

Towards ab initio calculation of electron energies in semiconductors

Citation for published version (APA):

Farid, B. (1989). *Towards ab initio calculation of electron energies in semiconductors*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Applied Physics and Science Education]. Technische Universiteit Eindhoven. <https://doi.org/10.6100/IR306688>

DOI:

[10.6100/IR306688](https://doi.org/10.6100/IR306688)

Document status and date:

Published: 01/01/1989

Document Version:

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

Please check the document version of this publication:

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

[Link to publication](#)

General rights

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

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

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

www.tue.nl/taverne

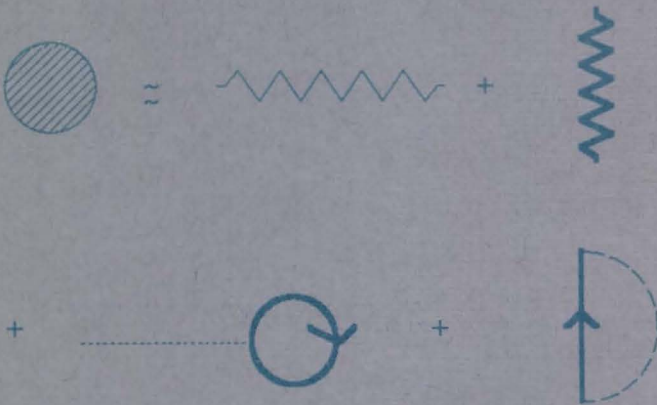
Take down policy

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

openaccess@tue.nl

providing details and we will investigate your claim.

**TOWARDS *AB INITIO* CALCULATION OF
ELECTRON ENERGIES
IN
SEMICONDUCTORS**



BEHNAM FARID

**TOWARDS *AB INITIO* CALCULATION OF
ELECTRON ENERGIES
IN
SEMICONDUCTORS**

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit
Eindhoven, op gezag van de Rector Magnificus, Prof. ir. M. Tels, voor een
commissie aangewezen door het college van dekanen in het openbaar te
verdedigen op dinsdag 21 maart 1989 te 16.00 uur

door

BEHNAM FARID

geboren te Teheran (Iran)

Dit proefschrift is goedgekeurd door
de promotor prof. dr. W. van Haeringen
en de co-promotor dr. D. Lenstra

Aan mijn ouders
Aan mijn broers



Green's windmill stands above Sneinton Dale, Nottingham. An upper room was used as a study by George Green (1793-1841), once its owner and one of Britain's most distinguished mathematical physicists. Built in 1807, it fell into disuse in 1860 and was gutted by fire in 1947. The mill was restored (1979-1986) as a memorial to George Green.

CONTENTS

	Page	
CHAPTER 1	INTRODUCTION	1
1.1	Historical Background	1
1.2	Green-Function Approach	5
1.3	Outline	8
CHAPTER 2	GREEN-FUNCTION APPROACH TO MANY-ELECTRON SYSTEMS	13
2.0	Introduction	13
2.1	General Aspects of the Green Function Approach	13
2.2	Diagrammatic Representation for M	17
2.3	General Connection Between the One-Particle Green Function and Excitation Energies of the Many-Electron System	20
2.4	Expressing G in terms of Quasi-Particle Wave Functions	21
2.5	Approximation Schemes	26
2.6	The Bubble Approximation	29
CHAPTER 3	PROPERTIES OF THE SELF-ENERGY FUNCTION IN THE GW SCHEME	33
3.0	Introduction	33
3.1	Relating the Self-Energy Function and the Quasi-Particle Excitation Structure in the GW Approximation	33
3.2	Difficulties in the Calculation of the GW Self-Energy Function	36
3.3	Contour-Deformation Procedure	38
3.4	Analytic Properties of the GW Self-Energy Function	45
3.5	Other Properties of the GW Self-Energy Function	50
3.6	Some Approximation Methods within the GW Scheme	53

CHAPTER 4	TOWARDS ACTUAL CALCULATION OF THE GW SELF-ENERGY FUNCTION IN THE QUASI-PARTICLE APPROXIMATION	61
4.0	Introduction	61
4.1	Integral Relations for the Bubble Polarization and the Screened Interaction Function	62
4.2	Quasi-Particle Approximation of the Bubble Polarization Function	68
4.3	Exploitation of Symmetry Properties	71
4.4	The Analytic Linear Tetrahedron Method	73
4.5	Numerical Determination of the Bubble Polarization Function (RPA) at Real Energies: A Novel Method of Calculating the RPA Dielectric Function for a Semiconductor at Real Energies	83
4.6	Back to M^{GW}	95
CHAPTER 5	CALCULATIONAL PROSPECTS	103
5.0	Introduction	103
5.1	Basic Expressions	104
5.2	Computational Aspects	110
APPENDIX A	DIAGRAMMATIC APPROACH TO THE CALCULATION OF THE ONE-PARTICLE GREEN AND SELF-ENERGY FUNCTION	119
APPENDIX B	LINEAR-RESPONSE APPROACH TO THE DETERMINATION OF THE ONE-PARTICLE GREEN AND SELF-ENERGY FUNCTION	133
REFERENCES		147
SUMMARY		157
SAMENVATTING		161
CURRICULUM VITAE		165

CHAPTER 1

INTRODUCTION

1.1 Historical Background

According to quantum theory, the physical properties of a system such as a solid can in principle be obtained from its wave function which, in a non-relativistic approximation, satisfies the Schrödinger wave equation. Observable quantities are then obtained by taking the expectation values of the corresponding operators in the state of the system. The wave function corresponding to a state depends on the coordinates of all composing particles.

From this point of view, it should be sufficient to know the constituent elements of the system and the kind of interaction between them in order to deduce all the properties of the system. It is well-known that solving the Schrödinger equation for a many-particle system is impossible if interparticle interaction is completely accounted for. In fact, due to the interaction, we do not succeed in separating the Schrödinger equation into independent equations for each particle. Strictly speaking, it is only the system as a whole which can be considered an individual entity. In view of these considerations, it is surprising that one-particle schemes have appeared in which several physical properties of interacting many-electron systems could be adequately dealt with.

Soon after its successful application in describing the energy spectrum of the hydrogen atom in 1926, the Schrödinger equation found application in clarifying the spectra of more complex atoms. In 1928, Hartree [1] suggested, on the basis of plausibility, that the behavior of each electron in a many-electron system may be described by means of a wave function satisfying a Schrödinger equation (Hartree equation) in which the effect of all other electrons as well as the charged nuclei are taken into account by means of a classical electrostatic field. In 1930 both Fock [2] and Slater [3] suggested the use of many-electron wave functions of the determinant type in order to take proper account of the antisymmetry requirement for spin $\frac{1}{2}$ particles. In combination with a variational principle, this led to the well-known Hartree-Fock equations [4,5]. Hartree-Fock equations distinguish themselves from Hartree equations by an extra non-

classical term containing an exchange interaction energy. This term can be shown to have the effect of a repulsive potential preventing electrons of equal spin from approaching each other [6,7]. The problem with the Hartree or Hartree-Fock equations for real systems is that *a priori* knowledge of the one-electron wave functions or at least the electron density is essential. This problem was circumvented in the self-consistent field method, in which first some initial guess concerning the potential felt by an individual electron is made whereafter the problem is solved selfconsistently.

Until two decades ago the activity of physicists involved in band structure calculations of solids was focused on the construction of appropriate "guess" potentials, so that a reliable band structure could be calculated without having to carry through the self-consistency procedure. The main reason for this kind of pragmatic approach should be found in the restricted computational possibilities. A full calculation within the Hartree-Fock scheme for a crystal is a task which is as yet unattainable; only in the last fifteen years have some Hartree-Fock kind of calculations been done in combination with the local-basis-function method (LBF) [8], the linear combination of atomic orbitals (LCAO) [9,10,11], the local-density approximation (LDA) [12] and the muffin-tin-orbital method [13], to name a few.

In 1964, Hohenberg and Kohn (HK) [14] introduced a theorem which forms the foundation of what is now known as the density-functional theory (DFT). This theorem states that the total energy of a many-electron system in some external potential is the minimum value of an energy functional of the electron density, which apart from a classical electrostatic contribution due to the external potential, is a universal functional. The minimum value of this energy functional corresponds to the *exact* electron density of the system. The explicit form of the universal functional was, and still is, unknown. Nevertheless, Kohn and Sham in 1965 [15], proposed a reasonable *Ansatz* concerning the energy functional to derive one-electron wave equations similar to the HF equations. These equations are referred to as Kohn-Sham (KS) equations. In principle, the effective potential present in these equations is a non-local functional of the electron density, but in the local density approximation one assumes a local dependence on density. Due to the known and simple forms of the functionals in this approximation, the self-consistent calculation of the one-electron wave functions in the LDA scheme is equally simple as in the Hartree

scheme; it has the advantage, however, that both exchange and correlation effects have been taken into account.

The KS-LDA one-electron scheme has proved its value in accurate calculation of *ground-state* properties of atoms and molecules [16], metals [17], semiconductors [18-20], surfaces [21] and defects [22]. However, the electron band structure in semiconductors as calculated in the KS scheme, shows energy gaps which are too small. The underestimation of gap energies is sometimes as large as fifty percent of the experimental values. This is completely in contrast with the result of HF calculations, which yield band gaps that are too large, the overestimation varying from two to five times the actual values [11-13,23-26]. Despite the lack of any theoretical justification as to the validity of assuming the eigenvalues of the KS equations to be the one-particle excitation spectrum of a many-electron system, the failure of LDA in correctly describing the energy gap structures in semiconductors was first thought to be a consequence of its local character. On the basis of successful applications in energy calculations for atoms and wide-gap insulators such as rare gas solids (Ne, Ar, etc.), Perdew and Zunger [27] suggested that a kind of correction potential, which they termed self-interaction correction (SIC), should be added to the LDA potential in order to achieve energy gap structures far better in agreement with experimental values. They have pointed out, however, that the self-interaction correction would not be the cure for all ills concerning the band gap structure. Heaten *et al* [28], reported a remarkable improvement in band gap and core levels of the semiconductor LiCl by application of the SIC method. To our knowledge, there are no reports on the application of SIC to other semiconductors. For completeness we mention that Perdew and Norman [29], after simplifying the self-interaction correction, were able to connect the latter with the real part of a self-energy function. It is this function that will play a central role in the present thesis.

Another approximation beyond LDA, is what is known as weighted density approximation (WDA) [30,31]. This approximation essentially introduces some non-locality effects. Kerker [32], reported in this connection on a substantial improvement over LDA in the bandgap energy of silicon. However, Wang and Pickett [33] noted that the success of WDA in Kerker's calculation could be due to a questionable change of a prefactor in the Kohn-Sham exchange potential. Indeed, Perdew and Levy [34] refer to unpublished calculations on

silicon by von Barth and Car, in which only small improvement over LDA is obtained. Perdew and Levy [34], and at the same time, Sham and Schlüter [35], concluded that at least a substantial fraction of the energy band "error" in LDA must be found in the Kohn-Sham formalism itself. Sham and Schlüter [35], derive an expression in which the LDA gap deficit in semiconductors is related to a discontinuity in the exchange-correlation potential-energy functional at the ground-state density. This discontinuity is peculiar for many-electron systems with an energy gap; it is completely disregarded in the LDA scheme. Sham [36,37] derives within the exact many-body theory, along with some assumptions, the above-mentioned expression for the gap-energy correction.

In order to understand the role of the discontinuity in the exchange-correlation potential of a semiconductor, with respect to the energy-gap deficit, one should note that the gap energy is the difference in ground-state energy differences $E_{N+1} - E_N$ and $E_N - E_{N-1}$, where $N+1$, N and $N-1$ refer to the number of electrons in the semiconductor. By N we mean the number of electrons for which the electron system is fully charge compensated by the ionic background. It seems that LDA (or DFT) is unable to describe an $(N+1)$ -electron system as a result of the discontinuous behavior of the exact exchange-correlation potential. A different view point is adopted by Gunnarsson and Schönhamer [38,39], who argue that the gap discrepancy is mainly due to the *approximate* (local) nature of LDA rather than to the discontinuity in the exchange-correlation potential, which in their opinion is *not* substantial in general. However, the numerical calculations by Manghi *et al* [40] on bulk GaAs indicate that the gap energy in a non-local DFT framework still remains too small.

In all of the above-mentioned approximation schemes, drastically different values for the energy gap are found, none of which coincide with experimental values. This leads one to believe that no self-consistent independent-particle scheme exists from which a reliable gap structure follows for all semiconductors. In other words, all these procedures of incorporating the mutual Coulomb interaction between electrons in an effective one-electron potential appear to be too rough as far as *excitation* properties are concerned. One therefore has to return to the original many-electron Hamiltonian and to reconsider the effect of the electron-electron interaction.

1.2 Green-Function Approach

The treatment of many-body systems with mutual interaction between the particles is most conveniently given in terms of Green functions. More specifically, the gap structure in a semiconductor should follow from the properties of the *one-electron* Green function $G(\mathbf{r},t;\mathbf{r}',t')$, where the latter is defined as the expectation value in the many-electron ground state of a time-ordered product of a *creation* operator for an electron at the space-time point \mathbf{r},t and an *annihilation* operator at \mathbf{r}',t' . This Green function probes, in a way, what an *additional* electron or hole in the many-electron system experiences. As the behavior of an electron, when added to a system in the ground state, will largely reflect the excitation structure of the system, it follows that detailed knowledge of the one-electron Green function will reveal this structure. Aside from this rather phenomenological explanation, there exists a formal correspondence between the spectral structure of the one-particle Green function and the energies of all $(N+1)$ - and $(N-1)$ -particle states of a many-electron system. This correspondence is provided by the Lehmann representation [41].

Green functions were originally introduced by the British (miller and) mathematical physicist George Green [42] in the context of the theory of electricity and magnetism. Nowadays, all functions satisfying an inhomogeneous (integro-) differential equation with a Dirac delta function as the source term, are called Green functions. In our case, this integro-differential equation, which is referred to as the equation of motion of the one-particle Green function, can be obtained directly from the defining relation of the Green function by application of the equation of motion for the above-mentioned creation and annihilation field operators. The occurrence of a complicated function in this equation of motion for G , which is referred to as the self-energy function M , hampers a straightforward solution of the one-particle Green function from its equation of motion. A general method of attack to this problem is offered by perturbation theory. Expansion of both G and M in terms of increasing powers of the Coulomb interaction can formally be given. In this connection use is made of a theorem due to Wick [43]. The various terms in the series can be uniquely represented by means of Feynman diagrams [44]. Incidentally we mention that, by introducing the one-particle Green function belonging to some unperturbed Hamiltonian, the equation of motion of G can be transposed into an integral

equation, the so-called Dyson equation [45], the iterated (formal) solution of which is just the above-mentioned perturbation series. An essential problem in using the perturbation expansion of G is that, due to the long range of bare Coulomb interaction, the term-by-term summation of the series may be questioned.

In 1948, Feynman [46] formulated the laws of quantum mechanics by means of action principles, an idea (Lagrangian formulation of quantum mechanics) suggested earlier by Dirac [47]. A subsequent progress was made by Schwinger in 1951 [48], who introduced the so-called dynamical principle. This principle can be shown to be a differential form of Feynman's principle and, like the latter, gives an alternative formulation of quantum mechanics [49]. Using his dynamical principle, Schwinger in 1951 [50-54] derived an explicit relation for G in terms of some variational derivative.

Martin and Schwinger, in 1959, presented a paper [55] in which they dealt with many-body systems from a unified non-perturbative point of view. This work has formed a solid basis for much theoretical work concerning many-electron systems. It was Hedin [56] who, by employing the ideas of Martin and Schwinger in combination with those of Hubbard concerning the dynamical screening processes in many-electron systems [57], introduced a new method of calculating the one-particle Green function and the corresponding self-energy function. It should be mentioned that the theory of Hedin does not employ the dynamical principle of Schwinger, but rather is based on the Hamilton formulation of quantum mechanics and utilizes the theory of linear response (for the formulation of Hedin's theory according to Schwinger's principle see [54]). In the formulation of Hedin the relationship between G and M is expressed in terms of four equations which, in combination with Dyson's equation, form the basis for the self-consistent solution of G and M . Utilizing Hedin's equations one can obtain a perturbation series for the self-energy function M with the dynamically screened interaction W as the perturbation function, rather than the bare Coulomb interaction. Due to screening effects, the screened interaction is substantially "weaker" than the bare interaction. Concerning the range of the interaction, it can be shown that in uniform systems, such as metals, the screened interaction is effectively of short range. As can be expected, in systems such as covalent semiconductors with strong charge inhomogeneity, the long range of interaction cannot completely be screened away. The first-order term in

the expansion of M in terms of W is written as a product of G and W and due to this it is referred to as the GW self-energy function of Hedin.

Concerning the existence or non-existence of one-particle excitations in a many-electron system, one can, using the formal bi-orthonormal representation of G [58], arrive at one-electron Schrödinger-like equations, the solutions of which, under some assumptions, can be viewed as one-particle excitation functions. The related particles are called quasi-particles; the Schrödinger-like equations are termed quasi-particle equations [59-61].

The first reported work on the quasi-particle band structure calculation of crystals (all of the calculations to be mentioned below are carried out within the GW approximation) is due to Strinati *et al* [62], who employ a minimal-basis tight-binding approach. Contrary to later workers in the field, these authors go beyond the so-called random phase approximation of the dielectric function (see below), but it appears that their calculated band gap for crystal silicon is too large. Wang and Pickett [33,63] obtain, using many simplifying approximations, the quasi-particle band structure of silicon which is very close to the experimentally established values. On account of the observation that their obtained corrections to the LDA band energies are clearly energy dependent, they conclude that the energy dependence of the self-energy function is crucial in band structure calculations. In 1985 Hybertsen and Louie [64,65] reported on successful "*ab initio*" calculations of the quasi-particle band structures of silicon and diamond. Their calculated gap structures are in excellent agreement with experimental results. Although they called their calculations "first principles", the dynamically screened interaction employed by them was based on a plasmon-pole approximation of this interaction. Their justification for designating their calculations as "first principles" lies in the fact that the parameters in their plasmon-pole model are not adjustable but fixed by demanding that the inverse of their model dielectric function satisfies both the causality condition (Kramers-Kronig relation) and a so-called f sum rule.

Subsequent numerical results are reported in a number of papers by Godby *et al*. The first in the series is a report on the quasi-particle band structure calculation of silicon [66]. Apart from excellent agreement with experimental values, they also justify numerically that the underestimation of the gap energy in semiconductors by LDA is to a large extent attributable to the discontinuity of the exact exchange-correlation potential. The second in the

series [67], gives a report on the success of (GW) quasi-particle band structure calculations in reproducing the band structures of GaAs and AlAs (there is, however, a discrepancy about the L conduction-band minimum in AlAs which is found to be 0.8 ± 0.2 eV above the X minimum rather than 0.3 eV). In a subsequent paper [68], the authors show that more than eighty percent of the LDA bandgap deficits in Si, GaAs and AlAs are indeed due to the DFT method itself. Godby *et al* were the first to present rather accurate numerical results for quasi-particle energies using neither any adjustable parameter nor plasmon pole type of approximation. However, it can be shown that the self-energy function on which their calculations are based is in fact incomplete [69]. Surprisingly, their accurate results are obtained by using a simple truncated Taylor series based on the incomplete expression for M . Von der Linden and Horsch [37], making use of some generalized plasmon-pole model, also report on the ability of the GW scheme to give reliable quasi-particle band structures.

In all of the above-mentioned papers two aspects concerning the screened interaction play an essential role in the correct calculation of the quasi-particle band structures: (i) its energy dependence, which should be properly incorporated (ii) its non-vanishing off-diagonal elements in the plane-wave representation, describing local-field effects, which should not be neglected.

We recall that band structure calculation techniques in which electrons are assumed to feel an effective real-valued one-electron potential, invariably lead to gap structures which deviate more or less severely from the experimentally established structures. Indeed, there is growing evidence that no self-consistent independent-particle scheme exists from which a reliable gap structure follows for all semiconductors. One is forced therefore to consider the effect of electron-electron interaction anew.

1.3 Outline

This thesis deals with the Green-function approach to the determination of the electronic structure in semiconductors. Both from the point of view of the insufficiency of effective-potential methods to describe one-particle excitation energies and from the fact that most of the existing GW treatments, although very successful and promising, still suffer from questionable assumptions, we

conclude that a general investigation into the prospects of first-principles calculation of band structures of semiconductors within the Green-function formalism is worthwhile. We will focus to a large extent on the GW approach as this approach, according to growing evidence in the literature, is very likely successful in correct predictions of electronic excitation energies, even though no *a priori* proof has been given as yet. We will not contribute to the fundamental question whether the GW scheme is sufficient, but we will investigate whether and to what extent the GW scheme may be carried through without further approximations.

In chapter 2, we present general aspects of the Green-function theory. The convenient use of Feynman diagrams leads to a systematic account of contributions to G and M to all orders in the electron-electron interaction. Presenting the material in this fashion leads in a natural way to the introduction of the *dynamically screened* interaction function $W(\mathbf{r}\mathbf{t},\mathbf{r}'\mathbf{t}')$ [instead of the unscreened static Coulomb interaction function $v(\mathbf{r}-\mathbf{r}')\delta(\mathbf{t}-\mathbf{t}') = (4\pi\epsilon_0|\mathbf{r}-\mathbf{r}'|)^{-1} \times \delta(\mathbf{t}-\mathbf{t}')$] and of the related polarization function $P(\mathbf{r}\mathbf{t},\mathbf{r}'\mathbf{t}')$. An important part of this material is given in appendix A, where we present the general principles of the diagrammatic expansions of the one-particle Green function and the self-energy function. New in this appendix is a diagrammatic derivation of the Hedin equations. The consequences of choosing an unperturbed Hamiltonian with a non-local potential are studied and new types of diagrams introduced. Also the concept of skeleton diagrams is completely dealt with in connection with one-particle Green function. Appendix B contains an alternative derivation of Hedin's equations, which are rigorously rederived using a variational technique.

One of the further aims of chapter 2 is to present the forms to which G and M reduce within various effective one-electron potential schemes, such as the Hartree (H), Hartree-Fock (HF) and local density functional (LDF) scheme. Also other possible effective-potential schemes will be discussed. By presenting matters in this way we gain insight into those contributions to G and M which are missing in each scheme. Crucial in our analysis in chapter 2 will be the observation that the functions G , M , W and P can all be expressed in terms of so-called quasi-particle wave functions. These satisfy quasi-particle wave equations, in which the self-energy function M , as distinct from an effective potential v_{eff} , plays an essential role. As M , unlike v_{eff} , is a non-Hermitian operator, the quasi-particle eigenvalues are complex-valued. The real parts form

the excitation spectrum; while $2/\hbar$ times the imaginary parts are interpreted as the inverse lifetimes of quasi-particles (electrons or holes). All effective-potential schemes are based on approximate self-energy functions $M(\mathbf{r},\mathbf{r}',t-t')$ which are written in the factorized form $A(\mathbf{r},\mathbf{r}')\delta(t-t')$, where $A(\mathbf{r},\mathbf{r}')$ is Hermitian. As mentioned before, effective-potential schemes generally lead to incorrect results for the energy spectrum. An important question therefore concerns non-factorizable and/or non-Hermitian parts of the exact function M that are essential in obtaining the correct energy spectrum. The chapter ends by introducing both the GW approximation of the self-energy function and the bubble approximation of the polarization function. In the remaining chapters we invariably disregard all contributions to M which are of second and higher order in W .

In chapter 3 we discuss the problems encountered in the direct evaluation of the GW self-energy function in the plane-wave representation. Subsequently, we introduce a contour-deformation procedure to avoid a number of computational difficulties. A thorough investigation of the GW self-energy function in the complex energy plane will be given. The last part of chapter 3 is devoted to some common and uncommon approximation schemes within GW. In addition to a formal presentation, we give the physical significance of each approximation and try to indicate the bounds of validity of each scheme.

In chapter 4, our investigation of the GW self-energy function is continued by deriving general integral relations for the bubble polarization and screened interaction function. Various useful relations will be presented, based on space-group symmetry. We discuss the analytic linear tetrahedron method, especially suited to deal with singularities in the integrand of an integral over wave vectors, but we advocate in a separate paper (section 4.5) the use of a properly adjusted special-points method together with a new method by which the polarization function can be obtained by solving a Fredholm integral equation [70]. An important section deals with the difficulties encountered in calculating M , connected with the occurrence of singular functions in the integrand of the involved $1/Bz$ \mathbf{k} integration.

In chapter 5 we present our final expression for the plane-wave matrix elements of the self-energy function M . This expression is free of numerical intricacies, and the extent to which actual energy band calculations based on this expression are indeed achievable, is discussed. The general conclusion is that

a fully self-consistent *ab initio* calculation of quasi-particle band structures is not feasible at present. If one restricts oneself to a first-iteration step of the self-consistency procedure, the prospects are somewhat better. We show the feasibility of such a calculation within a fifteen-band model of a semiconductor. Our estimates for computation time given in chapter 5 are somewhat conservative as we have taken a normal computing system without vector and parallel computing facilities as our reference. For that reason a first-iteration-step calculation for a more realistic model is *possibly* also within reach.

CHAPTER 2

GREEN-FUNCTION APPROACH TO MANY-ELECTRON SYSTEMS

2.0 Introduction

In this chapter a general description is given of the Green-function approach to many-electron systems such as met in (semiconducting) crystals. In section 2.1 we define the one-particle Green function G and show which integro-differential equation it satisfies. In section 2.2 we introduce the diagrammatic representation for the related self-energy function M . In this connection reference is made to two rather lengthy appendices A and B, in which the diagrammatic expansion and the general connection between M and G are discussed in full detail. Section 2.3 is devoted to a discussion on the connection between G and the excitation energies of the system. In section 2.4 we introduce the quasi-particle concept and show how the Green function G can approximately be expressed in terms of quasi-particle wave functions and energies. Section 2.5 is devoted to discussions of the Hartree, the Hartree-Fock, the Local Density Functional, the Slater $X\alpha$ and GW approximation schemes. As the subsequent chapters 3, 4 and 5 are exclusively devoted to the discussion of the GW scheme, we consider it worthwhile to clarify the latter's position in relation to the more commonly known one-electron effective-potential schemes. Section 2.6 introduces the bubble-approximation scheme for the polarization function. Use of the latter scheme implies a further approximation not conflicting with the GW scheme.

2.1 General Aspects of the Green-Function Approach

In describing the properties of a many-electron system it is most helpful to make use of operators that create or annihilate electrons at a given space-time point r,t . These operators are usually written $\hat{\psi}^\dagger(r,t)$ and $\hat{\psi}(r,t)$. They fulfill the Heisenberg equations of motion [61,71]

$$i\hbar \frac{\partial \hat{\psi}(\mathbf{r}t)}{\partial t} = [\hat{\psi}(\mathbf{r}t), \hat{H}]_{-}, \quad (2.1a)$$

$$i\hbar \frac{\partial \hat{\psi}^{\dagger}(\mathbf{r}t)}{\partial t} = [\hat{\psi}^{\dagger}(\mathbf{r}t), \hat{H}]_{-}, \quad (2.1b)$$

where \hat{H} is the Hamilton operator of the interacting N-electron system, which does not depend on time, and in which $[\ , \]_{-}$ stands for the commutation operation. The formalism to be described below will be applied to crystalline materials, such as semiconductors; N represents the number of electrons for which the whole crystalline system is charge neutral. The operators $\hat{\psi}^{\dagger}$ and $\hat{\psi}$ satisfy the usual equal-time anticommutation relations

$$[\hat{\psi}^{\dagger}(\mathbf{r}t), \hat{\psi}^{\dagger}(\mathbf{r}'t)]_{+} = [\hat{\psi}(\mathbf{r}t), \hat{\psi}(\mathbf{r}'t)]_{+} = 0, \quad (2.2a)$$

$$[\hat{\psi}^{\dagger}(\mathbf{r}t), \hat{\psi}(\mathbf{r}'t)]_{+} = \delta(\mathbf{r}-\mathbf{r}'). \quad (2.2b)$$

The above operators are defined in the so-called Heisenberg picture, and as such time dependent. In this picture the state vector of the system is time independent and is indicated $|\Psi_N\rangle_H$. We may alternatively write operators and states in the Schrödinger picture, leading to (S stands for Schrödinger)

$$|\Psi_N(t)\rangle_S = e^{-i\hat{H}t/\hbar} |\Psi_N\rangle_H, \quad (2.3a)$$

$$\hat{\psi}^{\dagger}(\mathbf{r}) = e^{-i\hat{H}t/\hbar} \hat{\psi}^{\dagger}(\mathbf{r}t) e^{i\hat{H}t/\hbar}, \quad (2.3b)$$

$$\hat{\psi}(\mathbf{r}) = e^{-i\hat{H}t/\hbar} \hat{\psi}(\mathbf{r}t) e^{i\hat{H}t/\hbar}. \quad (2.3c)$$

The Schrödinger operators $\hat{\psi}^{\dagger}(\mathbf{r})$ and $\hat{\psi}(\mathbf{r})$ will be used below in the expression for the time-independent operator \hat{H} .

The one-particle Green function is defined as, up to a numerical factor $-i$ (i stands for the imaginary unit), the expectation value in the ground state

$|\Psi_N\rangle_H$ of the time-ordered product of operators $\hat{\psi}^\dagger(\mathbf{r}_1 t_1)$ and $\hat{\psi}(\mathbf{r}_2 t_2)$ at different space-time points [72,73],

$$G(1,2) = -i \langle \Psi_N | \mathcal{T} \{ \hat{\psi}(1) \hat{\psi}^\dagger(2) \} | \Psi_N \rangle_H, \quad (2.4)$$

where the arguments 1,2 are short-hand notations for the space-time points $\mathbf{r}_1 t_1$ and $\mathbf{r}_2 t_2$, respectively, and where $\mathcal{T} \{ \hat{\psi}(1) \hat{\psi}^\dagger(2) \}$ equals either $\hat{\psi}(1) \hat{\psi}^\dagger(2)$ or $-\hat{\psi}^\dagger(2) \hat{\psi}(1)$ depending on whether $t_1 > t_2$ or $t_1 < t_2$ (\mathcal{T} is the Fermion time-ordering operator). The function G owes its name Green function from the fact that it satisfies a (non-linear) Green-type equation, as will be shown below (see (2.7)).

The total Hamiltonian \hat{H} is given by

$$\hat{H} = \hat{T} + \hat{U} + \hat{V}, \quad (2.5a)$$

$$\hat{T} = \int d^3r \hat{\psi}^\dagger(\mathbf{r}) \left[-\frac{\hbar^2}{2m} \nabla^2 \right] \hat{\psi}(\mathbf{r}), \quad (2.5b)$$

$$\hat{U} = \int d^3r \hat{\psi}^\dagger(\mathbf{r}) u(\mathbf{r}) \hat{\psi}(\mathbf{r}), \quad (2.5c)$$

$$\hat{V} = \frac{1}{2} \int d^3r d^3r' \hat{\psi}^\dagger(\mathbf{r}) \hat{\psi}^\dagger(\mathbf{r}') v(\mathbf{r}-\mathbf{r}') \hat{\psi}(\mathbf{r}') \hat{\psi}(\mathbf{r}), \quad (2.5d)$$

where $u(\mathbf{r})$ is the "external" potential energy (in a crystal $u(\mathbf{r})$ is the potential energy due to the periodic array of nuclei), and $v(\mathbf{r}-\mathbf{r}') = e^2 / (4\pi\epsilon_0 |\mathbf{r}-\mathbf{r}'|)$ is the electron-electron Coulomb potential energy with e the electron charge and ϵ_0 the vacuum permittivity. With the purpose of establishing a transparent scheme which allows orderly comparison of the various existing approximation methods, we will write \hat{H} in a different way, as follows:

$$\hat{H} = \hat{H}_0 + \hat{H}_1, \quad (2.6a)$$

$$\hat{H}_0 = \hat{T} + \hat{U} + \hat{Z}_\ell + \hat{Z}_{n\ell} \quad (2.6b)$$

$$\hat{H}_1 = \hat{V} - \hat{Z}_\ell - \hat{Z}_{n\ell} \quad (2.6c)$$

where

$$\hat{Z}_\ell = \int d^3r \hat{\psi}^\dagger(\mathbf{r}) z_\ell(\mathbf{r}) \hat{\psi}(\mathbf{r}), \quad (2.6d)$$

$$\hat{Z}_{n\ell} = \int d^3r d^3r' \hat{\psi}^\dagger(\mathbf{r}) z_{n\ell}(\mathbf{r}, \mathbf{r}') \hat{\psi}(\mathbf{r}'). \quad (2.6e)$$

Here the real-valued functions $z_\ell(\mathbf{r})$ and $z_{n\ell}(\mathbf{r}, \mathbf{r}')$ may be arbitrarily chosen. They will later play a role as *local* and *non-local* one-electron effective potentials. The function $z_\ell(\mathbf{r})$ can, for instance, be taken equal to the well-known Hartree potential [74], the Slater $X\alpha$ potential [7], the Kohn-Sham local potential [15], or whatsoever. Similarly, the function $z_{n\ell}(\mathbf{r}, \mathbf{r}')$ can be taken equal to zero, the Hartree-Fock potential [75], or any other non-local potential. The idea is to bring first the formalism as far as possible with unspecified z_ℓ and $z_{n\ell}$ and then to make specific choices for z_ℓ and $z_{n\ell}$.

