fort.9`	Density matrix input to `properties`	NEWK	binary
`fort.10`	Eigenvectors in the IBZ	SCF	binary
`fort.30`	Eigenvalues in the IBZ	NEWK	binary
`fort.33`	Lattice data and symmetry operators	EXTPRT	ascii
`fort.34`	Accurate atomic positions	COORPRT	ascii
`dat.inp`	Gaussian orbital basis set	User	ascii
`dat.out`	CRYSTAL95 standard output	`integrals scf`, `properties`	ascii

Finally the short executable `init_data` is run, which gathers all the relevant data from each of these files and concatenates them into the portable ascii file `Xtal.dat`:

```
% ./init_data
```

The two files `fort.9` and `Xtal.dat` now contain all the material specific data needed for execution of the main optical properties code, and should be stored for subsequent use.

A.1.2 Input data structure: General parameters

The program's input file is characterised by an initial block of general parameters followed by a block of various optional tasks. Each block is arranged in lines of *records*, where a record, denoted by the symbol •, comprises a list of data on a single line. The symbol ○ indicates an optional record which is inserted when certain conditions (detailed in the tables) are met. The initial input block is specified as follows:

Record	Format	Options
•TITLE	`character*40`	
•MATSYM	`character*4`	`HALF, FULL`
•BRAVAIS	`character*3`	`FCC, SC, BCC`
•SHR SHG ITOL	`integer` (×3)	
•ISKI ISKH ISQ ISKM NPRO	`integer` (×5)	
If NPRO > 0 then insert record:		
∘(PRO(N), N = 1,NPRO)	`integer` (×NPRO)	

TITLE User comments: description of basis set, run-time parameters, etc.

MATSYM Use of `HALF` specifies the calculation of only the lower triangular section of the (Hermitian) dielectric matrix, which brings a considerable improvement in speed. `FULL` forces all the elements of the dielectric matrix to be calculated. This option is mainly provided for debugging purposes.

BRAVAIS Specifies the type of Bravais lattice; currently only `FCC, SC, BCC` lattice types are supported. This flag influences the special point generation algorithm as well as the form of the auxiliary parameter in equation (C.21).

SHR The number of shells summed over in real space to ensure correct normalization of the Bloch functions. **SHR** = 1 is usually sufficient.

SHG The number of **G**-vectors that determine the size of the dielectric matrix. This should be set to a value that defines a complete 'star' of **G** vectors. Setting **SHG** = 1 or 0 is equivalent to neglecting local fields; the special case **SHG** = 0 forces the optical matrix elements to be saved into memory, thus allowing a rapid calculation of the frequency dependent dielectric function.

ITOL The one-electron integrals may be evaluated more efficiently by avoiding the calculation of those terms which only contribute a negligible amount towards

the total integral. Integrals involving shells of Gaussian orbitals whose largest overlap is less than the cutoff $10^{-\mathbf{ITOL}}$ are neglected. **ITOL** = 8 is generally sufficient.

NPRO Number of printing options requested. This flag relates to *run-time* data only; printing of information from `Xtal.dat` may be invoked using the `CHECKDAT` input block(see next section). Values of **PRO(i)** currently supported are:

0 Prints full list of **G**-vector coordinates.

1 Printing of dielectric matrices

2 Printing of subroutine timing information

The remaining input flags in the initial input block relate to the various 'shrinking factors' of the reciprocal space points. The shrinking factor relates the oblique coordinates $[i1\, i2\, i3]$ of a vector $\mathbf{k}$ in reciprocal space to the cartesian coordinates via the expression

$$\mathbf{k} \equiv (k_1, k_2, k_3) = \frac{i1}{IS}\mathbf{b}_1 + \frac{i2}{IS}\mathbf{b}_2 + \frac{i3}{IS}\mathbf{b}_3$$

where $\mathbf{b}_1, \mathbf{b}_2$ and $\mathbf{b}_3$ are the reciprocal lattice vectors. For reasons that will be explained later, only multiples of 4 are permissible as shrinking factors in the input file. In addition, a common **k**-point mesh is generated for all $\mathbf{k}, \mathbf{q}$ and $\mathbf{k}+\mathbf{q}$. This means that all four of the following shrinking factors must be multiples of one another; for instance **ISKI** = 12 and **ISKH** = 8 are not allowed in the same input file, but shrinking factors of 12 and 4 are.

ISKI, ISKH Shrinking factors of the Monkhorst-Pack **k**-point mesh for the body (**ISKI**) and head and wing (**ISKH**) elements of the dielectric matrix.

ISQ Shrinking factor of the **q**-point mesh used in the computation of the self-energy, and of the individual **q**-points specified in the spectral/dielectric bandstructure calculations.

ISKM Shrinking factor of the **k**-points at which the self-energy corrections are calculated.

A.1.3 Input data structure: Optional blocks

The preliminary input section is immediately followed by an input block (or series of blocks) specifying a requested task. Each of these additional blocks is identified by a keyword inserted at the beginning of the block and is terminated by the word `ENDTASK`. The remainder of this section details the tasks currently supported by the code, and outlines the format of each input block and of any output files that are generated by a particular task.

Exchange energy: Calculates the total Hartree-Fock exchange energy. This option is provided primarily for testing the optical matrix elements and the corrected integration scheme of Appendix C.3. Note that carefully chosen numbers of **G**-vectors, **k**- and **q**-points are necessary to ensure a well converged calculation. The CRYSTAL95 value for the Hartree-Fock exchange energy is obtained by using a UHF or ROHF Hamiltonian, and inserting the keyword EXCHGENE into the `scf` part of the input.

Record	Format	Options
`EXCHANGE`		
•VB1 VB2	`integer` (×2)	
`ENDTASK`		

VB1, VB2 Lower and upper band indices for summing the exchange energy. The values should ideally encompass all of the valence bands.

Check Input: Prints selected output from `Xtal.dat` such as basis set, **k**-point information and eigenvalue/eigenvector data to the standard output stream. The routine is essentially independent of the rest of the code, and is thus a useful tool for debugging.

Record	Format	Options
`CHECKDAT`		
•PRINTOPT	`character*8`	See note.
If PRINTOPT = 'EIGVAL' then insert:		
◦KPT	`integer`	
If PRINTOPT = 'EIGVEC' then insert:		
◦KPT B1 B2	`integer` (×3)	
`ENDTASK`		

PRINTOPT The printing options available are:

`BASIS` Prints the full basis set.

`KNET` Prints the IBZ **k**-points in oblique coordinates.

`CRYS` Prints some material dependent information.

`EIGVAL` Prints eigenvalues at selected **k**-point, KPT.

`EIGVEC` Prints eigenvectors at selected **k**-point, KPT, and bands ranging from B1 to B2.

Note: Any number of **PRINTOPT** options may be requested, and a separate record assigned to each one.

Self Energy : Calculates the quasiparticle correction to the Hartree-Fock band-structure and the real part of the self-energy $\Sigma(\omega)$ within the *GW* approximation.

Record	Format	Options
`SELF_ENG`		
•DV1 DV2 DC1 DC2	`integer` (×4)	
•SM1 SM2 SN1 SN2	`integer` (×4)	
•NKM	`integer`	NKM ≥ 1
Insert NKM records:		
◦KM1 KM2 KM3	`integer` (×3)	
•FIT_TYPE	`character*4`	`REAL,IMAG`
•HWF	`real*8`	
•SPEC1 SPEC2 NFRE	`real*8` (×2) `integer`	NFRE ≥ 1
•DELTA	`real*8`	
•W_QL Z_QL	`real*8 real*8`	
•EFIT	`real*8`	
`ENDTASK`		

DV1, DV2, DC1, DC2 Lower and upper band indices of valence and conduction bands, respectively, for calculating the dielectric matrix.

SM1, SM2 Lower and upper band indices of **k**-points at which the self-energy is to be evaluated.

SN1, SN2 Lower and upper band indices involved in the summation of the self-energy, arising from the spectral representation of the Green's function [equation (5.2)]

NKM Number of **k**-points at which self-energy is to be calculated.

FIT_TYPE Selects whether to fit for plasmon-pole approximation along the real or imaginary frequency axis. The use of `IMAG` is generally more reliable.

HWF Fitting frequency for the plasmon-pole model (in eV). The dielectric matrix is fully calculated at this frequency as well as at static frequency. Where **FIT_TYPE** = `REAL`, HWF should be set to a value less than the direct gap.

SPEC1, SPEC2 Lower and upper frequency limits, in eV, for the frequency dependent calculation of $\Sigma(\omega)$.

NFRE Number of frequencies at which $\Sigma(\omega)$ is calculated. NFRE divides the frequency range delimited by SPEC1 and SPEC2 into equal intervals.

DELTA Phenomenological broadening parameter which may be used when plotting $\Sigma(\omega)$, in eV.

W_QL Estimated plasmon-pole frequency at $\mathbf{q} \to 0$ for the highest dielectric matrix eigenvalue, in eV.

Z_QL Weighting parameter at $\mathbf{q} \to 0$ for the same eigenvalue. This is simply given by $1 - \varepsilon_n^{-1}(\mathbf{q} \to 0, 0)$.

EFIT Defines size of interval used when fitting the self-energy to a quadratic curve.

Output filename	Format
`hf.dat`	$\lvert\mathbf{k}_m\rvert$ $(E^{HF}_{m\mathbf{k}}, m =$ SM1,SM2)
`qp.dat`	$\lvert\mathbf{k}_m\rvert$ $(E^{QP}_{m\mathbf{k}}, m =$ SM1,SM2)
`sekX.dat`(X=1,NKM)	eV $\langle\Sigma_{m\mathbf{k}}(E)\rangle$

Full Dielectric Matrix: Computes the frequency dependent dielectric matrix. The (causal) matrix is calculated, diagonalized and inverted over a range of frequencies, thus yielding the macroscopic dielectric function with and without local fields, as well as the frequency dependence of the eigenvalues. If local fields are not required, the computation time may be significantly reduced if **SHG**=0. In addition, sum rules are evaluated over the specified frequency interval.

Record	Format	Options
`DMDIRECT`		
•VB1 VB2 CB1 CB2	`integer` (×4)	
•Q1 Q2 Q3	`integer` (×3)	
•SPEC1 SPEC2 NFRE	`real*8` (×2) `integer`	NFRE ≥ 1
•DELTA	`real*8`	
`ENDTASK`		

VB1, VB2, CB1, CB2 Minimum and maximum valence (VB) and conduction (CB) band indices, respectively, involved in the calculation of the dielectric matrix.

Q1, Q2, Q3 Oblique coordinates of **q**-vector, in units of $\mathbf{b}_i/ISQ$.

SPEC1, SPEC2 Lower and upper frequency limits (in eV).

NFRE Number of frequency points at which the dielectric matrix is evaluated. Sum rules are only calculated if NFRE is an odd number.

DELTA Phenomenological broadening parameter used when generating output for plotting (in eV).

Output filename	Format
`dirspec.dat`	eV Re[ε_{00}] Im[ε_{00}] Re[ε_M] Im[ε_M]
`direigs.dat`	eV ($\varepsilon_n, n = 1, \mathrm{SHG}$)
`direigi.dat`	eV ($\varepsilon_n^{-1}, n = 1, \mathrm{SHG}$)

Plasmon-pole model Primarily meant for testing the plasmon-pole approximation to the frequency dependent dielectric matrix. In addition, it may be used to estimate the $\mathbf{q} \rightarrow 0$ pole position of the highest dielectric eigenvalue as required for self-energy calculations.

Record	Format	Options
`DMPLASPO`		
•VB1 VB2 CB1 CB2	`integer` (×4)	
•FIT_TYPE	`character*4`	`REAL,IMAG`
•Q1 Q2 Q3	`integer` (×3)	
•HWF	`real*8`	
•SPEC1 SPEC2 NFRE	`real*8` (×2) `integer`	NFRE ≥ 1
•DELTA	`real*8`	
`ENDTASK`		

VB1, VB2, CB1, CB2 Minimum and maximum valence (VB) and conduction (CB) band indices, respectively, involved in the calculation of the dielectric matrices.

FIT_TYPE Selects whether to fit using a real or imaginary energy parameter for the plasmon-pole approximation. **FIT_TYPE**=`IMAG` type is generally more reliable.

Q1, Q2, Q3 Oblique coordinates of **q**-vector, in units of $\mathbf{b}_i/ISQ$.

HWF Fitting frequency for the plasmon-pole model (in eV). The dielectric matrix is fully calculated at this frequency as well as at static frequency. Where **FIT_TYPE** = `REAL`, HWF should be set to a value less than the direct gap.

SPEC1, SPEC2 Lower and upper frequency limits (in eV).

NFRE Number of frequency points at which the plasmon-pole model of the dielectric matrix is evaluated.

DELTA Broadening parameter for plotting the model eigenvalues (in eV). Note that **DELTA**=0 during the calculation of the dielectric matrices at the two fitting frequencies.

Output filename	Format
`polspec.dat`	eV $(\varepsilon_n, n = 1, \text{SHG})$
`poleigs.dat`	eV $(\varepsilon_n, n = 1, \text{SHG})$
`poleigi.dat`	eV $(\varepsilon_n^{-1}, n = 1, \text{SHG})$

Dielectric Bandstructure Calculates the dielectric bandstructure, the 'empty lattice' bandstructure and the real-space dielectric eigenpotentials. This routine may produce a lot of output.

Record	Format	Options
`DMDBS`		
•VB1 VB2 CB1 CB2	`integer` (×4)	
•NQ GRID	`integer` (×2)	GRID = 0,1
Insert NQ records:		
◦Q1 Q2 Q3	`integer` (×3)	
If NQ = 1 and *GRID = 1 then insert following records:*		
◦RECALC	`character*4`	`READ, CALC`
◦NG1 NG2	`integer` (×2)	≤ SHG
◦O1 O2 O3	`real*8` (×3)	
◦X1 X2 X3 NXPT	`real*8` (×3) `integer`	NXPT ≥ 1
◦Y1 Y2 Y3 NYPT	`real*8` (×3) `integer`	NYPT ≥ 1
◦Z1 Z2 Z3 NZPT	`real*8` (×3) `integer`	NZPT ≥ 0
`ENDTASK`		

VB1, VB2, CB1, CB2 Minimum and maximum valence (VB) and conduction (CB) band indices, respectively, involved in the calculation of the dielectric matrix.

NQ Number of **q**-points in the bandstructure.

GRID Toggle for selecting the generation of the eigenpotentials for the specified **q**-point. GRID = 1 if the eigenpotentials are required, otherwise GRID = 0.

Q1, Q2, Q3 Oblique coordinates of required **q**-vector, in units of $\mathbf{b}_i/ISQ$.

RECALC Recalculation of the dielectric matrix may be unnecessary, for example when examining eigenpotentials inside different cells. If RECALC = `CALC` the matrix is calculated as normal and saved into the file `dm.dat`. Setting RECALC = `READ` will generate eigenpotentials from dielectric matrix data stored in `dm.dat` from a previous run.

Dielectric eigenpotentials are generated on a mesh of points in a three dimensional cell, which are defined by three vectors and an origin. All four vectors are defined with respect to the CRYSTAL95 basis, but are not restricted to lie inside the CRYSTAL95 unit cell. Each cell vector is divided into a number of equal size intervals. Note that if NZPT= 0, a two dimensional plane is defined which may be viewed as a contour plot.

NG1, NG2 Lower and upper indices of DBS eigenvalues for which potentials are plotted.

O1, O2, O3 Cartesian coordinates of cell origin.

X1...Z3 Cartesian coordinates of vectors defining sides of cell.

NXPT, NYPT, NZPT Determines the number of points along the three cell vectors which define the 3-D grid.

Output filename	Format
`dbs.dat`	$\|\mathbf{q}\|\ n\ \varepsilon_n\ \varepsilon_n^{-1}$
`diag.dat`	$\|\mathbf{q}\|\ n\ \bar{\varepsilon}_n\ \bar{\varepsilon}_n^{-1}$
`econ.datX`(X=NG1,NG2)	$x\ y\ z\ V_n(\mathbf{r})$
`epot.datX`(X=NG1,NG2)	See below.

Note on format:

The data files `dbs.dat` and `diag.dat` are formatted for viewing with the `plot` command of the GNUPLOT program, such that groups of points to be

connected by lines are arranged in separate blocks, followed by a blank line. The contour data-files `econ.datX` are similarly formatted for use with `splot`. The isosurface data contained in the `epot.datX` files is organised as 'General Array Format' for use in IBM's *Visual Data Explorer* (DX).

Macroscopic Dielectric Function: Calculates the macroscopic ($\mathbf{q} \to 0$) frequency dependent dielectric function, without local fields. This option is characterised by the use of matrix elements of $\mathbf{p}$, rather than $e^{i\mathbf{q}\cdot\mathbf{r}}$, which makes it possible to restrict the $\mathbf{k}$-point mesh to lie within the irreducible Brillouin zone. In addition, the more accurate integration scheme of Gilat-Raubenheimer [32] may be used instead of the usual 'root-sampling' technique.

Record	Format	Options
`SPECTRUM`		
•DV1 DV2 DC1 DC2	`integer` (×4)	
•SPEC1 SPEC2 NFRE	`real*8` (×2) `integer`	NFRE ≥ 1
•DELTA	`real*8`	
•INT_TYPE	`character*4`	`RS, GR`
`ENDTASK`		

VB1, VB2, CB1, CB2 Minimum and maximum valence (VB) and conduction (CB) band indices, respectively, involved in the calculation of the dielectric function.

DV1, DV2, DC1, DC2 Lower and upper band indices, respectively, of valence and conduction bands, respectively for calculating the dielectric matrix.

SPEC1, SPEC2 Lower and upper frequency limits (in eV).

NFRE Number of frequency points at which the plasmon-pole model of the dielectric matrix is evaluated.

INT_TYPE Selects the **k**-point integration technique: `RS` specifies the 'root-sampling' scheme, `GR` selects the Gilat-Raubenheimer scheme.

Output filename	Format
`spectrum.dat`	eV Re[ε_{00}] Im[ε_{00}] Re[ε_M] Im[ε_M]

A.2 Programming Guidelines

A.2.1 General Structure of the Source Code

The program is written in FORTRAN77, mainly for ease of integration with the CRYSTAL95 code. No machine specific libraries have been used in order to promote portability, and the code has been tested successfully on Linux, IBM-AIX and HP-UX operating systems. Subroutines and functions which deal with related tasks are grouped into several different source files:

initdata.f Source code for the preliminary data manipulation, creating the link between CRYSTAL95 and the optical properties code via the data file `Xtal.dat`.

main.f The primary source code for the program, which controls the reading of the input file, the declaration and initialization of common blocks and the temporary `WORK` arrays and the designation of the computational tasks.

tools.f General purpose subroutines for matrix manipulation and printing of matrix elements, which are called from many of the main subroutines.

kpoint.f Routines generating the general parallelipiped **k**-point mesh as well as the Monkhorst-Pack special points meshes for $\mathbf{k}$ and $\mathbf{k}+\mathbf{q}$.

dbs.f Subroutines which directly involve the calculation of the finite $\mathbf{q}$ dielectric matrix, such as those for generating the dielectric bandstructure, the plasmon-pole testing routine and the frequency dependent calculations.

self.f Source code dealing with the computation of the self-energy and the generation of the quasiparticle bandstructure.

exchange.f Calculation of the Hartree-Fock exchange energy.

integrals.f Contains the main algorithms for the Gaussian orbital integration schemes.

zeroq.f All subroutines and functions dealing explicitly with the momentum operator formulation of the optical matrix elements. Often these routines simply mirror those of `dbs.f` for a **k**-mesh in the irreducible Brillouin zone.

numeric.f Linear algebra routines such as matrix inversion and diagonalization which are not specific to the optical properties code, and have been assembled from external sources.

A schematic of the workings of the code and the interconnectivity of its source files is presented in Figure A-1.

A.2.2 Memory Allocation

One of the main limitations of the FORTRAN77 programming language is the lack of dynamic memory allocation. This imposes major structural limitations on the code, as will now be explained. The principle structural entity of the code is the three-dimensional array `matrix(mxbf,mxbf,mxk)`, where `mxbf` denotes the number of basis functions and `mxk` the number of **k**-points. Use of this structure turns out to be an efficient method of computing the integrals over the Gaussian orbitals and then over the Bloch sums, as it means that terms independent of **k** need not be recalculated or stored onto disk. However if large basis sets ($\sim$ 80 AO's) and dense **k**-meshes ($\sim$ 2048 **k**-points) have to be used, the memory requirements for each occurence of the structure exceeds 1000MBytes, which is prohibitively large. In order to minimise the use of these arrays, two temporary `work(mxbf,mxbf,mxk)` arrays are declared in `main.f` and are passed into all subroutines, where they are initialized and used locally, but do not construe true local data. Large basis sets were not found to be a critical requirement for the calculation of the self-energy, as

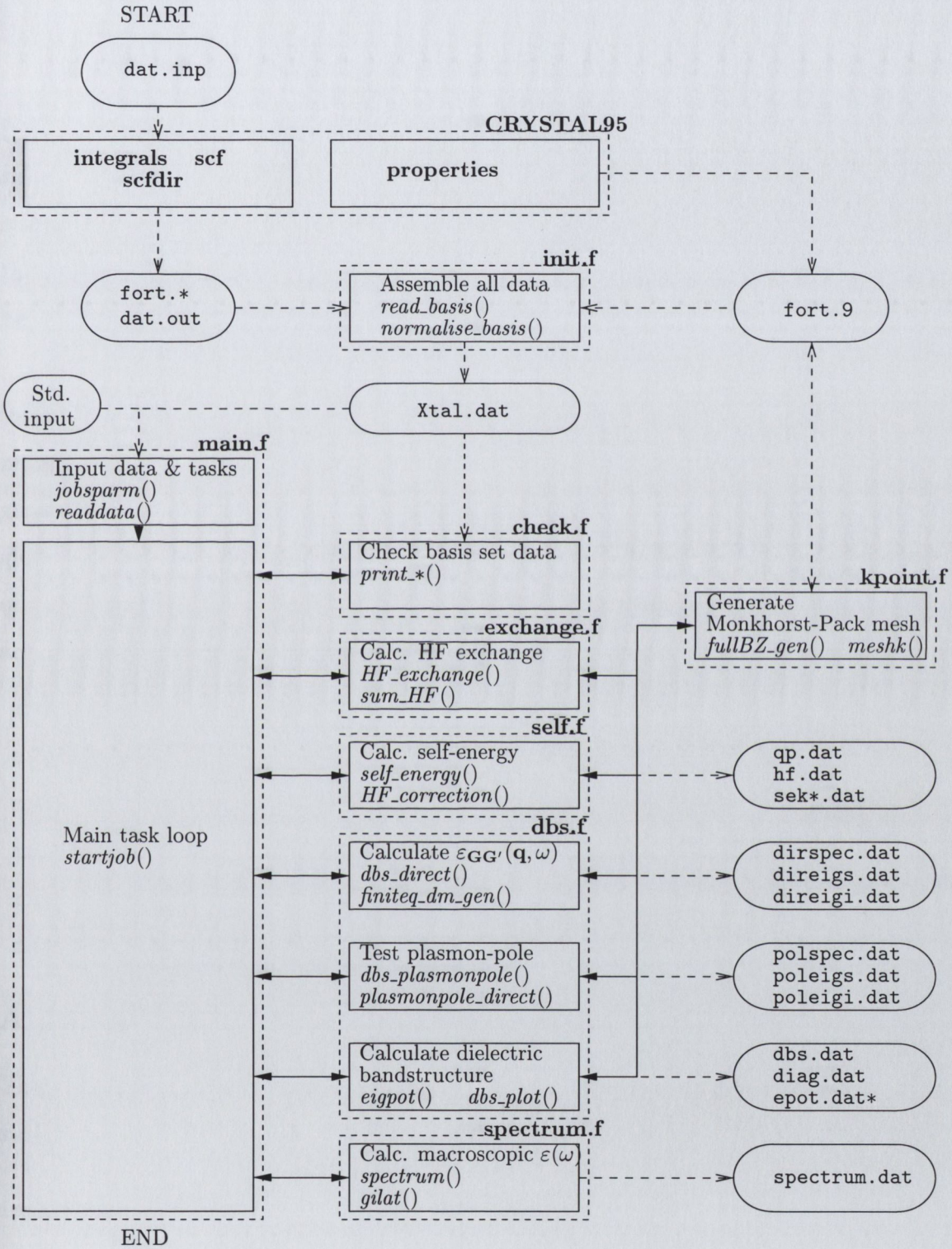


Figure A-1: Flow diagram of the computer code.

was mentioned in Section 5.5, but are required for highly accurate calculations of the dielectric function.

With the exception of the `work` arrays, all global variables and arrays derive from `Xtal.dat` and are stored in named `common` blocks. These are initialized in subroutine `readdata`, in the file `main.f`. Since the purpose of most of these data structures can be easily derived by inspecting the source code or the comments in the code, only `common` blocks directly associated with the Gaussian orbital basis will be considered with here.

Use is made throughout the source code of the `IMPLICIT NONE` declaration, which forces all variables and array to be declared explicitly. Although this significantly increases the size of the source code, it is actually a useful feature when debugging. A general file `param.h` is included at the start of each subroutine, which initializes the global parameters controlling the array sizes. This enables the memory requirements of the code to be easily determined at compile time. The main array dimension sizes are defined as follows:

`mxbf`	Maximum number of basis functions/atomic orbitals
`mxgf`	Maximum number of Gaussian primitives
`mxsh`	Maximum number of shells of Gaussian primitives
`mxat`	Maximum number of atoms in primitive cell
`mxk`	Maximum number of **k**-points in grid
`mxg`	Maximum number of **G**-vectors

A.2.3 Global Strucures

The most frequently accessed structure is that containing the basis set dimensions, stored as:

```
integer ndf,ndf1,nkf
common/gen/ndf,ndf1,nkf
```

where:

ndf is the number of basis (Bloch) functions.

ndf1 is the dimension of the atomic orbital array, which differs from the number of basis functions if d-orbitals are present. It is computationally more transparent to expand the 5 d-orbitals into 6 linearly dependent functions, which are then recombined at the inclusion of the Bloch eigenvectors.

nkf is the number of **k**-points in the irreducible Brillouin zone, as determined by the initial CRYSTAL95 calculation. The number of points in the Full Brillouin zone is determined at run-time.

Information concerning the Gaussian basis is stored in the following data structures:

```
real*8  x(mxgf),y(mxgf),z(mxgf)
real*8  sc(mxgf),pc(mxgf),dc(mxgf),expo(mxgf)
integer nat,ng(mxsh),shtype(mxsh),gpos(mxsh),ns
real*8  bi(3,3),ai(3,3)

common/xyz/x,y,z
common/basis/sc,pc,dc,expo
common/info/nat,ng,shtype,gpos,ns
common/lattice/ai,bi
```

where:

x, y, z are the centres of the Gaussian orbitals, usually situated at atomic sites.

sc, pc, dc are the values of the contraction coefficents for the s, p and d Gaussian primitives.

expo contains the value of the exponent of each Gaussian primitive.

nat is the number of atoms in the primitive cell.

ng is the number of Gaussian primitives in each shell.

shtype labels the orbital type; $0 = s, 1 = sp, 2 = p, 3 = d$

gpos labels the position of the first primitive in a shell in the position, coefficient and exponent arrays.

ns is the total number of shells.

bi, ai are 3×3 vector arrays containing the reciprocal and real space lattice vectors, respectively.

The above data is called directly into the Gaussian integration routines such as in `integrals.f`

A.2.4 k-Point Mesh Generation.

The integration of quantities over the Brillouin zone is a process central to many of the calculations in this thesis. Ideally, the generation of Monkhorst-Pack meshes would be done directly using a routine in CRYSTAL95; unfortunately, however, only the mesh of points in the irreducible Brillouin zone may be obtained directly from the code. Special points meshes are therefore generated using a rather convoluted procedure, which will now be outlined. All the relevant subroutines and functions are located in the source file `kpoint.f`

1. **Selection of parallelipiped indices**

   The calculation of the optical matrix elements generally involves the use of eigenvalues and eigenvectors on a special-points **k**-mesh, as well as at all **k** + **q**, where **q** does not necessarily lie on the **k**-point mesh. Although quite inefficient, the easiest way to obtain all such points is to generate a sufficiently dense grid that encompasses all **k** and **q**, or in other words, to rewrite all **k** and **q** in terms of the same shrinking factor, ISKQ. This grid is set up as a parallelipiped in the three oblique coordinates, such that the indices span the

range $-\text{ISKQ}/2, \ldots, -1, 0, 1, \ldots, \text{ISKQ}/2$, thus yielding $(\text{ISKQ}+1)^3$ **k**-points. This procedure is very inefficient for the SC lattice (Table A.1), but is used nevertheless for programming simplicity.

2. **Generation of parallelipiped eigenfunctions**

   Eigenvalues and eigenvectors are obtained at each point on the grid by using the `BAND` option of the `properties` program. The purpose of this option is actually for plotting bandstructures but it may be adapted to generate the required quantities along a line in **k**-space. A `BAND` calculation is thus repeatedly performed at runtime for several lines such as to span the whole parallelipiped. Use is made of the FORTRAN77 `system` function which allows shell commands to be invoked from within a running program:

   ```
   temp = system($PROPDIR/k_properties < band.gen >! band.out)
   ```

   where the file `band.gen` is generated by the code for each line in **k**-space. This command writes eigenvectors and eigenvalues to `fort.10` and `band.out` respectively, and these are read in and stored directly in memory. The parameter `$PROPDIR` may be altered in `param.h`.