By using (2.1a) it is straightforward (but tedious) to derive the following integro-differential equation for the function $G(1,2)$:

$$\begin{aligned} & [i\hbar \frac{\partial}{\partial t_1} + \frac{\hbar^2}{2m} \nabla_1^2 - u(\mathbf{r}_1)] G(\mathbf{r}_1 t_1, \mathbf{r}_2 t_2) \\ & - \int d^3r_3 dt_3 \left[\delta(t_1 - t_3) \{z_\ell(\mathbf{r}_1) \delta(\mathbf{r}_1 - \mathbf{r}_3) + z_{n\ell}(\mathbf{r}_1, \mathbf{r}_3)\} \right. \\ & \left. + \hbar M(\mathbf{r}_1 t_1, \mathbf{r}_3 t_3) \right] G(\mathbf{r}_3 t_3, \mathbf{r}_2 t_2) = \hbar \delta(t_1 - t_2) \delta(\mathbf{r}_1 - \mathbf{r}_2), \end{aligned} \quad (2.7)$$

where the function M is defined through the relation

$$\begin{aligned} & {}_H \langle \Psi_N | \mathcal{S} \left\{ \left[\hat{\psi}(\mathbf{r}_1 t_1), \hat{V} - \hat{Z}_\ell - \hat{Z}_{n\ell} \right]_- \hat{\psi}^\dagger(\mathbf{r}_2 t_2) \right\} | \Psi_N \rangle_H \\ & \equiv i\hbar \int d^3r_3 dt_3 M(\mathbf{r}_1 t_1, \mathbf{r}_3 t_3) G(\mathbf{r}_3 t_3, \mathbf{r}_2 t_2). \end{aligned} \quad (2.8)$$

The function M is called the self-energy function or mass operator. It incorporates all interaction effects which have not yet effectively been taken into

account by the potentials z_ℓ and $z_{n\ell}$. From (2.8) it is clear that M depends on z_ℓ and $z_{n\ell}$. It is furthermore observed from (2.7) that those contributions to $M(r_1, t_1, r_3, t_3)$ which factorize with $\delta(t_1 - t_3)$ or $\delta(t_1 - t_3)\delta(r_1 - r_3)$ can equally well be accounted for through a simple redefinition of the functions $z_{n\ell}(r_1, r_3)$ or $z_\ell(r_1)$.

Many approximation schemes exist in which M is put equal to zero, while certain choices for z_ℓ and $z_{n\ell}$ are made. In such cases we refer to the solution of (2.7) as the "unperturbed" Green function G^0 , though by the choice of z_ℓ and $z_{n\ell}$ the function G^0 may already incorporate some electron-electron interaction effects. The equation for G^0 reads

$$\begin{aligned}
 & [i\hbar \frac{\partial}{\partial t_1} + \frac{\hbar^2}{2m} \nabla_1^2 - u(r_1) - z_\ell(r_1)] G^0(r_1, t_1, r_2, t_2) \\
 & - \int d^3r_3 z_{n\ell}(r_1, r_3) G^0(r_3, t_1, r_2, t_2) = \hbar \delta(t_1 - t_2) \delta(r_1 - r_2). \quad (2.9)
 \end{aligned}$$

From (2.7) and (2.9) we may obtain Dyson's equation

$$G(1,2) = G^0(1,2) + \int d(3)d(4) G^0(1,3) M(3,4) G(4,2), \quad (2.10)$$

or, symbolically, $G = G^0 + G^0 M G$. The validity of (2.10) can easily be verified by substitution in (2.7). Note that contrary to (2.7), equation (2.9) is a linear equation.

2.2 Diagrammatic Representation for M

A central result of the general theory of many-electron systems is that the Green function $G(1,2)$ is expressible as a sum of an infinite number of terms in increasing powers of the electron-electron Coulomb interaction [71,73]. Each term corresponds to a multiple space-time integral of a specific series of products of functions G^0 , v , z_ℓ and $z_{n\ell}$. By identifying this expression for G with Dyson's expression (2.10), one can also extract an expansion for M in terms of functions G^0 , v , z_ℓ and $z_{n\ell}$. It is of great help to represent each contribution to M by means of a corresponding Feynman diagram. The natural ordering of these

diagrams is in accordance with increasing orders in v , z_ℓ and $z_{n\ell}$. A well-known complication is that various contributing diagrams lead to divergent contributions. It is believed, however, that proper summation of the infinite series of diagrams leads to finite, meaningful results. The reader is referred to Appendix A where the expansions of G and M in terms of diagrams are explained in some detail, while furthermore a diagrammatic derivation is presented of a set of four coupled equations, the so-called Hedin equations [56,76], which, together with Dyson's equation (2.10) determine both G and M . A more rigorous derivation of Hedin's equations, in which no reference is made to divergent contributions, is given in Appendix B.

The above process of expanding M (see Appendix A) leads to

$$\begin{aligned}
 M(1,2) = & \text{Diagram 1} = \text{Diagram (a1)} + \text{Diagram (a2)} \\
 & + \text{Diagram (b)} + \text{Diagram (c)} \\
 & + \text{Diagram (d)} + \text{Diagram (e)}
 \end{aligned}$$

* All skeleton diagrams of higher order
in the interaction $v(i,j) = i \text{-----} j$.

(2.11)

Here a skeleton diagram is defined by the requirement that none of the thick lines contains an internal M diagram itself. The calculational prescription is as follows: a thick line with an arrow directed from 2 to 1 represents $G(1,2)$; a dotted line attached to 1 and 2 gives $v(1,2) = v(\mathbf{r}_1, \mathbf{r}_2) \delta(t_1 - t_2)$; diagram (a1) gives $-\hbar^{-1} z_\ell(\mathbf{r}_1) \delta(\mathbf{r}_1 - \mathbf{r}_2) \delta(t_1 - t_2)$ while diagram (a2) gives $-\hbar^{-1} z_{n\ell}(\mathbf{r}_1, \mathbf{r}_2) \delta(t_1 - t_2)$.

Sometimes, as in diagrams (b) and (c) of (2.11), the prescription leads to equal-time arguments in Green functions. In that case the prescription is to take $G(r_1 t_1, r_2 t_1^*)$ where $t_1^* = t_1 + \eta$, with η positive and infinitesimally small. Each closed loop consisting of thick lines gives rise to a factor -2 (see the end of appendix a). Furthermore, each diagram has an additional factor $(i/\hbar)^n$ where n is the number of dotted lines. Finally one has to integrate over all *internal* space-time variables.

Dyson's equation (2.10) is diagrammatically represented by means of

$$\begin{array}{c} 1 \\ | \\ \uparrow \\ | \\ 2 \end{array} = \begin{array}{c} 1 \\ | \\ \uparrow \\ | \\ 2 \end{array} + \begin{array}{c} 1 \\ | \\ \uparrow \\ \bullet \\ \uparrow \\ | \\ 2 \end{array} \quad (2.12)$$

where the thin line with an arrow denotes $G^0(1,2)$. Equations (2.11) and (2.12) are the basic equations necessary to find both G and M in terms of G^0 . In order to complete the diagrammatic notions needed in what follows we also introduce the *dynamically screened* interaction function

$$W(1,2) = \begin{array}{c} 1 \\ | \\ | \\ | \\ 2 \end{array} = \begin{array}{c} 1 \\ | \\ | \\ | \\ 2 \end{array} + \begin{array}{c} 1 \\ | \\ \text{shaded diamond} \\ | \\ 2 \end{array} \quad (2.13)$$

where the shaded insertion stands for all topologically allowed subdiagrams (see (A.12)) which together stand for the polarization function $P(3,4)$. The introduction of W makes it possible to put together very specific subsets of M diagrams in (2.11) in a much more compact form. Unlike $v(1,2)$, the function $W(1,2)$ is no longer proportional to $\delta(t_1 - t_2)$; the function $P(3,4)$ accounts for *static* as well as *dynamic* screening of the interaction function $v(1,2)$.

2.3 General Connection Between the One-Particle Green Function and Excitation Energies of the Many-Electron System

We start by mentioning that $G(\mathbf{r}_1, \mathbf{r}_2; t_1, t_2) = G(\mathbf{r}_1, \mathbf{r}_2; t_1 - t_2)$. The Fourier transform of such a function $F(1,2)$ with respect to $t_1 - t_2$ at frequency ϵ/\hbar is given by

$$F(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = \int d(t_1 - t_2) \exp[i\epsilon(t_1 - t_2)/\hbar] F(1,2). \quad (2.14)$$

If the function $F(1,2)$ is continuous for all values of $t_1 - t_2$, the inverse relation reads

$$F(1,2) = \int \frac{d\epsilon}{2\pi\hbar} \exp[-i\epsilon(t_1 - t_2)/\hbar] F(\mathbf{r}_1, \mathbf{r}_2; \epsilon). \quad (2.15)$$

In case the function $F(1,2)$ has a finite discontinuity for some specific value of $t_1 - t_2$, the right-hand side of (2.15) is given the value $[F(1^+, 2) + F(1, 2^+)]/2$ at $t_1 - t_2$. Note that by construction (see (2.4)), G has such a discontinuity at $t_1 - t_2 = 0$.

By Fourier transforming the one-particle Green function according to (2.14), it can easily be shown [41,77,78] that

$$G(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = \hbar \sum_s f_s(\mathbf{r}_1) f_s^*(\mathbf{r}_2) \left\{ \frac{\Theta(\epsilon_s - \mu)}{\epsilon - \epsilon_s + i\eta} + \frac{\Theta(\mu - \epsilon_s)}{\epsilon - \epsilon_s - i\eta} \right\}, \quad (2.16)$$

where η is an infinitesimally small positive quantity, and where $\Theta(x) = 0$ or 1 depending on whether $x < 0$ or $x > 0$. For semiconductors, μ equals *some* energy in the energy-gap region of the system. Furthermore

$$f_s(\mathbf{r}) = \begin{cases} \text{H} \langle \Psi_{N-1,s} | \hat{\psi}(\mathbf{r}) | \Psi_N \rangle_{\text{H}}, & \text{if } \epsilon_s < \mu, \end{cases} \quad (2.17a)$$

$$\begin{cases} \text{H} \langle \Psi_N | \hat{\psi}(\mathbf{r}) | \Psi_{N+1,s} \rangle_{\text{H}}, & \text{if } \epsilon_s > \mu, \end{cases} \quad (2.17b)$$

in which $\{|\Psi_{N-1,s}\rangle_{\text{H}}\}_s$ and $\{|\Psi_{N+1,s}\rangle_{\text{H}}\}_s$ indicate the complete set of $(N-1)$ - and $(N+1)$ -particle eigenstates of H , with energy eigenvalues $E_{N-1,s}$ and $E_{N+1,s}$, respectively, and where

$$\epsilon_s = \begin{cases} E_N - E_{N-1, s} , & \text{if } \epsilon_s < \mu, \\ E_{N+1, s} - E_N , & \text{if } \epsilon_s > \mu, \end{cases} \quad (2.18a)$$

$$(2.18b)$$

E_N being the ground-state energy of the N -electron system. The representation (2.16) is the well-known Lehmann representation [41]. In this thesis we shall refer to the functions $f_s(\mathbf{r})$ as the Lehmann amplitudes, while the quantities ϵ_s will be referred to as the Lehmann energies. From (2.16) we observe that the excitation energies of the system show up as singularities of $G(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$ on the real energy axis. It should be noted that the Lehmann representation, although a correct representation, is of limited value in actual calculations as its evaluation requires complete knowledge of the many-electron wave functions of the interacting system.

Consistent with our assumption concerning the existence of an energy gap, we will define the gap energy [79] as $E_g \equiv (E_{N+1} - E_N) - (E_N - E_{N-1}) = E_{N+1} + E_{N-1} - 2E_N$, where $E_{N\pm 1}$ are the $(N\pm 1)$ -particle groundstate energies. Note that these states are not charge neutral, as they refer to many-electron wave functions with one electron more or less than the charge neutral N -electron ground state. Actually, in semiconductors the quantity $E_N - E_{N-1}$ is the smallest value one can choose for μ in (2.16); the largest value one can choose for μ is $E_{N+1} - E_N$. All values in between can be chosen as well. It remains to be shown in what sense the above-introduced gap can be related to the band-gap concept appearing in the conventional one-electron theory of energy band structures. This problem will be touched upon several times in the sequel.

In the next section we will introduce another representation, which resembles the Lehmann representation but is more accessible for a practical approach. The resemblance with the Lehmann representation is striking and we again identify its singularities with certain excitation energies of the system, the so-called quasi-particle energies.

2.4 Expressing G in terms of Quasi-Particle Wave Functions

Owing to the time independence of the Hamiltonian, we first note that the dependence of $M(1,2)$ on time is, similar to $G(1,2)$, via $t_1 - t_2$. If we Fourier

transform (2.7) with respect to t_1-t_2 we obtain

$$\begin{aligned}
 & [\epsilon + \frac{\hbar^2}{2m} \nabla_1^2 - u(r_1) - z_\ell(r_1)] G(r_1, r_2; \epsilon) \\
 & - \int d^3 r_3 z_n \ell(r_1, r_3) G(r_3, r_2; \epsilon) \\
 & - \hbar \int d^3 r_3 M(r_1, r_3; \epsilon) G(r_3, r_2; \epsilon) = \hbar \delta(r_1 - r_2). \tag{2.19}
 \end{aligned}$$

We will first generally prove that the solution of this equation may be written as [58,59,61,80]

$$G(r_1, r_2; \epsilon) = \hbar \sum_n \frac{\varphi_n(r_1; \epsilon) \psi_n^*(r_2; \epsilon)}{\epsilon - E_n(\epsilon)}, \tag{2.20}$$

where $\varphi_n(r; \epsilon)$ and $\psi_n(r; \epsilon)$ are solutions of

$$\begin{aligned}
 & [E_n(\epsilon) + \frac{\hbar^2}{2m} \nabla^2 - u(r) - z_\ell(r)] \varphi_n(r; \epsilon) - \int d^3 r' z_n \ell(r, r') \varphi_n(r'; \epsilon) \\
 & - \hbar \int d^3 r' M(r, r'; \epsilon) \varphi_n(r'; \epsilon) = 0, \tag{2.21}
 \end{aligned}$$

and

$$\begin{aligned}
 & [E_n^*(\epsilon) + \frac{\hbar^2}{2m} \nabla^2 - u(r) - z_\ell(r)] \psi_n(r; \epsilon) - \int d^3 r' z_n \ell(r, r') \psi_n(r'; \epsilon) \\
 & - \hbar \int d^3 r' M^\dagger(r, r'; \epsilon) \psi_n(r'; \epsilon) = 0, \tag{2.22}
 \end{aligned}$$

respectively. In (2.21) and (2.22) the eigenvalues $E_n(\epsilon)$ are generally complex-valued. Here M^\dagger is the Hermitian adjoint of M , that is,

$$M^\dagger(r, r'; \epsilon) = M(r', r; \epsilon)^*. \tag{2.23}$$

Writing (2.21) and (2.22) in short-hand notation as $(E_n(\epsilon) - \mathcal{L}(\epsilon))\varphi_n = 0$

and $(E_n^*(\epsilon) - \mathcal{L}^\dagger(\epsilon))\psi_n = 0$, respectively, we easily obtain the following identities

$$\begin{aligned} \langle \psi_m, \mathcal{L}(\epsilon)\varphi_n \rangle &= E_n(\epsilon) \langle \psi_m, \varphi_n \rangle \\ &= \langle \mathcal{L}^\dagger(\epsilon)\psi_m, \varphi_n \rangle = E_m(\epsilon) \langle \psi_m, \varphi_n \rangle, \end{aligned} \quad (2.24)$$

where the scalar product \langle , \rangle is defined by

$$\langle \psi, \varphi \rangle = \int d^3r \psi^*(r; \epsilon) \varphi(r; \epsilon). \quad (2.25)$$

Equation (2.24) gives

$$(E_n(\epsilon) - E_m(\epsilon)) \langle \psi_m, \varphi_n \rangle = 0, \quad (2.26)$$

implying

$$\langle \psi_m, \varphi_n \rangle = 0, \quad \text{if } E_n(\epsilon) \neq E_m(\epsilon). \quad (2.27)$$

The case $E_n(\epsilon) = E_m(\epsilon)$ deserves additional attention in case of degeneracy. It can be proven [81] that the freedom in choice of functions φ_n and ψ_m in case of *degeneracy* makes it possible to choose them such that *all* functions φ_n and ψ_m are *bi-orthonormal* in the sense that

$$\langle \psi_m, \varphi_n \rangle = \delta_{m,n}. \quad (2.28)$$

Assuming completeness of the functions $\{\varphi_n\}_n$, we may now write any function $F(\mathbf{r})$ as

$$F(\mathbf{r}) = \sum_n \alpha_n \varphi_n(\mathbf{r}; \epsilon) \quad (2.29)$$

which, owing to (2.28), implies the closure relation

$$\sum_n \varphi_n(\mathbf{r}; \epsilon) \psi_n^*(\mathbf{r}'; \epsilon) = \delta(\mathbf{r}-\mathbf{r}'). \quad (2.30)$$

In order to prove that (2.20) is the solution of (2.19) we first substitute (2.20) in (2.19) and apply (2.21) whereafter the identity (2.30) is readily recognized and the proof is completed.

It is easily observed, by taking the complex conjugate of (2.22), that the functions $\psi_n^*(\mathbf{r}; \epsilon)$ satisfy an equation like (2.21) in which, however, M is replaced by $M^{\dagger*}$. This leads us to the conclusion that, quite generally, in the above-considered case of degeneracy, one has to find (i.e., to construct), according to (2.20), at each couple of energy levels $E_n(\epsilon)$ and $E_n^*(\epsilon)$ in equations (2.21) and (2.22), two sets of functions $\varphi_n^{(i)}$ and $\psi_n^{(j)}$ with $i, j=1, 2, \dots, n_1$, such that $\langle \psi_n^{(i)}, \varphi_n^{(j)} \rangle = \delta_{i,j}$. Here n_1 is the number of linearly independent eigenfunctions of equation (2.21) at energy $E_n(\epsilon)$, or, equivalently, the number of linearly independent eigenfunctions $\psi_n^{(i)}$ of equation (2.22) at energy $E_n^*(\epsilon)$. Only if the property $M(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = M(\mathbf{r}_2, \mathbf{r}_1; \epsilon)$ holds, which can indeed be shown for crystals with a Hamiltonian \hat{H} as in (2.5), [62], φ_n and ψ_n^* are both solutions of the same equation (2.21).

We now want to discuss the quasi-particle interpretation of $G(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$. We emphasize that (2.20), although an *exact* representation of G , is not yet very useful in an actual numerical scheme. The expression as it stands asks for the determination of functions φ_n , ψ_n^* and E_n for all values of ϵ , which is an enormous task. In the quasi-particle approximation we *assume* G in (2.20) to have simple poles ϵ_n for which holds $\epsilon_n = E_n(\epsilon_n)$; the (possible) singularities due to non-analyticity of $\varphi_n(\mathbf{r}_1; \epsilon)$, $\psi_n^*(\mathbf{r}_2; \epsilon)$ or $E_n(\epsilon)$ will be discarded. The corresponding approximation to G is then obtained by putting

$$G(\mathbf{r}_1, \mathbf{r}_2; \epsilon) \simeq \hbar \sum_n g_n \frac{\varphi_n(\mathbf{r}_1; \epsilon_n) \psi_n^*(\mathbf{r}_2; \epsilon_n)}{\epsilon - \epsilon_n}, \quad (2.31)$$

where

$$g_n = \left[1 - \frac{dE_n(\epsilon)}{d\epsilon} \Big|_{\epsilon=\epsilon_n} \right]^{-1}. \quad (2.32)$$

The set of complex energy values ϵ_n defined through $\epsilon_n = E_n(\epsilon_n)$ is interpreted as

the quasi-particle spectrum. The general believe is that the values of $|\text{Im}(\epsilon_n)|$ are small with respect to $|\text{Re}(\epsilon_n)|$, such that the real parts of the poles ϵ_n can be interpreted as single quasi-particle energies, provided that the following relations are satisfied: $\text{Im}(\epsilon_n) > 0$ when $\text{Re}(\epsilon_n) < \mu$ and $\text{Im}(\epsilon_n) < 0$ when $\text{Re}(\epsilon_n) > \mu$, where μ is the quantity defined in section 2.3. Loosely speaking, we can refer to μ as the chemical potential separating the occupied quasi-particle states from the empty ones, i.e. the valence electrons from the conduction electrons. The numerical advantage of (2.31) over (2.20) is that the functions φ_n , ψ_n^* and E_n have to be determined at the above values ϵ_n only.

In the case of a crystal here considered, the self-energy function has the property

$$M(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = M(\mathbf{r}_1 + \mathbf{R}, \mathbf{r}_2 + \mathbf{R}; \epsilon), \quad (2.33)$$

where \mathbf{R} is any lattice vector belonging to the underlying Bravais lattice of the crystal. The solutions of (2.21) can therefore be chosen of the Bloch type, to be denoted by $\varphi_{\ell, \mathbf{k}}(\mathbf{r}; \epsilon)$; here \mathbf{k} is a wave vector in the first Brillouin zone (1Bz), and ℓ is a band index. Owing to the time-inversion symmetry [82], it follows that if $\varphi_{\ell, \mathbf{k}}$ is an eigenfunction at energy eigenvalue E , there exists also an eigenfunction $\varphi_{\ell, -\mathbf{k}}$ at the same eigenvalue [83]. Starting from the set of all Bloch functions $\varphi_{\ell, \mathbf{k}}$ at eigenvalue E , let us consider the *dual* set of functions ψ satisfying (2.22) at eigenvalue E^* . We note that due to $M^{\dagger*} = M$, the complex conjugate functions ψ^* satisfy (2.21). The functions ψ^* are therefore linear combinations of the $\varphi_{\ell, \mathbf{k}}$ functions. As $\langle \varphi_{\ell, \mathbf{k}'}^*, \varphi_{\ell, \mathbf{k}} \rangle = 0$ for any $\mathbf{k}' \neq -\mathbf{k}$ (due to the Bloch property), it follows directly from the proof leading to (2.28) that $\langle \varphi_{\ell, -\mathbf{k}}^*, \varphi_{\ell, \mathbf{k}} \rangle$ is necessarily different from zero. Assuming now the functions $\varphi_{\ell, \mathbf{k}}$ and $\varphi_{\ell, -\mathbf{k}}$ to be properly normalized, we rewrite the quasi-particle approximation (2.31) as

$$G(\mathbf{r}_1, \mathbf{r}_2; \epsilon) \simeq \hbar \sum_{\ell, \mathbf{k}} g_{\ell, \mathbf{k}} \frac{\varphi_{\ell, \mathbf{k}}(\mathbf{r}_1; E_{\ell}(\mathbf{k})) \varphi_{\ell, -\mathbf{k}}(\mathbf{r}_2; E_{\ell}(\mathbf{k}))}{\epsilon - E_{\ell}(\mathbf{k})}, \quad (2.34)$$

where $E_{\ell}(\mathbf{k}) = E_{\ell}(\mathbf{k}; E_{\ell}(\mathbf{k}))$.

The problem of determining the energy spectrum (more specifically the gap structure) in a semiconductor now consists of selfconsistently solving equations (2.21), (2.11) and (2.34). This is, however, as yet a formidable task. In the next section we will therefore discuss a number of simplifying approximation schemes, and in comparing these schemes, try to gain insight in the most promising schemes to approach the gap structure problem in a semiconductor.

2.5 Approximation Schemes

In comparing the various approximation schemes that have been (or may be) employed in order to describe or predict the electronic excitation structure of a given crystalline material, in particular the gap structure in a semiconductor, it will be clear from the considerations in the preceding sections that a given scheme is completely specified only (i) if the functions $z_{\ell}(\mathbf{r})$, $z_{n\ell}(\mathbf{r}, \mathbf{r}')$ are given (this fixes the effective one-electron potential in (2.21)) and (ii) if a complete prescription is given how to calculate the function $M(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$ occurring in (2.21). The latter prescription for M may be obtained either by explicitly stating which diagrams are to be retained in (2.11), or by giving some analytical expression for M in terms of the functions z_{ℓ} , $z_{n\ell}$, G and v .

It should be emphasized that those parts of $M(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$ that are Hermitian (this means here *real*) and independent of ϵ , may alternatively be represented in terms of functions $z'_{\ell}(\mathbf{r})$ or $z'_{n\ell}(\mathbf{r}, \mathbf{r}')$. This follows directly from (2.21) by observing that such terms in M can be written either in the form $\hbar^{-1}z'_{\ell}(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}')$ or $\hbar^{-1}z'_{n\ell}(\mathbf{r}, \mathbf{r}')$. This ambiguity is essential in our classification of existing approximation schemes: the various one-electron effective potentials that can be proposed, may as well be accounted for in terms of *some* M function. In approximation schemes such as the H, HF, Slater $X\alpha$, LDF scheme, this function M , by construction, is Hermitian and independent of ϵ . Conversely, in these cases the z_{ℓ} and $z_{n\ell}$ functions can all be *chosen* such that $M(1,2)=0$.

In the *Hartree* scheme the function $z_{\ell}(\mathbf{r})$ is chosen equal to

$$z_{\ell}^H(\mathbf{r}_1) = -2i \int d^3r' v(\mathbf{r}_1 - \mathbf{r}') G(\mathbf{r}'t, \mathbf{r}'t^*). \quad (2.35)$$

It can easily be shown that the electron density $\rho(\mathbf{r}') = -2iG(\mathbf{r}'t, \mathbf{r}'t^+)$, where the factor 2 originates from summation over spins (see the end of appendix A). The potential (2.35) is therefore easily recognized as the Coulomb potential due to all electrons. The choice (2.35) is such that the M diagrams (a1) and (b) in (2.11) automatically cancel. The prescription in the Hartree scheme is furthermore to discard all other diagrams in (2.11).

In the *Hartree-Fock* scheme we take z_ℓ as in (2.35) and $z_{n\ell}$ equal to

$$z_{n\ell}^{\text{HF}}(\mathbf{r}_1, \mathbf{r}_2) = i v(\mathbf{r}_1 - \mathbf{r}_2) G(\mathbf{r}_1 t, \mathbf{r}_2 t^+), \quad (2.36)$$

such that apart from the diagrams (a1) and (b), also the diagrams (a2) and (c) in (2.11) compensate. All other diagrams in (2.11) are discarded. In fact, it is known that the selfconsistent G (see (2.34)) emerging from the HF scheme will be totally different from the one emerging from the H scheme, due to the extra exchange effects involved in diagram (c) of (2.11).

A natural extension of the above diagrammatic approach would now be to investigate whether, or in how far, the contributions to $M(1,2)$ originating from diagrams (d), (e), ... in (2.11) may also be, or may partially be, swallowed in terms of a more refined choice of either $z_\ell(\mathbf{r}_1)$ or $z_{n\ell}(\mathbf{r}_1, \mathbf{r}_2)$. To this end we easily deduce that such $M(1,2)$ contributions should have their $(t_1 - t_2)$ dependence entirely in a factor $\delta(t_1 - t_2)$ (see the text following equation (2.11)). In order to investigate the possibility of existence of such a term, let us denote the contribution of all the diagrams (d), (e), ..., occurring in (2.11) by $M'(1,2)$. Considering the Fourier transform of $M'(1,2)$ for $|\epsilon| \rightarrow \infty$, it can be easily shown that $|M'(\mathbf{r}_1, \mathbf{r}_2; \epsilon)| \leq |\epsilon|^{-\gamma}$, with $\gamma > 1$. Therefore, owing to the fact that a δ -function behavior of a function is the indication that its Fourier transform for large values of $|\epsilon|$ does approach a *non-vanishing* constant value, we conclude that $M'(1,2)$ cannot contain a part with $\delta(t_1 - t_2)$ time behavior. This means that any refinement of z_ℓ and $z_{n\ell}$ in the above sense cannot lead to acceptable band structures for *all* energies, but at most for *limited energy intervals*. As this is the general objective in most (if not all) band structure calculations, the procedure of searching convenient z_ℓ and $z_{n\ell}$ functions might very well be of value.

The procedure of approximating M by an effective, possibly nonlocal, potential has been advocated by Pratt [84], but has not yet been systematically

investigated. If the evidence persists that the ϵ dependence of M is essential in obtaining correct bandstructures, the Pratt scheme will obviously be insufficient. An important point yet to be mentioned here is that the self-energy function $M(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$ is complex-valued outside some energy region around the midgap energy, as will be shown in chapter 3. Obviously no *real* effective potential, such as z_ℓ and $z_{n\ell}$ can account for the imaginary part of the self-energy function, which causes the quasi-particles to have finite life-times.

In the LDF scheme the choice for z_ρ , $z_{n\ell}$ and M is as follows:

$$z_\ell^{\text{LDF}}(\mathbf{r}_1) = -2i \int d^3r' v(\mathbf{r}_1 - \mathbf{r}') G(\mathbf{r}', \mathbf{r}' t^*)$$

$$+ \left. \frac{\delta E_{xc}[\rho]}{\delta \rho} \right|_{\rho=\rho(\mathbf{r}_1)} = -2i G(\mathbf{r}_1 t, \mathbf{r}_1 t^*) ,$$

$$z_{n\ell}(\mathbf{r}_1, \mathbf{r}_2) = 0; \quad M(1,2) = 0. \quad (2.37)$$

Here the exchange-correlation-energy functional $E_{xc}[\rho]$ is approximated by the expression $E_{xc}[\rho] = \int d^3r \rho(\mathbf{r}) \epsilon_{xc}(\rho(\mathbf{r}))$, where a uniform-density expression for ϵ_{xc} is used. The choice in (2.37) is such that the contribution of diagram (b) in (2.11) is compensated by a part of (a1), while diagram (a2) contributes zero. It is not at all clear, however, which set of M diagrams is compensated for by adding the extra exchange-correlation term to $z_\ell(\mathbf{r}_1)$ in (2.37). The specific choice of this term is motivated by the Hohenberg-Kohn-Sham theory [14,15,85], which shows that the exact $E_{xc}[\rho]$ leads to the exact groundstate electronic density if the selfconsistent Kohn-Sham procedure [15,85] is followed. As is well known, however, the *excitation* structure in a semiconductor following from the choice (2.37), is at variance with the experimental findings [19,20,86].

Also the Slater $X\alpha$ method [7], in which some of the elements of HF and DF theory are brought together [15], and in which $z_\ell(\mathbf{r}_1)$ is chosen equal to $z_\ell^{\text{H}}(\mathbf{r}_1) - 3\alpha e^2 [3\rho(\mathbf{r}_1)/\pi]^{1/3} / (8\pi\epsilon_0)$ with $2/3 \leq \alpha \leq 1$, does not lead to correct excitation structures in semiconductors [87-90]. It therefore seems that none of the existing schemes in which the ϵ dependence of $M(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$ is neglected, leads to the correct gap structure.

In the GW scheme, introduced by Hedin [56,76], the approximation is as

follows: (i) Take $z_{nl}(\mathbf{r}_1)$ as in the Hartree scheme, and take $z_{nl}(\mathbf{r}_1, \mathbf{r}_2) = 0$, hence the diagrams (a1), (a2) and (b) in (2.11) do not contribute to M. (ii) Take into account the contribution to M originating from all those diagrams in the expansion (2.11) which can be comprised in *one* renormalized diagram

$$M^{GW}(1,2) = \begin{array}{c} 1 \\ | \\ \uparrow \\ | \\ 2 \end{array} \cdot \quad (2.38)$$

where the dashed line has been defined in (2.13). It leads effectively to replacing in diagram (c) in (2.11) the bare Coulomb interaction function $v(\mathbf{r}_1 - \mathbf{r}_2)\delta(t_1 - t_2)$ by the screened interaction function $W(1,2)$ of (2.13), (see also (A.15c) for this diagram). The thus-renormalized diagram (c) leads to the contribution $M^{GW}(1,2) = (i/\hbar)G(1,2)W(1^*,2)$ which explains the name GW-approximation scheme. We will approximate $W(1,2)$ by restricting the polarization insertion in (2.13) to the first term in the diagrammatic expansion given in (A.12). Because of its typical topological character this term is called the "bubble" approximation term. This approximation will be discussed in the next section.

The idea behind the GW scheme is that the function $W(1,2)$ will turn out to be a much "weaker" interaction function than $v(1,2)$, such that all diagrams of higher order in W in the expansion (A.15) can be neglected [91]. Unlike $v(1,2)$, the function $W(1,2)$ does not factorize with $\delta(t_1 - t_2)$. Schemes in which only this factorized part is taken into account are not believed to lead to correct semiconductor excitation structures. On the contrary, the full time dependence plays an essential role [37,40,62,64-66,68].

The GW approach clearly *anticipates* the "weakness" of $W(1,2)$, but there is no *a priori* proof that the function $W(1,2)$ is indeed a much "weaker" interaction function than $v(1,2)$. Related to this, there is no *a priori* reason why the bubble approximation to the renormalized interaction should suffice.

2.6 The Bubble Approximation

In this scheme the polarization function is approximated by taking only the

zeroth-order term in the perturbation expansion of the polarization function in the screened interaction. In fact we have (see (A.12)):

$$P(1,2) \sim \text{bubble diagram} \quad (2.39)$$

Physically, this approximation describes the polarization effects due to the excitation of electron-hole pairs in the system. In this approximation the mutual interaction between the electrons and holes in the electron-hole pairs are missing. Calculation of the exact static polarization function within the framework of LDA, for material Si, has shown [92] that this bubble scheme actually underestimates the static polarization function by an amount of twelve percent, however there are strong indications [37,93] that reliable band structure results can be obtained within this bubble scheme. In view of (A.19) the vertex function $\Gamma(i,j;k)$ reduces to

$$\Gamma(i,j;k) = \delta(i,j)\delta(i,k). \quad (2.40)$$

This scheme is also referred to as the Random Phase Approximation (RPA) scheme, a terminology which has originally been employed by Pines and Bohm in analyzing the behavior of interacting electrons in a *dense* electron gas [94-96]. Incidentally, we note that in case of a degenerate electron gas, the RPA approximation gives rise to a negative pair correlation function at small distances [97-100], the latter function is, however, non-negative by construction. The screened interaction function $W(1,2)$ in this bubble scheme is formally given by

$$W(1,2) = \text{vertical dashed line} + \text{bubble diagram} \quad (2.41)$$

or [cf. (A.19), (B.35)]

$$W(1,2) = v(1,2) - \frac{2i}{\hbar} \int d(3)d(4) v(1,3)G(3,4)G(4,3^+)W(4,2),$$

(2.42)

in which the additional factor 2 in front of the integral sign is due to the spin summation.