3. **Selection of special points**

   Determining the set of special points from the parallelipiped grid is achieved by selecting points which satisfy some criterion that depends on the Bravais lattice type. For the FCC lattice, for example, this criterion is given by $|k_x| + |k_y| + |k_z| \leq \frac{3}{2}\text{ISKQ}$, where $\mathbf{k} = (k_x, k_y, k_z)$. The special point set, in oblique coordinates, is stored in the array `kfnet(mxk,3)`. Eigenvectors at a particular **k**-point are then accessed by using a simple pointer array [`kfpointer(mxk)`] which relates the oblique coordinates $(k_1\, k_2\, k_3)$ of a special **k**-point to its position in the large grid, via the function:

$$\begin{aligned} \text{kindex} &= (k_1 + \tfrac{ISKQ}{2}) \times (2\tfrac{ISKQ}{2} + 1)^2 \\ &+ (k_2 + \tfrac{ISKQ}{2}) \times (2\tfrac{ISKQ}{2} + 1) \\ &+ (k_3 + \tfrac{ISKQ}{2}) \\ &+ 1 \end{aligned}$$

This enables the eigenvectors at $\mathbf{k}$ and $\mathbf{k}+\mathbf{q}$ to be accessed rapidly during the calculation of the optical matrix elements. As an indication of the inefficiency of this approach, Table A.1 compares the actual size of the parallelipiped grid with the number of special points for a particular value of shrinking factor.

IS	`mxk`	$N_{\mathbf{k}}$(SC)	$N_{\mathbf{k}}$(FCC)
4	125	8	32
8	729	64	256
12	2197	216	864
16	4913	512	2048

Table A.1: Number of **k**-points in the special point mesh at different values of shrinking factor, IS, for the simple cubic (SC) and face-centred centred (FCC) Bravais lattices. Also shown is the corresponding dimension with respect to **k** of the array containing the eigenvectors.

Appendix B

Integral Evaluation

The Gaussian Type Function has emerged as the dominant type of basis function used in quantum chemistry, largely superceding the original Slater type orbital due to its suitability for large scale integral evaluation. In a seminal work, Boys [89] presented an analytical scheme for the evaluation of four-centre integrals over s and p functions. This approach is not, in general, useful for treating functions of higher order quantum number, such as d and f shells, which are often necessary to realistically describe transition metal states or the subtleties of electronic polarisation. Since 1975, algorithms and recursion techniques have developed sufficiently to allow the introduction of the Gaussian orbital scheme into the realm of solid state physics, i.e. periodic systems, which traditionally have been modelled using plane-wave or muffin-tin schemes.

In this section we will discuss the basic features, properties and definitions of Gaussian Type Functions and show how they can be used to evaluate quantum mechanical integrals, in particular those of interest in the study of optical properties of solids. For a more complete treatment, such as details of the more involved two-electron integrals, the reader is referred to the works of Saunders [90].

B.1 Properties of Gaussian Type Functions

The most basic Gaussian Type function (GTF) to be considered is the three-dimensional unnormalised *simple* GTF:

$$\mathrm{G}[a, \mathbf{A}] = e^{-a|\mathbf{r}-\mathbf{A}|^2} = e^{-ar_{\mathbf{A}}^2}, \tag{B.1}$$

centred at the fixed point $\mathbf{A} = (A_x, A_y, A_z)$ (usually a nuclear centre) with exponent $a (a > 0)$. The usefulness of the Gaussian function generally stems from two properties: its separability in the coordinates x, y and z, and the fact that the product of two simple GTFs is another simple GTF, multiplied by a constant factor:

$$\mathrm{G}[a, \mathbf{A}]\,\mathrm{G}[b, \mathbf{B}] = e^{-ab|\mathbf{A}-\mathbf{B}|^2/p}\,\mathrm{G}[p, \mathbf{P}] = K_{ab}\,\mathrm{G}[p, \mathbf{P}] \tag{B.2}$$

where $p = a + b$ and $\mathbf{P} = (a\mathbf{A} + b\mathbf{B})/p$.

The simple GTF can be extended to basis functions of higher quantum number (e.g. p_x, d_{yz}) by constructing *Gaussian lobe orbitals*, as used for example by Hanke [91]. These are essentially linear combinations of simple GTFs situated at different centres. We will however concentrate on a more general class of Gaussian type function, the *Cartesian Gaussian Type Function* (CGTF). First proposed for use in quantum chemistry by Boys [89], it is the most widely used system of basis function in quantum chemistry and molecular physics. It has the form:

$$\mathrm{G}[a, \mathbf{A}; i, j, k] = \mathrm{g}[\mathbf{A}; i, j, k]\; e^{-ar_{\mathbf{A}}^2} = {x_{\mathbf{A}}}^i\, {y_{\mathbf{A}}}^j\, {z_{\mathbf{A}}}^k\; e^{-ar_{\mathbf{A}}^2} \tag{B.3}$$

where i, j, k are non-negative integers; it is clearly the product of an exponential part and a monomial in $x_{\mathbf{A}}, y_{\mathbf{A}}, z_{\mathbf{A}}$ of degree $i + j + k$ ($x_{\mathbf{A}} = x - A_x$).

We describe a further type of Gaussian function here, the *Hermite* GTF(HGTF), defined by:

$$\Lambda[a, \mathbf{A}; i, j, k] = \left(\frac{\partial}{\partial A_x}\right)^i \left(\frac{\partial}{\partial A_y}\right)^j \left(\frac{\partial}{\partial A_z}\right)^k e^{-ar_{\mathbf{A}}^2} \tag{B.4}$$

Although it was originally proposed as a basis function[92], the HGTF is more use-

fully employed as an intermediate function for the evaluation of quantum integrals[93]. Note that it is separable in $x_{\mathbf{A}}, y_{\mathbf{A}}$ and $z_{\mathbf{A}}$:

$$\Lambda[a, \mathbf{A}; i, j, k] = \Lambda_x[a, \mathbf{A}; i]\ \Lambda_y[a, \mathbf{A}; j]\ \Lambda_z[a, \mathbf{A}; k], \tag{B.5}$$

$$\text{where} \quad \Lambda_x[a, \mathbf{A}; i] = \left(\frac{\partial}{\partial A_x}\right)^i e^{-ax_{\mathbf{A}}^2}. \tag{B.6}$$

The HGTFs can be related to the Hermite polynomials $\mathrm{H}_i[\epsilon]$ by substituting $\epsilon = a^{1/2}(x - A_x) = a^{1/2}x_{\mathbf{A}}$ into the standard expression $\mathrm{H}_i[\epsilon] = (-1)^i\, e^{\epsilon^2}(\partial/\partial\epsilon)^i e^{-\epsilon^2}$, having noted that $\partial\epsilon/\partial A_x = -a^{1/2}$:

$$\Lambda_x[a, \mathbf{A}; i] = a^{i/2}\mathrm{H}_i[a^{1/2}x_{\mathbf{A}}]e^{-ax_{\mathbf{A}}^2}$$

However, if we examine the explicit form for the Hermite polynomial,

$$\mathrm{H}_i[\epsilon] = \sum_{t=0}^{i/2} \frac{(-1)^t i!}{t!(i-2t)!}(2\epsilon)^{i-2t}$$

we can see that $\Lambda_x[a, \mathbf{A}; i]$ may also be factored into the product of a polynomial in x_A of degree i, and an exponential:

$$\Lambda_x[a, \mathbf{A}; i] = \lambda_x[a, \mathbf{A}; i]\ e^{-ax_{\mathbf{A}}^2} \tag{B.7}$$

$$\text{with} \quad \lambda_x[a, \mathbf{A}; i] = a^{1/2}\mathrm{H}_i[a^{1/2}x_{\mathbf{A}}] \tag{B.8}$$

The Hermite polynomials are linearly independent, which implies that any polynomial in $x_{\mathbf{A}}, y_{\mathbf{A}}$ and $z_{\mathbf{A}}$ may be expanded as a linear combination of the λ_x polynomials of equal or lower degree.

Finally we mention two recurrence relations obeyed by the GTFs. The Cartesian GTF obeys the following trivial relation:

$$\mathrm{g}[\mathbf{A}; i+1, j, k] = x_{\mathbf{A}}\ \mathrm{g}[\mathbf{A}; i, j, k] \tag{B.9}$$

while a recursion formula for the λ_x polynomials can be derived from the standard

relation for the Hermite polynomials, $\epsilon \mathrm{H}_i[\epsilon] = \frac{1}{2}\mathrm{H}_{i+1}[\epsilon] + i\mathrm{H}_{i-1}[\epsilon]$, to give:

$$x_{\mathbf{A}}\lambda[a, \mathbf{A}, i, j, k] = \frac{1}{2a}\lambda[a, \mathbf{A}; i+1, j, k] + i\lambda[a, \mathbf{A}; i-1, j, k] \tag{B.10}$$

The combination of expression B.10 with the linear independence of the Hermite polynomials provides a powerful mathematical tool for evaluating molecular integrals, as will be discussed in the following section.

B.2 Integral Evaluation

We will now show how an integral over the crystal wavefunctions reduces to a sum of integrals over the GTFs, by considering as an example the overlap integral $\langle i\mathbf{k}|i'\mathbf{k}\rangle$. Writing the crystalline orbital $\psi_i(\mathbf{r};\mathbf{k}) = \langle \mathbf{r}|i\mathbf{k}\rangle$ as a linear combination of Bloch functions [equations (2.25) and (2.26)), we obtain the following:

$$\langle i\mathbf{k}|i'\mathbf{k}\rangle = \sum_{\mu,\mu'} a_{\mu,i}(\mathbf{k})a_{\mu',i'}(\mathbf{k}) \sum_{\mathbf{R},\mathbf{R}'} e^{i\mathbf{k}\cdot(\mathbf{R}'-\mathbf{R})} \int \mathrm{d}\mathbf{r}\ \varphi^*_\mu(\mathbf{r} - \mathbf{A}_\mu - \mathbf{R})\varphi_{\mu'}(\mathbf{r} - \mathbf{A}_{\mu'} - \mathbf{R}'), \tag{B.11}$$

where as before, φ_μ is a local function (atomic orbital) centred (with respect to the unit cell origin) at $\mathbf{A}_\mu$ and $\sum_{\mathbf{R}}$ spans the set of all lattice vectors $\mathbf{R}$. This expression may be simplified for computational purposes by using the substitutions $\mathbf{r}' = \mathbf{r} - \mathbf{R}$ and $\mathbf{R}'' = \mathbf{R} - \mathbf{R}'$ which has the desired effect of making one of the $\mathbf{R}$ summations redundant:

$$\langle i\mathbf{k}|i'\mathbf{k}\rangle = \sum_{\mu,\mu'} a_{\mu,i}(\mathbf{k})a_{\mu',i'}(\mathbf{k}) \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \int \mathrm{d}\mathbf{r}\ \varphi^*_\mu(\mathbf{r} - \mathbf{A}_\mu)\varphi_{\mu'}(\mathbf{r} - [\mathbf{A}_{\mu'} + \mathbf{R}]). \tag{B.12}$$

Finally, if we set $\mathbf{B} = \mathbf{A}_{\mu'} + \mathbf{R}$ and take a CGTF form for the atomic orbitals, the remaining spatial integral term reduces to

$$\int \mathrm{d}\mathbf{r}\ \mathrm{g}[\mathbf{A}; i, j, k]\ \mathrm{g}[\mathbf{B}; i', j', k']\ e^{-ar^2_{\mathbf{A}}}\ e^{-br^2_{\mathbf{B}}}, \tag{B.13}$$

such that the orbital label μ is now implicitly contained in the indices i, j, k. Often, (in particular for core states) the atomic orbitals will be a contraction of primitive gaussians, such that the integral could more generally be written as

$$\sum_{l}^{\eta_l} \sum_{l'}^{\eta_{l'}} d_l, d_{l'} \int \mathrm{d}\mathbf{r}\ \mathrm{g}[\mathbf{A}; i, j, k]\ \mathrm{g}[\mathbf{B}; i', j', k']\ e^{-ar^2_{\mathbf{A},l}}\ e^{-br^2_{\mathbf{B},l'}},$$

where l labels a particular primitive in the contraction of length η_l. Ultimately, therefore, the crystalline orbital integral reduces to the level of an integral over Gaussian primitives.

For the case of overlap between s-orbitals (the simple GTF) the integration is trivial, since the monomials $\mathrm{g}[\mathbf{A}; i, j, k] = \mathrm{g}[\mathbf{B}; i', j', k'] = 1$, and the integration may be evaluated by using the Gaussian product theorem, Eq. (B.2). Difficulties arise when dealing with products like ${x_{\mathbf{A}}}^i {x_{\mathbf{B}}}^{i'}$, (termed the overlap distribution), which occur at least once in the integrals we are interested in. There have been several methods suggested for tackling this problem and they generally use the same idea of removing the polynomial part from the integration completely. Boys[89] proposed using the following expression in CGTF integrals:

$$G[a, \mathbf{A}; 1, 0, 0] = x_{\mathbf{A}} e^{-ar^2_{\mathbf{A}}} = \frac{1}{2a} \frac{\partial}{\partial A_x} e^{-ar^2_{\mathbf{A}}} = \frac{1}{2a} \frac{\partial}{\partial A_x} G[a, \mathbf{A}; 0, 0, 0]$$

thus reducing the integration problem to the level of the simple GTF. Unfortunately the proliferation of partial derivatives leads to an increasingly complex analytical expression as the order of $\mathrm{g}[\mathbf{A}; i, j, k]$ increases, which is inefficient with regard to execution time and data storage.

An alternative approach is to consider expanding the ${x_{\mathbf{A}}}^i {x_{\mathbf{B}}}^{i'}$ term explicitly. In an extension of the Gaussian product theorem for simple GTFs, we write the product of two polynomial GTFs, of degree μ and ν and located at $\mathbf{A}$ and $\mathbf{B}$ respectively, as a polynomial GTF centred at $\mathbf{P}$ of degree $\mu + \nu$ in $x_{\mathbf{P}}, y_{\mathbf{P}}$ and $z_{\mathbf{P}}$. This is because $x_{\mathbf{A}} = x_{\mathbf{P}} + P_x - A_x = x_{\mathbf{P}} + \overline{PA}_x$. Saunders[94] has investigated in depth the evaluation of Gaussian integrals expanded over Cartesian monomials located at $\mathbf{P}$; however it turns out to be more efficient to complete the expansion over the Hermite

GTF basis. Thus we can write:

$$x_{\mathbf{A}}{}^{i}\, x_{\mathbf{B}}{}^{i'} = \sum_{\tau=0}^{i+i'} \mathrm{e}[i,i';\tau]\ \lambda_x[p,\mathbf{P};\tau] \tag{B.14}$$

where the $\mathrm{e}[i,i';\tau]$ are the linear expansion coefficients. Now if we consider the trivial extension

$$x_{\mathbf{A}}{}^{i+1}\, x_{\mathbf{B}}{}^{i'} = \sum_{\tau=0}^{i+i'} \mathrm{e}[i,i';\tau]\ x_{\mathbf{A}}\ \lambda_x[p,\mathbf{P};\tau] \tag{B.15}$$

and combine it with Eq. (B.10) and $x_{\mathbf{A}} = x_{\mathbf{P}} + \overline{PA}_x$, we find that:

$$x_{\mathbf{A}}{}^{i+1}\, x_{\mathbf{B}}{}^{i'} = \sum_{\tau=0}^{i+i'} \mathrm{e}[i,i';\tau]\Big\{\lambda_x[p,\mathbf{P};\tau+1]/2p + \overline{PA}_x\lambda_x[p,\mathbf{P};\tau] + \tau\lambda_x[p,\mathbf{P};\tau-1]\Big\}, \tag{B.16}$$

and from this we deduce a recurrence relation for the expansion coefficients:

$$\mathrm{e}[i+1,i';\tau] = \mathrm{e}[i,i';\tau-1]/2p + \overline{PA}_x\mathrm{e}[i,i';\tau] + (\tau+1)\mathrm{e}[i,i';\tau+1]. \tag{B.17}$$

Note that the e[0,0,0] coefficient is equal to unity since $\lambda_x[p,\mathbf{P};0] = 1$; and that $\mathrm{e}_x[i,i';\tau]$ is zero for $\tau < 0$ or $\tau > i + i'$. Finally, due to the separability of the HGTFs we can express the full product of two Cartesian monomials as

$$\mathrm{g}[\mathbf{A};i,j,k]\,\mathrm{g}[\mathbf{B};i',j',k'] = \sum_{\tau,\mu,\nu} \mathrm{E}[i,j,k,i',j',k';\tau,\mu,\nu]\lambda[p,\mathbf{P};\tau,\mu,\nu] \tag{B.18}$$

$$\text{where}\quad \mathrm{E}[i,j,k,i',j',k';\tau,\mu,\nu] = \mathrm{e}_x[i,i';\tau]\mathrm{e}_y[j,j';\mu]\mathrm{e}_z[k,k';\nu]. \tag{B.19}$$

The e-coefficients turn out to be reasonably simple functions of $\overline{PA}_x$,$\overline{PB}_x$ and p; for example $\mathrm{e}_x[0,1,1] = 1/(2p)$, $\mathrm{e}_y[2,0,0] = \overline{PA}_y(\overline{PA}_y) + 1/(2p)$, and so on.

B.2.1 Overlap integral

Let us now fully evaluate a few integrals using the formalism developed above, starting with the overlap integral, $\mathbf{S}_{mn}$. To avoid overuse of indices we have let m and n represent a composite of unprimed and primed indices respectively, so that

the overlap integral takes the form:

$$\mathbf{S}_{mn} = \int \phi_m \phi_n \mathrm{d}\mathbf{r} = \sum_{\tau,\mu,\nu} \mathrm{E}[m,n;\tau,\mu,\nu] \int \Lambda[p,\mathbf{P};\tau,\mu,\nu]\mathrm{d}\mathbf{r} \tag{B.20}$$

Focussing on the integral over x, a typical term in the summation above is (using Eq. (B.8)):

$$\mathrm{I}_x[p,t] = \int_{-\infty}^{+\infty} e^{-px^2} p^{t/2} \mathrm{H}_t[p^{1/2}x]\mathrm{d}x \tag{B.21}$$

Due to the fact that $\mathrm{H}_0[\epsilon] = 1$ and noting the orthogonality properties of the Hermite polynomials, it is evident that $\mathrm{I}_x[p,t] = 0$ unless $t = 0$. The whole summation therefore reduces to a single term,

$$\mathrm{I}_x[p,t] = \mathrm{I}_x[p,0] = (\pi/p)^{1/2},$$

and after similar integration over y and z the overlap integral becomes:

$$\mathbf{S}_{mn} = \mathrm{E}[m,n;0,0,0] K_{ab} (\pi/p)^{3/2} \tag{B.22}$$

It is immediately evident that an overlap integral of arbitrary quantum number can now be completely evaluated with sole use of Eq. (B.17).

B.2.2 Optical matrix elements

Optical matrix elements may be determined by evaluating the 'dipole integral' $\mathbf{D}_{mn}$ or the 'momentum integral' $\mathbf{M}_{mn}$. The x component of the dipole integral has the form

$$\mathbf{D}^x_{mn} = \int \phi_m \phi_n \; x_C \; \mathrm{d}\mathbf{r} = \sum_{\tau,\mu,\nu} \mathrm{E}[m,n;\tau,\mu,\nu] \int x_C \; \Lambda[p,\mathbf{P};\tau,\mu,\nu]\mathrm{d}\mathbf{r} \tag{B.23}$$

The integral is easily evaluated if we use Eq. (B.10) and the fact that $x_\mathbf{C} = x_\mathbf{P} + \overline{PC}_x$ to write it as a sum of overlap integrals, hence:

$$\mathbf{D}^x_{mn} = \Big\{ \mathrm{E}[m,n;1,0,0] + \overline{PC}_x \mathrm{E}[m,n;0,0,0] \Big\} K_{ab} (\pi/p)^{3/2}. \tag{B.24}$$

By using relation (B.17) to show that $\mathrm{e}_x[i,i',1] = \mathrm{e}_x[i,i'+1,0] - \overline{PA}_x \mathrm{e}_x[i,i',0]$, we can ultimately show

$$\mathbf{D}^x_{mn} = \mathrm{e}_y[j,j';0]\mathrm{e}_z[k,k';0]\Big\{\mathrm{e}_x[i,i'+1;0] + \overline{AC}_x \mathrm{e}_x[i,i',0]\Big\} K_{ab}(\pi/p)^{3/2}. \quad \text{(B.25)}$$

The x component of momentum $\mathbf{M}^x_{mn}$ may expressed as

$$\mathbf{M}^x_{mn} = -i\,\hbar \int \phi_m \frac{\partial}{\partial x}\phi_n \mathrm{d}\mathbf{r} = -i\,\hbar \int \mathrm{G}[a,\mathbf{A};i,j,k]\frac{\partial}{\partial x}\mathrm{G}[a,\mathbf{A};i,j,k]\mathrm{d}\mathbf{r} \quad \text{(B.26)}$$

Applying the following relation for the derivative of the CTGF:

$$\frac{\partial}{\partial x}\mathrm{G}[b,\mathbf{B};i',j',k'] = i'\mathrm{G}[b,\mathbf{B};i'-1,j',k'] - 2b\mathrm{G}[b,\mathbf{B};i'+1,j',k'],$$

and using the same techniques as for the overlap integral, we find:

$$\begin{aligned}
\mathbf{M}^x_{mn} &= -i\,\hbar\,\{i'\mathrm{E}[i,j,k,i'-1,j',k';0,0,0] - 2b\ \mathrm{E}[i,j,k,i'+1,j',k';0,0,0]\} \\
&= -i\,\hbar\ \mathrm{e}_y[j,j';0]\mathrm{e}_z[k,k';0] \times \\
&\quad \Big\{i'\mathrm{e}_x[i,i'-1;0] - 2b\ \mathrm{e}_x[i,i'+1;0]\Big\} K_{ab}(\pi/p)^{3/2}. \qquad \text{(B.27)}
\end{aligned}$$

B.2.3 Matrix elements of $e^{i\mathbf{q}\cdot\mathbf{r}}$

The plane wave integral $\mathbf{E}_{mn}$ directly enters the expression for the dielectric matrix. In terms of the Gaussian primitives, the integral becomes:

$$\mathbf{E}_{mn} = \int \mathrm{d}\mathbf{r}\ \phi_m e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}\phi_n = \int \mathrm{d}\mathbf{r}\ \mathrm{G}[a,\mathbf{A};i,j,k]\ \ e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}\ \mathrm{G}[b,\mathbf{B};i',j',k']. \quad \text{(B.28)}$$

Relabelling $(\mathbf{q}+\mathbf{G})$ as $\mathbf{Q}$, the x component of the integral may be written as

$$\begin{aligned}
\mathbf{E}^x_{mn} &= \int_{-\infty}^{+\infty} \mathrm{d}x\ \mathrm{g}_x[a,\mathbf{A};i]\ e^{iQ_x x}\ e^{-px^2_{\mathbf{P}}}\mathrm{g}_x[b,\mathbf{B};i] \qquad \text{(B.29)} \\
&= K_{ab}\sum_{\tau}\mathrm{e}[i,i';\tau]\int_{-\infty}^{+\infty}\lambda_x[p,\mathbf{P};\tau]e^{iQ_x x}\mathrm{d}x \qquad \text{(B.30)}
\end{aligned}$$

This may be further evaluated by substituting equation (B.6) for the Hermite coefficient $\lambda_x[p, \mathbf{P}; \tau]$, such that

$$\begin{aligned}
\int_{-\infty}^{+\infty} \lambda_x[p, \mathbf{P}, \tau] e^{iQ_x x} \mathrm{d}x &= \int_{-\infty}^{+\infty} \left(\frac{\partial}{\partial P_x}\right)^{\tau} e^{-px_{\mathbf{P}}^2} e^{iQ_x x} \mathrm{d}x \\
&= \left(\frac{\partial}{\partial P_x}\right)^{\tau} \int_{-\infty}^{+\infty} e^{-px_{\mathbf{P}}^2 + iQ_x x} \mathrm{d}x \\
&= \sqrt{\frac{\pi}{p}} \left(\frac{\partial}{\partial P_x}\right)^{\tau} e^{-(\frac{Q_x^2}{4p} - iQ_x P_x)} \\
&= \sqrt{\frac{\pi}{p}} e^{-\frac{Q_x^2}{4p}} (iQ_x)^{\tau} e^{iQ_x P_x}
\end{aligned}$$

where the spatial integral has been evaluated evaluated by completing the square on x. Finally, what remains is a rather convoluted expression for the total integral:

$$\begin{aligned}
\mathbf{E}_{mn} &= \left(\frac{\pi}{p}\right)^{\frac{3}{2}} e^{-(Q_x^2 + Q_y^2 + Q_z^2)/4p} K_{ab} \\
&\quad \times \sum_{\tau,\mu,\nu} \mathrm{E}[m, n; \tau, \mu, \nu] \; i^{\tau+\mu+\nu} \; Q_x^{\tau} Q_y^{\mu} Q_z^{\nu} \; e^{i(Q_x P_x + Q_y P_y + Q_z P_z)}
\end{aligned} \tag{B.31}$$

Recall the sum over τ, μ, ν runs over all values of τ up to $i + i'$, etc.

Although equation B.31 is computationally feasible, it is somewhat more attractive in this case to use a recursive technique from the outset. This consists of deriving analytical expressions for integrals involving higher order atomic orbitals based on the result for the integral over s functions:

$$\begin{aligned}
\langle s_{\mathbf{A}} || s_{\mathbf{B}} \rangle &= \int \mathrm{d}\mathbf{r} \; e^{-ar_{\mathbf{A}}^2} \; e^{i\mathbf{Q}\cdot\mathbf{r}} \; e^{-br_{\mathbf{B}}^2} \\
&= e^{-ab|\mathbf{A}-\mathbf{B}|^2/p} \int \mathrm{d}\mathbf{r} \; e^{-pr_{\mathbf{P}}^2 + i\mathbf{Q}\cdot\mathbf{r}} \\
&= \left(\frac{\pi}{p}\right)^{\frac{3}{2}} e^{-ab|\mathbf{A}-\mathbf{B}|^2/p} \; e^{-\frac{|\mathbf{Q}|^2}{4p} + i\mathbf{Q}\cdot\mathbf{P}}
\end{aligned} \tag{B.32}$$

where again the integral has been evaluated by completing the square.

By defining the two functions

$$\mathrm{D}_{\mathbf{A}i} = \overline{PA}_i + \frac{iQ_i}{2p} = -\frac{b}{p}\overline{AB}_i + \frac{iQ_i}{2p} \tag{B.33}$$

$$\mathrm{D}_{\mathbf{B}j} = \overline{PB}_j + \frac{iQ_j}{2p} = +\frac{a}{p}\overline{AB}_j + \frac{iQ_j}{2p}, \tag{B.34}$$

where i, j, k are dummy indices running over x, y, z, and by using the standard formulae of Boys,

$$\frac{1}{2a}\frac{\partial}{\partial A_i}|s_{\mathbf{A}}\rangle = |p_{\mathbf{A}}i\rangle \tag{B.35}$$

$$\frac{1}{2a}\frac{\partial}{\partial A_k}|s_{\mathbf{A}}\rangle = |d_{\mathbf{A}ik}\rangle - \frac{\delta_{ik}}{2a}|s_{\mathbf{A}}\rangle, \tag{B.36}$$

we may derive a set of partially recursive expressions for all orders of function required (i.e., up to d character). A complete list of all such expressions is presented in Table B.1 overleaf.

$$
\begin{aligned}
\langle s_{\mathbf{A}} || s_{\mathbf{B}} \rangle &= \left(\frac{\pi}{p}\right)^{\frac{3}{2}} e^{-ab|\mathbf{A}-\mathbf{B}|^2/p}\; e^{-\frac{|\mathbf{Q}|^2}{4p}+i\mathbf{Q}\cdot\mathbf{P}} \\
\langle s_{\mathbf{A}} || p_{\mathbf{B}j} \rangle &= \left[\overline{PB}_j + \frac{iQ_j}{2p}\right] \langle s_{\mathbf{A}} || s_{\mathbf{B}} \rangle = \mathrm{D}_{\mathbf{B}j} \langle s_{\mathbf{A}} || s_{\mathbf{B}} \rangle \\
\langle p_{\mathbf{A}i} || s_{\mathbf{B}} \rangle &= \left[\overline{PA}_i + \frac{iQ_i}{2p}\right] \langle s_{\mathbf{A}} || s_{\mathbf{B}} \rangle = \mathrm{D}_{\mathbf{A}i} \langle s_{\mathbf{A}} || s_{\mathbf{B}} \rangle \\
\langle p_{\mathbf{A}i} || p_{\mathbf{B}j} \rangle &= \left[\frac{\delta_{ij}}{2p} + \mathrm{D}_{\mathbf{A}i}\mathrm{D}_{\mathbf{B}j}\right] \langle s_{\mathbf{A}} || s_{\mathbf{B}} \rangle \\
\langle d_{\mathbf{A}ik} || s_{\mathbf{B}} \rangle &= \left[\frac{\delta_{ik}}{2p} + \mathrm{D}_{\mathbf{A}i}\mathrm{D}_{\mathbf{A}k}\right] \langle s_{\mathbf{A}} || s_{\mathbf{B}} \rangle \\
\langle s_{\mathbf{A}} || d_{\mathbf{B}jk} \rangle &= \left[\frac{\delta_{jk}}{2p} + \mathrm{D}_{\mathbf{B}j}\mathrm{D}_{\mathbf{B}k}\right] \langle s_{\mathbf{A}} || s_{\mathbf{B}} \rangle \\
\langle p_{\mathbf{A}i} || d_{\mathbf{B}jk} \rangle &= \left[\mathrm{D}_{\mathbf{A}i}\mathrm{D}_{\mathbf{B}j}\mathrm{D}_{\mathbf{B}k} + \frac{1}{2p}\Big(\delta_{ik}\mathrm{D}_{\mathbf{B}j} + \delta_{ij}\mathrm{D}_{\mathbf{B}k} + \delta_{jk}\mathrm{D}_{\mathbf{A}i}\Big)\right] \langle s_{\mathbf{A}} || s_{\mathbf{B}} \rangle \\
\langle d_{\mathbf{A}ij} || p_{\mathbf{B}k} \rangle &= \left[\mathrm{D}_{\mathbf{A}i}\mathrm{D}_{\mathbf{A}j}\mathrm{D}_{\mathbf{B}k} + \frac{1}{2p}\Big(\delta_{jk}\mathrm{D}_{\mathbf{A}i} + \delta_{ik}\mathrm{D}_{\mathbf{A}j} + \delta_{ij}\mathrm{D}_{\mathbf{B}k}\Big)\right] \langle s_{\mathbf{A}} || s_{\mathbf{B}} \rangle \\
\langle d_{\mathbf{A}ij} || d_{\mathbf{B}kl} \rangle &= \left[\frac{1}{4p^2}\Big(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{jk}\delta_{il}\Big) + \mathrm{D}_{\mathbf{A}i}\mathrm{D}_{\mathbf{A}j}\mathrm{D}_{\mathbf{B}k}\mathrm{D}_{\mathbf{B}l}\right. \\
&\quad + \frac{1}{2p}\Big(\delta_{ik}\mathrm{D}_{\mathbf{A}j}\mathrm{D}_{\mathbf{B}l} + \delta_{jk}\mathrm{D}_{\mathbf{A}i}\mathrm{D}_{\mathbf{B}l} + \delta_{jl}\mathrm{D}_{\mathbf{A}i}\mathrm{D}_{\mathbf{B}k} \\
&\quad \left. + \delta_{il}\mathrm{D}_{\mathbf{A}j}\mathrm{D}_{\mathbf{B}k} + \delta_{ij}\mathrm{D}_{\mathbf{B}k}\mathrm{D}_{\mathbf{B}l} + \delta_{kl}\mathrm{D}_{\mathbf{A}i}\mathrm{D}_{\mathbf{A}j}\Big)\right] \langle s_{\mathbf{A}} || s_{\mathbf{B}} \rangle .
\end{aligned}
$$

Table B.1: Analytical expressions for integals of $e^{i\mathbf{Q}\cdot\mathbf{r}}$ over Gaussian primitive functions of s, p and d character. See main text for meaning of symbols.