CHAPTER 3

PROPERTIES OF THE SELF-ENERGY FUNCTION IN THE GW SCHEME

3.0 Introduction

In this chapter various properties of the self-energy function M within the GW-approximation scheme are discussed. Section 3.1 deals with the relation between the plane-wave matrix elements of the function $M^{\text{GW}}(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$ on the one hand and the quasi-particle wave functions and energies on the other. In section 3.2 we discuss the difficulties met when one intends to determine M^{GW} , starting from the expression for M^{GW} given in section 3.1 by directly performing the involved energy integration. In section 3.3 a contour-deformation procedure is outlined by means of which a number of these difficulties can be circumvented. Section 3.4 deals with the analytic continuation of M^{GW} in the complex energy plane, while section 3.5 is devoted to (non-) Hermiticity properties of M^{GW} . In section 3.6 a discussion is given of various approximations to the GW scheme. In this connection we pay special attention to the COHSEX scheme of Hedin [56] and the Plasmon-Pole scheme [64,65].

3.1 Relating the Self-Energy Function and the Quasi-Particle Excitation Structure in the GW Approximation

The self-energy function $M(1,2)$ in the GW approximation, given by equation (2.38), can easily be Fourier transformed with respect to t_1-t_2 , leading to [cf. (2.14)]

$$M^{\text{GW}}(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = \frac{i}{\hbar} \int_{-\infty}^{+\infty} \frac{d\epsilon'}{2\pi\hbar} G(\mathbf{r}_1, \mathbf{r}_2; \epsilon - \epsilon') W(\mathbf{r}_1, \mathbf{r}_2; \epsilon') e^{-i\epsilon' \eta / \hbar}, \quad (3.1)$$

where η is a positive infinitesimally small quantity. The presence of the exponential function $\exp[-i\epsilon' \eta / \hbar]$ is a consequence of the "+" sign in the

argument of the screened interaction in the expression of M^{GW} in the time domain (see the text following (2.38)). The functions G , W and M have the translational property [cf. (2.33)]

$$L(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = L(\mathbf{r}_1 + \mathbf{R}, \mathbf{r}_2 + \mathbf{R}; \epsilon), \quad (3.2)$$

where \mathbf{R} is a lattice vector. For that reason, any of these functions can be expressed as [101]

$$L(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = \frac{1}{\Omega} \sum_{\mathbf{k}} \sum_{\mathbf{K}, \mathbf{K}'} e^{i(\mathbf{k} + \mathbf{K}) \cdot \mathbf{r}_1} L_{\mathbf{K}, \mathbf{K}'}(\mathbf{k}; \epsilon) e^{-i(\mathbf{k} + \mathbf{K}') \cdot \mathbf{r}_2}, \quad (3.3)$$

where $L_{\mathbf{K}, \mathbf{K}'}(\mathbf{k}; \epsilon)$ is the Fourier transform of L with respect to \mathbf{r}_1 and \mathbf{r}_2 ,

$$L_{\mathbf{K}, \mathbf{K}'}(\mathbf{k}; \epsilon) = \frac{1}{\Omega} \int_{\Omega} d^3 r_1 d^3 r_2 e^{-i(\mathbf{k} + \mathbf{K}) \cdot \mathbf{r}_1} L(\mathbf{r}_1, \mathbf{r}_2; \epsilon) e^{i(\mathbf{k} + \mathbf{K}') \cdot \mathbf{r}_2}, \quad (3.4)$$

and where Ω is the volume of the crystal; \mathbf{k} is a wave vector in the first Brillouin zone (1Bz); \mathbf{K} and \mathbf{K}' are reciprocal lattice vectors. The Fourier transform of (3.1) can thus be expressed as

$$M_{\mathbf{G}, \mathbf{G}'}^{GW}(\mathbf{k}; \epsilon) = \frac{i}{\hbar \Omega} \sum_{\mathbf{k}'} \sum_{\mathbf{K}, \mathbf{K}'} \int_{-\infty}^{+\infty} \frac{d\epsilon'}{2\pi\hbar} \times G_{\mathbf{K}, \mathbf{K}'}(\mathbf{k}', \epsilon - \epsilon') W_{\mathbf{G} - \mathbf{K}, \mathbf{G}' - \mathbf{K}'}(\mathbf{k} - \mathbf{k}', \epsilon') e^{-i\epsilon' \eta / \hbar}, \quad (3.5)$$

where \mathbf{G} , \mathbf{G}' are reciprocal lattice vectors.

Now, suppose the matrix elements $M_{\mathbf{G}, \mathbf{G}'}(\mathbf{k}; \epsilon)$ have all been calculated. In order to show then how these matrix elements are related to the excitation spectrum, we consider the eigenvalue equations (2.21) in which we set $z_{\ell}(\mathbf{r})$ equal to $z_{\ell}^H(\mathbf{r})$ of (2.35). This leads to

$$\begin{aligned}
& \left[-\frac{\hbar^2}{2m} \nabla_1^2 + u(r_1) + z_\ell^H(r_1) - E_\ell(\mathbf{k}; \epsilon) \right] \varphi_{\ell, \mathbf{k}}(r_1; \epsilon) \\
& + \frac{\hbar}{\Omega} \int d^3 r_3 \sum_{\mathbf{k}'} \sum_{\mathbf{G}, \mathbf{G}'} e^{i(\mathbf{k}'+\mathbf{G}) \cdot \mathbf{r}_1} M_{\mathbf{G}, \mathbf{G}'}^{\text{GW}}(\mathbf{k}'; \epsilon) e^{-i(\mathbf{k}'+\mathbf{G}') \cdot \mathbf{r}_3} \varphi_{\ell, \mathbf{k}}(r_3; \epsilon) \\
& = 0, \tag{3.6}
\end{aligned}$$

which defines for each given ϵ an equation for the set of Bloch functions $\{\varphi_{\ell, \mathbf{k}}(r; \epsilon)\}$ and corresponding set of eigenenergies $\{E_\ell(\mathbf{k}; \epsilon)\}$ with ℓ the band index and \mathbf{k} in 1Bz. For ϵ values at which M^{GW} is non-Hermitian, the obtained set of functions will not be orthonormal, while the eigenvalues will generally be complex valued. We will return to this point in more detail in section 3.4.

According to the general theory of chapter 2, the quasi-particle approximation consists of finding those values ϵ such that $\epsilon = E_\ell(\mathbf{k}; \epsilon)$. If we from now on drop the argument ϵ in $E_\ell(\mathbf{k}; \epsilon)$, the quasi-particle eigenfunctions are given by $\varphi_{\ell, \mathbf{k}}(r; E_\ell(\mathbf{k}))$ with $E_\ell(\mathbf{k})$ the corresponding eigenvalue. If we expand $\varphi_{\ell, \mathbf{k}}(r; E_\ell(\mathbf{k}))$ in plane waves $\exp[i(\mathbf{k}+\mathbf{G}) \cdot \mathbf{r}]/\sqrt{\Omega}$, i.e.

$$\varphi_{\ell, \mathbf{k}}(r; E_\ell(\mathbf{k})) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{G}} d_{\ell, \mathbf{k}}(\mathbf{G}; E_\ell(\mathbf{k})) e^{i(\mathbf{k}+\mathbf{G}) \cdot \mathbf{r}}, \tag{3.7}$$

it is easily shown that for each given combination ℓ, \mathbf{k} the coefficients $d_{\ell, \mathbf{k}}(\mathbf{G}; E_\ell(\mathbf{k}))$ fulfill the set of coupled equations

$$\begin{aligned}
& \left[\frac{\hbar^2}{2m} (\mathbf{k}+\mathbf{G})^2 - E_\ell(\mathbf{k}) \right] d_{\ell, \mathbf{k}}(\mathbf{G}; E_\ell(\mathbf{k})) \\
& + \sum_{\mathbf{G}'} \left[u_{\mathbf{G}, \mathbf{G}'} + z_\ell^H(\mathbf{G}-\mathbf{G}') + \hbar M_{\mathbf{G}, \mathbf{G}'}^{\text{GW}}(\mathbf{k}; E_\ell(\mathbf{k})) \right] d_{\ell, \mathbf{k}}(\mathbf{G}'; E_\ell(\mathbf{k})) = 0. \tag{3.8}
\end{aligned}$$

This system of equations may be diagonalized by standard means, leading to the quasi-particle spectrum $E_\ell(\mathbf{k})$; the real part of $E_\ell(\mathbf{k})$ defines the quasi-particle band structure within this scheme. Note, that in principle the equations (3.8) and (3.5) have to be solved self-consistently. There is growing evidence, however, that the first-iteration step in a self-consistency cycle suffices if one starts with LDA wave functions [64,65]. If so, the amount of work in finding energy band

structures is considerably reduced.

3.2 Difficulties in the Calculation of the GW Self-Energy Function

The calculation of $M_{\mathbf{G}, \mathbf{G}'}^{\text{GW}}(\mathbf{k}, \epsilon)$ according to (3.5) involves two discrete summations over reciprocal lattice vectors, one energy integration along the whole real axis, and one \mathbf{k}' summation over the set of wave vectors in 1Bz. Due to the dense distribution of \mathbf{k}' vectors, the latter summation can equally well be performed as an integration over \mathbf{k}' , provided the density of \mathbf{k}' points $\Omega/(2\pi)^3$ is taken into account.

Let us enumerate the various difficulties encountered in the evaluation of M^{GW} , and see how this will give us a guide-line to devise techniques which make the evaluation of M tractable. Firstly, from the fact that the quasi-particle energies, which are all real or at least *almost real*, constitute the pole-structure of the one-particle Green function, we observe that the Green-function part of the integrand on the right-hand side of (3.5) will show very large variations along the path of integration, i.e. the real axis. Also, the determination of the screened-interaction part of the integrand at real ϵ' is not an easy task *a priori* as it involves the determination of the polarization function P and subsequent inversion of the the matrix $\epsilon=1-vP$ (see (A.13) and (A.14)). The obtained functions $W_{\mathbf{G}-\mathbf{K}, \mathbf{G}'-\mathbf{K}'}(\mathbf{k}-\mathbf{k}'; \epsilon')$ considered as functions of ϵ' are expected to have a *resonant* structure as well [102-106], but contrary to the one-particle Green functions $G_{\mathbf{K}, \mathbf{K}'}(\mathbf{k}'; \epsilon-\epsilon')$, their dependence on ϵ' will be much smoother [107-109]. Anyhow, performing the ϵ' integration first is no promising procedure.

Secondly, when discussing the possibility of performing the \mathbf{k}' integration over 1Bz first, the prospects are even worse. Consider in this respect the analytic linear tetrahedron method [110,111] and several versions of the special-point method [112-114], both being well-known methods in this type of problems. None of these methods can directly be applied in evaluating the \mathbf{k}' integration in (3.5). To illustrate this, let us for simplicity consider, instead of the exact Green-function matrix elements in (3.5), those of the one-particle Green function of some "unperturbed" Hamiltonian,

$$G_{\mathbf{K}, \mathbf{K}'}^{\circ}(\mathbf{k}; \epsilon) = \hbar \sum_{\ell} \frac{d_{\ell, \mathbf{k}}(\mathbf{K}) d_{\ell, \mathbf{k}}^*(\mathbf{K}')}{\epsilon - \epsilon_{\ell}(\mathbf{k}) - i\eta \operatorname{sgn}(\mu - \epsilon_{\ell}(\mathbf{k}))}, \quad (3.9)$$

in which $\operatorname{sgn}(x) = +1$ or -1 depending on whether $x > 0$ or $x < 0$. Here the earlier introduced $d_{\ell, \mathbf{k}}(\mathbf{K}, E_{\ell}(\mathbf{k}))$ and $E_{\ell}(\mathbf{k})$ have been reduced to $d_{\ell, \mathbf{k}}(\mathbf{K})$ and real-valued eigenvalues $\epsilon_{\ell}(\mathbf{k})$, respectively. It is not unreasonable to consider (3.9), as in an actual calculation procedure we will anyhow start an iteration cycle with some G° instead of G .

After substitution of (3.9) in (3.5), the \mathbf{k}' integration appears to be of the type

$$h(\mathbf{k}) \equiv \int_{1\text{Bz}} d^3\mathbf{k}' f(\mathbf{k}-\mathbf{k}') \frac{g(\mathbf{k}')}{\epsilon - \epsilon' - \epsilon_{\ell}(\mathbf{k}') \pm i\eta}, \quad (3.10)$$

where a simplified notation has been used in which the symbols \mathbf{K} , \mathbf{K}' , G , G' , ϵ , ϵ' and ℓ are suppressed except if they occur in the denominator. Let us for the time being *naively* assume that $f(\mathbf{k}-\mathbf{k}')$, considered as a function of \mathbf{k}' , is free from singularities (the function $g(\mathbf{k}')$ is harmless in this connection). Then we can use the standard relation $1/(x \pm i\eta) = \mathcal{P}(1/x) \mp i\pi\delta(x)$, in which \mathcal{P} stands for principal value, to obtain

$$h(\mathbf{k}) = \mathcal{P} \left\{ \int d^3\mathbf{k}' f(\mathbf{k}-\mathbf{k}') \frac{g(\mathbf{k}')}{\epsilon - \epsilon' - \epsilon_{\ell}(\mathbf{k}')} \right\} \\ \mp i\pi \int_{\epsilon_{\ell}(\mathbf{k}') = \epsilon - \epsilon'} d^2\mathbf{k}' f(\mathbf{k}-\mathbf{k}') \frac{g(\mathbf{k}')}{|\nabla_{\mathbf{k}'} \epsilon_{\ell}(\mathbf{k}')|}. \quad (3.11)$$

Although the first integral on the right-hand side of (3.11) is one in which the poles from the integrand are excluded, in the sense of the Cauchy principal-value concept, the integrand still shows significant variations [115] in 1Bz through the dependence of ϵ_{ℓ} on \mathbf{k}' . Therefore, an accurate evaluation of this integral requires the knowledge of the integrand at quite a large number of \mathbf{k}' points within 1Bz ,

which makes a numerical evaluation difficult.

A more serious problem is due to the second integral on the right-hand side of (3.11). Here, the use of the commonly applied special-point methods is ruled out, for the integration is a two-dimensional one. Furthermore, the denominator $|\nabla_{\mathbf{k}} \epsilon_{\lambda}(\mathbf{k}')|$ may have zeros – also called *critical points* or van Hove singularities [107-109]. The linear tetrahedron method can be applied here, albeit that due to the wild variations of the integrand and the occurrence of van Hove singularities accurate numerical results can be obtained only if the integrand is known at a large number of points on the surface of constant energy, $\epsilon_{\lambda}(\mathbf{k}') = \epsilon - \epsilon'$. This will make the evaluation of the integral a very time-consuming procedure [116-119].

So far we have *naively* assumed that the function $f(\mathbf{k}-\mathbf{k}')$ is regular. However, the actual situation is worse in that $f(\mathbf{k}-\mathbf{k}')$ represents the screened interaction which, as we shall see in chapter 4, has at least a $|\mathbf{k}-\mathbf{k}'|^{-1}$ singularity [120]. This kind of singularity is related to the remaining long range of the screened interaction in the spatial domain, which is a peculiarity of semiconductors and not of metals. The explanation is that in a semiconductor the density of valence electrons may be quite low in some regions in the primitive cell. Therefore the electron-electron interaction is not as effectively screened out as usually is the case in metals [120].

By the above considerations we come to the conclusion that a great deal of the above problems are due to the fact that we have to deal with an energy integration along the *real* axis. Hence, it turns out worthwhile to investigate a method in which integration over real energies can be avoided. This can be achieved by the contour deformation to be discussed in the next section. This procedure is no remedy for handling the singularity problem in the above function $f(\mathbf{k}-\mathbf{k}')$. This will be discussed and resolved in chapter 4.

3.3 Contour-Deformation Procedure

The problems which will be encountered when evaluating expression (3.5) for the GW self-energy function M^{GW} can to a large extent be circumvented by applying a contour-deformation procedure [66,68,69,121]. Before we will be able to deform the contour of the ϵ' integration in (3.1) into the complex energy

plane, we first need to have analytic continuations [122,123] of the functions G and W occurring in the integrand. As far as G is concerned, it is most convenient to use the Lehmann representation (2.16), in which we let ϵ assume complex values as well, as the analytic continuation of $G(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$ in the complex ϵ plane.

In order to obtain the analytic continuation of the screened interaction W we employ the relation [cf. (B.32)]

$$W(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = \int d^3 r' \epsilon^{-1}(\mathbf{r}_1, \mathbf{r}'; \epsilon) v(\mathbf{r}' - \mathbf{r}_2), \quad (3.12)$$

in which ϵ^{-1} stands for the inverse of the Fourier transform of the time-ordered dielectric function. The function ϵ^{-1} can be expressed in terms of the Fourier transform of the time-ordered density-density correlation function, D, [124-127], [cf.(B.44) and (B.47)], as

$$\epsilon^{-1}(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = \delta(\mathbf{r}_1, \mathbf{r}_2) - \frac{i}{\hbar} \int d^3 r' v(\mathbf{r}_1 - \mathbf{r}') D(\mathbf{r}', \mathbf{r}_2; \epsilon). \quad (3.13)$$

In the time domain D is defined by [cf.(B.44)]

$$D(\mathbf{r}_1 t_1, \mathbf{r}_2 t_2) = {}_H \langle \Psi_N | \mathcal{T} \{ \hat{\rho}'(\mathbf{r}_1 t_1) \hat{\rho}'(\mathbf{r}_2 t_2) \} | \Psi_N \rangle_H, \quad (3.14)$$

in which $\hat{\rho}'(\mathbf{r}t)$ represents the density-deviation operator

$$\hat{\rho}'(\mathbf{r}t) = \hat{\psi}^\dagger(\mathbf{r}t) \hat{\psi}(\mathbf{r}t) - {}_H \langle \Psi_N | \hat{\psi}^\dagger(\mathbf{r}t) \hat{\psi}(\mathbf{r}t) | \Psi_N \rangle_H. \quad (3.15)$$

It is not difficult, just as for the one-particle Green function, to obtain a Lehmann-type of representation for the density-density correlation function [125,127-129], which can be expressed as,

$$D(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = 2i\hbar \sum_s \rho'_s(\mathbf{r}_1) \rho'_s{}^*(\mathbf{r}_2) \left\{ \frac{1}{\epsilon - \epsilon'_s + i\eta} - \frac{1}{\epsilon + \epsilon'_s - i\eta} \right\}, \quad (3.16)$$

in which

$$\rho'_s(\mathbf{r}) = \text{H} \langle \Psi_N | \hat{\rho}'(\mathbf{r}0) | \Psi_{N,s} \rangle \text{H}, \quad (3.17a)$$

$$\epsilon'_s = E_{N,s} - E_N \geq 0. \quad (3.17b)$$

The factor 2 on the right-hand side of (3.16) is due to spin summation. Note that, in view of the definition (3.15), the s summation in (3.16) involves *excited* N -particle states only, i.e., there is no ground-state contribution. From (3.16) and (3.17) it is apparent that the analytic continuation of D , and thus W , is given by just the same expression as the one for real energies.

As one observes from the Lehmann-type representation (3.16), the analytic continuation of the screened interaction is analytic in the whole complex energy plane, except possibly for real energy values outside the interval $(-\tilde{E}_g, \tilde{E}_g)$, where $\tilde{E}_g \equiv \min_s \{\epsilon'_s; \epsilon'_s \neq 0\}$. Note that \tilde{E}_g is the lowest excitation energy of the N -particle system itself and therefore has to be identified with the lowest *exciton* energy. In an exact theory \tilde{E}_g should be distinguished from our earlier-defined gap energy $E_g \equiv (E_{N+1} - E_N) - (E_N - E_{N-1})$. In the thermodynamic limit, i.e. $\Omega \rightarrow \infty$, $N \rightarrow \infty$, $N/\Omega \rightarrow C$, with C a constant, one may think of the excitation energies $\{\epsilon_s\}_s$ of (2.18a,b) and $\{\epsilon'_s\}_s$ of (3.17b) to form dense sets of points on the real energy axis, implying that both the one-particle Green function and the screened interaction would possess two branch cuts [122,123] along the real energy axis.

Substituting (3.12) in (3.1) and making use of (3.13) we obtain

$$\begin{aligned} M^{\text{GW}}(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = & -\hbar^{-1} \left\{ \frac{1}{2} \rho(\mathbf{r}_1, \mathbf{r}_2) v(\mathbf{r}_1 - \mathbf{r}_2) + i \int d^3 r' d^3 r'' v(\mathbf{r}_1 - \mathbf{r}') \right. \\ & \left. \cdot \Xi(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}', \mathbf{r}''; \epsilon) v(\mathbf{r}'' - \mathbf{r}_2) \right\}, \end{aligned} \quad (3.18)$$

where $\rho(\mathbf{r}_1, \mathbf{r}_2)$ is the density matrix [130] given by

$$\begin{aligned} \rho(\mathbf{r}_1, \mathbf{r}_2) = & -2iG(\mathbf{r}_1 t, \mathbf{r}_2 t^+) \\ \equiv & -2i \int_{-\infty}^{\infty} \frac{d\epsilon'}{2\pi\hbar} G(\mathbf{r}_1, \mathbf{r}_2; \epsilon') \exp[+i\epsilon'\eta/\hbar], \quad \eta \downarrow 0, \end{aligned} \quad (3.19)$$

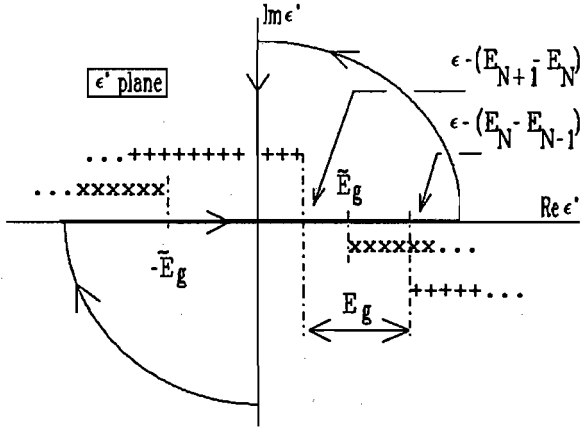


Fig.1. Schematic representation of the complex ϵ' plane. Indicated are the singular points of the integrand of (3.20). The original integration path as well as the one to be employed in (3.21) are depicted. The + sign denotes a pole due to the one-particle Green function while x denotes a pole due to the density-density correlation function.

and Ξ is an auxiliary function,

$$\Xi(r_1, r_2; r', r''; \epsilon) = -i\hbar^{-1} \int_{-\infty}^{\infty} \frac{d\epsilon'}{2\pi\hbar} G(r_1, r_2; \epsilon - \epsilon') D(r', r''; \epsilon'). \quad (3.20)$$

The ϵ dependence of M is through the function Ξ .

From the Lehmann (-type) representations of G and D it can immediately be seen that the integrand on the right-hand side of (3.20) may have singularities in each quadrant of the complex ϵ' plane. The function $D(r', r''; \epsilon')$ has singularities in the second and fourth quadrant only, while $G(r_1, r_2; \epsilon - \epsilon')$ may have, depending on ϵ , singularities in each quadrant. A possible configuration of singularities for some given value of ϵ is depicted in Fig. 1. Here the positions of poles of D are independent of ϵ ; the G poles shift horizontally when ϵ varies along the real axis.

We will deform the integration contour as follows: First we split the integration on the right-hand side of (3.20) into two integrals, one from minus infinity to zero, and the other one from zero to plus infinity. Then, the first integration range is extended with an integral along the negative imaginary axis and closed by a circle segment in the third quadrant as indicated in Fig.1. The second integration range is extended with an integral along the positive imaginary energy axis and closed, again by a circle segment, in the first quadrant as indicated. Concerning the behavior of the integrand for large values of ϵ' , we note that it shows an asymptotic behavior $\mathcal{O}(|\epsilon'|^{-3})$, as can be deduced from the Lehmann (-type) representations. Therefore, there will be no contribution to the integral along the circle segments.

The integrals along the real axis can now be completely expressed in terms of residue contributions and an energy integral along the imaginary ϵ' axis. Thus we obtain for (3.20)

$$\begin{aligned} \Xi(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}', \mathbf{r}''; \epsilon) &= \hbar^{-1} \sum_{\mathbf{s}} f_{\mathbf{s}}(\mathbf{r}_1) f_{\mathbf{s}}^*(\mathbf{r}_2) \\ &\times \left\{ [\Theta(\epsilon_{\mathbf{s}} - \epsilon) \Theta(\mu - \epsilon_{\mathbf{s}}) - \Theta(\epsilon - \epsilon_{\mathbf{s}}) \Theta(\epsilon_{\mathbf{s}} - \mu)] D(\mathbf{r}', \mathbf{r}''; \epsilon - \epsilon_{\mathbf{s}}) \right. \\ &\quad \left. + \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} d\epsilon' \frac{D(\mathbf{r}', \mathbf{r}''; \epsilon')}{\epsilon - \epsilon' - \epsilon_{\mathbf{s}}} \right\}. \end{aligned} \quad (3.21)$$

This result is now substituted in (3.18) in which the density matrix $\rho(\mathbf{r}_1, \mathbf{r}_2)$ is expressed as [130]

$$\rho(\mathbf{r}_1, \mathbf{r}_2) = 2 \sum_{\mathbf{s}} \Theta(\mu - \epsilon_{\mathbf{s}}) f_{\mathbf{s}}(\mathbf{r}_1) f_{\mathbf{s}}^*(\mathbf{r}_2). \quad (3.22)$$

We then end up with the following expression for the GW self-energy function

$$\begin{aligned}
M^{\text{GW}}(\mathbf{r}_1, \mathbf{r}_2; \epsilon) &= -\hbar^{-1} \sum_{\mathbf{s}} f_{\mathbf{s}}(\mathbf{r}_1) f_{\mathbf{s}}^*(\mathbf{r}_2) \\
&\times \left\{ \Theta(\mu - \epsilon_{\mathbf{s}}) v(\mathbf{r}_1 - \mathbf{r}_2) + \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} d\epsilon' \frac{\tilde{W}(\mathbf{r}_1, \mathbf{r}_2; \epsilon')}{\epsilon - \epsilon' - \epsilon_{\mathbf{s}}} \right. \\
&\left. + [\Theta(\epsilon_{\mathbf{s}} - \epsilon) \Theta(\mu - \epsilon_{\mathbf{s}}) - \Theta(\epsilon - \epsilon_{\mathbf{s}}) \Theta(\epsilon_{\mathbf{s}} - \mu)] \tilde{W}(\mathbf{r}_1, \mathbf{r}_2; \epsilon - \epsilon_{\mathbf{s}}) \right\}, \quad (3.23)
\end{aligned}$$

where

$$\tilde{W}(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = W(\mathbf{r}_1, \mathbf{r}_2; \epsilon) - v(\mathbf{r}_1 - \mathbf{r}_2) \quad (3.24)$$

is the part of the electron-electron interaction which corresponds to the screening effects. In obtaining (3.23), use has also been made of (3.12) and (3.13). We mention that in case $\epsilon = \epsilon_{\mathbf{s}}$, the integral on the right-hand side of (3.23) is to be considered as a Cauchy principal-value integral. Consistent with this we put $\Theta(0) = 1/2$.

Note that, (3.23) has only a formal character, for the functions $f_{\mathbf{s}}(\mathbf{r})$ defined by (2.17) are matrix elements between unknown many-electron wave functions, and as such not accessible to direct calculation. Therefore, for practical aims, one should work with an expression similar to (3.23), however obtained by substitution in (3.20) of the quasi-particle approximation (2.31) of the bi-orthonormal representation of G , rather than the exact Lehmann representation of G . In principle if one wants to attack the problem by following the required self-consistency scheme, starting from an unperturbed Hamiltonian, one has to use in each step of iteration, say m , the results of the foregoing step, $m-1$. In so doing one would employ in the m th iteration step the quasi-particle approximation of the Green function in which the quasi-particle wave functions and energies follow from the $(m-1)$ st step. It is easily seen that in the m th step of this scheme, (3.23) looks like as $-\hbar^{-1} \sum_{\mathbf{n}} \tilde{g}_{\mathbf{n}} \tilde{\varphi}_{\mathbf{n}}(\mathbf{r}_1; \tilde{E}_{\mathbf{n}}) \tilde{\psi}_{\mathbf{n}}^*(\mathbf{r}_2; \tilde{E}_{\mathbf{n}}) \{ \Theta(\mu - \text{Re}\{\tilde{E}_{\mathbf{n}}\}) \times v(\mathbf{r}_1 - \mathbf{r}_2) + \dots \}$, in which the symbols $\tilde{g}_{\mathbf{n}}$, $\tilde{\varphi}_{\mathbf{n}}$, $\tilde{\psi}_{\mathbf{n}}$ and $\tilde{E}_{\mathbf{n}}$ indicate that the approximate quantities pertain to the $(m-1)$ st step of iteration. By choosing (2.34) rather than (2.31), making use of a plane-wave representation for the quasi-particle wave functions of the $(m-1)$ st step, (see (3.7)), the Fourier

transform of (3.23) can be written as follows:

$$M_{\mathbf{G}, \mathbf{G}'}^{\text{GW}}(\mathbf{k}; \epsilon) = \frac{-1}{\hbar\Omega} \sum_{\mathbf{k}} \sum_{\mathbf{K}, \mathbf{K}'} F_{\mathbf{K}, \mathbf{K}'}(\mathbf{G}, \mathbf{G}', \mathbf{k}, \mathbf{k}'; \epsilon), \quad (3.25)$$

in which

$$\begin{aligned} F_{\mathbf{K}, \mathbf{K}'}(\mathbf{G}, \mathbf{G}', \mathbf{k}, \mathbf{k}'; \epsilon) &= \sum_{\ell} \tilde{g}_{\ell, \mathbf{k}-\mathbf{k}'} \tilde{d}_{\ell, \mathbf{k}-\mathbf{k}'}(\mathbf{K}) \tilde{d}_{\ell, -(\mathbf{k}-\mathbf{k}')}(-\mathbf{K}') \\ &\times \left\{ \Theta(\mu - \text{Re}\{\tilde{E}_{\ell}(\mathbf{k}-\mathbf{k}')\}) v_{\mathbf{G}-\mathbf{K}, \mathbf{G}'-\mathbf{K}'}(\mathbf{k}') \right. \\ &+ \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} d\epsilon' \frac{\tilde{W}_{\mathbf{G}-\mathbf{K}, \mathbf{G}'-\mathbf{K}'}(\mathbf{k}'; \epsilon')}{\epsilon - \epsilon' - \tilde{E}_{\ell}(\mathbf{k}-\mathbf{k}')} \\ &+ [\Theta(\text{Re}\{\tilde{E}_{\ell}(\mathbf{k}-\mathbf{k}')\}) - \epsilon] \Theta(\mu - \text{Re}\{\tilde{E}_{\ell}(\mathbf{k}-\mathbf{k}')\}) \\ &- \Theta(\epsilon - \text{Re}\{\tilde{E}_{\ell}(\mathbf{k}-\mathbf{k}')\}) \Theta(\text{Re}\{\tilde{E}_{\ell}(\mathbf{k}-\mathbf{k}')\} - \mu)] \\ &\left. \times \tilde{W}_{\mathbf{G}-\mathbf{K}, \mathbf{G}'-\mathbf{K}'}(\mathbf{k}'; \epsilon - \tilde{E}_{\ell}(\mathbf{k}-\mathbf{k}')) \right\}, \quad (3.26) \end{aligned}$$

where the Fourier transform of the Coulomb interaction $v_{\mathbf{K}, \mathbf{K}'}(\mathbf{k})$ is given by

$$v_{\mathbf{K}, \mathbf{K}'}(\mathbf{k}) = \frac{e^2}{\epsilon_0 |\mathbf{k} + \mathbf{K}|^2} \delta_{\mathbf{K}, \mathbf{K}'}. \quad (3.27)$$

In order to obtain a convenient notation, we have suppressed in (3.26) the energy arguments of the plane-wave coefficients $\tilde{d}_{\ell, \pm}(\mathbf{k}-\mathbf{k}')$. In a first iteration step, in which some unperturbed G^0 is used, rather than the general quasiparticle representation of G , it will be obvious that (3.26) reduces to a simple form in which the \tilde{g} factors equal to 1, and in which the energy eigenvalues are real.