Appendix C

Reciprocal Space Integration

Many of the calculations in this thesis involve the integration of a periodic function of wavevector $\mathbf{k}$ over the Brillouin zone (BZ):

$$\bar{f} = \frac{\Omega}{(2\pi)^3} \int_{\mathrm{BZ}} \mathrm{d}\mathbf{k}\, f(\mathbf{k}) \tag{C.1}$$

This integration essentially represents an averaging over the electronic states; the dielectric function, for example, consists of an averaging over optical matrix elements and transition energies. In practice, integration may be performed by evaluating the function at a carefully selected set of points in reciprocal space, termed special-points. Several such sets of special points exist; the original concept is due to Baldereschi [95] who demonstrated how a single special-point could be used to compute the charge density. However, as was shown in Section 4.1.3, a dense set of points is necessary for the evaluation of quantities such as the static dielectric matrix. Chadi and Cohen [96] have developed a suitable procedure for generating these special-points sets, although we have chosen to adopt the more systematic procedure of Monkhorst and Pack. [97]

Before this procedure is described in more detail, it should be noted that the special-point technique is only suitable for the integration of smooth functions. Accurate calculation of spectral functions, such as the frequency dependent dielectric matrix, require a modified procedure based on the analytic continuation integration

scheme of Gilat and Raubenheimer [32], outlined in Appendix C.2. Finally, the computation of the self-energy involves the integration over the BZ of a singular function, for which the special point theorem does not hold. A method for handling this is presented in Appendix C.3 along with the results of some test calculations.

C.1 Special-Point Technique

C.1.1 Theory

The rationale behind the special-point technique is best explained by making use of symmetry arguments. Although not strictly necessary, it is assumed that the integrand $f(\mathbf{k})$ is a periodic function of $\mathbf{k}$ whose period is given by the reciprocal lattice vectors. The function may therefore be expanded over the direct lattice vectors $\mathbf{R}$ as a Fourier sum:

$$f(\mathbf{k}) = \sum_{\mathbf{R}} f(\mathbf{R}) e^{i\mathbf{k}\cdot\mathbf{R}}. \tag{C.2}$$

Since $f(\mathbf{k})$ is invariant with respect to the point group operations of the crystal, the Fourier components $f(\mathbf{R})$ for symmetry related $\mathbf{R}$-vectors must be equivalent. This allows us to divide the sum over the lattice vectors into a sum over 'shells', C_n, where a shell is defined to contain all symmetry related lattice vectors of equal length. The index n labels the shells according to increasing modulus. Within a particular shell, the Fourier transforms have the same value, which will be denoted by $f(\mathbf{R}) = f_n$. Therefore, equation (C.2) may be rewritten as follows:

$$f(\mathbf{k}) = \sum_{n=0}^{\infty} f_n \sum_{|\mathbf{R}|\in C_n} e^{i\mathbf{k}\cdot\mathbf{R}} = f_0 + \sum_{n=1}^{\infty} f_n A_n(\mathbf{k}) \tag{C.3}$$

Returning to the BZ integration, from the definition of $A_n(\mathbf{k})$ it is clear that

$$\int_{\mathrm{BZ}} \mathrm{d}\mathbf{k}\; A_n(\mathbf{k}) = 0 \quad \text{if } |\mathbf{R}| \neq 0, \tag{C.4}$$

which implies that the full integral simply reduces to

$$\frac{\Omega}{(2\pi)^3}\int_{\mathrm{BZ}} \mathrm{d}\mathbf{k}\; f(\mathbf{k}) = f_0. \tag{C.5}$$

By the mean-value-point theorem, it is always possible to find a point $\mathbf{k}$ inside the integration zone such that $A_n(\mathbf{k}) = 0$ for each $n \neq 0$, which implies that the integral is simply given by $f_0 = f(\mathbf{k})$. However, such a point would depend on the integrand function itself, whereas we require a scheme which is generally applicable to all smooth functions $f(\mathbf{k})$. Equation C.3 shows that the value of $f(\mathbf{k})$ differs from f_0 by a quantity which depends on the values of $A_n(\mathbf{k})$. The best approximation to f_0 for a generic $\mathbf{k}$ is therefore achieved by choosing a $\mathbf{k}$ which gives $A_n(\mathbf{k}) = 0$ for $n = 1, 2, \ldots, N_p$ with the largest possible N_p.

In general a single point does not suffice for integration, and it is more suitable to use a set of $N_{\mathbf{k}}$ points, $\{\mathbf{k}_i\}$, with relative weights w_i such that:

$$\sum_{i=1}^{N_{\mathbf{k}}} w_i = 1 \tag{C.6}$$

$$\sum_{i=1}^{N_{\mathbf{k}}} w_i A_n(\mathbf{k}_i) = 0 \quad n = 1, \ldots, N_p. \tag{C.7}$$

Combining these equations with the Fourier expansion of $f(\mathbf{k})$ yields the final expression for the integral:

$$\frac{\Omega}{(2\pi)^3}\int_{\mathrm{BZ}} \mathrm{d}\mathbf{k}\; f(\mathbf{k}) \approx \sum_{i=1}^{N_{\mathbf{k}}} w_i f(\mathbf{k}_i). \tag{C.8}$$

C.1.2 Special-Point Generation

Equations C.6 and C.7 turn out to be sufficient to specify the coordinates of a special-point set. The Monkhorst-Pack scheme uses the following method:

(a) The size of the mesh is determined by choosing a parameter q. In the present code, q is equal to half the shrinking factor ISKQ.

(b) A sequence of q numbers is then defined by

$$u_r = \frac{2r - q - 1}{2q} \quad \text{where } r = 1, 2, \ldots, q. \tag{C.9}$$

and similarly for u_p and u_s.

(c) These factors specify a regularly spaced mesh of q^3 $\mathbf{k}$ points by:

$$\mathbf{k}_{prs} = u_p \mathbf{b}_1 + u_r \mathbf{b}_2 + u_s \mathbf{b}_3 \tag{C.10}$$

where $\mathbf{b}_i$ are the reciprocal lattice vectors.

(d) The set of special-points is then chosen by mapping this set inside an irreducible zone of the reciprocal lattice, taking care to ensure the proper weighting factors for each $\mathbf{k}$, such that equation (C.6) is obeyed.

Proof that the above scheme actualy satisfies the criteria for a special-point set may be found in [97], and a comprehensive discussion of the special-point scheme may be found in [98].

The above procedure is only sufficient for generating a complete set of special points in the irreducible Brillouin zone. For quantities such as the dielectric matrix $\varepsilon_{\mathbf{GG'}}(\mathbf{q}, \omega)$, a set of special points in the full zone is required. These may be generated by a slightly different procedure:

(a) Again, a parameter q is chosen which determines the grid density.

(b) The sequence of $2q$ numbers is defined by

$$u_r = \frac{2r - 2q + 1}{2q} \quad \text{where } r = 0, 1, \ldots, 2q - 1. \tag{C.11}$$

and similarly for u_p and u_s.

(c) A regularly spaced mesh of $(2q)^3$ $\mathbf{k}$ points is thus generated:

$$\mathbf{k}_{prs} = u_p \mathbf{b}_1 + u_r \mathbf{b}_2 + u_s \mathbf{b}_3 \tag{C.12}$$

(d) Finally, only points which lie inside the full BZ are selected, and the corresponding weights w_i are all set to 1.

n	$\mathbf{k}(\times\frac{2\pi}{a})$	$\widehat{\mathbf{k}}(\times\mathbf{b}_i)$	n	$\mathbf{k}(\times\frac{2\pi}{a})$	$\widehat{\mathbf{k}}(\times\mathbf{b}_i)$
1	$(-\frac{3}{4},-\frac{1}{4},-\frac{1}{4})$	$[-\frac{1}{4}\ -\frac{2}{4}\ -\frac{2}{4}]$	17	$(\ \frac{1}{4},-\frac{3}{4},-\frac{1}{4})$	$[-\frac{2}{4}\ \ \frac{0}{4}\ -\frac{1}{4}]$
2	$(-\frac{3}{4},-\frac{1}{4},\ \frac{1}{4})$	$[\ \frac{0}{4}\ -\frac{1}{4}\ -\frac{2}{4}]$	18	$(\ \frac{1}{4},-\frac{3}{4},\ \frac{1}{4})$	$[-\frac{1}{4}\ \ \frac{1}{4}\ -\frac{1}{4}]$
3	$(-\frac{3}{4},\ \frac{1}{4},-\frac{1}{4})$	$[\ \frac{0}{4}\ -\frac{2}{4}\ -\frac{1}{4}]$	19	$(\ \frac{1}{4},-\frac{1}{4},-\frac{3}{4})$	$[-\frac{2}{4}\ -\frac{1}{4}\ \ \frac{0}{4}]$
4	$(-\frac{3}{4},\ \frac{1}{4},\ \frac{1}{4})$	$[\ \frac{1}{4}\ -\frac{1}{4}\ -\frac{1}{4}]$	20	$(\ \frac{1}{4},-\frac{1}{4},-\frac{1}{4})$	$[-\frac{1}{4}\ \ \frac{0}{4}\ \ \frac{0}{4}]$
5	$(-\frac{1}{4},-\frac{3}{4},-\frac{1}{4})$	$[-\frac{2}{4}\ -\frac{1}{4}\ -\frac{2}{4}]$	21	$(\ \frac{1}{4},-\frac{1}{4},\ \frac{1}{4})$	$[\ \frac{0}{4}\ \ \frac{1}{4}\ \ \frac{0}{4}]$
6	$(-\frac{1}{4},-\frac{3}{4},\ \frac{1}{4})$	$[-\frac{1}{4}\ \ \frac{0}{4}\ -\frac{2}{4}]$	22	$(\ \frac{1}{4},-\frac{1}{4},\ \frac{3}{4})$	$[\ \frac{1}{4}\ \ \frac{2}{4}\ \ \frac{0}{4}]$
7	$(-\frac{1}{4},-\frac{1}{4},-\frac{3}{4})$	$[-\frac{2}{4}\ -\frac{2}{4}\ -\frac{1}{4}]$	23	$(\ \frac{1}{4},\ \frac{1}{4},-\frac{3}{4})$	$[-\frac{1}{4}\ -\frac{1}{4}\ \ \frac{1}{4}]$
8	$(-\frac{1}{4},-\frac{1}{4},-\frac{1}{4})$	$[-\frac{1}{4}\ -\frac{1}{4}\ -\frac{1}{4}]$	24	$(\ \frac{1}{4},\ \frac{1}{4},-\frac{1}{4})$	$[\ \frac{0}{4}\ \ \frac{0}{4}\ \ \frac{0}{4}]$
9	$(-\frac{1}{4},-\frac{1}{4},\ \frac{1}{4})$	$[\ \frac{0}{4}\ \ \frac{0}{4}\ -\frac{1}{4}]$	25	$(\ \frac{1}{4},\ \frac{1}{4},\ \frac{1}{4})$	$[\ \frac{1}{4}\ \ \frac{1}{4}\ \ \frac{1}{4}]$
10	$(-\frac{1}{4},-\frac{1}{4},\ \frac{3}{4})$	$[\ \frac{1}{4}\ \ \frac{1}{4}\ -\frac{1}{4}]$	26	$(\ \frac{1}{4},\ \frac{1}{4},\ \frac{3}{4})$	$[\ \frac{2}{4}\ \ \frac{2}{4}\ \ \frac{1}{4}]$
11	$(-\frac{1}{4},\ \frac{1}{4},-\frac{3}{4})$	$[-\frac{1}{4}\ -\frac{1}{4}\ \ \frac{0}{4}]$	27	$(\ \frac{1}{4},\ \frac{3}{4},-\frac{1}{4})$	$[\ \frac{1}{4}\ \ \frac{0}{4}\ \ \frac{2}{4}]$
12	$(-\frac{1}{4},\ \frac{1}{4},-\frac{1}{4})$	$[\ \frac{0}{4}\ -\frac{1}{4}\ \ \frac{0}{4}]$	28	$(\ \frac{1}{4},\ \frac{3}{4},\ \frac{1}{4})$	$[\ \frac{2}{4}\ \ \frac{1}{4}\ \ \frac{2}{4}]$
13	$(-\frac{1}{4},\ \frac{1}{4},\ \frac{1}{4})$	$[\ \frac{1}{4}\ \ \frac{0}{4}\ \ \frac{0}{4}]$	29	$(\ \frac{3}{4},-\frac{1}{4},-\frac{1}{4})$	$[-\frac{1}{4}\ \ \frac{1}{4}\ \ \frac{1}{4}]$
14	$(-\frac{1}{4},\ \frac{1}{4},\ \frac{3}{4})$	$[\ \frac{2}{4}\ \ \frac{1}{4}\ \ \frac{0}{4}]$	30	$(\ \frac{3}{4},-\frac{1}{4},\ \frac{1}{4})$	$[\ \frac{0}{4}\ \ \frac{2}{4}\ \ \frac{1}{4}]$
15	$(-\frac{1}{4},\ \frac{3}{4},-\frac{1}{4})$	$[\ \frac{1}{4}\ -\frac{1}{4}\ \ \frac{1}{4}]$	31	$(\ \frac{3}{4},\ \frac{1}{4},-\frac{1}{4})$	$[\ \frac{0}{4}\ \ \frac{1}{4}\ \ \frac{2}{4}]$
16	$(-\frac{1}{4},\ \frac{3}{4},\ \frac{1}{4})$	$[\ \frac{2}{4}\ \ \frac{0}{4}\ \ \frac{1}{4}]$	32	$(\ \frac{3}{4},\ \frac{1}{4},\ \frac{1}{4})$	$[\ \frac{1}{4}\ \ \frac{2}{4}\ \ \frac{2}{4}]$

Table C.1: Special-points mesh in the full Brillouin zone for the FCC lattice. Both cartesian ($\mathbf{k}$) and oblique ($\widehat{\mathbf{k}}$) coordinates are given. The parameter $q = 2$ defines the density of the mesh and the corresponding shrinking factor IS = 4. All weighting factors are equal to 1.