By observing (3.25) and (3.26) we note that part of the problems enumerated in the previous section are indeed resolved by application of the contour deformation procedure of this section. As to the energy integration in (3.26), contrary to the real-axis ϵ' integration of previous section, we have to do here with an integrand which is *regular* for all ϵ' on the path of integration (the imaginary ϵ' axis), except for the point $\epsilon' = \epsilon - \tilde{E}_\ell(\mathbf{k}-\mathbf{k}')$. As the anticipated imaginary part of $\tilde{E}_\ell(\mathbf{k}-\mathbf{k}')$ is much smaller than its real part and ϵ is real, we note that this ϵ' point is located very close to 0. Let $(-i\Delta, i\Delta)$ be the ϵ' interval in which Δ is somewhat larger than the largest absolute value of the imaginary part of the quasi-particle energies under consideration, we then have, in view of the varying ℓ and $\mathbf{k}-\mathbf{k}'$ values to be considered, still to take special measures to treat this region of integration. In section 4.6 this issue is discussed further, and a simple integration procedure for this region is proposed there. The main point to note is that the energy integration regions $(-i\infty, -i\Delta)$ and $(i\Delta, i\infty)$ are harmless. In this connection we also note that \tilde{W} along the imaginary energy axis does not have any resonant structure at all, as will be pointed out in section 3.6. The summation in (3.23), which in (3.25) and (3.26) has been replaced by a summation over ℓ and \mathbf{k}' , still requires attention, but the \mathbf{k}' -summation problem is much less troublesome, because of the absence of vanishing denominators in the ϵ' -integration region.

3.4 Analytical Properties of the GW Self-Energy Function

In this section we will investigate the analytical properties of the self-energy function M^{GW} as a function of the complex energy variable ϵ . Thorough knowledge of the behavior of M^{GW} in the complex ϵ plane will turn out to be most helpful in discussing various numerical evaluation methods for M^{GW} , [68,69].

We will start with expression (3.23) for real energy values ϵ , and introduce two auxiliary functions, $g_s(z)$ and $h_s(z)$ of complex variable z , given by

$$g_s(z) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} d\epsilon' \frac{\tilde{W}(r_1, r_2; \epsilon')}{z - \epsilon' - \epsilon_s}, \quad (3.28)$$

$$h_s(z) = g_s(z) + [\Theta(\epsilon_s - \text{Re}(z))\Theta(\mu - \epsilon_s) - \Theta(\text{Re}(z) - \epsilon_s)\Theta(\epsilon_s - \mu)] \\ \times \tilde{W}(r_1, r_2; z - \epsilon_s), \quad (3.29)$$

in which we have suppressed the dependence of the functions h_s and g_s on the spatial variables. Note that for $\text{Re}(z)$ within the gap region, i.e. $E_N - E_{N-1} < \text{Re}(z) < E_{N+1} - E_N$, the second term on the right-hand side of (3.29) always vanishes, implying that for each ϵ_s the functions h_s and g_s coincide, that is,

$$h_s(z) = g_s(z), \text{ if } E_N - E_{N-1} < \text{Re}(z) < E_{N+1} - E_N. \quad (3.30)$$

We first want to discuss that the function $g_s(z)$ is discontinuous, and hence non-analytic, at the line $\text{Re}(z) = \epsilon_s$. Let us therefore introduce the functions $g_s^-(z)$ and $g_s^+(z)$,

$$g_s^\pm(z) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} d\epsilon' \frac{\tilde{W}(r_1, r_2; \epsilon')}{z - \epsilon' - \epsilon_s}, \text{ Re}(z) \gtrless \epsilon_s. \quad (3.31)$$

For $\text{Re}(z) = \epsilon_s$ we will define the value of $g_s^\pm(z)$ by the corresponding limit procedure. As \tilde{W} is both analytic (therefore differentiable) along the imaginary energy axis and behaves asymptotically like $\mathcal{O}(|\epsilon'|^{-2})$ for large values of $|\epsilon'|$, it can be shown [132] that g_s^+ and g_s^- are analytic in complex half-planes where they are defined [133,134]. Moreover, from the Plemelj relations [133,135], satisfied by these two functions, it follows that on the line $\text{Re}(z) = \epsilon_s$

$$g_s^+(\epsilon_s + iy) - g_s^-(\epsilon_s + iy) = \tilde{W}(r_1, r_2; iy), \quad y \in \mathbb{R}. \quad (3.32)$$

This is precisely the magnitude of the discontinuity of $g_s(z)$ at $z = \epsilon_s + iy$.

It can readily be verified that in the function $h_s(z)$ the discontinuity (3.32) in $g_s(z)$ is precisely canceled through the last term in (3.29) with Θ functions. Therefore, using Riemann's principle [136] it follows that $h_s(z)$ is analytic across the line $\text{Re}(z)=\epsilon_s$ as well as analytic in the entire z plane, except for singular points of $\check{W}(r_1, r_2; z - \epsilon_s)$ itself. The analytical properties of $\check{W}(r_1, r_2; z)$ were investigated in the previous section and from the results obtained there we conclude that $h_s(z)$ is analytic in the whole complex z plane except possibly on the lines $\{z = \epsilon + i\eta; \epsilon < (E_N - E_{N-1}) - \check{E}_g\}$, and $\{z = \epsilon - i\eta; \epsilon > (E_{N+1} - E_N) + \check{E}_g\}$, depending on whether $\epsilon_s < \mu$ or $\epsilon_s > \mu$, respectively. This follows directly by considering the possible pole positions of the function $\check{W}(r_1, r_2; z - \epsilon_s)$ as a function of z .

We will now consider the Taylor expansion of $h_s(z)$ around the midpoint ν of the interval $(E_N - E_{N-1} - \check{E}_g, E_{N+1} - E_N + \check{E}_g)$, i.e.,

$$\nu = (E_{N+1} - E_{N-1})/2, \quad (3.33)$$

where ν is in fact the real midgap-energy value. It is then convenient to go over to the function $\check{h}_s(z)$, defined by

$$\check{h}_s(z) = h_s(z + \nu), \quad (3.34)$$

and consider the Taylor expansion of $\check{h}_s(z)$ around $z=0$. The simple relationship between the functions \check{h}_s at real z and the self-energy function (3.23) can be expressed as

$$M^{GW}(r_1, r_2; \epsilon + \nu) = l(\epsilon), \quad (3.35)$$

where $l(z)$ is the function of complex variable z defined as

$$l(z) = -\hbar^{-1} \sum_s f_s(r_1) f_s^*(r_2) \{ \Theta(\mu - \epsilon_s) v(r_1 - r_2) + \check{h}_s(z) \}. \quad (3.36)$$

The singularities of $\check{h}_s(z)$ nearest to the origin $z=0$ are at $\pm z_s$, where $z_s \geq E_g/2 + \check{E}_g$, in which E_g is the gap energy (see section 2.3) and \check{E}_g the exciton gap (see section 3.3). This means that the Taylor expansion of $l(z)$ around the

origin has a radius of convergence equal to $\mathcal{R} = E_g/2 + \tilde{E}_g$, [69]. Note that this radius \mathcal{R} is also the largest convergence radius one can have if the expansion is around some point on the real axis. This is due to the special choice of ν such that $z=0$ is symmetric with respect to the nearest singularities.

The above-introduced $l(z)$ is the analytic continuation of M^{GW} into the complex energy plane. Hence a Taylor expansion of $l(z)$, i.e.,

$$l(z) = \sum_{n=0}^{\infty} \alpha_n z^n, \quad (|z| < \mathcal{R} = E_g/2 + \tilde{E}_g), \quad (3.37)$$

is also a Taylor expansion of M^{GW} , i.e.,

$$M^{\text{GW}}(r_1, r_2; z+\nu) = \sum_{n=0}^{\infty} \alpha_n z^n, \quad |z| < \mathcal{R}, \quad (3.38)$$

where [cf. (3.30), (3.34) and (3.36)]

$$\begin{aligned} \alpha_n &= \frac{1}{n!} \left. \frac{\partial^n l(z)}{\partial z^n} \right|_{z=0} \\ &\equiv \frac{1}{n!} \left. \frac{\partial^n}{\partial z^n} \left\{ -\hbar^{-1} \sum_s f_s(r_1) f_s^*(r_2) [\Theta(\mu - \epsilon_s) v(r_1 - r_2) + g_s(z+\nu)] \right\} \right|_{z=0}. \end{aligned} \quad (3.39)$$

The merit of (3.38) lies in the fact that, as long as we limit ourselves to $|z| < E_g/2 + \tilde{E}_g$, we do not have to evaluate pole contributions to (3.23) explicitly. The direct calculation of \hat{W} at such real energies and the execution of the corresponding s summation in the residue part are thus avoided [cf. (3.28) and (3.29)]. Furthermore, we may expect that only a few terms in the Taylor expansion need to be evaluated [68]. In Fig.2 some relevant information as to the analytical structure of $M^{\text{GW}}(r_1, r_2; \epsilon)$ in the complex ϵ plane, is schematically summarized.

In actual calculations, especially in the first iteration cycle of a self-

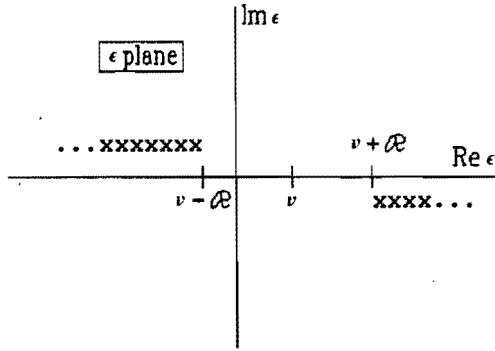


Fig.2. Schematic representation of the analytical structure of $M^{\text{GW}}(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$ in the complex ϵ plane. The sign x indicates a singular point of $M^{\text{GW}}(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$. The quantity ν is defined in (3.33) while \mathcal{R} is defined in (3.37).

consistent calculation procedure, the starting values for \tilde{E}_g , $(E_N - E_{N-1})$ and $(E_{N+1} - E_N)$ are usually different from the corresponding exact quantities. For instance, if the first iteration cycle is based on results for "non-interacting" particles, the values \tilde{E}_g , $(E_N - E_{N-1})$ and $(E_{N+1} - E_N)$ will be equal to ϵ_g , ϵ_N and ϵ_{N+1} , respectively. Here ϵ_j stands for the j th one-electron energy eigenvalue of the "unperturbed" system while ϵ_g represents the (one-electron) energy gap, i.e.,

$$\epsilon_g \equiv \epsilon_{N+1} - \epsilon_N, \quad (3.40)$$

which is nothing but the value to which the energy-gap value $E_g = E_{N+1} - E_{N-1} - 2E_N$ reduces in the unperturbed case. Here we have assumed that the one-electron eigenenergies are ordered according to the relation $\epsilon_{j+1} \geq \epsilon_j$. The above-defined radius of convergence \mathcal{R} , will now be equal to $3\epsilon_g/2$. Note that, depending on the choice of the unperturbed Hamiltonian, the quantity ϵ_g may substantially deviate from the exact energy gap E_g . For instance, it is well known that the LDA Hamiltonian in a semiconductor gives rise to an energy gap which is too small, the underestimation being sometimes about fifty percent of

the actual gap [19,20,86]. Contrary to this, ϵ_g when calculated in the HF approximation may amount to about 2-5 times the exact energy gap [11,12,13,23-26].

3.5 Other Properties of the GW Self-Energy Function

In this section we will investigate conditions under which the GW self-energy function $M^{\text{GW}}(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$ is Hermitian, i.e. $M^{\text{GW}}(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = M^{\text{GW}}(\mathbf{r}_2, \mathbf{r}_1; \epsilon)^*$. This investigation is relevant, since non-Hermiticity of the self-energy function at real ϵ directly leads to finite quasi-particle lifetimes. The reverse assertion, namely that the Hermiticity of M^{GW} at real ϵ guarantees the quasi-particle life-times to be infinite, is as yet an open question to which we come furtheron. It can generally be demonstrated that the exact self-energy function satisfies the symmetry relation $M(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = M(\mathbf{r}_2, \mathbf{r}_1; \epsilon)$. As this relation is also satisfied by the GW self-energy function (this will incidentally be shown below), we have to find out under which conditions $M^{\text{GW}}(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$ is a *real* function of real ϵ . To this end we consider (3.23). We first note that the function $f_s(\mathbf{r}_1)f_s^*(\mathbf{r}_2)$ can be chosen real [77]. As the bare Coulomb interaction is real (and symmetric) as well, we have only to concentrate on the second and the last term within the braces on the right-hand side of (3.23); to the former term we refer as the *integral part* and to the latter term as the *residue part*.

As to the integral part, it follows directly from the real-valuedness of iD along the imaginary energy axis (see (3.12)-(3.17)), that \tilde{W} along the imaginary energy axis is real-valued. We remark in this connection that, just as $f_s(\mathbf{r}_1)f_s^*(\mathbf{r}_2)$, the function $\rho_s'(\mathbf{r}_1)\rho_s'^*(\mathbf{r}_2)$ can be chosen real too. Moreover, \tilde{W} and D are both even functions of energy. This, as well as the reality of \tilde{W} along the imaginary energy axis, results in the fact that for real values of ϵ the contribution of the ϵ' integral along the negative imaginary axis is the complex conjugate of one along the positive imaginary axis. Therefore, for real energies ϵ the *integral part* is real.

Apparently, if the self-energy function is to have a complex contribution, this can only be due to the residue part. To investigate this further, we substitute the right-hand side of (3.16) in (3.13); subsequently substitute the obtained expression in (3.12) and make use of (3.24) as well as the

standard relation $1/(x \pm i\eta) = \mathcal{P}(1/x) \mp i\pi\delta(x)$, with \mathcal{P} the principal value. We then obtain

$$\tilde{W}(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = \tilde{W}'(\mathbf{r}_1, \mathbf{r}_2; \epsilon) + i \tilde{W}''(\mathbf{r}_1, \mathbf{r}_2; \epsilon), \quad (3.41)$$

with

$$\tilde{W}'(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = 2 \sum_{\mathbf{s}} \mathcal{P} \left\{ \frac{1}{\epsilon - \epsilon'_s} - \frac{1}{\epsilon + \epsilon'_s} \right\} w_{\mathbf{s}}(\mathbf{r}_1) w_{\mathbf{s}}^*(\mathbf{r}_2), \quad (3.42a)$$

$$\tilde{W}''(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = -2\pi \sum_{\mathbf{s}} \{ \delta(\epsilon - \epsilon'_s) + \delta(\epsilon + \epsilon'_s) \} w_{\mathbf{s}}(\mathbf{r}_1) w_{\mathbf{s}}^*(\mathbf{r}_2), \quad (3.42b)$$

and

$$w_{\mathbf{s}}(\mathbf{r}) = \int d^3r' v(\mathbf{r} - \mathbf{r}') \rho_{\mathbf{s}}'(\mathbf{r}'). \quad (3.43)$$

Note that, as $w_{\mathbf{s}}(\mathbf{r}_1) w_{\mathbf{s}}^*(\mathbf{r}_2)$ can be chosen real [125] and is symmetric, the above-mentioned symmetry property of M^{GW} , i.e. $M^{\text{GW}}(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = M^{\text{GW}}(\mathbf{r}_2, \mathbf{r}_1; \epsilon)$, is incidentally established. As we have $\epsilon_s' > \tilde{E}_{\mathbf{g}}$, the implication is that $\tilde{W}''(\mathbf{r}_1, \mathbf{r}_2; \epsilon) \equiv 0$ if ϵ is in the energy range $(-\tilde{E}_{\mathbf{g}}, \tilde{E}_{\mathbf{g}})$. Hence, for real ϵ the *residue part* is always real as long as $|\epsilon - \epsilon_s'| < \tilde{E}_{\mathbf{g}}$, [137, 138].

It immediately follows that $M^{\text{GW}}(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$ is *real*, and thus Hermitian, whenever $E_{\mathbf{N}} - E_{\mathbf{N}-1} - \tilde{E}_{\mathbf{g}} < \epsilon < E_{\mathbf{N}+1} - E_{\mathbf{N}} + \tilde{E}_{\mathbf{g}}$. This result is fully consistent with and, in fact, can also be derived from, the Taylor expansion (3.38) by realizing that all Taylor coefficients α_n are real. Therefore, M^{GW} has to be real for real ϵ *within* the circle of convergence.

The non-Hermiticity of M^{GW} at real energies outside the above-mentioned region is closely related to the finite probability for quasi-particles at such energies to decay through the excitation of electron-hole pairs (via Coulomb interaction). On account of energy conservation such excitations are impossible if ϵ lies in the interval $(E_{\mathbf{N}} - E_{\mathbf{N}-1} - \tilde{E}_{\mathbf{g}}, E_{\mathbf{N}+1} - E_{\mathbf{N}} - \tilde{E}_{\mathbf{g}})$, i.e., in the energy region where M^{GW} is real. However, the implication of the Hermitian or non-Hermitian

character of M at real ϵ with regard to the quasi-particle life-times is *not* immediately obvious: firstly, one has to solve the Schrödinger-like quasi-particle equations (3.8) in which $M_{\mathbf{G},\mathbf{G}'}(\mathbf{k};\epsilon)$ matrix elements play their role, leading to eigenvalues $E_{\lambda}(\mathbf{k};\epsilon)$; and secondly, one has to solve the equation $E_{\lambda,\mathbf{k}}(\epsilon)=\epsilon$. If we call $\epsilon=E_{\lambda}(\mathbf{k})$ the solution of this equation, there is quite generally a real part $\text{Re}(E_{\lambda}(\mathbf{k}))$, being the quasi-particle energy and an imaginary part $\text{Im}(E_{\lambda}(\mathbf{k}))$, $\hbar/(2\text{Im}(E_{\lambda}(\mathbf{k})))$ being the life-time [139]. A complex eigenvalue $E_{\lambda}(\mathbf{k};\epsilon)$ at real ϵ directly implies the complexity of the solution $\epsilon=E_{\lambda}(\mathbf{k})$. However, if $E_{\lambda}(\mathbf{k};\epsilon)$ is real for real ϵ (which is the case if M is Hermitian) it can not *a priori* be excluded that the solution of the equation $E_{\lambda,\mathbf{k}}(\epsilon)=\epsilon$ is, nevertheless, complex.

Let us consider a consequence of the reality of M^{GW} on a finite part of the real energy axis. As M^{GW} is analytic (strictly speaking we mean the analytical continuation of M^{GW} , i.e. the function $l(z)$ defined in (3.36)) in parts of the half-planes $\text{Im}(z)>0$ and $\text{Im}(z)<0$, while real-valued in the interval $-E_g/2-\tilde{E}_g<\epsilon-\nu<E_g/2+\tilde{E}_g$, with $\nu=(E_{N+1}-E_{N-1})/2$, we can conclude on the basis of the Riemann-Schwarz reflection principle [140] that M^{GW} takes conjugate values for conjugate values of z , i.e.

$$M^{\text{GW}}(\mathbf{r}_1,\mathbf{r}_2;z^*) = M^{\text{GW}}(\mathbf{r}_1,\mathbf{r}_2;z)^* \quad (3.44)$$

This shows again that the Taylor coefficients in (3.39) are real valued. In general, the coefficients of Taylor expansions around complex conjugate points are to all orders complex conjugate of each other.

From the above considerations, we know that the region of validity of the Taylor expansion of $M^{\text{GW}}(\mathbf{r}_1,\mathbf{r}_2;z)$ around $z=\nu$ is directly related to the distance of the nearest singularity of the dynamically screened interaction to the origin, \tilde{E}_g . The radius of convergence for $M^{\text{GW}}(\mathbf{r}_1,\mathbf{r}_2;\epsilon)$ was shown to be $E_g/2+\tilde{E}_g$ (see below (3.39)). On the other hand it is known that GW self-energy functions evaluated with a rather crude approximation of W , the Plasmon-Pole method, give rise to very satisfactory quasi-particle band structures [64,65,141]. This is in a way surprising since in the Plasmon-Pole (PP) model W has singularities, the plasmon poles, which are located at distances from the origin much larger than \tilde{E}_g , i.e., there is no such \tilde{E}_g -related singularity in the plasmon-pole approximation of \tilde{W} , \tilde{W}^{PP} , at all. Hence, in the framework of a plasmon-

pole model the circle of convergence for M^{GW} is far more larger than in the exact theory. One may hope, therefore, that by truncating the Taylor series expansion of M^{GW} around $z=\nu$, and extrapolating the thus-obtained polynomial outside the region of convergence, one may still obtain a good approximation for the self-energy. Indeed, Godby, Schlüter and Sham have shown that this strategy works very well for several semiconductors [68]. However, one has to realize that, due to the reality of the Taylor coefficients, this approximation is not capable of describing an eventual non-Hermitian part of the self-energy function at real energies [69].

3.6 Some Approximation Methods within the GW Scheme

In this section we will rubricate a number of common and uncommon approximations to M within the GW scheme. At first sight these approximations may seem to be based on mathematical manipulations rather than on physical considerations; however it can and will be argued that their merits have a physical content as well. We will distinguish two classes of approximations, namely the *static* and *dynamic* approximations. Static approximations for M^{GW} have in common that they show no ϵ dependence; dynamic approximations, on the other hand, show non-trivial dependence of M^{GW} on ϵ .

In considering (3.23), we may as a first crude approximation, neglect all terms involving the function \tilde{W} . This yields about the simplest static approximation [69] to M one can think of [see also (3.18), (3.19) and (3.22)]

$$\begin{aligned} M^{\text{HF}}(\mathbf{r}_1, \mathbf{r}_2) &= -\hbar^{-1} v(\mathbf{r}_1 - \mathbf{r}_2) \sum_s \Theta(\mu - \epsilon_s) f_s(\mathbf{r}_1) f_s^*(\mathbf{r}_2) \\ &= i\hbar^{-1} v(\mathbf{r}_1 - \mathbf{r}_2) G(\mathbf{r}_1, t, \mathbf{r}_2, t^+). \end{aligned} \quad (3.45)$$

Here the superscript HF explicitly refers to the fact that in (3.45) a Hartree-Fock exchange-type approximation is recognized. We must admit, however, that (3.45) leads to the exact Hartree-Fock potential only if the functions $\{f_s(\mathbf{r})\}_s$ are the self-consistently obtained Hartree-Fock wave functions.

We realize that the result (3.45) is simply to be expected on the basis of

the *trivial* fact that by neglecting \tilde{W} , the screened interaction W becomes just equal to the bare Coulomb interaction, (see (2.13) and (3.24)), so that diagram (2.38), representing M^{GW} , goes over into the diagram (c) of (2.11), being the exchange part of the Hartree-Fock self-energy function. We have introduced (3.45) mainly for the sake of uniformity of presentation; we shall not expound the peculiarities of M^{HF} further [7].

As a second approximation to M^{GW} we replace the screening-correction functions \tilde{W} on the right-hand side of (3.23) by the static function $\tilde{W}(\mathbf{r}_1, \mathbf{r}_2; 0)$. After evaluating the remaining integral and employing the relation (3.24) we arrive at an expression which is known in the literature as the COHSEX (COulomb-Hole with Screened EXchange) approximation [56,143]. We have

$$\begin{aligned} M^{\text{COHSEX}}(\mathbf{r}_1, \mathbf{r}_2; \epsilon) &= -\hbar^{-1} \sum_{\mathbf{s}} \{ \Theta(\mu - \epsilon_{\mathbf{s}}) W(\mathbf{r}_1, \mathbf{r}_2; 0) - \frac{1}{2} \tilde{W}(\mathbf{r}_1, \mathbf{r}_2; 0) \} \\ &\quad \times f_{\mathbf{s}}(\mathbf{r}_1) f_{\mathbf{s}}^*(\mathbf{r}_2) \\ &= i\hbar W(\mathbf{r}_1, \mathbf{r}_2; 0) G(\mathbf{r}_1 t, \mathbf{r}_2 t^+) + \frac{1}{2\hbar} \tilde{W}(\mathbf{r}_1, \mathbf{r}_2; 0) \delta(\mathbf{r}_1 - \mathbf{r}_2), \end{aligned} \quad (3.46)$$

where in the last step the closure relation

$$\sum_{\mathbf{s}} f_{\mathbf{s}}(\mathbf{r}_1) f_{\mathbf{s}}^*(\mathbf{r}_2) = \delta(\mathbf{r}_1 - \mathbf{r}_2), \quad (3.47)$$

has been used. Equation (3.47) can readily be verified by making use of the equal-time anti-commutation relation (2.2b).

The name COHSEX is very indicative of the physical content of the above-described approximation; COH refers to the contribution of the second term on the right-hand side of (3.46), whereas SEX refers to the first term. The reason why the latter contribution is called SEX is the following: By comparing this contribution with M^{HF} , we see that we have to do with an EXchange contribution (EX) in which, however, the role of the bare Coulomb interaction is now played by the Screened interaction (S). The reason for referring to the second term as COulomb-Hole contribution, stems from the fact that each electron, by means of its Coulombic repulsion, lowers the charge density due to

other electrons in its immediate vicinity. In other words, the electron in question creates a sort of *hole* surrounding itself. This phenomenon, which is completely absent in case of non-interacting particles, is known as polarization. The deviation of the charge density with respect to that of a non-interacting system, which in all other aspects is identical with the interacting system, is called the polarization charge density. This polarization charge will induce some force acting on the electrons themselves, modifying their motion in the crystal. Let us make this plausible by repeating and extending a reasoning due to Hedin [56].

Consider a classical electron at \mathbf{r} . If the electrons in the system were non-interacting, then the potential energy of the electron at \mathbf{r} , as a result of the presence of another electron at \mathbf{r}' , would be $v(\mathbf{r}-\mathbf{r}')$. However, because of polarization effects in the interacting system, the latter energy, in a *static* approximation, is equal to $W(\mathbf{r},\mathbf{r}';0)$. The difference between these two energies, i.e. $\tilde{W}(\mathbf{r},\mathbf{r}';0)$, see (3.24), is thus the *induced* potential energy of the electron in question. The corresponding force exerted on the electron, equals $-\nabla_{\mathbf{r}}\tilde{W}(\mathbf{r},\mathbf{r}';0)$, where the gradient operator acts on the \mathbf{r} variable only. Now the *induced* force exerted on the electron at \mathbf{r} , as a result of its *own* interaction with other electrons is given by

$$\mathcal{F} = -\lim_{\mathbf{r}' \rightarrow \mathbf{r}} \nabla_{\mathbf{r}} \tilde{W}(\mathbf{r},\mathbf{r}';0). \quad (3.48)$$

This can also be written, because of the symmetry property $\tilde{W}(\mathbf{r},\mathbf{r}';\epsilon) = \tilde{W}(\mathbf{r}',\mathbf{r};\epsilon)$, as

$$\begin{aligned} \mathcal{F} &= -\frac{1}{2} \lim_{\mathbf{r}' \rightarrow \mathbf{r}} \{ \nabla_{\mathbf{r}} \tilde{W}(\mathbf{r},\mathbf{r}';0) + \nabla_{\mathbf{r}} \tilde{W}(\mathbf{r}',\mathbf{r};0) \} \\ &\equiv -\frac{1}{2} \nabla_{\mathbf{r}} \tilde{W}(\mathbf{r},\mathbf{r};0). \end{aligned} \quad (3.49)$$

Hence, the part of the potential energy of a particle due to its own presence in a polarizable surrounding is equal to $\frac{1}{2} \tilde{W}(\mathbf{r},\mathbf{r};0)$. This explains the second contribution to the self-energy on the right-hand side of (3.46). In COHSEX the energy dependence of the screened interaction is completely neglected. This makes COHSEX less useful for actual calculations in semiconductors. The band

gap energies in semiconductors calculated within the framework of this approximation generally deviate substantially from the experimental values [64,65,141]. Apparently, *dynamical* screening effects have to be taken into account in order to obtain acceptable results for excitation properties.

The third approximation method we introduce here, does take account, in a way, of *dynamical* screening effects. In this method it is assumed that the plane-wave matrix elements of \tilde{W} have the following form [64,65,141]:

$$\tilde{W}_{\mathbf{K},\mathbf{K}'}^{pp}(\mathbf{k};\epsilon) = \frac{-1}{2} e_{\mathbf{K},\mathbf{K}'}(\mathbf{k}) \tilde{W}_{\mathbf{K},\mathbf{K}'}(\mathbf{k};0) \times \left\{ \frac{1}{\epsilon - e_{\mathbf{K},\mathbf{K}'}(\mathbf{k}) + i\eta} - \frac{1}{\epsilon + e_{\mathbf{K},\mathbf{K}'}(\mathbf{k}) - i\eta} \right\}, \quad \eta \downarrow 0, \quad (3.50)$$

in which the points $\epsilon = \pm e_{\mathbf{K},\mathbf{K}'}(\mathbf{k}) \mp i\eta$, with $e_{\mathbf{K},\mathbf{K}'}(\mathbf{k}) > 0$ ($\geq E_g$) and $\eta \downarrow 0$, are the so-called *plasmon poles* (see the discussion of the last section concerning the plasmon poles); the function $\tilde{W}_{\mathbf{K},\mathbf{K}'}(\mathbf{k};0)$ is assumed to be *exact*. The function $\tilde{W}^{pp}(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$ can readily be obtained by using the relation (3.3).

As we observe, expression (3.50) is strongly motivated by the plane-wave representation of the exact relation (3.41), (see also (3.42)), however, contrary to the plane-wave representation of the latter relation, the above expression consists of two terms only (notice that the s summations in (3.42a) and (3.42b) are maintained after Fourier transformation). Although by this oversimplification the fine structure of $\tilde{W}_{\mathbf{K},\mathbf{K}'}(\mathbf{k};\epsilon)$ as a function of energy is completely neglected, it turns out that $\tilde{W}_{\mathbf{K},\mathbf{K}'}^{pp}(\mathbf{k};\epsilon)$ is able to describe some global features of the exact function, specifically for energies not too close to the poles, very well [65]. Note that $\tilde{W}_{\mathbf{K},\mathbf{K}'}^{pp}(\mathbf{k};0)$, and equally well $\tilde{W}^{pp}(\mathbf{r}_1, \mathbf{r}_2; 0)$, are exact by construction. Moreover, \tilde{W}^{pp} is, just as the exact function, an even function of energy. The asymptotic behavior of \tilde{W} for large values of $|\epsilon|$, namely $\mathcal{O}(|\epsilon|^{-2})$, is correctly described by \tilde{W}^{pp} as well. Note that since $\tilde{W}_{\mathbf{K},\mathbf{K}'}(\mathbf{k};0)$ is real-valued, the imaginary part of $\tilde{W}_{\mathbf{K},\mathbf{K}'}^{pp}(\mathbf{k};\epsilon)$ is easily seen to consist of two δ functions if it is assumed that $e_{\mathbf{K},\mathbf{K}'}(\mathbf{k})$ is real; the δ functions are symmetrically located with respect to the origin.

The behavior of $\tilde{W}_{\mathbf{K},\mathbf{K}'}^{\text{PP}}(\mathbf{k};\epsilon)$ is governed by two "parameters", namely $\tilde{W}_{\mathbf{K},\mathbf{K}'}(\mathbf{k};0)$ and $e_{\mathbf{K},\mathbf{K}'}(\mathbf{k})$; the former is the *exact* value, which is to be evaluated directly from \tilde{W} 's defining equation, while the latter can be solved, for instance, by equating the right-hand side of (3.50) and the exact function $\tilde{W}_{\mathbf{K},\mathbf{K}'}(\mathbf{k};\epsilon')$ at *some* energy ϵ' not equal to zero. As the behavior of $\tilde{W}_{\mathbf{K},\mathbf{K}'}(\mathbf{k};\epsilon')$ is *not* exactly described by $\tilde{W}_{\mathbf{K},\mathbf{K}'}^{\text{PP}}(\mathbf{k};\epsilon')$, the value of $e_{\mathbf{K},\mathbf{K}'}(\mathbf{k})$ to be obtained in this way, does depend on the value of ϵ' . To circumvent this ambiguity, one can employ one of the existing *sum rules* fixing the value of $e_{\mathbf{K},\mathbf{K}'}(\mathbf{k})$. A commonly used sum rule in this connection is the so-called *f* sum rule [64,65,141,144-148] which follows from gauge-invariance requirements of the employed theory [148].

It is obvious that the expression for the analytic continuation of \tilde{W}^{PP} into the complex energy plane is the same as the one for real energies. At purely imaginary energy values $i\epsilon$ (ϵ real) we have

$$\tilde{W}_{\mathbf{K},\mathbf{K}'}^{\text{PP}}(\mathbf{k};i\epsilon) = \tilde{W}_{\mathbf{K},\mathbf{K}'}(\mathbf{k};0) \frac{e_{\mathbf{K},\mathbf{K}'}^2(\mathbf{k})}{\epsilon^2 + e_{\mathbf{K},\mathbf{K}'}^2(\mathbf{k})}. \quad (3.51)$$

From this expression we observe that $\tilde{W}_{\mathbf{K},\mathbf{K}'}^{\text{PP}}(\mathbf{k};i\epsilon)$ is a very smooth function of ϵ . Numerical calculations of \tilde{W} within the framework of the bubble approximation, described in this thesis, show that the thus-obtained $\tilde{W}_{\mathbf{K},\mathbf{K}'}(\mathbf{k};i\epsilon)$ functions agree surprisingly well [70] with the simple expressions (3.51). Because of the apparent lack of structure in \tilde{W} along the imaginary energy axis, and due to the fact that a reliable evaluation of the "exact" \tilde{W} along the imaginary energy axis, e.g. within bubble scheme, is a relatively simple task, we propose to obtain $e_{\mathbf{K},\mathbf{K}'}(\mathbf{k})$ as the solution of

$$\tilde{W}_{\mathbf{K},\mathbf{K}'}(\mathbf{k};ie_{\mathbf{K},\mathbf{K}'}(\mathbf{k})) = 1/2 \tilde{W}_{\mathbf{K},\mathbf{K}'}(\mathbf{k};0), \quad (3.52)$$

which is an alternative for using the above-mentioned *f* sum rule.

Now, as to the self-energy function, substitution of (3.51) in (3.23), and

evaluation of the ϵ' integral by, e.g., using Cauchy's residue theorem [123], leads to

$$\begin{aligned}
M^{\text{PP}}(\mathbf{r}_1, \mathbf{r}_2; \epsilon) &= -\hbar^{-1} \sum_{\mathbf{s}} f_{\mathbf{s}}(\mathbf{r}_1) f_{\mathbf{s}}^*(\mathbf{r}_2) \left\{ \Theta(\mu - \epsilon_{\mathbf{s}}) v(\mathbf{r}_1 - \mathbf{r}_2) \right. \\
&+ \frac{1}{\Omega} \left[\sum_{\mathbf{k}} \sum_{\mathbf{K}, \mathbf{K}'} e^{i(\mathbf{k} + \mathbf{K}) \cdot \mathbf{r}_1} \xi_{\mathbf{s}}(\epsilon, e_{\mathbf{K}, \mathbf{K}'}(\mathbf{k})) \tilde{W}_{\mathbf{K}, \mathbf{K}'}(\mathbf{k}; 0) e^{-i(\mathbf{k} + \mathbf{K}') \cdot \mathbf{r}_2} \right] \\
&\left. + [\Theta(\epsilon_{\mathbf{s}} - \epsilon) \Theta(\mu - \epsilon_{\mathbf{s}}) - \Theta(\epsilon - \epsilon_{\mathbf{s}}) \Theta(\epsilon_{\mathbf{s}} - \mu)] \tilde{W}^{\text{PP}}(\mathbf{r}_1, \mathbf{r}_2; \epsilon - \epsilon_{\mathbf{s}}) \right\}, \quad (3.53)
\end{aligned}$$

in which

$$\xi_{\mathbf{s}}(\epsilon, e) = \frac{e/2}{\epsilon - \epsilon_{\mathbf{s}} + e \operatorname{sgn}(\operatorname{Re}(\epsilon) - \epsilon_{\mathbf{s}})}. \quad (3.54)$$

We remark that, due to the fact that the function \tilde{W}^{PP} can approximate \tilde{W} much better along the imaginary energy axis than along the real axis, an alternative to (3.53) would be to replace \tilde{W} by \tilde{W}^{PP} only in the integral along the imaginary axis in (3.23) and not in the residue term.

In the previous section we discussed the possibility of Taylor expanding the GW self-energy function around $\epsilon = \nu$ and employing a truncated Taylor series as an extrapolation polynomial for all energies. In this connection, we would like to mention here that the evaluation of the first few Taylor coefficients can be greatly facilitated if one uses $M^{\text{PP}}(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$.

The last approximation method that we would like to discuss in this section is a mixture of a dynamic and a static approximation. It consists of replacing $\tilde{W}(\mathbf{r}_1, \mathbf{r}_2; \epsilon - \epsilon_{\mathbf{s}})$ in the residue term of (3.23) by its static value $\tilde{W}(\mathbf{r}_1, \mathbf{r}_2; 0)$. This method will be referred to as the Static-Pole Approximation (SPA). In this method one might also consider the possibility of replacing \tilde{W} along the imaginary energy axis by its plasmon-pole approximation \tilde{W}^{PP} , but in principle the exact evaluation of the integral term in (3.23) is assumed. Despite the approximate nature of M^{SPA} , it yields exact results for energies ϵ within the gap, i.e., satisfying $E_N - E_{N-1} < \epsilon < E_{N+1} - E_N$, (see (3.30)). The justification for using M^{SPA} outside the latter energy interval follows from the fact that in view of the

factor $[\Theta(\epsilon_s - \epsilon)\Theta(\mu - \epsilon_s) - \Theta(\epsilon - \epsilon_s)\Theta(\epsilon_s - \mu)] \equiv \Theta(\mu - \epsilon_s) - \Theta(\epsilon - \epsilon_s)$, for energies not far from the gap region the arguments of the contributing functions $\tilde{W}(\mathbf{r}_1, \mathbf{r}_2; \epsilon - \epsilon_s)$, for all s , are close to zero. As \tilde{W} is an *even* function of energy, the deviations of $\tilde{W}(\mathbf{r}_1, \mathbf{r}_2; \epsilon - \epsilon_s)$ with respect to $\tilde{W}(\mathbf{r}_1, \mathbf{r}_2; 0)$ will be of second order in $\epsilon - \epsilon_s$.

CHAPTER 4

TOWARDS ACTUAL CALCULATION OF THE GW SELF-ENERGY FUNCTION IN THE QUASI-PARTICLE APPROXIMATION

4.0 Introduction

This chapter starts with a closer examination of the bubble polarization function P , as the determination of this function is crucial for the determination of the self-energy function M^{GW} . In section 4.1 we first derive Kramers-Kronig (KK) type of relations for P in the bubble approximation, relations that are shown to hold as well for the *exact* screening-interaction function $\tilde{W}=W-v$. The importance of these KK relations lies in the possibility they offer to obtain the function P at real energy values by means of its determination at purely imaginary energy values (which is much less cumbersome). In section 4.2 we derive an explicit expression for plane-wave matrix elements $P_{\mathbf{K},\mathbf{K}'}(\mathbf{q};\epsilon)$ in the quasi-particle approximation scheme introduced in chapter 2. In the framework of an iteration procedure for the determination of M^{GW} , the obtained expression for P is useful in each iteration step, in the sense that the involved energies and wave-function coefficients refer to the solution of the quasi-particle wave equations in some intermediate step. In section 4.3 we present a number of symmetry properties that may reduce the computational work of obtaining P . In section 4.4 we outline details of the (linear) analytic tetrahedron method, a method that can be used in performing the necessary 1Bz integration in the expression of P . Though a special-point technique or, alternatively, a technique based on the KK relations obtained in 4.1, is to be preferred, both methods being less time-consuming, the analytic tetrahedron method is nevertheless advantageous, as it offers the possibility to check the accuracy of these methods. In section 4.5 we present a novel way of calculating of $P_{0,0}(\mathbf{q}=0;\epsilon)$ in which use is made of both the special-point and the KK technique. Section 4.6 is devoted to the discussion of some important calculational aspects in the determination of M^{GW} . A general discussion is devoted in this connection to the $\mathbf{q}=0$ behavior of $P_{\mathbf{K},\mathbf{K}'}(\mathbf{q};\epsilon)$ and $\tilde{W}_{\mathbf{K},\mathbf{K}'}(\mathbf{q};\epsilon)$ and measures are discussed which have to be taken

to perform the involved 1Bz integrations in the process of determining M^{GW} . These measures are necessary in order to cope, in numerical evaluations, with the singular behavior of matrix elements $\tilde{W}_{\mathbf{K},\mathbf{K}'}(\mathbf{q};\epsilon)$ at $\mathbf{q}=0$.

4.1 Integral Relations for the Bubble Polarization and Screened Interaction Function

At the end of chapter 2 we introduced the bubble approximation of the polarization function. As we will extensively make use of this approximation in the next sections, the present section will be devoted to a further discussion of the bubble polarization.

By referring to (2.39), we can write [cf. (2.42) and (A.13)]

$$P(1,2) = -\frac{2i}{\hbar} G(1,2^*)G(2,1), \quad (4.1)$$

which after Fourier transformation can be written as (see (2.14))

$$P(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = -\frac{2i}{\hbar} \int_{-\infty}^{\infty} \frac{d\epsilon'}{2\pi\hbar} G(\mathbf{r}_1, \mathbf{r}_2; \epsilon') G(\mathbf{r}_2, \mathbf{r}_1; \epsilon' - \epsilon) e^{i\eta_0 \epsilon' / \hbar}, \quad (4.2)$$

in which $\eta_0 > 0$ is infinitesimally small. By making use of the Lehmann representation (2.16), $P(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$ can be written

$$P(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = 2 \sum_{s, s'} I_{s, s'}(\epsilon) A_{s, s'}(\mathbf{r}_1, \mathbf{r}_2), \quad (4.3)$$

in which

$$I_{s, s'}(\epsilon) \equiv \frac{1}{2\pi i} \int_{-\infty}^{\infty} d\epsilon' \left\{ \frac{\Theta(\epsilon_s - \mu)}{\epsilon' - \epsilon_s + i\eta} + \frac{\Theta(\mu - \epsilon_s)}{\epsilon' - \epsilon_s - i\eta} \right\} \times \left\{ \frac{\Theta(\epsilon_s' - \mu)}{\epsilon' - \epsilon - \epsilon_s' + i\eta} + \frac{\Theta(\mu - \epsilon_s')}{\epsilon' - \epsilon - \epsilon_s' - i\eta} \right\} e^{i\eta_0 \epsilon' / \hbar}, \quad (4.4)$$

and

$$A_{s,s'}(\mathbf{r}_1, \mathbf{r}_2) \equiv f_s(\mathbf{r}_1) f_s^*(\mathbf{r}_2) f_{s'}(\mathbf{r}_2) f_{s'}^*(\mathbf{r}_1). \quad (4.5)$$

By choosing the combination $f_s(\mathbf{r}_1) f_s^*(\mathbf{r}_2)$ to be real-valued [149], $A_{s,s'}(\mathbf{r}_1, \mathbf{r}_2)$ becomes real-valued and we have the following identities

$$A_{s,s'}(\mathbf{r}_1, \mathbf{r}_2) = A_{s,s'}(\mathbf{r}_2, \mathbf{r}_1) = A_{s',s}(\mathbf{r}_1, \mathbf{r}_2). \quad (4.6)$$

Inspection of (4.4) reveals that the exponential $\exp[i\eta_0 \epsilon' / \hbar]$ can as well be replaced by 1. By deforming the contour of integration in (4.4) into the complex ϵ' plane and making use of the Cauchy residue theorem we obtain

$$\begin{aligned} I_{s,s'}(\epsilon) &= \frac{\Theta(\epsilon_s - \mu) \Theta(\mu - \epsilon_{s'})}{\epsilon + \epsilon_{s'} - \epsilon_s + i\eta} - \frac{\Theta(\mu - \epsilon_s) \Theta(\epsilon_{s'} - \mu)}{\epsilon + \epsilon_{s'} - \epsilon_s - i\eta} \\ &\equiv \frac{\Theta(\mu - \epsilon_{s'}) - \Theta(\mu - \epsilon_s)}{\epsilon + \epsilon_{s'} - \epsilon_s + i\eta \operatorname{sgn}(\epsilon)}. \end{aligned} \quad (4.7)$$

The bubble polarization function can thus be expressed as

$$P(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = 2 \sum_{s,s'} \frac{\Theta(\mu - \epsilon_{s'}) - \Theta(\mu - \epsilon_s)}{\epsilon + \epsilon_{s'} - \epsilon_s + i\eta \operatorname{sgn}(\epsilon)} A_{s,s'}(\mathbf{r}_1, \mathbf{r}_2). \quad (4.8)$$

From (4.8) and (4.6) it can be deduced that the bubble polarization function satisfies the following symmetry relations

$$P(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = P(\mathbf{r}_2, \mathbf{r}_1; \epsilon) = P(\mathbf{r}_1, \mathbf{r}_2; -\epsilon). \quad (4.9)$$

Let us define $\sum_s^{\geq} \{\dots\}$ as being the summation over those terms for which $\epsilon_s \geq \mu$. If we then introduce the function $p(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$ by

$$p(r_1, r_2; \epsilon) = 2 \sum_s^> \sum_{s'}^< \frac{A_{s, s'}(r_1, r_2)}{\epsilon + \epsilon_{s'} - \epsilon_s + i\eta}, \quad (4.10)$$

it can, by using (4.6), readily be shown that

$$P(r_1, r_2; \epsilon) = p(r_1, r_2; \epsilon) + p(r_1, r_2; -\epsilon). \quad (4.11)$$

Obviously, only those functions $A_{s, s'}(r_1, r_2)$ have to be evaluated for which $\epsilon_s > \mu$ and $\epsilon_{s'} < \mu$. Note that, the singularities of the analytic continuation of $p(r_1, r_2; \epsilon)$ in the complex ϵ plane lay all in the lower half plane. Therefore $p(r_1, r_2; \epsilon)$ is the Fourier transform of a causal function, which is known as the *causal polarization* of Adler and Wiser [101,150]. Making use of (2.15), it can be readily verified that this function in the time domain is identically vanishing for $t_1 - t_2 < 0$. It therefore immediately follows that the real and imaginary parts of $p = p' + ip''$ satisfy the Kramers-Kronig (KK) relations [151]

$$p'(r_1, r_2; \epsilon) = \frac{-1}{\pi} \mathcal{P} \int_{-\infty}^{\infty} d\epsilon' \frac{p''(r_1, r_2; \epsilon')}{\epsilon - \epsilon'}, \quad (4.12a)$$

$$p''(r_1, r_2; \epsilon) = \frac{1}{\pi} \mathcal{P} \int_{-\infty}^{\infty} d\epsilon' \frac{p'(r_1, r_2; \epsilon')}{\epsilon - \epsilon'}. \quad (4.12b)$$

The functions p' and p'' can be written as

$$p'(r_1, r_2; \epsilon) = 2 \sum_s^> \sum_{s'}^< \mathcal{P} \left(\frac{1}{\epsilon + \epsilon_{s'} - \epsilon_s} \right) A_{s, s'}(r_1, r_2), \quad (4.13a)$$

$$p''(r_1, r_2; \epsilon) = -2\pi \sum_s^> \sum_{s'}^< \delta(\epsilon + \epsilon_{s'} - \epsilon_s) A_{s, s'}(r_1, r_2). \quad (4.13b)$$

where it is understood that the involved \mathcal{P} and δ "functions" stand for

$$\mathcal{P} \left(\frac{1}{\epsilon} \right) \rightarrow \frac{\epsilon}{\epsilon^2 + \eta^2}, \quad \eta \downarrow 0, \quad (4.14a)$$

$$\delta(\epsilon) \rightarrow \frac{\eta/\pi}{\epsilon^2 + \eta^2}, \quad \eta \downarrow 0, \quad (4.14b)$$

Equations (4.14) will be used in order to define analytic continuations of $p'(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$ and $p''(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$ into the complex ϵ plane. Note incidentally, that the explicit representations (4.14) can also be used for defining *generalized* Kramers-Kronig relations between two constituent parts of, more generally, a causal function, in which, unlike (4.13) these parts are not the real and imaginary parts. Such situations may indeed occur; as an example of such a function, one may think, e.g., of one of the *Fourier components* with respect to \mathbf{r}_1 and \mathbf{r}_2 of the function $p(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$ in (4.10). This point was noted first by Johnson [145].

From (4.13b) we conclude that $p''(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$ is identically vanishing for energy values satisfying $\epsilon < E_g$. Consequently, (4.12a) reduces to

$$p'(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = \frac{-1}{\pi} \mathcal{P} \int_{E_g}^{\infty} d\epsilon' \frac{p''(\mathbf{r}_1, \mathbf{r}_2; \epsilon')}{\epsilon - \epsilon'}. \quad (4.15)$$

For the same reason $p''(\mathbf{r}_1, \mathbf{r}_2; -\epsilon')$ vanishes on the integration interval in (4.15), so that in view of (4.11) we can replace p'' in (4.15) by P'' of the bubble polarization function, or

$$p'(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = \frac{-1}{\pi} \mathcal{P} \int_{E_g}^{\infty} d\epsilon' \frac{P''(\mathbf{r}_1, \mathbf{r}_2; \epsilon')}{\epsilon - \epsilon'}. \quad (4.16)$$

Using (4.14a), the analytic continuation of (4.16) for purely imaginary energy can be written

$$p'(\mathbf{r}_1, \mathbf{r}_2; i\epsilon) = \frac{-1}{\pi} \int_{E_g}^{\infty} d\epsilon' \frac{P''(\mathbf{r}_1, \mathbf{r}_2; \epsilon')}{i\epsilon - \epsilon'}. \quad (4.17)$$

In view of (4.11) we then obtain immediately

$$\begin{aligned}
P'(\mathbf{r}_1, \mathbf{r}_2; i\epsilon) &= \frac{2}{\pi} \int_{E_g}^{\omega} d\epsilon' \frac{\epsilon'}{\epsilon^2 + \epsilon'^2} P''(\mathbf{r}_1, \mathbf{r}_2; \epsilon') \\
&= P(\mathbf{r}_1, \mathbf{r}_2; i\epsilon),
\end{aligned} \tag{4.18}$$

where the last equality follows from the fact that the analytic continuation of P'' along the imaginary energy axis vanishes identically. This relation is of practical significance in view of the following considerations: First of all we note that because of absence of vanishing denominators, the direct evaluation of the polarization function along the imaginary energy axis will be much simpler than along the real energy axis. Let us therefore suppose that $P(\mathbf{r}_1, \mathbf{r}_2; i\epsilon)$ has been calculated. Then, relation (4.18) can be considered as a Fredholm integral equation of the first kind [152], the solution of which gives the imaginary part of the polarization function, P'' , or p'' , along the real energy axis. Subsequent application of (4.15) results in the values of p' , or P' , for all real energy values. This procedure will facilitate the determination of the polarization function along the real energy axis enormously, as will be discussed in more detail in section 4.5.

In the remainder of this section we will be concerned with the screened interaction function \tilde{W} . We will derive integral relations for \tilde{W} which are similar to the above-derived relations for the bubble polarization function. A noticeable difference, however, is that the relations to be derived pertain to the *exact* screened interaction function and not necessarily as in the case of P , to some special approximation of \tilde{W} . Let us therefore consider the exact screened interaction function W and particularly the screening-correction part of it, i.e., $\tilde{W}(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = W(\mathbf{r}_1, \mathbf{r}_2; \epsilon) - v(\mathbf{r}_1 - \mathbf{r}_2)$. From (3.41) and (3.42) we deduce that \tilde{W} can be written as

$$\tilde{W}(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = \tilde{w}(\mathbf{r}_1, \mathbf{r}_2; \epsilon) + \tilde{w}(\mathbf{r}_1, \mathbf{r}_2; -\epsilon), \tag{4.19}$$

in which

$$\tilde{w}(\mathbf{r}_1, \mathbf{r}_2; \epsilon) = 2 \sum_s \frac{1}{\epsilon - \epsilon'_s + i\eta} w_s(\mathbf{r}_1) w_s^*(\mathbf{r}_2). \tag{4.20}$$

As $\tilde{w}(r_1, r_2; \epsilon)$ is *causal*, we have the KK relations

$$\tilde{w}'(r_1, r_2; \epsilon) = \frac{-1}{\pi} \mathcal{P} \int_{-\infty}^{\infty} d\epsilon' \frac{\tilde{w}''(r_1, r_2; \epsilon')}{\epsilon - \epsilon'}, \quad (4.21a)$$

$$\tilde{w}''(r_1, r_2; \epsilon) = \frac{1}{\pi} \mathcal{P} \int_{-\infty}^{\infty} d\epsilon' \frac{\tilde{w}'(r_1, r_2; \epsilon')}{\epsilon - \epsilon'}, \quad (4.21b)$$

in which the primed functions are defined as

$$\tilde{w}'(r_1, r_2; \epsilon) = 2 \sum_s \mathcal{P} \left(\frac{1}{\epsilon - \epsilon'_s} \right) w_s(r_1) w_s^*(r_2), \quad (4.22)$$

$$\tilde{w}''(r_1, r_2; \epsilon) = -2\pi \sum_s \delta(\epsilon - \epsilon'_s) w_s(r_1) w_s^*(r_2). \quad (4.23)$$

The functions $w_s(r_1) w_s^*(r_2)$ can be chosen real [153], just as the earlier functions $f_s(r_1) f_s^*(r_2)$. Reality of $w_s(r_1) w_s^*(r_2)$ implies that \tilde{W}' and \tilde{W}'' are the real and imaginary parts of \tilde{W} . The same remarks as made below (4.14) can be made here.

As in the case of the bubble polarization function, we can, making use of (4.14a), (4.19) and (4.21a), obtain

$$\tilde{W}(r_1, r_2; i\epsilon) = \frac{2}{\pi} \int_{-\infty}^{\infty} d\epsilon' \frac{\epsilon'}{\epsilon^2 + \epsilon'^2} \tilde{w}''(r_1, r_2; \epsilon'), \quad (4.24)$$

which, similar to (4.18), can again be considered as a Fredholm integral equation of the first kind for the "unknown" function $\tilde{w}''(r_1, r_2; \epsilon)$ along the real ϵ' axis.

We should add an important remark here concerning the calculation of \tilde{W} along these lines at real energies in the framework of some effective-potential approximation. In such a scheme, one could directly use (4.19) and (4.20), in which the functions w_s are calculated from independent-particle density deviation functions (see (3.17) and (3.43)), to obtain an approximation for W . However, it can be shown that this approximation leads precisely to

$W=(1+vP)v$, where P is the bubble polarization function, approximated by using the corresponding independent-particle wave functions in the Green functions G . On the other hand, if one calculates W via $W=\epsilon^{-1}v$ with $\epsilon=1-vP$, we have in fact the formal relationship $W=(1+vP +vPvP+ \dots)v$. Thus, the former calculation is just the first-order approximation of the latter one. Both results are approximate W functions; one cannot draw conclusions concerning the superiority of the one above the other *a priori*. Results obtained by inverting $(1-vP)$ are believed to be much more reliable [154-156]. Apparently, only the complete series is capable of producing the collective excitations of the systems (plasmons).

In short, we have introduced two pairs of KK relations, one for p' , p'' and the other for \tilde{w}', \tilde{w}'' . The function $p=p'+ip''$ concerns the bubble *approximation* of the dielectric function while $\tilde{w}=\tilde{w}'+i\tilde{w}''$ is related to the *exact* screened interaction function. Whether the *exact* polarization function can also be related to a function similar to p , say ρ , whose real and imaginary part satisfy KK relations, has not been investigated here. We mention in this connection that according to Kirzhnitz [157,158], and Dolgov and Maksimov [158], the *exact* function $\rho(r_1, r_2; \epsilon)$, contrary to the one in (4.11) and contrary to the *exact* $\tilde{w}(r_1, r_2; \epsilon)$ function may in general *not* be causal.

4.2 Quasi-Particle Approximation of the Bubble Polarization Function

If in the expression (4.3) for $P(r_1, r_2; \epsilon)$ in the bubble approximation, the quasi-particle approximation for the function G , as given in (2.34), had been used instead of the Lehmann representation, it is straightforward to show that the expression for the function $p(r_1, r_2; \epsilon)$ as given in (4.10) would change into

$$p(r_1, r_2; \epsilon) = 2 \sum_{\mathbf{k}, \ell_c} \sum_{\mathbf{k}', \ell_v} \tilde{g}_{\ell_c, \mathbf{k}} \tilde{g}_{\ell_v, \mathbf{k}'} \frac{A_{\mathbf{k}, \ell_c; \mathbf{k}', \ell_v}(r_1, r_2)}{\epsilon + \tilde{E}_{\ell_c}(\mathbf{k}') - \tilde{E}_{\ell_c}(\mathbf{k}) + i\eta}, \quad (4.25)$$

in which $\tilde{E}_{\ell_v}(\mathbf{k}')$ and $\tilde{E}_{\ell_c}(\mathbf{k})$ are the generally complex-valued valence- and conduction-band energies, respectively (see (2.34)), and

$$\begin{aligned}
& A_{\mathbf{k},\ell,\mathbf{k}',\ell'}(r_1,r_2) \\
&= \tilde{\varphi}_{\ell,\mathbf{k}}(r_1)\tilde{\varphi}_{\ell,-\mathbf{k}}(r_2)\tilde{\varphi}_{\ell',\mathbf{k}'}(r_2)\tilde{\varphi}_{\ell',-\mathbf{k}'}(r_1), \tag{4.26}
\end{aligned}$$

where the quasi-particle wave function $\tilde{\varphi}_{\ell,\mathbf{k}}(r;\tilde{E}_{\ell}(\mathbf{k}))$ have been given in short-hand notation (see (2.34) and below it). Moreover (see (2.32))

$$\tilde{g}_{\ell,\mathbf{k}} = \left[1 - \frac{\partial E_{\ell}(\mathbf{k};\epsilon)}{\partial \epsilon} \Big|_{\epsilon=\tilde{E}_{\ell}(\mathbf{k})} \right]^{-1}. \tag{4.27}$$

The vectors \mathbf{k}, \mathbf{k}' occurring in (4.25) are vectors in 1Bz. We remark that by definition (see chapter 2, the text between (2.32) and (2.33)) the imaginary part of $\tilde{E}_{\ell_v}(\mathbf{k})$ is positive and that of $\tilde{E}_{\ell_c}(\mathbf{k})$ negative so that $\text{Im}\{\tilde{E}_{\ell_v}(\mathbf{k}')-\tilde{E}_{\ell_c}(\mathbf{k})\}>0$. This is consistent with $i\eta$ in the denominator of (4.25) which in case of real energies guarantees the proper position of the poles of p in the complex ϵ plane. Real energies show up if $\tilde{\varphi}_{\ell,\mathbf{k}}$ functions in (4.26) are the solutions of an approximate quasi-particle equation (2.21) with a Hermitian self-energy function M .

Starting from the general expression (4.25) we want to arrive at plane-wave matrix elements $p_{\mathbf{K},\mathbf{K}'}(\mathbf{q};\epsilon)$. The reason is that from this the matrix elements of the dielectric matrix $\epsilon=1-vP$ can be constructed. Subsequent inversion of ϵ (see (A.14)) leads to matrix elements $W_{\mathbf{K},\mathbf{K}'}(\mathbf{q};\epsilon)$ needed in the evaluation of $M_{\mathbf{G},\mathbf{G}'}^{\text{GW}}(\mathbf{k};\epsilon)$. According to (4.26) and (3.4), we need the Fourier transform:

$$\begin{aligned}
& \frac{1}{\Omega} \int d^3r_1 d^3r_2 e^{-i(\mathbf{q}+\mathbf{K})\cdot\mathbf{r}_1} A_{\mathbf{k},\ell,\mathbf{k}',\ell'}(r_1,r_2) e^{i(\mathbf{q}+\mathbf{K}')\cdot\mathbf{r}_2} \\
&= \frac{1}{\Omega} \delta_{\mathbf{k},\mathbf{k}'+\mathbf{q}+\mathbf{K}_0} a_{\mathbf{K}}(-\mathbf{k},\mathbf{k}'+\mathbf{q};\ell,\ell) a_{-\mathbf{K}'}(\mathbf{k},-\mathbf{k}'-\mathbf{q};\ell,\ell), \tag{4.28}
\end{aligned}$$

in which

$$\begin{aligned}
a_{\mathbf{K}}(-\mathbf{k}', \mathbf{k}' + \mathbf{q}; \ell', \ell) &= \int_{\Omega} d^3r \tilde{u}_{\ell', -\mathbf{k}'}(\mathbf{r}) e^{-i\mathbf{K} \cdot \mathbf{r}} \tilde{u}_{\ell, \mathbf{k}' + \mathbf{q}}(\mathbf{r}) \\
&\equiv \sum_{\mathbf{G}} \tilde{d}_{\ell', -\mathbf{k}'}(-\mathbf{G}) \tilde{d}_{\ell, \mathbf{k}' + \mathbf{q}}(\mathbf{K} + \mathbf{G}).
\end{aligned} \tag{4.29}$$

Here $\tilde{u}_{\ell, \mathbf{k}}(\mathbf{r})$ represents the periodic part of the Bloch function $\tilde{\varphi}_{\ell, \mathbf{k}}(\mathbf{r})$, i.e.

$$\tilde{\varphi}_{\ell, \mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k} \cdot \mathbf{r}} \tilde{u}_{\ell, \mathbf{k}}(\mathbf{r}). \tag{4.30}$$

The plane-wave coefficients $\tilde{d}_{\ell, \mathbf{k}}(\mathbf{G})$ have been introduced in (3.7). The reciprocal lattice vector \mathbf{K}_0 in (4.28) is added to ensure that $\mathbf{k}' + \mathbf{q} + \mathbf{K}_0$ lays in 1Bz. In deriving (4.28) use has been made of $\tilde{u}_{\ell, \mathbf{k} + \mathbf{K}_0}(\mathbf{r}) = \exp(-i\mathbf{K}_0 \cdot \mathbf{r}) \tilde{u}_{\ell, \mathbf{k}}(\mathbf{r})$. We note that only if the self-energy function is Hermitian, the Bloch functions have the property that $\varphi_{\ell, -\mathbf{k}}(\mathbf{r}) = \varphi_{\ell, \mathbf{k}}^*(\mathbf{r})$, so that only in that case we have $d_{\ell, -\mathbf{k}}(-\mathbf{G}) = d_{\ell, \mathbf{k}}^*(\mathbf{G})$ and the related property $a_{-\mathbf{K}'}(\mathbf{k}', -\mathbf{k}' - \mathbf{q}; \ell', \ell') = a_{\mathbf{K}'}^*(-\mathbf{k}', \mathbf{k}' + \mathbf{q}; \ell', \ell')$.

Substitution of the right-hand side of (4.28) in the expression for the Fourier transform of $p(\mathbf{r}_1, \mathbf{r}_2; \epsilon)$, and performing the \mathbf{k} summation (after which \mathbf{k}' is changed to \mathbf{k}), we arrive at

$$\begin{aligned}
P_{\mathbf{K}, \mathbf{K}'}(\mathbf{q}; \epsilon) &= \frac{2}{\Omega} \sum_{\mathbf{k}} \sum_{\ell_{\mathbf{v}}, \ell_{\mathbf{c}}} \tilde{g}_{\ell_{\mathbf{v}}, \mathbf{k}} \tilde{g}_{\ell_{\mathbf{c}}, \mathbf{k} + \mathbf{q}} \\
&\times \frac{a_{\mathbf{K}}(-\mathbf{k}, \mathbf{k} + \mathbf{q}; \ell_{\mathbf{v}}, \ell_{\mathbf{c}}) a_{-\mathbf{K}'}(\mathbf{k}, -\mathbf{k} - \mathbf{q}; \ell_{\mathbf{v}}, \ell_{\mathbf{c}})}{\epsilon + \tilde{E}_{\ell_{\mathbf{v}}}(\mathbf{k}) - \tilde{E}_{\ell_{\mathbf{c}}}(\mathbf{k} + \mathbf{q}) + i\eta},
\end{aligned} \tag{4.31}$$

in which use has been made of the periodicity properties $\tilde{E}_{\ell}(\mathbf{k} + \mathbf{q} + \mathbf{K}_0) = \tilde{E}_{\ell}(\mathbf{k} + \mathbf{q})$ and $\tilde{g}_{\ell, \mathbf{k} + \mathbf{q} + \mathbf{K}_0} = \tilde{g}_{\ell, \mathbf{k} + \mathbf{q}}$. In case of a Hermitian self-energy function, e.g. if we are dealing with the first iteration cycle of some self-consistent calculation scheme, equation (4.31) may be written

$$P_{\mathbf{K},\mathbf{K}'}(\mathbf{q};\epsilon) = \frac{2}{\Omega} \sum_{\mathbf{k}} \sum_{\ell_{\mathbf{v}},\ell_{\mathbf{c}}} \frac{a_{\mathbf{K}}(-\mathbf{k},\mathbf{k}+\mathbf{q};\ell_{\mathbf{v}},\ell_{\mathbf{c}}) a_{\mathbf{K}'}^*(-\mathbf{k},\mathbf{k}+\mathbf{q};\ell_{\mathbf{v}},\ell_{\mathbf{c}})}{\epsilon + \epsilon_{\ell_{\mathbf{v}}}(\mathbf{k}) - \epsilon_{\ell_{\mathbf{c}}}(\mathbf{k}+\mathbf{q}) + i\eta} \quad (4.32)$$

in which the eigenvalues $\epsilon_{\ell_{\mathbf{v}}}$ and $\epsilon_{\ell_{\mathbf{c}}}$ are real and the \tilde{g} factors are equal to unity.

The plane-wave representation of the dielectric function can now be written

$$\epsilon_{\mathbf{K},\mathbf{K}'}(\mathbf{q};\epsilon) = \delta_{\mathbf{K},\mathbf{K}'} - v_{\mathbf{K},\mathbf{K}'}(\mathbf{q}) P_{\mathbf{K},\mathbf{K}'}(\mathbf{q};\epsilon), \quad (4.33)$$

where $v_{\mathbf{K},\mathbf{K}'}(\mathbf{q})$ is given by (3.27) and where $P_{\mathbf{K},\mathbf{K}'}(\mathbf{q};\epsilon)$ is the Fourier transform of $P(\mathbf{r}_1,\mathbf{r}_2;\epsilon)$, i.e.

$$P_{\mathbf{K},\mathbf{K}'}(\mathbf{q};\epsilon) = p_{\mathbf{K},\mathbf{K}'}(\mathbf{q};\epsilon) + p_{\mathbf{K},\mathbf{K}'}(\mathbf{q};-\epsilon). \quad (4.34)$$

4.3 Exploitation of Symmetry Properties

This section is devoted to symmetry properties which may help to facilitate the actual calculation of matrix elements $P_{\mathbf{K},\mathbf{K}'}(\mathbf{q};\epsilon)$ of $P(\mathbf{r}_1,\mathbf{r}_2;\epsilon)$. First we remark that $P(\mathbf{r}_1,\mathbf{r}_2;\epsilon)$ is invariant under the operations of the space group [159-162] of the system. This implies

$$\{ \mathcal{R}_j | \mathbf{R} \} P(\mathbf{r}_1,\mathbf{r}_2;\epsilon) = P(\underline{\beta}_j \mathbf{r}_1 + \tau_j + \mathbf{R}, \underline{\beta}_j \mathbf{r}_2 + \tau_j + \mathbf{R}) = P(\mathbf{r}_1,\mathbf{r}_2;\epsilon). \quad (4.35)$$

where the 3×3 matrix $\underline{\beta}_j$ denotes the matrix representation of the point-group part of the operation $\{ \mathcal{R}_j | \mathbf{R} \}$, τ_j is the accompanying non-primitive translation (if any) and \mathbf{R} denotes a lattice translation. By going over to integration variables $\mathbf{r}_1' = \underline{\beta}_j \mathbf{r}_1 + \tau_j + \mathbf{R}$ and $\mathbf{r}_2' = \underline{\beta}_j \mathbf{r}_2 + \tau_j + \mathbf{R}$ in the expression $P_{\mathbf{K},\mathbf{K}'}(\mathbf{q};\epsilon)$, as given in (3.4), it is straightforward to find

$$P_{\mathbf{K},\mathbf{K}'}(\underline{\beta}_j \mathbf{q};\epsilon) = e^{i \underline{\beta}_j^{-1}(\mathbf{K}-\mathbf{K}') \cdot \tilde{\tau}_j} P_{\underline{\beta}_j^{-1} \mathbf{K}, \underline{\beta}_j^{-1} \mathbf{K}'}(\mathbf{q};\epsilon), \quad (4.36)$$

in which $\tilde{\tau}_j$ belongs to the operation $\underline{\beta}_j^{-1}$. This relation is of great practical importance for cases in which the wave vector \mathbf{q} is a symmetry point, i.e. a point for which holds $\underline{\beta}_j \mathbf{q} = \mathbf{q}$, for (some of) the point-group operations in $\{\mathcal{R}_j | \mathbf{R}\}$. It is then obvious that quite a number of P-matrix elements can directly be obtained just by multiplication of some other matrix elements by a simple phase factor. In cases in which $\underline{\beta}_j \mathbf{q} = \mathbf{q} + \mathbf{K}_0$, where \mathbf{K}_0 is some reciprocal lattice vector, it is helpful to realize that $P_{\mathbf{K}, \mathbf{K}'}(\mathbf{q} + \mathbf{K}_0; \epsilon) = P_{\mathbf{K} + \mathbf{K}_0, \mathbf{K}' + \mathbf{K}_0}(\mathbf{q}; \epsilon)$.