The actual computational implementation of this scheme is discussed in some detail in Appendix A.2.4. An example of the full zone special point mesh is presented in Table C.1 for the FCC lattice with $q = 2$.

C.2 Analytic Continuation Method

The analytic continuation method has been mentioned previously in the thesis in connection with the accurate calculation of spectral functions, in particular the frequency dependent dielectric function, $\varepsilon(\mathbf{q},\omega)$. The relevant function to be integrated over the BZ is of the form

$$I_{c,v}(\mathbf{q},\omega) = \int_{\mathrm{BZ}} \mathrm{d}\mathbf{k}\ \frac{F_{c,v}(\mathbf{k},\mathbf{q})}{W_{c,v}(\mathbf{k},\mathbf{q}) + \omega + i\eta}, \tag{C.13}$$

where $F_{c,v}(\mathbf{k},\mathbf{q})$ represents the matrix elements of the momentum operator $\hat{\mathbf{p}}$ between Bloch states, and $W_{c,v}(\mathbf{k},\mathbf{q}) = E_{v\mathbf{k}} - E_{c,\mathbf{k}+\mathbf{q}}$. The numerical integration scheme to be used must be more accurate than the ordinary special point scheme used during the calculation of the static dielectric matrix, due to the erratic behaviour of the integrand when the frequency ω approaches a possible transition frequency, i.e., $|W_{c,v}(\mathbf{k},\mathbf{q}) + \omega| \leq 1\text{eV}$. Away from this frequency, however, the usual special-point ('root sampling') techniques suffice. [99].

In the analytic continuation technique due to Gilat and Raubenheimer [32], the special-point scheme is modified as follows: the BZ is subdivided into a set of equally spaced cells of volume Ω_i, which are centred around the special points $\mathbf{k}_i$. In contrast to the usual approach, the integral over each small cell is not just simply sampled at the point $\mathbf{k}_i$ but is explicitly evaluated. Rewriting equation C.13 as

$$I_{c,v}(\mathbf{q},\omega) = \sum_{i=1}^{N_\mathbf{k}} w_i \int_{\Omega_i} \mathrm{d}\mathbf{k}\, \frac{F_{c,v}(\mathbf{k}_i,\mathbf{q})}{W_{c,v}(\mathbf{k}_i,\mathbf{q}) + \omega + i\eta} = \sum_{i=1}^{N_\mathbf{k}} w_i I_{c,v}(\mathbf{q},\mathbf{k}_i;\omega), \tag{C.14}$$

the functions $F_{c,v}(\mathbf{k}_i,\mathbf{q})$ and $W_{c,v}(\mathbf{k}_i,\mathbf{q})$ are expanded to first order in a Taylor series about $\mathbf{k}_i$, after which the integral over the cell may be solved analytically. Following some algebraic manipulation the solution is obtained thus:

$$\begin{aligned} I_{c,v}(\mathbf{q},\mathbf{k}_i;\omega) &= F_{c,v}(\mathbf{k}_i,\mathbf{q}) I^0_{c,v}(\mathbf{q},\mathbf{k}_i;\omega) && \text{(C.15)} \\ &\quad + \sum_{p=1}^{3} \nabla_p F_{c,v}(\mathbf{k}_i,\mathbf{q}) I^i_{c,v}(\mathbf{q},\mathbf{k}_i;\omega) && \text{(C.16)} \end{aligned}$$

where

$$\begin{aligned} I^0_{c,v}(\mathbf{q},\mathbf{k}_i;\omega) &= \frac{1}{2\Lambda} \sum_{n_{1,2,3}=-1}^{1} n_1 n_2 n_3 W^2_{n_{1,2,3}} \ln W_{n_{1,2,3}} && \text{(C.17)} \\ I^i_{c,v}(\mathbf{q},\mathbf{k}_i;\omega) &= \frac{1}{6\Lambda} \sum_{n_{1,2,3}=-1}^{1} n_1 n_2 n_3 W^3_{n_{1,2,3}} \ln W_{n_{1,2,3}} && \text{(C.18)} \\ &\quad \times \ln W_{n_{1,2,3}} - i I^0_{c,v}(\mathbf{q},\mathbf{k}_i;\omega) - \frac{8\Omega_i}{l_i} && \text{(C.19)} \\ \text{and} \quad W_{n_{1,2,3}} &= \frac{W_{c,v}(\mathbf{k}_i,\mathbf{q}) + \omega + i\eta}{|\nabla W_{c,v}(\mathbf{k}_i,\mathbf{q})|} + \sum_i n_i l_i a_i. && \text{(C.20)} \end{aligned}$$

In the above expressions l_i represent the direction cosines of $\nabla W_{c,v}(\mathbf{k}_i, \mathbf{q})$, $\Lambda = l_1 l_2 l_3 |\nabla W_{c,v}(\mathbf{k}_i, \mathbf{q})|$ and a_i denote the dimensions of the small integration cell about $\mathbf{k}_i$.

The analysis may be simplfied without loss of accuracy with the assumption that the $F_{c,v}(\mathbf{k}_i, \mathbf{q})$ function varies smoothly over the small cell, an approximation which can be arbitrarily improved upon by increasing the $\mathbf{k}$-point density. This has the effect of setting the second term in equation (C.17) to zero, which significantly reduces the computational complexity. An example of the use of the analytic continuation technique is presented in Figure C-1, which compares a calculation using the more accurate method with the corresponding result using just the standard special point scheme.

C.3 Auxiliary Function Technique

The special point theorem does not hold in the presence of divergent terms in the integrand, which presents a problem for the calculation of the self-energy [equation (5.13)]. In this case a modified scheme must be used, which will now be discussed with particular reference to the self-energy.

Inspection of equation (5.13) reveals that divergences occur both for the case of $\mathbf{q} = 0, \mathbf{G} = \mathbf{G}' = 0$ and also for $\mathbf{q} = 0, \mathbf{G} \neq \mathbf{G}' = 0$. The former case leads to a $1/q^2$ divergence, which is more severe than the $1/q$ behaviour arising from the latter, and is the only divergent term which will be corrected for explicitly. The method is as follows: noting that the integrand $f(\mathbf{q})$ diverges as A/q^2 for small $\mathbf{q}$, equation (5.13) may be rewritten as

$$\sum_{\mathbf{q}} f(q) \equiv \sum_{\mathbf{q}} [f(\mathbf{q}) - Ag(\mathbf{q})] + A \sum_{\mathbf{q}} g(\mathbf{q}) \tag{C.21}$$

where

$$A = \lim_{q \to 0} [q^2 f(\mathbf{q})],$$

and $g(\mathbf{q})$ is an auxiliary function which has the same divergent character as $f(\mathbf{q})$

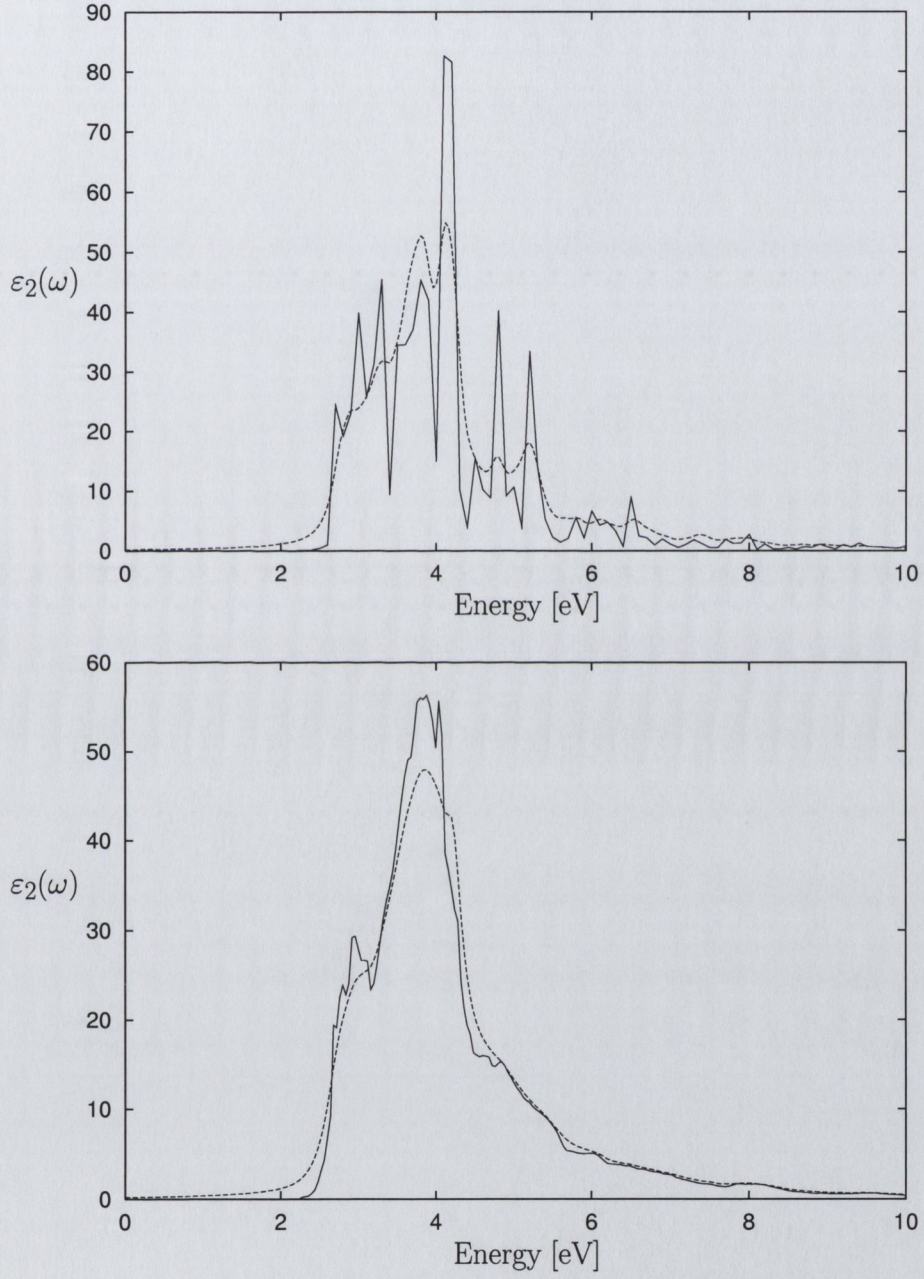


Figure C-1: Imaginary part of the dielectric function of silicon calculated using the root-sampling (RS) scheme (top) and the Gilat-Raubenheimer (GR) scheme (bottom) of reciprocal space integration. A phenomenological broadening parameter η is used in generating both sets of data — solid line: $\eta = 0.005$ eV; dashed line: $\eta = 0.1$ eV — although its use is not necessary in the GR scheme. 195 $\mathbf{k}$-points in the irreducible BZ were used in both cases.

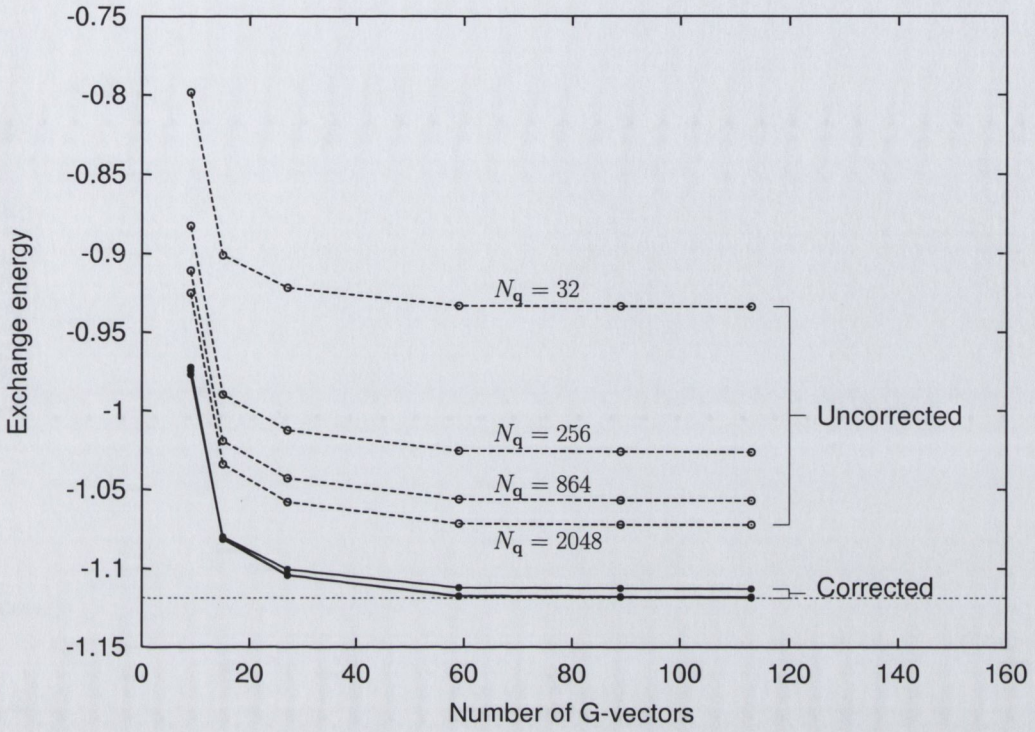


Figure C-2: Accuracy of the improved **q**-integration scheme for singular functions. The dotted lines show the Hartree-Fock exchange energies for various mesh densities, calculated using the uncorrected special points scheme. Solid lines denote the results using the auxiliary term in the integration, for $N_{\mathbf{q}} = 32$ and the fully converged calculation for $N_{\mathbf{q}} \geq 256$. The correct value, represented by a single horizontal line, is obtained directly as output from CRYSTAL95 by using an unrestricted Hartree-Fock Hamiltonian.

for $\mathbf{q} \to 0$, i.e. $g(\mathbf{q}) \sim 1/q^2$. This has the desired effect that the $f(\mathbf{q}) - Ag(\mathbf{q})$ term does not contain any divergences and thus can be evaluated within the special points scheme. In choosing the $g(\mathbf{q})$ we follow Ref. [68] which presents functions with the correct divergent behaviour which have the symmetry of the Bravais lattice. Explicit formulae for $g(\mathbf{q})$ suitable for the SC, FCC and BCC lattices are given in Table C.2 along with their integrated values over the BZ as required for the $\sum_{\mathbf{q}} g(\mathbf{q})$ term.