Secondly, in view of the expression (4.31) for the function $p_{\mathbf{K}, \mathbf{K}'}(\mathbf{q}; \epsilon)$, being a summation (integration) over the wave vector \mathbf{k} , we may investigate the transformation properties with respect to point-group operations of the functions occurring in the (integrand) summand. To this end we first rewrite (4.31) as

$$p_{\mathbf{K}, \mathbf{K}'}(\mathbf{q}; \epsilon) = \frac{1}{4\pi^3} \sum_{\ell_v, \ell_c} \sum_{\underline{\alpha}_j \in \mathcal{G}_L} \int_{\text{IBz}} d^3k g_{\ell_v, \underline{\alpha}_j \mathbf{k}} g_{\ell_c, \underline{\alpha}_j \mathbf{k} + \mathbf{q}} \times \frac{a_{\mathbf{K}}(-\underline{\alpha}_j \mathbf{k}, \underline{\alpha}_j \mathbf{k} + \mathbf{q}; \ell_v, \ell_c) a_{-\mathbf{K}'}(\underline{\alpha}_j \mathbf{k}, -\underline{\alpha}_j \mathbf{k} - \mathbf{q}; \ell_v, \ell_c)}{\epsilon + \tilde{\mathbb{E}}_{\ell_v}(\underline{\alpha}_j \mathbf{k}) - \tilde{\mathbb{E}}_{\ell_c}(\underline{\alpha}_j \mathbf{k} + \mathbf{q}) + i\eta}, \quad (4.37)$$

where \mathcal{G}_L is the point group leaving the underlying Bravais lattice invariant, the so-called lattice point group. Here IBz is the *irreducible wedge* of 1Bz. It is that part of 1Bz which by application of all point-group operations of the underlying lattice fills 1Bz without producing overlapping parts and without leaving voids. In (4.37) we have replaced $\sum_{\mathbf{k}}(\dots)$ by $\Omega/8\pi^3 \int d^3k(\dots)$. We have generally

$$\tilde{\mathbb{E}}_{\ell}(\underline{\alpha}_j \mathbf{k} + \mathbf{K}_0) = \tilde{\mathbb{E}}_{\ell}(\mathbf{k}), \quad (4.38)$$

$$\tilde{\mathbb{E}}_{\ell, \underline{\alpha}_j \mathbf{k} + \mathbf{K}_0} = \tilde{\mathbb{E}}_{\ell, \mathbf{k}}, \quad (4.39)$$

for all $\underline{\alpha}_j \in \mathcal{G}_L$. As far as the transformation properties of the functions $a_{\mathbf{K}}$ and $a_{-\mathbf{K}'}$ is concerned, we may make use of the relation [20]

$$\mathcal{R}_j \tilde{\psi}_{\ell, \mathbf{k}}(\mathbf{x}) = \tilde{\psi}_{\ell, \underline{\beta}_j^{-1} \mathbf{k}}(\mathbf{x}), \quad (4.40)$$

for all elements $\underline{\beta}_j$ occurring in the space group of the system. A closely related property reads [160]

$$\tilde{d}_{\ell, \underline{\beta}_j \mathbf{k}}(\mathbf{G}) = e^{i(\mathbf{k} + \underline{\beta}_j \cdot \mathbf{G})} \cdot \tilde{\tau}_j \tilde{d}_{\ell, \mathbf{k}}(\underline{\beta}_j \cdot \mathbf{G}). \quad (4.41)$$

Let us, from now on, for simplicity assume that the point-group element $\underline{\beta}_j$ just make up the lattice point group \mathcal{P}_L . This is not generally the case, but it holds for the important example of the semiconductor Si, which has diamond structure, and $\mathcal{P}_L = O_h$. Thus identifying the point-group operations $\underline{\alpha}_j$ and $\underline{\beta}_j$, we may now consider the expression

$$a_{\mathbf{K}}(-\underline{\alpha}_j \mathbf{k}, \underline{\alpha}_j \mathbf{k} + \mathbf{q}; \ell, \ell') = \sum_{\mathbf{G}} \tilde{d}_{\ell, -\underline{\alpha}_j \mathbf{k}}(-\mathbf{G}) \tilde{d}_{\ell, \underline{\alpha}_j \mathbf{k} + \mathbf{q}}(\mathbf{K} + \mathbf{G}), \quad (4.42)$$

and conclude from (4.41) that it is always possible to rewrite the right-hand side of (4.42) in terms of $\tilde{d}_{\ell, \mathbf{k}'}$ coefficients in which the occurring \mathbf{k}' lay all in $1Bz$. As a matter of course a number of phase factors will be generated in this way, containing wave vectors in the whole $1Bz$, but from a computational point of view this poses no problem at all. The important point to note is, that it is possible to rewrite the integrand in (4.37) in such a way that occurring quantities, such as $\tilde{d}_{\ell, \mathbf{k}'}(\mathbf{G})$, $\tilde{E}_{\ell}(\mathbf{k}')$ or $\tilde{g}_{\ell, \mathbf{k}'}$ all have their \mathbf{k}' in $1Bz$. The advantage is obvious: the dense packing of those \mathbf{k}' points in $1Bz$ makes possible to calculate $\tilde{d}_{\ell, \mathbf{k}'}$, $\tilde{E}_{\ell}(\mathbf{k}')$ and $\tilde{g}_{\ell, \mathbf{k}'}$ in a few points in $1Bz$ only and then make use of some interpolation procedure.

4.4 The Analytic Linear Tetrahedron Method

In the previous section we have discussed a number of essential properties of the polarization function in the bubble approximation. In the present section we will discuss a Brillouin-zone integration technique which can be used in an actual calculation of P . The advantage of the method to be presented is that it can deal, without any problem, with the singularities of the integrand in (4.37) as a

function of \mathbf{k} . In fact, the well-known alternative method of special point integration is argued to be less adequate if not inadequate [115] due to this singular behavior of the integrand. Nevertheless, the special-point method need not to be rejected completely, for it will be demonstrated in section 4.5 how such a method can successfully be used in a direct calculation.

The analytic linear tetrahedron method [110,111] to be discussed here is specially suited to deal with the singularities in the integrand. We present for the first time some intermediate formulas in sufficient detail to be helpful when using this method. This section is rather technical and lengthy while containing many details of less interest to the general reader who may skip the remainder of this section and jump to section 4.5 without losing the general line.

In order to describe the analytic linear tetrahedron method, we start from the quasi-particle expression (4.37) for the quantity $p_{\mathbf{K},\mathbf{K}'}(\mathbf{q};\epsilon)$. We notice that in an independent particle approximation (i.e., in a first iteration step) the energies in the denominator become real, implying possible singularities of the integrand as a function of \mathbf{k} for real energy ϵ exceeding the gap energy. As the tetrahedron method can successfully deal with the singularities, we will consider (4.37) in the case of real energies.

As the evaluation of ℓ_v , ℓ_c and $\underline{\alpha}$ summations are trivial, we shall only deal with the \mathbf{k} integral over IBz. By making use of the relation $1/(x \pm i\eta) = \mathcal{P}(1/x) \mp i\pi\delta(x)$ we write the contribution of the IBz integral as follows:

$$I(\epsilon) = I'(\epsilon) + i I''(\epsilon), \quad (4.43)$$

in which

$$I'(\epsilon) = \mathcal{P} \left(\int_{\text{IBz}} d^3\mathbf{k} \frac{N(\mathbf{k})}{\epsilon - D(\mathbf{k})} \right), \quad (4.44)$$

$$I''(\epsilon) = -\pi \int_{\text{IBz}} d^3\mathbf{k} N(\mathbf{k}) \delta(\epsilon - D(\mathbf{k})), \quad (3.45)$$

in which we have introduced the short-hand notations

$$N(\mathbf{k}) = a_{\mathbf{K}}(-\underline{\alpha}\mathbf{k}, \underline{\alpha}\mathbf{k}+\mathbf{q}; \ell_{\mathbf{V}}, \ell_{\mathbf{C}}) a_{\mathbf{K}'}^*(-\underline{\alpha}\mathbf{k}, \underline{\alpha}\mathbf{k}+\mathbf{q}; \ell_{\mathbf{V}}, \ell_{\mathbf{C}}), \quad (4.46)$$

$$D(\mathbf{k}) = \epsilon_{\mathbf{C}}(\underline{\alpha}\mathbf{k}+\mathbf{q}) - \epsilon_{\mathbf{V}}(\mathbf{k}). \quad (4.47)$$

We will now consider the evaluation of $I'(\epsilon)$. Let IBz be subdivided into N_t small tetrahedrons Δ_j , with $j=1, 2, \dots, N_t$. The number N_t is chosen such that the variations of $N(\mathbf{k})$ and $D(\mathbf{k})$ within each tetrahedron are not substantial. We then replace $N(\mathbf{k})$ by a constant equal to its value at some point \mathbf{k}_j in Δ_j , e.g. the center of mass of Δ_j , and $D_j(\mathbf{k})$ by a linear function of \mathbf{k} , i.e.,

$$D_j(\mathbf{k}) = a_j + \mathbf{b}_j \cdot \mathbf{k}, \quad (4.48)$$

The function D_j in (4.48) is an approximation to D in (4.47) within the j th tetrahedron. In the present approximation $I'(\epsilon)$ reads

$$I'(\epsilon) \simeq \sum_j^{N_t} N(\mathbf{k}_j) \mathcal{P} \int_{\Delta_j} d^3\mathbf{k} \frac{1}{(\epsilon - a_j) - \mathbf{b}_j \cdot \mathbf{k}}. \quad (4.49)$$

Thus, each tetrahedron introduces four unknown quantities, the constant a_j and three coordinates of the vector \mathbf{b}_j . Knowledge of the exact $D(\mathbf{k})$ of (4.47) at the four vertices of each tetrahedron is sufficient to calculate the values of these quantities uniquely. This is one of the advantages of the analytic linear *tetrahedron* method as compared to other analytic linear methods [164] in which IBz is subdivided in elements such as cubes [165], triangular [166,167] or rectangular [167,168] prisms. In the latter methods there is some redundancy in the number of known quantities. As a consequence, in the latter methods, the interpolated function D is not a continuous function through IBz, introducing undesirable boundary effects. Moreover, not all IBz's can be composed by means of, e.g., cubes, contrary to tetrahedrons which can build up any IBz.

An obvious advantage of the linear tetrahedron method is that each integral in (4.49) can be evaluated analytically. However, although the integration is straightforward in principle, it is a tedious one in practice. In the sequel we will sketch the procedure to accomplish the analytic integration, give some intermediate formulas and present the final results. In view of its rather

technical and lengthy character, the reader may decide to skip this part and proceed directly to section 4.5. On the other hand, the formulas presented here may prove useful to potential further users of the analytic linear tetrahedron method.

We first define an orthonormal coordinate system $\{i_x', i_y', i_z'\}$, (not necessarily a right-handed system), with respect to which the vertices of the j th tetrahedron have the following form:

$$\mathbf{k}_1 = (0,0,0), \quad (4.50a)$$

$$\mathbf{k}_2 = (x_1', 0, 0), \quad (4.50b)$$

$$\mathbf{k}_3 = (x_2', y_2', 0), \quad (4.50c)$$

$$\mathbf{k}_4 = (x_3', y_3', z_3'), \quad (4.50d)$$

and in which the coordinates are all *positive*. In what follows we shall drop the subscript j in D_j , a_j and b_j ; there will be no confusion in doing so. If we then consider $D(\mathbf{k}_i)$'s, $i=1,2,3,4$, as known quantities, and write (4.48) in terms of the coordinates of the involved vectors with respect to the basis $\{i_x', i_y', i_z'\}$, we obtain a *triangular* system of four linear equations. After solution of this system of equations, we obtain (primed quantities refer to the basis $\{i_x', i_y', i_z'\}$)

$$a' = D(\mathbf{k}_1), \quad (4.51a)$$

$$b_x' = (D(\mathbf{k}_2) - D(\mathbf{k}_1)) / x_1', \quad (4.51b)$$

$$b_y' = (D(\mathbf{k}_3) - D(\mathbf{k}_1)) / y_2' + (D(\mathbf{k}_1) - D(\mathbf{k}_2)) x_2' / (x_1' y_2'), \quad (4.51c)$$

$$b_z' = (D(\mathbf{k}_4) - D(\mathbf{k}_1)) / z_3' + (D(\mathbf{k}_1) - D(\mathbf{k}_2)) x_3' / (x_1' z_3') \\ + (D(\mathbf{k}_1) - D(\mathbf{k}_3)) y_3' / (y_2' z_3') + (D(\mathbf{k}_2) - D(\mathbf{k}_1)) x_2' y_3' / (x_1' y_2' z_3'). \quad (4.51d)$$

As the transformation from \mathbf{k} in the original coordinate system $\{i_x, i_y, i_z\}$ to \mathbf{k}' in the above primed coordinate system consists in general of a translation and an *orthogonal* rotation we have $d^3\mathbf{k} = d^3\mathbf{k}'$. In the primed system the region of integration Δ_j is defined by the following inequalities:

$$\xi_1(z') < x' < \xi_2(z'), \\ v_1(z') < y' < v_1(z') + y_2'(x' - \xi_1(z')) / x_2', \\ 0 < z' < z_3', \\ \text{and}$$

$$\begin{aligned}
& \xi_2(z') < x' < \xi_3(z'), \\
v_1(z') < y' < v_2(z') + y_2'(x' - \xi_2(z')) / (x_2' - x_1'), \\
& 0 < z' < z_3',
\end{aligned} \tag{4.52}$$

in which

$$\begin{aligned}
\xi_1(z') &= (x_3' / z_3') z' \\
\xi_2(z') &= x_2' + (x_3' - x_2') z' / z_3', \\
\xi_3(z') &= x_1' + (x_3' - x_1') z' / z_3', \\
v_1(z') &= (y_3' / z_3') z', \\
v_2(z') &= y_2' + (y_3' - y_2') z' / z_3'.
\end{aligned} \tag{4.53}$$

The evaluation of a k integral over Δ_j in (4.49) is now a trivial, but tedious, task. The three basic integral expressions

$$\begin{aligned}
\int dx / (px + q) &= (1/p) \ln |px + q|, \\
\int dx \ln |px + q| &= (1/p)(px + q) \{ \ln |px + q| - 1 \}, \\
\int dx (px + q) \ln |px + q| &= (1/2p)(px + q)^2 \{ \ln |px + q| - 1/2 \},
\end{aligned} \tag{4.54}$$

are to be used in this connection. In giving the final result for the principal-value integral over Δ_j , we must distinguish between five situations:

(i) $D(k_i) \neq D(k_j)$, $i, j = 1, 2, 3, 4$, $i \neq j$

$$\begin{aligned}
& \oint_{\Delta_j} d^3k \frac{1}{(\epsilon - a) - b \cdot k} \\
&= 3\Omega_j \left\{ \frac{(\epsilon - D(k_1))^2}{(D(k_1) - D(k_2))(D(k_1) - D(k_3))(D(k_1) - D(k_4))} \ln \left| \frac{\epsilon - D(k_4)}{\epsilon - D(k_1)} \right| \right. \\
&+ \frac{(\epsilon - D(k_2))^2}{(D(k_2) - D(k_1))(D(k_2) - D(k_3))(D(k_2) - D(k_4))} \ln \left| \frac{\epsilon - D(k_4)}{\epsilon - D(k_2)} \right| \\
&+ \left. \frac{(\epsilon - D(k_3))^2}{(D(k_1) - D(k_3))(D(k_2) - D(k_3))(D(k_3) - D(k_4))} \ln \left| \frac{\epsilon - D(k_4)}{\epsilon - D(k_3)} \right| \right\}, \tag{4.55}
\end{aligned}$$

in which $\Omega_j = x_1' y_2' z_3' / 6$ is the volume of the j th tetrahedron.

$$\underline{\text{(ii) } D(\mathbf{k}_1) = D(\mathbf{k}_2) = D(\mathbf{k}_3) = D(\mathbf{k}_4) \equiv D}$$

$$\varphi \int_{\Delta_j} d^3 \mathbf{k} \frac{1}{(\epsilon - a) - \mathbf{b} \cdot \mathbf{k}} = \Omega_j / (\epsilon - D). \quad (4.56)$$

The contribution of the integral in (4.56) should be neglected in case $D = \epsilon$. This rather rare situation is just a consequence of our crude assumption made in (4.48). As can be verified, using (4.51a)-(4.51d), the occurrence of this situation, according to (4.48), would imply $D_j(\mathbf{k}) = \epsilon$, $\mathbf{k} \in \Delta_j$, which cannot actually be the case.

$$\underline{\text{(iii) } D(\mathbf{k}_1) = D(\mathbf{k}_2) = D(\mathbf{k}_3) \equiv D \neq D(\mathbf{k}_4)}$$

$$\begin{aligned} & \varphi \int_{\Delta_j} d^3 \mathbf{k} \frac{1}{(\epsilon - a) - \mathbf{b} \cdot \mathbf{k}} \\ &= 3\Omega_j \left\{ \frac{(\epsilon - D(\mathbf{k}_4))^2}{(D(\mathbf{k}_4) - D)^3} \ln \left| \frac{\epsilon - D}{\epsilon - D(\mathbf{k}_4)} \right| \right. \\ & \left. + \frac{3(\epsilon - D(\mathbf{k}_4))^2 / 2 + (\epsilon - D)^2 / 2 - 2(\epsilon - D)(\epsilon - D(\mathbf{k}_4))}{(D(\mathbf{k}_4) - D)^3} \right\}. \quad (4.57) \end{aligned}$$

In case $D = \epsilon$, we neglect the contribution of the integral in (4.57). The occurrence of $D = \epsilon$ in this situation is again attributable to the crude assumption made in (4.48).

$$\underline{\text{(iv) } D(\mathbf{k}_1) = D(\mathbf{k}_2) \equiv D' \neq D(\mathbf{k}_3) = D(\mathbf{k}_4) \equiv D'}$$

$$\varphi \int_{\Delta_j} d^3 \mathbf{k} \frac{1}{(\epsilon - a) - \mathbf{b} \cdot \mathbf{k}}$$

$$= 3\Omega_j \left\{ \frac{2(\epsilon-D')(\epsilon-D'')}{(D'-D'')^3} \ln \left| \frac{\epsilon-D'}{\epsilon-D''} \right| + \frac{2\epsilon-D'-D''}{(D'-D'')^2} \right\}. \quad (4.58)$$

(v) $D(\mathbf{k}_1)=D(\mathbf{k}_2)\equiv D$, $D(\mathbf{k}_3)\neq D$, $D(\mathbf{k}_4)\neq D$

$$\begin{aligned} & \oint_{\Delta_j} d^3\mathbf{k} \frac{1}{(\epsilon-\mathbf{a})-\mathbf{b}\cdot\mathbf{k}} \\ &= 3\Omega_j \left\{ \frac{(\epsilon-D(\mathbf{k}_3))^2}{(D-D(\mathbf{k}_3))^2(D(\mathbf{k}_4)-D(\mathbf{k}_3))} \ln \left| \frac{\epsilon-D(\mathbf{k}_3)}{\epsilon-D} \right| \right. \\ & \quad + \frac{(\epsilon-D(\mathbf{k}_4))^2}{(D-D(\mathbf{k}_4))^2(D(\mathbf{k}_3)-D(\mathbf{k}_4))} \ln \left| \frac{\epsilon-D(\mathbf{k}_4)}{\epsilon-D} \right| \\ & \quad \left. + \frac{(\epsilon-D)}{(D-D(\mathbf{k}_3))(D-D(\mathbf{k}_4))} \right\}. \quad (4.59) \end{aligned}$$

The relations needed for other possible orderings of $D(\mathbf{k}_i)$'s, than considered above, are easily deducible from the relations given in (iii)-(v), by just renaming the $D(\mathbf{k}_i)$ quantities.

It is obvious that the convergence of (4.49) for increasing number N_t of tetrahedrons is dependent on both the smoothness of $N(\mathbf{k})$ in IBz and the fact in how far the expansion of $D(\mathbf{k})$, (4.48), is valid.

We will now discuss the evaluation of $I'(\epsilon)$. We rewrite (4.45) as follows:

$$\begin{aligned} I'(\epsilon) &= -\pi \int_{\substack{D(\mathbf{k})=\epsilon \\ \mathbf{k} \in \text{IBz}}} d^2\mathbf{k} \frac{N(\mathbf{k})}{|\nabla_{\mathbf{k}} D(\mathbf{k})|} \\ &\simeq -\pi \sum_j^{N_t} \frac{1}{|b_j|} \int_{\mathcal{S}_j} d^2\mathbf{k} (c + \mathbf{d}\cdot\mathbf{k}), \quad (4.60) \end{aligned}$$

in which use has been made of both the approximate relation (4.48) and

$$N(\mathbf{k}) \simeq c + \mathbf{d} \cdot \mathbf{k}. \quad (4.61)$$

Here \mathcal{S}_j denotes the part of the surface of constant energy difference $D(\mathbf{k})=\epsilon$, which lies within the j th tetrahedron. If it exists, this surface can be regarded, consistent with (4.48), as a plane surface. Note that in (4.61) we propose to use an expression for $N(\mathbf{k})$ which is linear in \mathbf{k} . This deviates from our earlier choice made in obtaining $I'(\epsilon)$ of (4.49), in which $N(\mathbf{k})$ was put equal to its value at some arbitrary point within the tetrahedron. The reason is just the fact that in the present case the contribution of this term to $I''(\epsilon)$ can be relatively easily evaluated analytically. In what follows we will choose the labeling of the \mathbf{k}_i vectors in such a way that

$$D(\mathbf{k}_i) \leq D(\mathbf{k}_{i+1}). \quad (4.62)$$

Concerning the geometrical forms the surface \mathcal{S}_j can assume, we have (i) if $D(\mathbf{k}_1) < \epsilon < D(\mathbf{k}_2)$ or $D(\mathbf{k}_3) < \epsilon < D(\mathbf{k}_4)$, then \mathcal{S}_j is a triangle; and (ii) if $D(\mathbf{k}_2) < \epsilon < D(\mathbf{k}_3)$, then \mathcal{S}_j is a tetragon. In the two cases where $\epsilon < D(\mathbf{k}_1)$ and $\epsilon > D(\mathbf{k}_4)$, there are no intersections of the surface of constant energy difference and the tetrahedron Δ_j at all. Consequently, in the latter cases, the j th tetrahedron has no contribution to $I''(\epsilon)$. We note incidentally that, in case \mathcal{S}_j is a tetragon, the contribution of the surface integral on the right-hand side of (4.60) is most easily calculated analytically if one considers the latter contribution as being the difference of two surface integrals each over two triangles.

By introducing a new basis $\{i_x', i_y', i_z'\}$, in which \mathcal{S}_j lies in the $\{i_x', i_y'\}$ plane, such that one of the sides of \mathcal{S}_j lies along the i_x' axis, one obtains after some algebra

$$\begin{aligned} I''(\epsilon) &\simeq -\pi \sum_j^{N_t} \{c + \mathbf{d} \cdot \boldsymbol{\kappa}\} S_j / |b| \\ &\simeq -\pi \sum_j^{N_t} N(\boldsymbol{\kappa}) S_j / |b|, \end{aligned} \quad (4.63)$$

in which S_j denotes the area of the plane surface \mathcal{S}_j ,

$$S_j = \begin{cases} f_1, & D(\mathbf{k}_1) \leq \epsilon \leq D(\mathbf{k}_2), \\ f_1 - f_2, & D(\mathbf{k}_2) \leq \epsilon \leq D(\mathbf{k}_3), \\ f_4, & D(\mathbf{k}_3) \leq \epsilon \leq D(\mathbf{k}_4), \end{cases} \quad (4.64)$$

and in which κ stands for the vector of the center of gravity of \mathcal{S}_j ,

$$\kappa = \begin{cases} \kappa_1, & D(\mathbf{k}_1) \leq \epsilon \leq D(\mathbf{k}_2), \\ (\kappa_1 f_1 - \kappa_2 f_2) / (f_1 - f_2), & D(\mathbf{k}_2) \leq \epsilon \leq D(\mathbf{k}_3), \\ \kappa_4, & D(\mathbf{k}_3) \leq \epsilon \leq D(\mathbf{k}_4), \end{cases} \quad (4.65)$$

where

$$\kappa_i = \mathbf{k}_i + \frac{1}{3} (\epsilon - D(\mathbf{k}_i)) \sum_{\substack{n=1 \\ n \neq i}}^4 \frac{\mathbf{k}_n - \mathbf{k}_i}{D(\mathbf{k}_n) - D(\mathbf{k}_i)}, \quad i=1,2,4. \quad (4.66)$$

The functions f_i , are defined as

$$f_1 = \frac{\Omega_j |b|}{2} \frac{(\epsilon - D(\mathbf{k}_1))^2}{(D(\mathbf{k}_2) - D(\mathbf{k}_1))(D(\mathbf{k}_3) - D(\mathbf{k}_1))(D(\mathbf{k}_4) - D(\mathbf{k}_1))}, \quad (4.67a)$$

$$f_2 = \frac{\Omega_j |b|}{2} \frac{(\epsilon - D(\mathbf{k}_2))^2}{(D(\mathbf{k}_2) - D(\mathbf{k}_1))(D(\mathbf{k}_3) - D(\mathbf{k}_2))(D(\mathbf{k}_4) - D(\mathbf{k}_2))}, \quad (4.67b)$$

$$f_4 = \frac{\Omega_j |b|}{2} \frac{(\epsilon - D(\mathbf{k}_4))^2}{(D(\mathbf{k}_4) - D(\mathbf{k}_1))(D(\mathbf{k}_4) - D(\mathbf{k}_2))(D(\mathbf{k}_4) - D(\mathbf{k}_3))}. \quad (4.67c)$$

Note that the last expression in (4.63) would also have been obtained if we had replaced $N(\mathbf{k})$ over Δ_j by the constant value $N(\kappa)$.

By the above relations one can straightforwardly calculate the bubble (RPA) polarization function within the framework of (analytic) linear tetrahedron method.

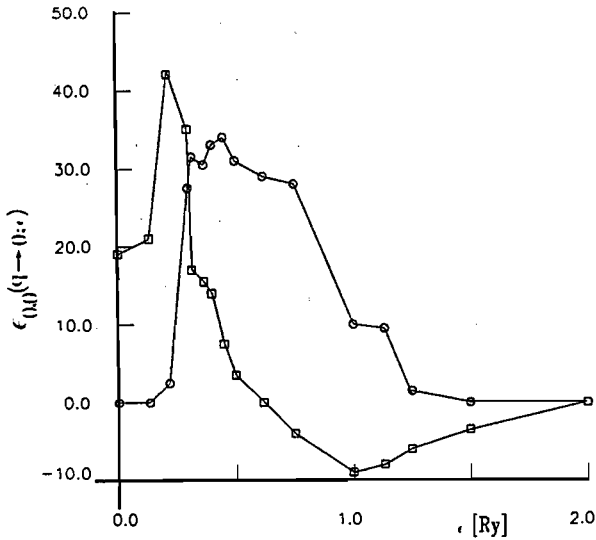


Fig.1. Real and imaginary parts of the dielectric response function $\epsilon_{0,0}(q;\epsilon)$, with $q \rightarrow 0$, as a function of energy. 1Bz has been subdivided into 8 tetrahedrons and the contributions of 10 bands have been incorporated. The marks \square (indicating the real part), and \circ indicate calculated values while the straight lines connecting them are merely a guide to the eye.

As an illustration of the above method we present in Fig. 1 numerical results concerning the real and imaginary parts of the dielectric-matrix element $\epsilon_{0,0}(q;\epsilon)$, with $q \rightarrow 0$, obtained within a simplified model of the semiconductor silicon, the description of which is given in section 4.5. The results given in Fig. 1 are far from being converged. From Fig. 2 it is observed that convergence in obtaining $P_{0,0}(q;0)$, with $q \rightarrow 0$, may be reached if the number of tetrahedrons is at least 500, say. This implies the use of a large storage capacity, while the computation time for obtaining this matrix element is excessively large. In the next section two alternative methods are discussed which do not suffer from these inconveniences. Nevertheless, the analytic linear tetrahedron method has its value as a reference method.

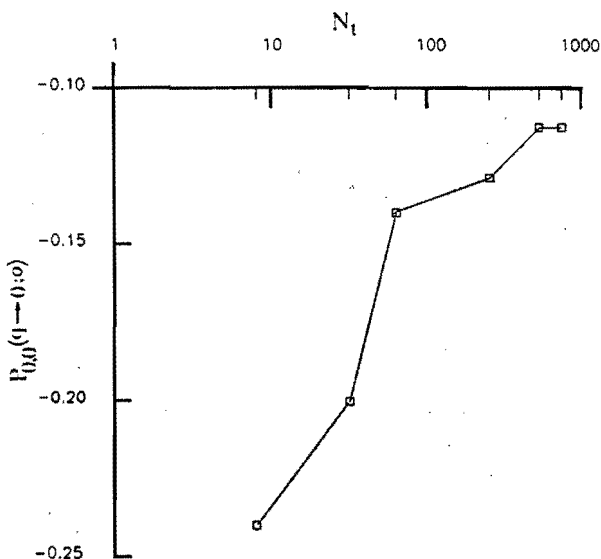


Fig.2. Convergence behavior of $P_{0,0}(q,0)$, with $q \rightarrow 0$, with regard to the increasing number of tetrahedron subdivisions of IBz. The marked points indicate the calculated $P_{0,0}(q,0)$ values (arbitrary units) obtained by subdividing IBz into $N_t = 8, 32, 64, 256, 512$ and 768 tetrahedrons. The straight lines connecting the marked points are guides to the eye. As in the case of Fig.1, the contributions of only 10 bands have been incorporated.

4.5 Numerical Determination of the Bubble Polarization Function (RPA) at Real Energies

The present section is an integral reprint of an article which has been published in Solid State Communication, Vol. 67, No. 1, pp. 7-11, (1988). The article describes a method to deal with the 1Bz integration for obtaining the polarization function at real energies. In this approach we use a special-point method, despite the singular character of the integrand. Therefore, the method to be presented in the article is a surprising alternative to the analytic linear tetrahedron method of the previous section, that is to say, surprising in the light

of certain remarks [115] which seem to claim that the special-point method cannot be used for a direct evaluation of P at real energies. Moreover, this section demonstrates the practical use of the integral relations established in section 4.1.

A NOVEL METHOD OF CALCULATING THE RPA DIELECTRIC
FUNCTION FOR A SEMICONDUCTOR
AT REAL ENERGIES

B. Farid, D. Lenstra and W. van Haeringen,
Department of Physics, Eindhoven University of Technology,
P.O.Box 513, 5600 MB Eindhoven, The Netherlands

(Received 20 april 1988 by M. Balkanski)

We demonstrate a novel method for direct calculation of the RPA dielectric function for a semiconductor at real energies based on a special point integration procedure. The method is relevant in the context of model-free *ab initio* calculations of dielectric properties and self-energies in semiconductors.

Calculation of the dielectric function for a semiconductor at real energies, even in the Random Phase Approximation (RPA), is known to be cumbersome at energies exceeding the energy gap. The reason is the occurrence of vanishing energy denominators which hampers the execution of the involved \mathbf{k} integration over the first Brillouin zone (1BZ). Direct calculation is nevertheless possible by applying some properly adjusted tetrahedron method^{1,2}, but such a procedure¹ asks for a fine mesh of \mathbf{k} points and is therefore very time consuming. It has been argued³ in this connection that special-point methods^{4,5,6} can not be used to evaluate the dielectric function at given real energies. Clearly, the reason for such remarks is the presence of singularities in the integrand.