As a means of testing the accuracy of the corrected integration scheme, we calculate the total Hartree-Fock exchange energy using equation (5.13) and compare it with the value obtained from the CRYSTAL95 code. In Hartree-Fock theory, the dielectric matrix is simply given by $\varepsilon_{\mathbf{GG'}} = \delta_{\mathbf{GG'}}$,[21] which means that the self-

Simple cubic (SC) lattice:

$$
\begin{aligned}
g(\mathbf{q}) &= \frac{\frac{a^2}{2}}{3-\cos(aq_x)-\cos(aq_y)-\cos(aq_z)} \\
\sum_{\mathbf{q}} g(\mathbf{q}) &= \left(\frac{a^2}{2\pi^3}\right) 4\pi \left[18+12\sqrt{2}-10\sqrt{3}-7\sqrt{6}\right] K^2\left[(2-\sqrt{3})(\sqrt{3}-\sqrt{2}\right] \\
&= \left(\frac{a^2}{2\pi^3}\right) 15.67249523473854 \qquad a=\Omega^{1/3}
\end{aligned}
$$

Face-centred cubic (FCC) lattice:

$$
\begin{aligned}
g(\mathbf{q}) &= \frac{\left(\frac{a}{2}\right)^2}{3-\cos\left(\frac{aq_x}{2}\right)\cos\left(\frac{aq_y}{2}\right)-\cos\left(\frac{aq_y}{2}\right)\cos\left(\frac{aq_z}{2}\right)-\cos\left(\frac{aq_x}{2}\right)\cos\left(\frac{aq_z}{2}\right)} \\
\sum_{\mathbf{q}} g(\mathbf{q}) &= \left(\frac{a^2}{4\pi^3}\right) \sqrt{3}\pi K^2\left[\sin\left(\frac{\pi}{12}\right)\right] \\
&= \left(\frac{a^2}{4\pi^3}\right) 13.89764556215925 \qquad a=(4\Omega)^{1/3}
\end{aligned}
$$

Body-centred cubic (BCC) lattice:

$$
\begin{aligned}
g(\mathbf{q}) &= \frac{\left(\frac{a}{2}\right)^2}{1-\cos\left(\frac{aq_x}{2}\right)\cos\left(\frac{aq_y}{2}\right)\cos\left(\frac{aq_z}{2}\right)} \\
\sum_{\mathbf{q}} g(\mathbf{q}) &= \left(\frac{a^2}{4\pi^3}\right) 4\pi K^2\left[\frac{\sqrt{2}}{2}\right] \\
&= \left(\frac{a^2}{4\pi^3}\right) 43.19806651591508 \qquad a=(2\Omega)^{1/3}
\end{aligned}
$$

Table C.2: Auxiliary functions $g(\mathbf{q})[\mathbf{q}=(q_x,q_y,q_z)]$ and their integrals over the Brillouin zone for the simple cubic, face-centred cubic and body-centred cubic lattices. K is the complete elliptic integral of the first kind, a is the lattice constant and Ω is the volume of the primitive cell.

energy matrix element becomes

$$\langle m\mathbf{k}|\Sigma(E)|m\mathbf{k}\rangle = -\frac{4\pi e^2}{V}\sum_{\mathbf{q},\mathbf{G}}\sum_{n\in VB}\frac{1}{|\mathbf{q}+\mathbf{G}|^2}|\langle m\mathbf{k}|e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}|n\mathbf{k}+\mathbf{q}\rangle|^2. \quad \text{(C.22)}$$

Since the wavefunctions used are the actual Hartree-Fock wavefunctions, the above expression is equal to the expectation value of the Hartree-Fock exchange operator, $\Sigma^{HF}_{m\mathbf{k}}$. The total exchange energy is obtained by summing $\Sigma^{HF}_{m\mathbf{k}}$ over all bands $m \in VB$ and $\mathbf{q}$-points in the first Brillouin zone.

Figure C-2 compares the calculated exchange energy of a model compound, the FCC helium lattice, for increasing number of $\mathbf{q}$-points with the expected value from CRYSTAL95. The uncorrected special points scheme shows a gradual improvement in exchange energy as the density of the $\mathbf{q}$-point mesh is increased; this is to be expected as the volume element that is associated with the $\mathbf{q} = 0$ point decreases in size for a denser mesh. Introducing the auxiliary correction dramatically improves the comparison with the CRYSTAL95 value, as well effectively limiting the number of $\mathbf{q}$-points that needs to be summed over. It is evident that substantial errors would be introduced into the computation of the self-energy correction were the method described here not used.

The full self-energy expression is made slightly more complicated by the sum over l in equation (5.13). In particular, the dielectric matrix eigenvectors $V_{l,\mathbf{G}}$ cannot be factored out of the sum if different plasmon-pole frequencies are used for each eigenvalue, as is the case in this work. The auxiliary term may be written as $A[E, m, \mathbf{k}]\sum_q f(\mathbf{q})$, where

$$A[E,m,\mathbf{k}] = \sum_l V_{l,\mathbf{G}=0}(\mathbf{q}\to 0)[V_{l,\mathbf{G}'=0}(\mathbf{q}\to 0)]^* \frac{z_{\mathbf{q}\to 0,l}\omega_{\mathbf{q}\to 0,l}}{2}\frac{1}{E - E_{m\mathbf{k}} \pm \omega_{\mathbf{q}\to 0,l}}. \quad \text{(C.23)}$$

A problem arises here in that the dielectric matrix cannot be easily evaluated in the limit as $\mathbf{q} \to 0$, which is needed to determine the eigenvectors $V_{l,\mathbf{G}=0}(\mathbf{q}\to 0)$ as well as the plasmon-pole parameters $z_{\mathbf{q}\to 0,l}$ and $\omega_{\mathbf{q}\to 0,l}$. Calculation of the $e^{i\mathbf{q}\cdot\mathbf{r}}$ matrix elements in the limit $\mathbf{q} \to 0$ requires the use of equation (3.20), which is

complicated by the presence of the nonlocal term in the Hartree-Fock Hamiltonian. In addition, equation (3.20) is strictly valid only in the limit of an infinite set of basis functions, which is a relevant problem in the Gaussian orbital formalism, as discussed in Section 4.1.2.

These restrictions require that a suitable approximation is made to the A term: assuming that the summation in C.23 is dominated by the l with the highest weight (i.e. $z^{max}_{\mathbf{q}\to 0,l}$), we retain only this term in the summation which effectively makes $z_{\mathbf{q}l}$ and $\omega_{\mathbf{q}l}$ independent of l. Using the orthogonality of the dielectric eigenvectors, $\sum_l V_{l,\mathbf{G}}(\mathbf{q})[V_{l,\mathbf{G}'}(\mathbf{q})]^* = 1$, A then becomes

$$A[E,m,\mathbf{k}] = \frac{z^{max}_{\mathbf{q}\to 0,l}\,\omega^{max}_{\mathbf{q}\to 0,l}}{2}\frac{1}{E - E_{m\mathbf{k}} \pm \omega^{max}_{\mathbf{q}\to 0,l}} \tag{C.24}$$

leaving only a single pole position and weight to be determined. These are simply estimated from a smooth interpolation of the Hartree-Fock dielectric bandstructure. Altering the values of these terms by 10% leads to changes in the self-energy correction to the direct band gap of the order of < 0.05 eV, which confirms that the precise evaluation of A is not critical. Note that a full summation over l is maintained in the actual self-energy term.

Appendix D

Basis Sets

The importance of using a large, quality basis set for calculating optical properties of solids was stressed in Chapter 4. In fact, it is often the case with any *ab initio* calculations using CRYSTAL95 that difficulties encountered during the initial `scf` step are due to a poorly chosen basis set. Extreme care must be taken when using or altering basis sets, and newly developed sets should be thoroughly tested before use.

In addition to the Pople STO-nG sets that are built into CRYSTAL95, basis sets for use in periodic solids are generally available at a number of sources, and have been adopted from standard atomic and molecular basis sets by removing or altering the gaussian exponents of the most diffuse orbitals. These sets are usually sufficient for most purposes, i.e., for ground state calculations. A comprehensive collection is maintained on the web at the following two locations:

The CRYSTAL development homepages at:

`http://www.dl.ac.uk/TCS/Software/CRYSTAL` and

`http://www.ch.unito.it/ifm/teorica/crystal.html`

Mike Towler's CRYSTAL resources page at:

`http://www.tcm.phy.cam.ac.uk/~mdt`

As discussed in some depth in Section 4.1.2, the standard basis sets are not suitable for calculating physical quantities that require an extended set of states,

such as the dielectric matrix. Improved sets may be constructed by carrying out the following procedures:

1. The outermost valence shells should be uncontracted and the exponents optimized with respect to the total energy. Optimization may be done by hand, or preferably by using a shell script such as the `Lopt` script supplied with CRYSTAL98, or Mike Towler's `billy` script available at the web-site listed above.

2. Polarization functions, if not already present, should be added, and the outermost functions reoptimized. A general rule of thumb is that successive exponents should differ by a factor of 2-3, or else problems associated with linear dependence may occur.

3. If the number of basis functions is still not sufficient, 'ghost' orbitals may be added at the interstities. Depending on the space group of the system this addition will result in a dramatic increase in the number of basis functions, so care must be taken to choose the correct coordinates for the new orbitals. If possible, this set should be reoptimized.

4. Finally, a `properties` calculation should be run with the new basis set and the bandstructure (particularly the band gap) should be checked. Optimization using the total energy as a sole criterion may sometimes result in erroneous band gaps.

A comprehensive guide to the use and optimization of basis sets in CRYSTAL95 may be found in [100].

Basis sets used in this thesis for calculating the dielectric matrix and the quasiparticle bandstructure are presented in the following pages for Si, diamond, NaCl, MgO and GaN. If interstitial functions are not required the outermost exponents must be re-optimized (as shown below for silicon). Note that the total energy may be poor for the sets with ghost atoms, as illustrated by Figure 4-1, which emphasizes the need for re-optimization after the addition of extra basis functions.

Notation: basis sets are referred to by the number of gaussian primitives in each contraction. A dash separates the core functions from the valence functions, and polarization functions are indicated by brackets. 'X[spd]' denotes a ghost atom with s, p and d orbitals. In the following basis set, for example, 8-1111G(d) indicates that there are eight primitives in each core function, one primitive (uncontracted) in each of the four valence functions and one polarization function of d character.

Silicon — Si

8-1111G(d): Highly optimized basis.

Fractional coordinates — Si: (0.125 0.125 0.125)

```
14  7
0 0   8   2.   1.
 149866.          0.0001215
  22080.6         0.0009770
   4817.5         0.0055181
   1273.5         0.0252
    385.11        0.0926563
    128.429       0.2608729
     45.4475      0.4637538
     16.2589      0.2952
0 1   8   8.   1.
    881.111      -0.0003         0.0006809
    205.840      -0.005          0.0059446
     64.8552     -0.0368         0.0312
      23.9       -0.1079         0.1084
      10.001      0.0134         0.2387
       4.4722     0.3675         0.3560066
       2.034      0.5685         0.341
       0.9079     0.2065         0.1326
0 1   1 4.   1.
       3.7505     1.0            1.0
0 1   1 0.   1.
       1.0714     1.0            1.0
0 1   1 0.   1.
       0.3451     1.0            1.0
0 1   1  0.   1.
       0.1132     1.0            1.0
0 3   1  0.   1.
       0.4533     1.0
```

Silicon — Si

8-1111G(d) + X[spd] with interstitial functions.

Fractional coordinates — Si: (0.125 0.125 0.125) X: (0.5 0.5 0.5)

```
14  7
0 0  8  2.  1.
 149866.          0.0001215
  22080.6         0.0009770
   4817.5         0.0055181
   1273.5         0.0252
    385.11        0.0926563
    128.429       0.2608729
     45.4475      0.4637538
     16.2589      0.2952
0 1  8  8.  1.
    881.111      -0.0003         0.0006809
    205.840      -0.005          0.0059446
     64.8552     -0.0368         0.0312
      23.9       -0.1079         0.1084
      10.001      0.0134         0.2387
       4.4722     0.3675         0.3560066
       2.034      0.5685         0.341
       0.9079     0.2065         0.1326
0 1  1 4.  1.
       3.7500     1.0            1.0
0 1  1 0.  1.
       1.5000     1.0            1.0
0 1  1 0.  1.
       0.7000     1.0            1.0
0 1  1  0.  1.
       0.3000     1.0            1.0
0 3  1  0.  1.
       0.45000    1.0
0 2
0 1  1 0.  1.
       0.3000     1.0            1.0
0 3  1 0.  1.
       0.4500     1.0            1.0
```

Diamond — C

6-111(d) + X[spd] with interstitial functions.

Fractional coordinates — C: (0.125 0.125 0.125) X: (0.5 0.5 0.5)

```
6   5
0 0 6 2.0 1.0
 3048.0      0.001826
  456.4      0.01406
  103.7      0.06876
   29.23     0.2304
    9.349    0.4685
    3.189    0.3628
0 1 1 4.0 1.0
    3.665    1.0         1.0
0 1 1 0.0 1.0
    1.4000   1.0         1.0
0 1 1 0.0 1.0
    0.4000   1.0         1.0
0 3 1 0.0 1.0
    0.8      1.0
0   2
0 1 1 0.0 1.0
    0.6      1.0         1.0
0 3 1 0.0 1.0
    1.00     1.0
```

Magnesium Oxide — MgO

85-11(d)/8-411(d) + X[spd] with interstitial functions.

Fractional coordinates — Mg: (0.0 0.0 0.0) O: (0.5 0.5 0.5) X: (0.25 0.25 0.25)

```
12 5
0 0 8 2.0 1.0
 68370.0     0.0002226
  9661.0     0.001901
  2041.0     0.011042
   529.6     0.05005
   159.17    0.1690
    54.71    0.36695
    21.236   0.4008
     8.791   0.1487
0 1 5 8.0 1.0
   143.7    -0.00671 0.00807
    31.27   -0.07927 0.06401
     9.661  -0.08088 0.2092
     3.726   0.2947  0.3460
     1.598   0.5714  0.3731
0 1 1 0.0 1.0
     0.688   1.0      1.0
0 1 1 0.0 1.0
     0.28    1.0      1.0
0 3 1 0.0 1.0
     0.65    1.0
```

```
8 5
0 0 8 2.0 1.0
 8020.0        0.00108
 1338.0        0.00804
  255.4        0.05324
   69.22       0.1681
   23.90       0.3581
    9.264      0.3855
    3.851      0.1468
    1.212      0.0728
0 1 4 8.0 1.0
   49.43      -0.00883     0.00958
   10.47      -0.0915      0.0696
    3.235     -0.0402      0.2065
    1.217      0.379       0.347
0 1 1 0.0 1.0
    0.500      1.0          1.0
0 1 1 0.0 1.0
    0.200      1.0          1.0
0 3 1 0.0 1.0
    0.65       1.0
0 2
0 1 1 0.0 1.0
    1.00       1.0     1.0
0 3 1 0.0 1.0
    1.00       1.0
```

Sodium Chloride — NaCl

86-311(d)/85-11(d) + X[spd] with interstitial functions.