In the present paper we will show how special-point methods can nevertheless successfully be used in *direct* calculations of the various matrix elements of the dielectric response function at real energies. Our method is based on two observations: Firstly, we note that the polarization function P at real energy can be written $P' + iP''$, in which P' and P'' fulfill Kramers-Kronig type of relations and in which P'' is a continuous function of real energy^{7,8}, to be expressed as a \mathbf{k} integration over 1BZ of weighted δ functions with energy arguments. Secondly, we note that in any discretization method for evaluating this \mathbf{k} integral, it is fully consistent to replace each δ function by a Lorentzian

with width roughly inversely proportional to the number of sampling points in 1BZ.

The method proposed, to be outlined and demonstrated below, by its simplicity, brings us a step closer to the model-free *ab initio* calculation of the self-energy in semiconductors. By using our method, the use of, e.g., plasmon-pole models⁹ can be avoided. We will confront the present special-point direct calculation of the above P'' at real energy with the directly related special-point calculation of the polarization function P at imaginary energies. This confrontation is of interest as P'' at real energy can also be determined by using a method for solving a Fredholm integral equation of the first kind¹⁰ in which both the function P'' at real energy and the function P at imaginary energy occur.

In RPA the matrix elements of the polarization function¹¹ in a plane-wave basis can easily be shown to be expressible as

$$P_{\mathbf{K}\mathbf{K}'}(\mathbf{k}_0; \epsilon) = P_{\mathbf{K}\mathbf{K}'}(\mathbf{k}_0; \epsilon) + P_{\mathbf{K}\mathbf{K}'}(\mathbf{k}_0; -\epsilon), \quad (1)$$

where

$$P_{\mathbf{K}\mathbf{K}'}(\mathbf{k}_0; \epsilon) = \frac{1}{4\pi^3} \int d^3k \sum_{lv,lc} \frac{A(l_v, l_c, \mathbf{k}, \mathbf{k}_0, \mathbf{K}, \mathbf{K}')}{\epsilon - \epsilon_{lc}(\mathbf{k} + \mathbf{k}_0) + \epsilon_{lv}(\mathbf{k}) + i\eta_1}. \quad (2)$$

Here \mathbf{k} denotes a wave vector in 1BZ, \mathbf{K} and \mathbf{K}' are reciprocal lattice vectors; lv and lc are valence- and conduction-band indices; $\epsilon_{lv}(\mathbf{k})$ and $\epsilon_{lc}(\mathbf{k})$ are band energies; \mathbf{k}_0 is the wave vector (in 1BZ) for which the matrix elements are to be determined. The numerators A are energy independent and can be completely given in terms of the plane-wave coefficients of the electron wave functions; these numerators are generally complex-valued except for $\mathbf{K} = \mathbf{K}'$ when they are real. The real quantity η_1 is supposed to be infinitesimally small and positive. It is easily seen from (2) that numerical problems due to vanishing denominators may arise if one performs the \mathbf{k} integration in case ϵ is larger than $\epsilon_g(\mathbf{k}_0) \equiv \min[\epsilon_{lc}(\mathbf{k} + \mathbf{k}_0) - \epsilon_{lv}(\mathbf{k})]$. The dielectric matrix element, related to (1) reads¹¹

$$\epsilon_{\mathbf{K}\mathbf{K}'}(\mathbf{k}_0; \epsilon) = \delta_{\mathbf{K}\mathbf{K}'} - \frac{e^2}{\epsilon_0 |\mathbf{k}_0 + \mathbf{K}|^2} P_{\mathbf{K}\mathbf{K}'}(\mathbf{k}_0; \epsilon). \quad (3)$$

The analytic continuation of (1) to purely imaginary energy $\epsilon=i\epsilon_1$ can be written

$$P_{\mathbf{K}\mathbf{K}'}(\mathbf{k}_0; i\epsilon_1) = -\frac{1}{(2\pi)^3} \int d^3k \Sigma_{l_v, l_c} \frac{A(l_v, l_c, \mathbf{k}, \mathbf{k}_0, \mathbf{K}, \mathbf{K}') \{ \epsilon_{l_c}(\mathbf{k}+\mathbf{k}_0) - \epsilon_{l_v}(\mathbf{k}) \}}{\epsilon_1^2 + \{ \epsilon_{l_c}(\mathbf{k}+\mathbf{k}_0) - \epsilon_{l_v}(\mathbf{k}) \}^2}. \quad (4)$$

As the right-hand side of (4) has no vanishing denominators (they are all larger than $(\epsilon_g(\mathbf{k}_0))^2$), we may consider the possibility of calculating (4) at given ϵ_1 by performing the \mathbf{k} integration using a special-point method. We will come to this furtheron. Let us first discuss the possibility of a *direct* evaluation of P at *real* energies.

For real energies, it makes sense to split the polarization into two contributions, using $1/(x+i\eta) = \mathcal{P}(1/x) - i\pi\delta(x)$, where \mathcal{P} stands for principal value. We may then write

$$P_{\mathbf{K}\mathbf{K}'}(\mathbf{k}_0; \epsilon) = P'_{\mathbf{K}\mathbf{K}'}(\mathbf{k}_0; \epsilon) + iP''_{\mathbf{K}\mathbf{K}'}(\mathbf{k}_0; \epsilon), \quad (5a)$$

with

$$P'_{\mathbf{K}\mathbf{K}'}(\mathbf{k}_0; \epsilon) = \mathcal{P}(P_{\mathbf{K}\mathbf{K}'}(\mathbf{k}_0; \epsilon)) \quad (5b)$$

and

$$P''_{\mathbf{K}\mathbf{K}'}(\mathbf{k}_0; \epsilon) = -\frac{1}{4\pi^2} \int d^3k \Sigma_{l_v, l_c} A(l_v, l_c, \mathbf{k}, \mathbf{k}_0, \mathbf{K}, \mathbf{K}') \cdot \{ \delta(\epsilon - \epsilon_{l_c}(\mathbf{k}+\mathbf{k}_0) + \epsilon_{l_v}(\mathbf{k})) - \delta(\epsilon + \epsilon_{l_c}(\mathbf{k}+\mathbf{k}_0) - \epsilon_{l_v}(\mathbf{k})) \}. \quad (5c)$$

It can directly be deduced from (4) and (5c) that the various matrix elements of $P(\mathbf{k}_0; i\epsilon_1)$ and $P''(\mathbf{k}_0; \epsilon)$ are related by

$$P_{\mathbf{K}\mathbf{K}'}(\mathbf{k}_0; i\epsilon_1) = \frac{2}{\pi} \int_0^\infty d\epsilon' \frac{\epsilon'}{\epsilon'^2 + \epsilon_1^2} P''_{\mathbf{K}\mathbf{K}'}(\mathbf{k}_0; \epsilon'). \quad (6)$$

In an attempt to evaluate the right-hand side of (5c) with a special-point method, we first replace all δ functions by Lorentzians with the same width η , i.e.

$$\delta(\epsilon - \tilde{\epsilon}) \longrightarrow \frac{\eta/\pi}{(\epsilon - \tilde{\epsilon})^2 + \eta^2}, \quad (7)$$

where it is noted that this replacement should lead to the original function P'' , strictly speaking, only when $\eta \rightarrow 0$, but that in practice it will be a good approximation for a finite value of η as well. After the substitution of (7) in (5c) there is no problem in performing the \mathbf{k} integration by making use of the special-point integration procedure. This means that the actual quantity to be calculated reads:

$$P''_{\mathbf{K}\mathbf{K}'}^{(n)}(\mathbf{k}_0; \epsilon) = -\frac{1}{2\pi^2 N_\alpha} \sum_{s=1}^n \sum_{l_V, l_C} \sum_{\alpha} w_s A(l_V, l_C, \alpha \mathbf{k}_s, \mathbf{k}_0, \mathbf{K}, \mathbf{K}') \\ \times \frac{\eta}{\pi} \left\{ \frac{1}{\{\epsilon - \epsilon_{l_C}(\alpha \mathbf{k}_s + \mathbf{k}_0) + \epsilon_{l_V}(\alpha \mathbf{k}_s)\}^2 + \eta^2} + \text{same expression with } \epsilon \rightarrow -\epsilon \right\}. \quad (8)$$

Here, α denotes all point-group operations of the crystal lattice with N_α the number of such operations; $\{\mathbf{k}_s; s=1, \dots, n\}$ is the set of special points to be considered in one irreducible wedge of 1BZ, and $\{w_s\}$ are the appropriate weighting factors. The numerical value for η should of course be consistent with the number of special points, that is, it should lead to sufficiently smooth dependence of $P''^{(n)}$ on ϵ .

In what follows we will also consider the \mathbf{K}, \mathbf{K}' elements of $P^{(n)}(\mathbf{k}_0; i\epsilon_1)$ which, according to (6) are related to the \mathbf{K}, \mathbf{K}' elements of $P''^{(n)}(\mathbf{k}_0; \epsilon)$ of (8) by

$$P_{\mathbf{K}\mathbf{K}'}^{(n)}(\mathbf{k}_0; i\epsilon_1) = \frac{2}{\pi} \int_0^\infty d\epsilon \frac{\epsilon}{\epsilon^2 + \epsilon_1^2} P''_{\mathbf{K}\mathbf{K}'}^{(n)}(\mathbf{k}_0; \epsilon). \quad (9)$$

To demonstrate our method, we have used a simplified model for the semiconductor Si with a (very small) energy cutoff of 2.57 Ry. Depending on the

Bloch-wave vector \mathbf{k} , this leads to only 15-22 plane waves in the expansion of the one-electron wave functions. The effective one-electron potential is borrowed from an earlier well-converged self-consistent calculation scheme¹² for Si with an energy cutoff of 11.5 Ry. The resulting model band structure has an indirect bandgap of 0.023 Ry, a direct bandgap of 0.237 Ry, and is qualitatively similar to the real Si bandstructure.

A series of numerical results obtained for $\text{Im } \epsilon_{00}(\mathbf{k}_0; \epsilon)$ along the real ϵ axis and evaluated with the use of (8) is given in Fig. 1. The number of Monkhorst-Pack⁵ (MP) special points varies from 10 to 60. The value of η in each case was determined as follows: A rough estimate tells us that approximately a number of $4 \times 11 \times q^3/2$ Lorentzians cover the energy interval of about 2 Ry, where $q = 2, 4, 6, 8, 10, 12, 14, 16, \dots$ depending on whether we have $n = 1, 2, 6, 10, 19, 28, 44, 60, \dots$ special points¹³. Here, the numbers 4 and 11 are the number of valence and conduction bands taken into account. It is important to realize that quite a large number of \mathbf{k} points lead to identical δ peaks for symmetry reasons, while only a few combinations of valence and conduction bands contribute significantly to the δ peak distribution pattern. We therefore come to the rough estimate of effectively $q^3/2$ Lorentzians covering 2 Ry, implying for the width η of each Lorentzian

$$\eta > 4/q^3 \text{ [Ry]}. \quad (10)$$

The η values chosen in Fig. 1 are 3.62 times larger than the minimum value implied by (10). Moreover, we have verified that the resulting curves are hardly sensitive to variations in the η values. Concerning the fine structures, one really needs to go as far as to 60 special points, but for global purposes (e.g., presumably in quasi-particle band structure calculations⁹) as few as 10 special points may already be good enough.

In view of the Fredholm integral equation (6) or (9), we may also consider a completely different way of calculating P'' at real energies by first calculating P along the imaginary axis with the use of the special-point method and after that solving the integral equation. Fig. 2 displays the dielectric matrix elements $\epsilon_{00}(\mathbf{k}_0; i\epsilon_1)$ for $n = 10, 19, 28, 44$ and 60 MP special points. We have observed that convergent results are hardest to obtain at $\epsilon_1=0$; from Table I it can be seen that convergence is already reached at 19 special points.

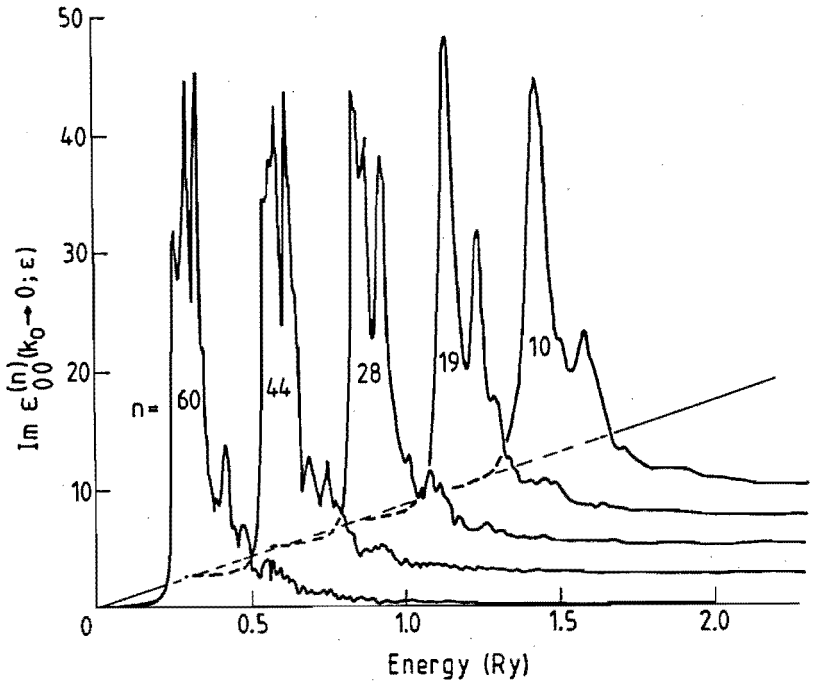


Fig.1. Numerical results for the imaginary part of the dielectric function for real energies obtained by a direct special-point calculation as described in the text. The five different curves correspond to five different numbers of MP special points as indicated.

Table I

Convergence of the static dielectric constant $\epsilon_{00}^{(n)}(\mathbf{k}_0 \rightarrow 0; 0)$. We take $\mathbf{k}_0 = 2\pi/a(1.25 \cdot 10^{-3}, 0, 0)$ while n is the number of MP special points used in the calculation.

n	1	2	6	10	19	28	44	60
$\epsilon_{00}^{(n)}$	35.71	18.20	14.00	12.66	12.07	12.07	11.94	11.97

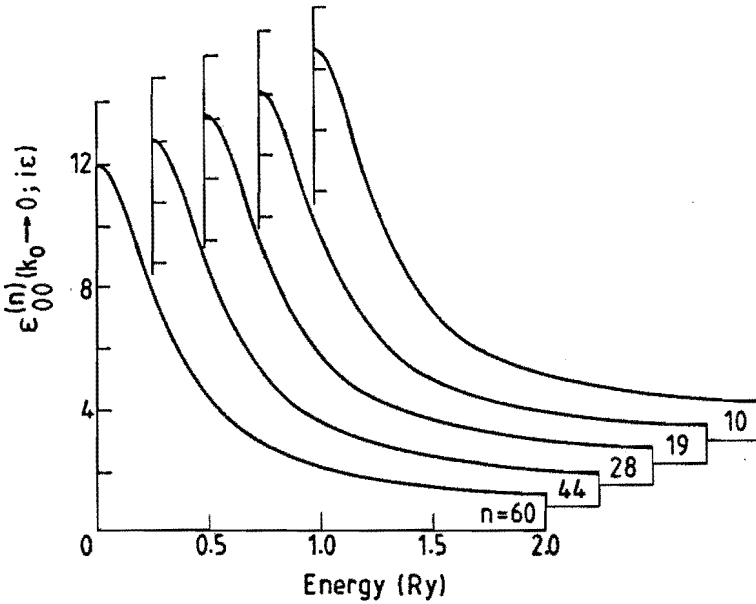


Fig.2. *Special-point calculations of the dielectric function for imaginary energies. Each curve is calculated using the indicated number of MP special points.*

Using the functions of Fig. 2 as input functions, an algorithm due to te Riele¹⁴, making use of the regularization method of Phillips¹⁵ and Tihonov¹⁶, was employed to solve the integral equation for P'' . The resulting functions $\text{Im } \epsilon_{00}(\mathbf{k}_0 \rightarrow 0; \epsilon)$ are given in Fig. 3 for the various numbers of special points used to calculate $P_{00}(i\epsilon_1)$. It is a well-known fact that integral equations of the type (6) or (9) are difficult to solve. Indeed, we observed that *minor* changes in $P(i\epsilon_1)$ can lead to enormous variations in $P''(\epsilon)$. Sensible results could only be obtained through the use of the regularization procedure that is employed in te Riele's procedure^{14,15,16,17}. As can be seen by comparing Figs. 3 and 1, this regularization washes out fine-structure details. In this connection we also mention that, since we know that $P''(\mathbf{k}_0; \epsilon) \equiv 0$, if $|\epsilon| < \epsilon_g(\mathbf{k}_0)$, the lower boundary of the integral is set equal to $\epsilon_g(\mathbf{k}_0)$. This is in agreement with the

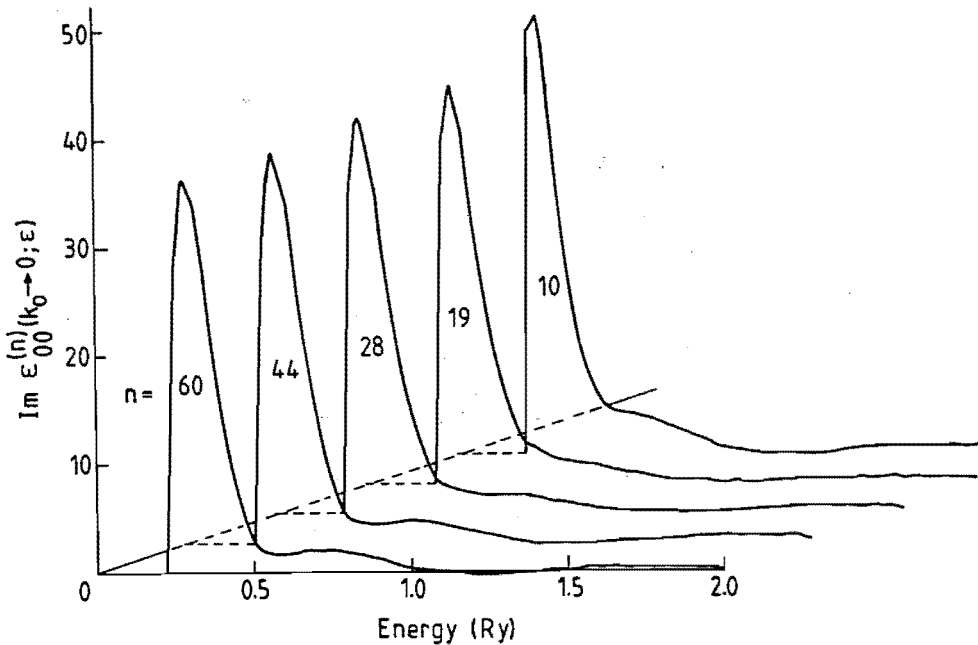


Fig.3. Numerical results for the imaginary part of the dielectric function for real energies obtained by solving the integral equation (6) for $P''_{00}(k_0 \rightarrow 0, \epsilon)$ using the respective input functions of Fig. 2 at 61 equidistant points covering the interval $[0, 2i]$. In applying to Riele's procedure¹⁴ the regularization parameter is 0.5×10^{-9} for all curves, while the linear functional L is chosen to be the identity functional.

general rule that in solving the integral equations of the first kind as much *a priori* information about the solution should be used as possible. Global properties can be obtained reasonably well in this way for as few as 10 special points. This is also illustrated in Table II, where the first four moments of $\epsilon''_{00}(k_0 \rightarrow 0; \epsilon) \equiv (-e^2/\epsilon_0 |k_0|^2) P''_{00}(k_0; \epsilon), |k_0| \rightarrow 0, \epsilon > 0$, are given as resulting from calculations within the framework of the above-described two methods and for various special point numbers. The zeroth moment, S , is just the area under the curve, $S = \int_0^\infty d\epsilon \epsilon''_{00}(k_0 \rightarrow 0, \epsilon)$, while the first moment, μ , is the mean value, $\mu = S^{-1} \int_0^\infty d\epsilon \epsilon \epsilon''_{00}(k_0 \rightarrow 0; \epsilon)$. All higher order moments are chosen to be central

Table II

The first four moments of the imaginary part of the dielectric function $\epsilon_{00}(\mathbf{k}_0 \rightarrow 0; \epsilon)$ for positive energies calculated in three different ways. In each column three numbers are presented. The first number is obtained by approximating the expressions (11), (12) or (13) by sums over special points, without the introduction of Lorentzians. The second number is obtained from the functions displayed in Fig. 1, i.e. it is obtained within the method in which the delta functions are smeared out to Lorentzians; the third number is obtained from the curves in Fig. 3, which are results of the integral-equation method.

n	10	19	28	44	60	n	10	19	28	44	60
S	6.13	5.94	5.94	5.91	5.92	μ	.374	.381	.381	.383	.382
	5.79	5.78	5.85	5.86	5.82		.395	.394	.391	.391	.391
	6.23	6.04	6.05	6.02	6.02		.381	.388	.389	.392	.392

n	10	19	28	44	60	n	10	19	28	44	60
I ⁽²⁾	.029	.029	.029	.029	.029	I ⁽³⁾	.016	.016	.016	.016	.016
	.041	.036	.032	.031	.031		.022	.019	.018	.017	.017
	.030	.031	.034	.036	.037		.012	.016	.022	.023	.024

moments, $I^{(n)} = S^{-1} \int_0^\infty d\epsilon (\epsilon - \mu)^n \epsilon''_{00}(\mathbf{k}_0 \rightarrow 0; \epsilon)$. The various moments may as well be determined from (5c) directly, i.e. from

$$S = \frac{e^2}{4\pi^2 \epsilon_0 |\mathbf{k}_0|^2} \sum_{lv,lc} \int d^3k A(lv,lc, \mathbf{k}, \mathbf{k}_0, K, K'), \quad (11)$$

$$\mu = \frac{e^2}{4\pi^2 \epsilon_0 |\mathbf{k}_0|^2 S} \sum_{lv,lc} \int d^3k A(lv,lc, \mathbf{k}, \mathbf{k}_0, K, K') [\epsilon_{lc}(\mathbf{k} + \mathbf{k}_0) - \epsilon_{lv}(\mathbf{k})], \quad (12)$$

$$I^{(n)} = \frac{e^2}{4\pi^2 \epsilon_0 |\mathbf{k}_0|^2 S} \sum_{lv,lc} \int d^3k A(lv,lc, \mathbf{k}, \mathbf{k}_0, K, K') [\epsilon_{lc}(\mathbf{k} + \mathbf{k}_0) - \epsilon_{lv}(\mathbf{k}) - \mu]^n, \quad (13)$$

where it is noted that the expressions (11), (12) and (13) can be evaluated by

replacing the 1BZ integrations by sums over special points; in doing so there is no need to introduce Lorentzians of width η . The moments calculated from (11), (12) and (13) for various MP special-point sets are also given in Table II. Comparison of these data with the others gives a rough idea of the global error introduced by the above methods.

In conclusion, we have outlined and demonstrated a novel method for direct calculation of the RPA dielectric function in a semiconductor at real energies using a special-point Brillouin-zone integration procedure. Fine-structure details such as the positions of sharp resonances can only be predicted accurately using in the order of 44 MP special points at least. Global properties can be obtained reasonably well by as few as 10 special points.

REFERENCES

1. G. Lehmann and M. Taut, *Phys. Stat. Sol. (b)* **54**, 469 (1972).
2. M.S. Methfessel, M.H. Boon and F.M. Mueller, *J. Phys. C: Solid State Phys.* **20**, 1069-1077 (1987).
3. A. Baldereschi and E. Tosatti, *Phys. Rev. B* **17**, 4710 (1978).
4. A. Baldereschi, *Phys. Rev. B* **7**, 5212 (1973).
5. D.J. Chadi and M.L. Cohen, *Phys. Rev. B* **8**, 5747 (1973).
6. H.J. Monkhorst and J.D. Pack, *Phys. Rev. B* **13**, 5188 (1976).
7. L. van Hove, *Phys. Rev.* **84**, 1189 (1953).
8. J.C. Phillips, *Phys. Rev.* **104**, 1263 (1956).
9. M.S. Hybertson and S.G. Louie, *Phys. Rev. B* **34**, 5390 (1986).
10. E.T. Whittaker and G.N. Watson, *A Course of Modern Analysis*, Cambridge, 1958 (Fourth Edition), pp. 211-231.
11. W. van Haeringen, B. Farid and D. Lenstra, *Physica Scripta* **T19**, 282 (1987).
12. P.J.H. Denteneer and W. van Haeringen, *J. Phys. C: Solid State Phys.* **18**, 4127 (1985).
13. P.J.H. Denteneer, Ph.D Thesis, Eindhoven University of Technology (1987).
14. H.J.J. te Riele, *Computer Physics Comm.* **36**, 423 (1985).
15. D.L. Phillips, *J. ACM* **9**, 84 (1962).
16. A.N. Tihonov, *Soviet Math. Dokl.* **4**, 1035, 1624 (1963).
17. H.J.J. te Riele, *Colloquium Numerical Treatment of Integral Equations*, MC Syllabus 41, Mathematisch Centrum, Amsterdam 1979.

4.6 Back to M^{GW}

In the preceding sections we have discussed a number of ingredients necessary for the calculation of the GW self-energy function. In this section we deal with some computational aspects of an actual calculation.

The invariance of the self-energy function on application of space-group operations $\{ \mathcal{R}_j | \mathbf{R} \}$, i.e.

$$M(\underline{\beta}_j \mathbf{r} + \boldsymbol{\tau}_j + \mathbf{R}, \underline{\beta}_j \mathbf{r}' + \boldsymbol{\tau}_j + \mathbf{R}; \epsilon) = M(\mathbf{r}, \mathbf{r}'; \epsilon), \quad (4.68)$$

leads, in a similar way as in the case of the polarization function P , to the symmetry relation

$$M_{\underline{\beta}_j^{-1}(\mathbf{K} + \mathbf{K}_0), \underline{\beta}_j^{-1}(\mathbf{K}' + \mathbf{K}_0)}(\mathbf{k}; \epsilon) = e^{-i \underline{\beta}_j^{-1}(\mathbf{K} - \mathbf{K}') \cdot \tilde{\boldsymbol{\tau}}_j} M_{\mathbf{K}, \mathbf{K}'}(\underline{\beta}_j \mathbf{k} + \mathbf{K}_0; \epsilon). \quad (4.69)$$

By means of this relation it is possible to obtain those M matrix elements, for which the reciprocal-lattice-vector indices are related to each other through the point-group operations of the group of the wave vector \mathbf{k} , in a very economic manner.

According to (3.25) and (3.26) we can express M^{GW} in terms of three terms, i.e.,

$$M_{\mathbf{G}, \mathbf{G}'}^{\text{GW}}(\mathbf{k}; \epsilon) = M_{\mathbf{G}, \mathbf{G}'}^{\text{Gv}}(\mathbf{k}) + M_{\mathbf{G}, \mathbf{G}'}^{\text{Int}}(\mathbf{k}; \epsilon) + M_{\mathbf{G}, \mathbf{G}'}^{\text{Res}}(\mathbf{k}; \epsilon), \quad (4.70)$$

in which M^{Gv} is a Hartree-Fock type of contribution involving the bare Coulomb interaction v , M^{Int} is an integral over imaginary energy involving the screened-interaction correction \tilde{W} and M^{Res} is the residue contribution, also involving \tilde{W} . Each term in (4.70) contains a Brillouin-zone integration $\int_{\text{Bz}} d^3k'(\dots)$, the evaluation of which requires special care because of the possible occurrence of singularities at $\mathbf{k}' = 0$.

The discussion of the typical problems encountered in the 1Bz integrations and the proposed strategy for solving them can best be given by

concentrating on the M^{Int} term in (4.70).

We write

$$M_{G,G}^{\text{Int}}(\mathbf{k};\epsilon) = \frac{-1}{(2\pi)^4\hbar} \int_0^\infty d\epsilon' \int_{1Bz} d^3k' \sum_{K,K'} \tilde{W}_{G-K,G'-K'}(\mathbf{k}';i\epsilon') \mathcal{P}_{K,K'}(\mathbf{k}-\mathbf{k}';\epsilon,i\epsilon'), \quad (4.71)$$

in which the function $\mathcal{P}_{K,K'}(\mathbf{k}-\mathbf{k}';\epsilon,i\epsilon')$ is given by

$$\mathcal{P}_{K,K'}(\mathbf{k}-\mathbf{k}';\epsilon,i\epsilon') \equiv \sum_{\ell} \tilde{\xi}_{\ell,\mathbf{k}-\mathbf{k}'} \tilde{d}_{\ell,\mathbf{k}-\mathbf{k}'}(K) \tilde{d}_{\ell,-\mathbf{k}+\mathbf{k}'}(-K') \cdot \left\{ \frac{1}{\epsilon + i\epsilon' - \tilde{E}_{\ell}(\mathbf{k}-\mathbf{k}')} + \frac{1}{\epsilon - i\epsilon' - \tilde{E}_{\ell}(\mathbf{k}-\mathbf{k}')} \right\}. \quad (4.72)$$

We will be concerned with the possibly singular behavior of the matrix elements of \tilde{W} near $\mathbf{k}'=0$ as this complicates the determination of the $1Bz$ integration over \mathbf{k}' . This behavior is directly related to the polarization matrix and the bare Coulomb-interaction matrix, through the sequence of (symbolic) relations $\tilde{W}=W-v$; $W=\epsilon^{-1}v$ and $\epsilon=1-vP$. As P is regular in the direct vicinity of $\mathbf{k}'=0$, we can generally express $P_{K,K'}(\mathbf{k}';i\epsilon)$ around $\mathbf{k}'=0$ as

$$P_{K,K'}(\mathbf{k}';i\epsilon) = A(K,K';i\epsilon) + \mathbf{k}' \cdot B(K,K';i\epsilon) + \mathbf{k}' \cdot \underline{C}(K,K';i\epsilon) \cdot \mathbf{k}' + \mathcal{O}(|\mathbf{k}'|^3). \quad (4.73)$$

For each given combination of K , K' and ϵ the scalar A , the vector B and the tensor \underline{C} are independent of \mathbf{k}' . Explicit expressions for these quantities can be derived by means of a straightforward perturbation technique relating Bloch functions $\tilde{\varphi}_{\ell,\mathbf{k}+\mathbf{k}'}(\mathbf{r})$, with $\mathbf{k}' \rightarrow 0$, to functions $\tilde{\varphi}_{\ell,\mathbf{k}}(\mathbf{r})$. As this procedure is rather

lengthy, complicated and cumbersome, we will not go through it here, but rather present the general results to the extent in which they are essentially needed in the sequel. A separate further publication on this topic including many details of the derivation is anticipated.

The first important result of the analysis is that we have to distinguish between the "head" element, i.e., the matrix element with $K=K'=0$; the "wing" elements, i.e., those with $K=0$ and $K'\neq 0$ or $K\neq 0$ and $K'=0$, and the "body" elements for which both K and $K'\neq 0$. One finds the general results

$$\text{Head element : } A(0,0;i\epsilon) = 0; B(0,0;i\epsilon) = 0, \quad (4.74)$$

$$\text{Wing elements : } A(0,K';i\epsilon) = A(K,0;i\epsilon) = 0.$$

This implies that the head and wing elements of $P_{K,K'}(\mathbf{k}';i\epsilon)$ vanish for $\mathbf{k}'\rightarrow 0$, the head element quadratically and the wing elements linearly. This typical behavior of P , combined with the singular behavior of $v_{0,0}(\mathbf{k}')$ for $\mathbf{k}'\rightarrow 0$, i.e., $v_{0,0}(\mathbf{k}') = e^2/(\epsilon_0|\mathbf{k}'|^2)$, determines the behavior of $\hat{W}_{K,K'}(\mathbf{k}';\epsilon)$ near $\mathbf{k}'\rightarrow 0$. It is straightforward, by employing an algebraic method of matrix inversion, the so-called method of "inversion by partitioning" [120,169], to arrive at the following formal expression for the elements of \hat{W} :

$$\begin{aligned} \hat{W}_{K,K'}(\mathbf{k}';i\epsilon) &= \hat{W}_{K,K'}(\mathbf{k}';i\epsilon) - \delta_{K,K'} v_{K',K'}(\mathbf{k}') \\ &+ \frac{1}{\mathbf{k}' \cdot \underline{\underline{D}}(i\epsilon) \cdot \mathbf{k}'} \left\{ \frac{\delta_{K,0} \delta_{0,K'}}{|\mathbf{k}'|^2} + \frac{\hat{\mathbf{k}}' \cdot \mathbf{F}(K,K';i\epsilon)}{|\mathbf{k}'|} (1 - \delta_{K,0} \delta_{0,K'}) \right\}, \quad (4.75) \end{aligned}$$

where $\hat{W}_{K,K'}(\mathbf{k}';i\epsilon) = \mathcal{O}(1)$ for $\mathbf{k}'\rightarrow 0$, and where explicit expressions can be given for the tensor $\underline{\underline{D}}$ and the vector \mathbf{F} in terms of the quantities A , B and $\underline{\underline{C}}$ introduced in (4.73). In fact (4.75) can be considered as the defining relation for the regular function $\hat{W}_{K,K'}(\mathbf{k}';i\epsilon)$. In (4.75) $\hat{\mathbf{k}}$ stands for $\mathbf{k}/|\mathbf{k}|$. It can be deduced from the explicit expressions, that for a cubic crystal $\underline{\underline{D}}$ is a multiple of the unit tensor of rank three. Quite generally, the tensor $\underline{\underline{D}}$ and the vector \mathbf{F}

ay be obtained in an actual calculation by evaluating the inverse of the dielectric matrix in few "suitable" \mathbf{k}' points, with $\mathbf{k}' \rightarrow 0$, so that the components of $\underline{\underline{D}}$ and \mathbf{F} follow from the solution of a linear system of equations. By "suitable" \mathbf{k}' points we mean that these points should not lead to a dependent system of equations. Note incidentally, that the direct evaluation of the inverse of the dielectric matrix in many \mathbf{k}' points may generally be avoided by employing the symmetry relation (4.36) which holds also for the inverse of the dielectric matrix.