Fractional coordinates — Na: (0.0 0.0 0.0) Cl: (0.5 0.5 0.5) X: (0.25 0.25 0.25)

```
17  6
0 0 8 2.0 1.0
135320.0     0.000225
 19440.0     0.00191
  4130.0     0.01110
  1074.0     0.04989
   323.4     0.1703
   111.1     0.3683
    43.4     0.4036
    18.18    0.1459
0 1 6 8.0 1.0
   324.8    -0.00763 0.00820
    73.00   -0.0829  0.0605
    23.71   -0.1046  0.2115
     9.138   0.2540  0.3765
     3.930   0.695   0.3967
     1.329   0.399   0.186
0 1 3 8.0 1.0
     4.755 -0.3740 -0.0340
     1.756 -0.4754  0.1617
     0.785  1.3400  0.9250
0 1 1 0.0 1.0
    0.2967  1.0      1.0
```

```
0 1 1 0.0 1.0
    0.1192  1.0      1.0
0 3 1 0.0 1.0
    0.2000  1.0
11  5
0 0 8 2.0 1.0
 56700.0      0.000225
  8060.0      0.00191
  1704.0      0.01105
   443.6      0.05006
   133.1      0.1691
    45.8      0.3658
    17.75     0.3998
     7.38     0.1494
0 1 5 8.0 1.0
   119.0     -0.00673 0.00803
    25.33    -0.0798  0.0639
     7.80    -0.0793  0.2074
     3.00     0.3056  0.3398
     1.289    0.5639  0.3726
0 1 1 0.0 1.0
     0.4879  1.0      1.0
0 1 1 0.0 1.0
     0.1814  1.0      1.0
0 3 1 0.0 1.0
     0.2500  1.0
0  2
0 1 1 0.0 1.0
     0.3     1.0      1.0
0 3 1 0.0 1.0
     0.5     1.0      1.0
```

Gallium Nitride — GaN

864-33111/5-111(d) + X[spd] with interstitial functions. DFT only.

Fractional coordinates — Ga: (0.25 0.25 0.25) N: (0.0 0.0 0.0)

X: (0.625 0.625 0.625)

```
31 8
0 0 8 2.0 1.0
444668.         0.00023
 64576.8        0.0019
 13935.2        0.0108
  3651.66       0.049
  1099.41       0.1672
   381.106      0.3643
   149.482      0.4028
    62.817      0.1477
0 1 6 8.0 1.0
  1155.65      -0.006     0.0088
   278.06      -0.07      0.063
    93.1902    -0.14      0.2196
    37.0948     0.2685    0.4083
    15.4816     0.6184    0.4154
```

```
      5.1517     0.3248    0.3536
0 1 4 8.0 1.0
     70.6212     0.0067   -0.0085
     27.1784    -0.0863   -0.035
     11.4394    -0.3455    0.0849
      4.6239     0.4086    0.5583
0 1 1 3.0 1.0
      1.8289     1.0       1.0
0 1 1 0.0 1.0
      0.66317    1.0       1.0
0 1 1 0.0 1.0
      0.17514    1.0       1.0
0 3 4 10.0 1.0
     67.4273     0.0259
     18.8367     0.1495
      6.3055     0.3805
      2.1298     0.4768
0 3 1 0. 1.0
      0.66949    1.0
7 5
0 0 7 2.0 1.0
 7590.0      0.000889
  991.2      0.008994
  190.1      0.05287
   52.69     0.1710
   18.10     0.3612
    7.048    0.4027
    2.922    0.1549
0 1 3 5.0 1.0
   18.40    -0.02807    0.01869
    4.242   -0.1146     0.10130
    1.347    0.1890     0.2394
0 1 1 0.0 1.0
     0.422    1.0        1.0
0 1 1 0.0 1.0
     0.140    1.0        1.0
0 3 1 0.0 1.0
     0.500    1.0
0 2
0 1 1 0.0 1.0
     0.800    1.0     1.0
0 3 1 0.0 1.0
     1.000    1.0
```

Bibliography

[1] A. Szabo and N. Ostlund. *Modern Quantum Chemistry.* McGraw-Hill, New York, (1989).

[2] J. Inkson. *Many-Body Theory of Solids.* Plenum Press, New York, (1984).

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

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

[5] R. Parr and W. Yang. *Density Functional Theory of Atoms and Molecules.* Oxford University Press, New York, (1989).

[6] M. Hybertson and S. Louie. (1986), *Phys. Rev. B*, **34**, 5390.

[7] R. Dreizler and E. Gross. *Density Functional Theory.* Springer, Berlin, (1990).

[8] G. Mahan. *Many-Particle Physics.* Plenum Press, New York, (1990).

[9] L. Hedin and S. Lundqvist. *Solid State Physics.* Academic Press, New York, (1969).

[10] F. Aryasetiawan and O. Gunnarsson. (1998), *Rep. Prog. Phys.*, **61**, 237.

[11] L. Hedin. (1965), *Phys. Rev.*, **139**, A796.

[12] S. Albrecht, L. Reining, R. Del Sole, and G. Onida. (1998), *phys. stat. sol. (a)*, **170**, 189.

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

[14] R. Daling and W. van Haeringen. (1989), *Phys. Rev. B*, **40**, 11659.

[15] R. Godby, M. Schlüter, and L. Sham. (1988), *Phys. Rev. B*, **37**, 10159.

[16] L. Hedin. (1995), *Int. J. Quantum Chem.*, **56**, 445.

[17] G. Strinati, H. Mattausch, and W. Hanke. (1980), *Phys. Rev. B*, **25**, 2867.

[18] R. Dovesi, V. Saunders, C. Roetti, M. Causà, N. Harrison, R. Orlando, and E. Aprà. CRYSTAL95 *User's Manual.* (1996).
See `http://www.dl.ac.uk/TCS/Software/CRYSTAL`.

[19] H. Rojas, R. Godby, and R. Needs. (1995), *Phys. Rev. Lett.*, **74**, 1827.

[20] M. Rieger, L. Steinbeck, I. White, H. Rojas, and R. Godby. (1999), *Comp. Phys. Commun.*, **117**, 211.

[21] M. Rohlfing, P. Krüger, and J. Pollman. (1993), *Phys. Rev. B*, **48**, 17791.

[22] F. Aryasetiawan and O. Gunnarsson. (1998), *Phys. Rev. Lett.*, **74**, 3221.

[23] M. Rohlfing, P. Krüger, and J. Pollman. (1997), *Phys. Rev. B*, **56**, R7065.

[24] M. Rohlfing, P. Krüger, and J. Pollmann. (1995), *Phys. Rev. B*, **52**, 1905.

[25] F. Wooten. *Optical Properties of Solids.* Academic Press, New York, (1972).

[26] E. Hecht and A. Zajac. *Optics.* Addison-Wesley, (1974).

[27] F. Bassani and G. Pastori Parravicini. *Electronic States and Optical Transitions in Solids.* Pergamon Press, (1975).

[28] S. Adler. (1962), *Phys. Rev.*, **126**, 413.

[29] H. Ehrenreich and M. Cohen. (1959), *Phys. Rev.*, **115**, 786.

[30] B. Adolph, V. Gavrilenko, K. Tenelsen, F. Bechstedt, and R. Del Sole. (1996), *Phys. Rev. B*, **53**, 9797.

[31] D. Aspnes and A. Studna. (1983), *Phys. Rev. B*, **27**, 985.

[32] G. Gilat and L. Raubenheimer. (1966), *Phys. Rev.*, **144**, 390.

[33] L. Kadanoff and G. Bahm. *Quantum Statistical Mechanics.* Addison-Wesley, (1989).

[34] N. Wiser. (1963), *Phys. Rev.*, **129**, 62.

[35] D. Johnson. (1974), *Phys. Rev. B*, **9**, 4475.

[36] M. Hybertson and S. Louie. (1987), *Phys. Rev. B*, **35**, 5585.

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

[38] M. Ehrnsperger and H. Bross. (1997), *J. Phys.: Condens. Matter*, **9**, 1225.

[39] H. Li. (1980), *J. Phys. Chem. Ref. Data*, **9**, 561.

[40] Z. Levine and D. Allan. (1991), *Phys. Rev. B*, **43**, 4187.

[41] J. Hughes and J. Sipe. (1996), *Phys. Rev. B*, **53**, 10751.

[42] Z. Levine and D. Allan. (1989), *Phys. Rev. Lett*, **63**, 1719.

[43] A. Dal Corso, S. Baroni, and R. Resta. (1994), *Phys. Rev. B*, **49**, 5323.

[44] V. Olevano, M. Palummo, G. Onida, and R. Del Sole. (1999), *Phys. Rev. B*, **60**, 14224.

[45] X. Gonze, P. Ghosez, and R. Godby. (1995), *Phys. Rev. Lett.*, **74**, 4035.

[46] D. Herrendörfer. *Quantum Mechanical Corrections of the Discrete Dipole Appoach to Optical Properties.* PhD thesis, University of Dublin, (1995).

[47] W. Mochán and R. Barrera. (1985), *Phys. Rev. Lett.*, **55**, 1192.

[48] G. Poppe, C. Wijers, and A. van Silfhout. (1991), *Phys. Rev. B*, **44**, 7917.

[49] P. de Boeij and C. Wijers. (2000), *Phys. Lett. A*, **272**, 264.

[50] G. Poppe. *Local Aspects of the Optical Reflection Problem.* PhD thesis, University of Twente, (1991).

[51] A. Lehnen and L. Bruch. (1980), *Physica*, **100A**, 215.

[52] M. Nicastro and C. Patterson. *Submitted to J. Phys.: Condensed Matter.*

[53] S. Sinha. (1969), *Phys. Rev.*, **177**, 1256.

[54] C. Hogan and C. Patterson. (1998), *Phys. Rev. B*, **57**, 14843.

[55] A. Baldereschi and E. Tosatti. (1978), *Sol. Stat. Commun.*, **29**, 131.

[56] A. Baldereschi and E. Tosatti. (1978), *Phys. Rev. B*, **17**, 4710.

[57] A. Baldereschi, R. Car, and E. Tosatti. (1979), *Sol. Stat. Comm.*, **32**, 757.

[58] R. Resta and A. Baldereschi. (1981), *Phys. Rev. B*, **23**, 6615.

[59] C. Hogan and C. Patterson. *To be published.*

[60] R. Car, E. Tosatti, S. Baroni, and S. Leelaprute. (1981), *Phys. Rev. B*, **24**, 985.

[61] G. Engel and B. Farid. (1993), *Phys. Rev. B*, **47**, 15931.

[62] M. Hybertson and S. Louie. (1985), *Phys. Rev. Lett.*, **55**, 1418.

[63] N. Hamada, M. Hwang, and A. Freeman. (1990), *Phys. Rev. B*, **41**, 3620.

[64] W. von der Linden and P. Horsch. (1988), *Phys. Rev. B*, **37**, 8351.

[65] R. Hott. (1991), *Phys. Rev. B*, **44**, 1057.

[66] M. Yu, S. Ulloa, and D. Drabold. (2000), *Phys. Rev. B*, **61**, 2626.

[67] R. Godby. *Private communication.*

[68] F. Gygi and A. Baldereschi. (1986), *Phys. Rev. B*, **34**, 4405.

[69] Landolt-Börnstein. *Numerical Data and Functional Relationships in Science and Technology.* Springer, New York, (1982).

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

[71] S. Dall'Olio, R. Dovesi, and R. Resta. (1997), *Phys. Rev. B*, **56**, 10105.

[72] R. Del Sole, L. Reining, and R. Godby. (1994), *Phys. Rev. B*, **49**, 8024.

[73] W.-D. Schöne and A. Eguiluz. (1998), *Phys. Rev. Lett.*, **81**, 1662.

[74] C. Patterson and D. Herrendörfer. (1997), *J. Vac. Sci. Technol. A*, **15**, 3036.

[75] N. Arzate and B. Mendoza. *Preprint.*

[76] A. Bagchi, R. Barrera, and R. Fuchs. (1982), *Phys. Rev. B*, **25**, 7086.

[77] B. Mendoza and W. Mochan. (1996), *Phys. Rev. B*, **53**, R10473.

[78] M. Born and K. Huang. *Dynamical Theory Of Crystal Lattices.* Clarendon Press, Oxford, (1988).

[79] B. Nijboer and F. De Wette. (1958), *Physica*, **24**, 422.

[80] D. Aspnes and A. Studna. (1985), *Phys. Rev. Lett.*, **54**, 1956.

[81] T. Yasuda, L. Mantese, U. Rossow, and D. Aspnes. (1995), *Phys. Rev. Lett.*, **74**, 3431.

[82] J. McIntyre and D. Aspnes. (1971), *Surf. Sci.*, **24**, 417.

[83] C. Hogan and C. Patterson. (1998), *Surf. Sci.*, **402-404**, 586.

[84] L. Kipp, D. Biegelsen, J. Northrup, L.-E. Swartz, and R. Bringans. (1996), *Phys. Rev. Lett.*, **76**, 2810.

[85] T. Yasuda, D. Aspnes, D. Lee, C. Bjorkman, and G. Lucovsky. (1994), *J. Vac. Sci. Technol. A*, **12**, 1152.

[86] K. Uwai and N. Kobayashi. (1997), *Phys. Rev. Lett.*, **78**, 959.

[87] R. Del Sole and G. Onida. (1999), *Phys. Rev. B*, **60**, 5523.

[88] C. Wijers and P. de Boeij. (2000), *Submitted to Phys. Rev. B.*

[89] S. Boys. (1950), *Proc. Roy. Soc.*, **A200**, 542.

[90] V. Saunders. *Methods in Computational Molecular Physics.* Reidel, (1983).

[91] W. Hanke and L. Sham. (1975), *Phys. Rev. B*, **12**, 4501.

[92] T. Zivkovic and Z. Maksic. (1968), *J. Chem. Phys.*, **49**, 3083.

[93] L. McMurchie and E. Davidson. (1978), *J. Comp. Phys.*, **26**, 218.

[94] V. Saunders. *Computational Techniques in Quantum Chemistry and Molecular Physics.* Reidel, (1975).

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

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

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

[98] A. Dal Corso. *Quantum-Mechanical Ab-initio Calculation of the Properties of Crystalline Materials.* Springer-Verlag, (1996).

[99] A. Forsyth, T. Josefsson, and A. Smith. (1996), *Phys. Rev. B*, **54**, 14355.

[100] M. Towler. *The use and optimization of Gaussian basis sets in periodic Hartree-Fock calculations.* University of Torino, (1995).