From (4.75) we deduce that the head element of \tilde{W} has a $1/|\mathbf{k}'|^2$ singularity, the wing elements have $1/|\mathbf{k}'|$ singularities and the body elements have no singularities at $\mathbf{k}'=0$. Furthermore, the singularities, except the one in the bare Coulomb-term $\delta_{0,0}^v v_{0,0}(\mathbf{k}')$, have *anisotropic* character, i.e., their strength depends on the orientation of $\mathbf{k}' \rightarrow 0$.

We want to stress that the singularities of \tilde{W} , as expressed in (4.75), can be integrated and pose no fundamental problem but only numerical inconvenience. The first inconvenience is that we have to split off the singular parts of the integrand such that the integration of these parts can be performed without any problem. This can be done, as will be discussed below, in an elegant way in which (time-consuming) evaluations of \tilde{W} and \mathcal{H} can be avoided. The second inconvenience has to do with the ϵ' dependence of the integrand in (4.71). The integrand may have singularities, depending also on the value of ϵ , for certain values of ϵ' in the interval $[0, \Delta]$, where Δ is somewhat larger than the largest imaginary part of all quasi-particle energies taken into account. These singularities, as a function of ϵ' , may occur for energies ϵ satisfying $|\text{Re}(\epsilon) - \tilde{\mu}| > \tilde{\epsilon}_g/2$ where $\tilde{\epsilon}_g \equiv \min_{\mathbf{k}} \{\text{Re}(\tilde{E}_{\ell_c}(\mathbf{k}) - \tilde{E}_{\ell_v}(\mathbf{k}))\}$ and $\tilde{\mu} = \max_{\mathbf{k}} \{\text{Re}(\tilde{E}_{\ell_v}(\mathbf{k}))\} + \tilde{\epsilon}_g/2$. In the special case where all one-particle energies are real, e.g., in the first iteration step, these singularities only occur at $\epsilon'=0$. We have demonstrated in chapter 3 that M^{Int} will have a discontinuity in each such case which is compensated precisely, for that matter, by the discontinuity in M^{Res} .

In view of these inconveniences, the general strategy in performing the integrations in (4.71) can be as follows: Split the ϵ' integration into two parts, an integral from $\epsilon'=0$ to Δ and an integral from $\epsilon'=\Delta$ to infinity. The former is performed analytically before the B_z integration is done by, e.g., approximating

\tilde{W} by its value at $\epsilon'=0$. Since Δ is usually small and $\tilde{W}(\epsilon')$ is very smooth (note moreover that \tilde{W} is an even function of energy), this will give an excellent approximation. The subsequent Bz integration of the resulting function can be carried out in the same spirit as the general approach for the remaining ϵ' integral, which will be discussed in the sequel.

We will now consider the second part of the ϵ' integration in (4.71), i.e., the integral from $\epsilon'=\Delta$ to ∞ . In this part we perform the Bz integration first and after that the ϵ' integration. As we anticipate smooth ϵ' dependence of the integrand after Bz integration, we need to calculate the integrand at a very restricted number of ϵ' values only, in order to perform the ϵ' integral.

Let us concentrate on a Brillouin-zone integration in (4.71), i.e.,

$$I \equiv \int_{1Bz} d^3k' \tilde{W}_{G-K, G'-K'}(k'; \epsilon') \mathcal{H}_{K, K'}(k-k'; \epsilon, i\epsilon'), \quad (4.76)$$

in which $\mathcal{H}_{K, K'}(k-k'; \epsilon, i\epsilon')$ is a regular function of $k-k'$. We write $\mathcal{H}_{K, K'}(k-k'; \epsilon, i\epsilon')$ in the form

$$\mathcal{H}_{K, K'}(k-k'; \epsilon, i\epsilon') = \alpha_{K, K'} + \beta_{K, K'} \cdot k' + \mathcal{H}_{K, K'}^{\#}(k, k'; \epsilon, i\epsilon'), \quad (4.77)$$

in which $\alpha_{K, K'} \equiv \mathcal{H}_{K, K'}(k; \epsilon, i\epsilon')$ and $\beta_{K, K'} \equiv -\nabla_k \mathcal{H}_{K, K'}(k; \epsilon, i\epsilon')$. By construction, the function $|\mathcal{H}_{K, K'}^{\#}(k, k'; \epsilon, i\epsilon')|$ approaches zero for $k' \rightarrow 0$, as $|k'|^2$. The scalar $\alpha_{K, K'}$ and the three components of the vector $\beta_{K, K'}$ can be easily determined numerically by solving a 4x4 linear system of equations, obtained by first calculating $\mathcal{H}_{K, K'}(k-k_j; \epsilon, i\epsilon')$ at four vertices of a regular tetrahedron of *small* dimensions (so that the effect of higher-order terms will be negligible) centered at k , and then equating the thus-obtained values with $\alpha_{K, K'} + \beta_{K, K'} \cdot k_j'$, $j=1,2,3,4$.

We substitute (4.77) in (4.76) and obtain

$$I = \alpha_{K, K'} \int_{1Bz} d^3k' \tilde{W}_{G-K, G'-K'}(k'; \epsilon')$$

$$\begin{aligned}
& + \beta_{\mathbf{K},\mathbf{K}'} \int_{1Bz} d^3\mathbf{k}' \mathbf{k}' \tilde{W}_{\mathbf{G}-\mathbf{K},\mathbf{G}'-\mathbf{K}'}(\mathbf{k}';i\epsilon') \\
& + \int_{1Bz} d^3\mathbf{k}' \tilde{W}_{\mathbf{G}-\mathbf{K},\mathbf{G}'-\mathbf{K}'}(\mathbf{k}';i\epsilon') \mathcal{A}_{\mathbf{K},\mathbf{K}'}^{\#}(\mathbf{k},\mathbf{k}';\epsilon,i\epsilon'). \quad (4.78)
\end{aligned}$$

The last integral on the right-hand side of (4.78) poses no problem at all, since the $|\mathbf{k}'|^2$ behavior of $|\mathcal{A}_{\mathbf{K},\mathbf{K}'}^{\#}(\mathbf{k},\mathbf{k}';\epsilon,i\epsilon')|$, for $\mathbf{k}' \rightarrow 0$ makes the whole integrand regular in $1Bz$. This integral can be performed by the method of special points. The same can be said of the first integral for all body elements, and the second integral for both body and wing elements. Hence, the only integrals which require closer examination are (we present below only the integral of one type of wing elements, since the wing elements of the other type can be treated identically):

$$I_1 = \int_{1Bz} d^3\mathbf{k}' \tilde{W}_{0,0}(\mathbf{k}';i\epsilon'), \quad (4.79a)$$

$$I_2 = \int_{1Bz} d^3\mathbf{k}' \tilde{W}_{0,\mathbf{G}'-\mathbf{K}'}(\mathbf{k}';i\epsilon'), \quad (4.79b)$$

$$I_3 = \int_{1Bz} d^3\mathbf{k}' \mathbf{k}' \tilde{W}_{0,0}(\mathbf{k}';i\epsilon'). \quad (4.79c)$$

We mention that for crystals with inversion symmetry, such as silicon, the property $\tilde{W}_{0,0}(\mathbf{k}';i\epsilon') = \tilde{W}_{0,0}(-\mathbf{k}';i\epsilon')$ holds so that I_3 vanishes. After substitution of the form (4.75) for $\tilde{W}_{\mathbf{G}-\mathbf{K},\mathbf{G}'-\mathbf{K}'}(\mathbf{k}';i\epsilon')$ in (4.79a,b,c) we can write

$$\begin{aligned}
I_1 = & \frac{e^2}{\epsilon_0} \int_{1Bz} d^3\mathbf{k}' \frac{1}{|\mathbf{k}'|^2} + \int_{1Bz} d^3\mathbf{k}' \frac{1}{\mathbf{k}' \cdot \underline{\mathbf{D}}(i\epsilon') \cdot \mathbf{k}'} \frac{1}{|\mathbf{k}'|^2} \\
& + \int_{1Bz} d^3\mathbf{k}' \tilde{W}_{0,0}(\mathbf{k}';i\epsilon'), \quad (4.80a)
\end{aligned}$$

$$I_2 = \int_{\text{Bz}} d^3k' \frac{1}{\hat{\mathbf{k}}' \cdot \underline{\underline{D}}(i\epsilon') \cdot \hat{\mathbf{k}}'} \frac{\hat{\mathbf{k}}' \cdot \mathbf{F}(0, \mathbf{G}' - \mathbf{K}'; i\epsilon')}{|\mathbf{k}'|} + \int_{\text{Bz}} d^3k' \hat{W}_{0, \mathbf{G}' - \mathbf{K}'}(\mathbf{k}'; i\epsilon'), \quad (4.80b)$$

$$I_3 = \frac{e^2}{\epsilon_0} \int_{\text{Bz}} d^3k' \frac{\hat{\mathbf{k}}'}{|\mathbf{k}'|} + \int_{\text{Bz}} d^3k' \frac{\hat{\mathbf{k}}'}{\hat{\mathbf{k}}' \cdot \underline{\underline{D}}(i\epsilon') \cdot \hat{\mathbf{k}}'} \frac{1}{|\mathbf{k}'|} + \int_{\text{Bz}} d^3k' \mathbf{k}' \hat{W}_{0,0}(\mathbf{k}'; i\epsilon'). \quad (4.80c)$$

None of the integrals involving \hat{W} pose any problem and therefore they all can be evaluated with the aid of special-point methods. The first integral in (4.80b) and the first two integrals in (4.80c) are seen to vanish because they have antisymmetric integrands. Hence, the remaining integrals to be discussed are the first and second ones on the right-hand side of (4.80a). As the first of these integrals is a special case of the second one, we will only consider the second integral in (4.80a) and write it into the form

$$I_4 = \int_{\text{Bz}} d^3k' \frac{1}{\hat{\mathbf{k}}' \cdot \underline{\underline{D}}(i\epsilon') \cdot \hat{\mathbf{k}}'} \frac{1}{|\mathbf{k}'|^2} = \int_0^\pi d\Theta \sin(\Theta) \int_0^{2\pi} d\varphi \kappa(\Theta, \varphi) / \sigma(\Theta, \varphi). \quad (4.81)$$

Here $\kappa(\Theta, \varphi)$ is the length of the \mathbf{k}' vector located on Bz surface in the polar directions Θ , φ and $\sigma(\Theta, \varphi) \equiv \hat{\mathbf{k}}' \cdot \underline{\underline{D}}(i\epsilon') \cdot \hat{\mathbf{k}}'$. In the specific case of a cubic crystal the tensor $\underline{\underline{D}}(i\epsilon')$ is a multiple of the unit tensor of rank three, so that the integrand in I_4 is simply proportional to $1/|\mathbf{k}'|^2$. The integral (4.81) involves a finite integration region as well as well-behaving integrand, free from singularities. Hence, the numerical execution of this integral should not pose any problem at all. In our final expression for $M_{\mathbf{G}, \mathbf{G}'}^{\text{GW}}(\mathbf{k}; \epsilon)$, to be presented in

chapter 5, some of the contributing terms will still contain Bz integrals over singular functions. However, in all these cases the integrands, though singular, are antisymmetric in the integration variable k' , such that these singular parts in fact do not contribute. The use of, e.g., a special-point integration technique in these cases is perfectly suited to deal with such integrals, as it leads to zero contribution to them. This concludes our brief and global discussion of the typical problems in the Bz integration present in (4.71). In chapter 5 we will discuss the feasibility of calculation of $M_{\mathbf{G}, \mathbf{G}'}^{\text{GW}}(\mathbf{k}; \epsilon)$, starting from expression (3.25) and (3.26), in which, however, we have carried through a number of rearrangements in order to cope, numerically, with the above-mentioned singularities.

CHAPTER 5

CALCULATIONAL PROSPECTS

5.0 Introduction

In this chapter we discuss the question whether a GW-bubble calculation of the energy bands in semiconductors is indeed achievable *without any further approximation*. The main obstacle in such a calculation is the determination of the self-energy function, which, strictly spoken, has to be performed in a self-consistent way. However, even the first iteration cycle in such a procedure is already an enormous task. We claim to be close to the achievement of calculating M^{GW} and the energy band structure *in the first iteration step*. Subsequent iteration steps do not pose principally new problems, but will lead to an enormous increase of computation time. Fortunately, however, calculations based on a plasmon-pole model [64,65] strongly suggest that a first iteration cycle is sufficient. Though this phenomenon is not yet fully understood, we consider it as realistic to anticipate its correctness. In view of this we will focus in what follows on the feasibility of an actual calculation of the GW self-energy matrix elements in the first iteration step, starting from our final expressions for $M_{\mathbf{G}, \mathbf{G}'}^{\text{GW}}(\mathbf{k}; \epsilon)$ to be given below. The particular form in which the matrix elements will be written makes it indeed possible to judge the feasibility of their numerical evaluation. However, there is also a price paid for that: We had to decompose the self-energy expression into no less than 32 different contributions, many of which need special attention and care.

Thus, the organization of matters in the present chapter has an advantage, but also a disadvantage. To start with the latter, the lengthy formulas of section 5.1 will certainly cause a lot of discomfort to those readers who are not interested in all the little details. We encourage these readers to jump directly to section 5.2 for the discussion of computational aspects. On the other hand, the advantage is that complete expressions for M^{GW} matrix elements are now for the first time available, in which all singularities, peculiarities and, from the computational point of view, less attractive

properties are unraveled and fully exposed. This expression will make the discussion of feasibility as realistic as possible and, hopefully, it will serve as a convenient starting point for further attempts towards *ab initio* band-structure calculations.

5.1 Basic Expressions

Starting from expressions (3.25) and (3.26) for $M_{\mathbf{G}, \mathbf{G}'}^{\text{GW}}(\mathbf{k}; \epsilon)$ and taking account of all intricacies such as the singular behavior of $v_{0,0}(\mathbf{k}')$ for $\mathbf{k}' \rightarrow 0$, the singular behavior of the head and wing elements of $\tilde{W}_{\mathbf{K}, \mathbf{K}'}(\mathbf{k}; \epsilon)$ for $\mathbf{k}' \rightarrow 0$ and the singularity in the imaginary energy range $(0, i\Delta)$ of the energy-dependent integrand in the expression for $M_{\mathbf{G}, \mathbf{G}'}^{\text{Int}}(\mathbf{k}; \epsilon)$, we finally come to an expression for $M_{\mathbf{G}, \mathbf{G}'}^{\text{GW}}(\mathbf{k}; \epsilon)$ in which matters have been organized in such a way that all occurring Bz integrations and summations over reciprocal lattice vectors and energy bands can be done numerically. We write M^{Gv} , M^{Int} and M^{Res} (see (4.70)) as follows:

$$M_{\mathbf{G}, \mathbf{G}'}^{\text{Gv}}(\mathbf{k}) = \frac{-1}{(2\pi)^3 \hbar} \sum_{n=1}^4 \mathcal{K}_{\mathbf{G}, \mathbf{G}'}^{(n)}(\mathbf{k}), \quad (5.1)$$

$$M_{\mathbf{G}, \mathbf{G}'}^{\text{Int}}(\mathbf{k}; \epsilon) = \frac{-1}{(2\pi)^4 \hbar} \sum_{n=1}^6 \left\{ \mathcal{K}_{\mathbf{G}, \mathbf{G}'}^{(n)}(\mathbf{k}; \epsilon, \Delta) + \int_{\Delta}^{\infty} d\epsilon' \mathcal{K}_{\mathbf{G}, \mathbf{G}'}^{(n)}(\mathbf{k}; \epsilon, \epsilon') \right\}, \quad (5.2)$$

$$M_{\mathbf{G}, \mathbf{G}'}^{\text{Res}}(\mathbf{k}; \epsilon) = \frac{-1}{(2\pi)^3 \hbar} \sum_{\ell} \sum_{n=1}^6 \mathcal{K}_{\mathbf{G}, \mathbf{G}'}^{\ell(n)}(\mathbf{k}; \epsilon), \quad (5.3)$$

where all contributing terms on the right-hand side can be calculated with the aid of simple algorithms. In below we indicate 1Bz integrals with *regular* integrands by boldface integral signs. Also, 1Bz integrals in which integrands occur with singular *parts* (which parts, in view of our analysis given in section 4.6 can be shown not to contribute) are indicated with boldface integral signs.

The reason why we mark these integrals in a special way is that they can all be performed in a relatively simple and inexpensive way, for instance by using a special-point method [112-114]. The remaining 1Bz integrals do have singularities, but can, in view of our considerations in section 4.6, be done as well in a relatively simple way. We will first give expressions for the functions $\mathcal{M}_{\mathbf{G},\mathbf{G}'}^{(n)}(\mathbf{k})$, $\mathcal{M}'_{\mathbf{G},\mathbf{G}'}^{(n)}(\mathbf{k};\epsilon,\Delta)$, $\mathcal{M}''_{\mathbf{G},\mathbf{G}'}^{(n)}(\mathbf{k};\epsilon,\epsilon')$ and $\mathcal{M}^{\Delta}_{\mathbf{G},\mathbf{G}'}^{(n)}(\mathbf{k};\epsilon)$, whereafter the respective symbols and functions occurring in them will be defined:

$$\mathcal{M}_{\mathbf{G},\mathbf{G}'}^{(1)}(\mathbf{k}) = \sum_{\mathbf{K} \neq 0} \int_{1\text{Bz}} d^3\mathbf{k}' f_{\mathbf{G}-\mathbf{K},\mathbf{G}'-\mathbf{K}}(\mathbf{k}') v_{\mathbf{K},\mathbf{K}}(\mathbf{k}-\mathbf{k}'), \quad (5.4)$$

$$\mathcal{M}_{\mathbf{G},\mathbf{G}'}^{(2)}(\mathbf{k}) = \int_{1\text{Bz}} d^3\mathbf{k}' \tilde{f}_{\mathbf{G},\mathbf{G}'}(\mathbf{k}',\mathbf{k}) v_{0,0}(\mathbf{k}-\mathbf{k}'), \quad (5.5)$$

$$\mathcal{M}_{\mathbf{G},\mathbf{G}'}^{(3)}(\mathbf{k}) = f_{\mathbf{G},\mathbf{G}'}(\mathbf{k}) \int_{1\text{Bz}} d^3\mathbf{k}' v_{0,0}(\mathbf{k}-\mathbf{k}'), \quad (5.6)$$

$$\mathcal{M}_{\mathbf{G},\mathbf{G}'}^{(4)}(\mathbf{k}) = [\nabla_{\mathbf{k}} f_{\mathbf{G},\mathbf{G}'}(\mathbf{k})] \cdot \int_{1\text{Bz}} d^3\mathbf{k}' (\mathbf{k}-\mathbf{k}') v_{0,0}(\mathbf{k}-\mathbf{k}'), \quad (5.7)$$

$$\mathcal{M}'_{\mathbf{G},\mathbf{G}'}^{(1)}(\mathbf{k};\epsilon,\Delta) \simeq \sum_{\mathbf{K},\mathbf{K}'} \int_{1\text{Bz}} d^3\mathbf{k}' \tilde{W}_{\mathbf{K},\mathbf{K}'}(\mathbf{k}';0) \mathcal{K}_{\mathbf{G}-\mathbf{K},\mathbf{G}'-\mathbf{K}'}^{\Delta}(\mathbf{k}-\mathbf{k}';\epsilon), \quad (5.8)$$

$$\mathcal{M}'_{\mathbf{G},\mathbf{G}'}^{(2)}(\mathbf{k};\epsilon,\Delta) \simeq \int_{1\text{Bz}} d^3\mathbf{k}' \hat{W}_{0,0}^{(0)}(\mathbf{k}';0) \mathcal{K}_{\mathbf{G},\mathbf{G}'}^{\Delta}(\mathbf{k}-\mathbf{k}';\epsilon), \quad (5.9)$$

$$\mathcal{M}'_{\mathbf{G},\mathbf{G}'}^{(3)}(\mathbf{k};\epsilon,\Delta) \simeq \int_{1\text{Bz}} d^3\mathbf{k}' \Gamma_{0,0}^{(0)}(\mathbf{k}';0) \mathcal{K}_{\mathbf{G},\mathbf{G}'}^{\Delta}(\mathbf{k},\mathbf{k}';\epsilon), \quad (5.10)$$

$$\mathcal{M}'_{\mathbf{G},\mathbf{G}'}^{(4)}(\mathbf{k};\epsilon,\Delta) \simeq - \int_{1\text{Bz}} d^3\mathbf{k}' \mathcal{K}_{\mathbf{G},\mathbf{G}'}^{\Delta}(\mathbf{k},\mathbf{k}';\epsilon) v_{0,0}(\mathbf{k}'), \quad (5.11)$$

$$\mathcal{M}'_{\mathbf{G},\mathbf{G}'}^{(5)}(\mathbf{k};\epsilon,\Delta) \simeq \mathcal{K}_{\mathbf{G},\mathbf{G}'}^{\Delta}(\mathbf{k};\epsilon) \int_{1\text{Bz}} d^3\mathbf{k}' \Gamma_{0,0}^{(0)}(\mathbf{k}';0), \quad (5.12)$$

$$\mathcal{M}'_{\mathbf{G},\mathbf{G}}(6)(\mathbf{k};\epsilon,\Delta) \simeq -\mathcal{H}_{\mathbf{G},\mathbf{G}}^{\Delta}(\mathbf{k};\epsilon) \int_{1\text{Bz}} d^3\mathbf{k}' v_{0,0}(\mathbf{k}'), \quad (5.13)$$

$$\mathcal{M}_{\mathbf{G},\mathbf{G}}^{n(1)}(\mathbf{k};\epsilon,\epsilon') = \sum_{\mathbf{K},\mathbf{K}'} \int_{1\text{Bz}} d^3\mathbf{k}' \hat{W}_{\mathbf{K},\mathbf{K}'}(\mathbf{k}';i\epsilon') \mathcal{H}_{\mathbf{G}-\mathbf{K},\mathbf{G}'-\mathbf{K}'}(\mathbf{k}-\mathbf{k}';\epsilon,i\epsilon'), \quad (5.14)$$

$$\mathcal{M}_{\mathbf{G},\mathbf{G}}^{n(2)}(\mathbf{k};\epsilon,\epsilon') = \int_{1\text{Bz}} d^3\mathbf{k}' \hat{W}_{0,0}^{(0)}(\mathbf{k}';i\epsilon') \mathcal{H}_{\mathbf{G},\mathbf{G}'}(\mathbf{k}-\mathbf{k}';\epsilon,i\epsilon'), \quad (5.15)$$

$$\mathcal{M}_{\mathbf{G},\mathbf{G}}^{n(3)}(\mathbf{k};\epsilon,\epsilon') = \int_{1\text{Bz}} d^3\mathbf{k}' \Upsilon_{0,0}^{(0)}(\mathbf{k}';i\epsilon') \mathcal{H}_{\mathbf{G},\mathbf{G}'}(\mathbf{k},\mathbf{k}';\epsilon,i\epsilon'), \quad (5.16)$$

$$\mathcal{M}_{\mathbf{G},\mathbf{G}}^{n(4)}(\mathbf{k};\epsilon,\epsilon') = - \int_{1\text{Bz}} d^3\mathbf{k}' \mathcal{H}_{\mathbf{G},\mathbf{G}'}(\mathbf{k},\mathbf{k}';\epsilon,i\epsilon') v_{0,0}(\mathbf{k}'), \quad (5.17)$$

$$\mathcal{M}_{\mathbf{G},\mathbf{G}}^{n(5)}(\mathbf{k};\epsilon,\epsilon') = \mathcal{H}_{\mathbf{G},\mathbf{G}'}(\mathbf{k};\epsilon,i\epsilon') \int_{1\text{Bz}} d^3\mathbf{k}' \Upsilon_{0,0}^{(0)}(\mathbf{k}';i\epsilon'), \quad (5.18)$$

$$\mathcal{M}_{\mathbf{G},\mathbf{G}}^{n(6)}(\mathbf{k};\epsilon,\epsilon') = - \mathcal{H}_{\mathbf{G},\mathbf{G}'}(\mathbf{k};\epsilon,i\epsilon') \int_{1\text{Bz}} d^3\mathbf{k}' v_{0,0}(\mathbf{k}'), \quad (5.19)$$

$$\begin{aligned} \mathcal{M}_{\mathbf{G},\mathbf{G}}^{\ell(1)}(\mathbf{k};\epsilon) &= \sum_{\mathbf{K},\mathbf{K}'} \int_{1\text{Bz}} d^3\mathbf{k}' \{ \mathcal{S}^{\ell}(\mathbf{k}-\mathbf{k}';\mu,\epsilon) h_{\mathbf{G}-\mathbf{K},\mathbf{G}'-\mathbf{K}'}^{\ell}(\mathbf{k}-\mathbf{k}') \\ &\quad \times \hat{W}_{\mathbf{K},\mathbf{K}'}(\mathbf{k}';\epsilon-\hat{\mathbf{E}}_{\ell}(\mathbf{k}-\mathbf{k}')) \}, \end{aligned} \quad (5.20)$$

$$\begin{aligned} \mathcal{M}_{\mathbf{G},\mathbf{G}}^{\ell(2)}(\mathbf{k};\epsilon) &= \sum_{\mathbf{K} \neq 0} \int_{1\text{Bz}} d^3\mathbf{k}' \{ \mathcal{S}^{\ell}(\mathbf{k}-\mathbf{k}';\mu,\epsilon) h_{\mathbf{G}-\mathbf{K},\mathbf{G}'}^{\ell}(\mathbf{k}-\mathbf{k}') \\ &\quad \times \hat{W}_{\mathbf{K},0}^{(1)}(\mathbf{k}';\epsilon-\hat{\mathbf{E}}_{\ell}(\mathbf{k}-\mathbf{k}')) \}, \end{aligned} \quad (5.21)$$

$$\begin{aligned} \mathcal{M}_{\mathbf{G},\mathbf{G}}^{\ell(3)}(\mathbf{k};\epsilon) &= \sum_{\mathbf{K}' \neq \mathbf{0}} \int_{\text{IBz}} d^3\mathbf{k}' \{ \mathcal{S}^{\ell}(\mathbf{k}-\mathbf{k}';\mu,\epsilon) \tilde{h}_{\mathbf{G},\mathbf{G}'-\mathbf{K}'}^{\ell}(\mathbf{k}-\mathbf{k}') \\ &\quad \cdot \hat{W}_{\mathbf{0},\mathbf{K}'}^{(1)}(\mathbf{k}';\epsilon-\tilde{\mathbb{E}}_{\Lambda}(\mathbf{k}-\mathbf{k}')) \}, \end{aligned} \quad (5.22)$$

$$\begin{aligned} \mathcal{M}_{\mathbf{G},\mathbf{G}}^{\ell(4)}(\mathbf{k};\epsilon) &= \sum_{\mathbf{K}' \neq \mathbf{0}} \int_{\text{IBz}} d^3\mathbf{k}' \{ \mathcal{S}^{\ell}(\mathbf{k}-\mathbf{k}';\mu,\epsilon) \tilde{h}_{\mathbf{G}-\mathbf{K},\mathbf{G}'}^{\ell}(\mathbf{k},\mathbf{k}') \\ &\quad \cdot \Upsilon_{\mathbf{K},\mathbf{0}}^{(1)}(\mathbf{k}';\epsilon-\tilde{\mathbb{E}}_{\Lambda}(\mathbf{k}-\mathbf{k}')) \}, \end{aligned} \quad (5.23)$$

$$\begin{aligned} \mathcal{M}_{\mathbf{G},\mathbf{G}}^{\ell(5)}(\mathbf{k};\epsilon) &= \sum_{\mathbf{K}' \neq \mathbf{0}} \int_{\text{IBz}} d^3\mathbf{k}' \{ \mathcal{S}^{\ell}(\mathbf{k}-\mathbf{k}';\mu,\epsilon) \tilde{h}_{\mathbf{G},\mathbf{G}'-\mathbf{K}'}^{\ell}(\mathbf{k},\mathbf{k}') \\ &\quad \cdot \Upsilon_{\mathbf{0},\mathbf{K}'}^{(1)}(\mathbf{k}';\epsilon-\tilde{\mathbb{E}}_{\Lambda}(\mathbf{k}-\mathbf{k}')) \}, \end{aligned} \quad (5.24)$$

$$\mathcal{M}_{\mathbf{G},\mathbf{G}}^{\ell(6)}(\mathbf{k};\epsilon)$$

$$= - \sum_{\mathbf{K}' \neq \mathbf{0}} [\nabla_{\mathbf{k}} \tilde{h}_{\mathbf{G}-\mathbf{K},\mathbf{G}'}^{\ell}(\mathbf{k})] \cdot \int_{\text{IBz}} d^3\mathbf{k}' \mathbf{k}' \mathcal{S}^{\ell}(\mathbf{k}-\mathbf{k}';\mu,\epsilon) \Upsilon_{\mathbf{K},\mathbf{0}}^{(1)}(\mathbf{k}';\epsilon-\tilde{\mathbb{E}}_{\Lambda}(\mathbf{k}-\mathbf{k}')), \quad (5.25)$$

$$\mathcal{M}_{\mathbf{G},\mathbf{G}}^{\ell(7)}(\mathbf{k};\epsilon)$$

$$= - \sum_{\mathbf{K}' \neq \mathbf{0}} [\nabla_{\mathbf{k}} \tilde{h}_{\mathbf{G},\mathbf{G}'-\mathbf{K}'}^{\ell}(\mathbf{k})] \cdot \int_{\text{IBz}} d^3\mathbf{k}' \mathbf{k}' \mathcal{S}^{\ell}(\mathbf{k}-\mathbf{k}';\mu,\epsilon) \Upsilon_{\mathbf{0},\mathbf{K}'}^{(1)}(\mathbf{k}';\epsilon-\tilde{\mathbb{E}}_{\Lambda}(\mathbf{k}-\mathbf{k}')), \quad (5.26)$$

$$\mathcal{M}_{\mathbf{G},\mathbf{G}}^{\ell(8)}(\mathbf{k};\epsilon) = - \int_{\text{IBz}} d^3\mathbf{k}' \mathcal{S}^{\ell}(\mathbf{k}-\mathbf{k}';\mu,\epsilon) \tilde{h}_{\mathbf{G},\mathbf{G}'}^{\ell}(\mathbf{k},\mathbf{k}') v_{\mathbf{0},\mathbf{0}}(\mathbf{k}'), \quad (5.27)$$

$$\mathcal{M}_{\mathbf{G},\mathbf{G}}^{\ell(9)}(\mathbf{k};\epsilon) = \int_{\text{IBz}} d^3\mathbf{k}' \mathcal{S}^{\ell}(\mathbf{k}-\mathbf{k}';\mu,\epsilon) \tilde{h}_{\mathbf{G},\mathbf{G}'}^{\ell}(\mathbf{k}-\mathbf{k}') \hat{W}_{\mathbf{0},\mathbf{0}}^{(1)}(\mathbf{k}';\epsilon-\tilde{\mathbb{E}}_{\Lambda}(\mathbf{k}-\mathbf{k}')), \quad (5.28)$$

$$\mathcal{M}_{G,G'}^{\ell(10)}(\mathbf{k}; \epsilon) = \int_{1Bz} d^3k' \mathcal{S}^{\ell}(\mathbf{k}-\mathbf{k}'; \mu, \epsilon) h_{G,G'}^{\ell}(\mathbf{k}, \mathbf{k}') \Upsilon_{0,0}^{(1)}(\mathbf{k}'; \epsilon - \hat{\mathbf{E}}_{\ell}(\mathbf{k}-\mathbf{k}')), \quad (5.29)$$

$$\mathcal{M}_{G,G'}^{\ell(11)}(\mathbf{k}; \epsilon) = h_{G,G'}^{\ell}(\mathbf{k}) \int_{1Bz} d^3k' \mathcal{S}^{\ell}(\mathbf{k}-\mathbf{k}'; \mu, \epsilon) \Upsilon_{0,0}^{(1)}(\mathbf{k}'; \epsilon - \hat{\mathbf{E}}_{\ell}(\mathbf{k}-\mathbf{k}')), \quad (5.30)$$

$$\mathcal{M}_{G,G'}^{\ell(12)}(\mathbf{k}; \epsilon) = -[\nabla_{\mathbf{k}} h_{G,G'}^{\ell}(\mathbf{k})] \cdot \int_{1Bz} d^3k' \mathbf{k}' \mathcal{S}^{\ell}(\mathbf{k}-\mathbf{k}'; \mu, \epsilon) \Upsilon_{0,0}^{(1)}(\mathbf{k}'; \epsilon - \hat{\mathbf{E}}_{\ell}(\mathbf{k}-\mathbf{k}')), \quad (5.31)$$

$$\mathcal{M}_{G,G'}^{\ell(13)}(\mathbf{k}; \epsilon) = \sum_{\mathbf{K}' \neq 0} h_{G-\mathbf{K},G'}^{\ell}(\mathbf{k}) \int_{1Bz} d^3k' \mathcal{S}^{\ell}(\mathbf{k}-\mathbf{k}'; \mu, \epsilon) \Upsilon_{\mathbf{K},0}^{(1)}(\mathbf{k}'; \epsilon - \hat{\mathbf{E}}_{\ell}(\mathbf{k}-\mathbf{k}')), \quad (5.32)$$

$$\mathcal{M}_{G,G'}^{\ell(14)}(\mathbf{k}; \epsilon) = \sum_{\mathbf{K}' \neq 0} h_{G,G'-\mathbf{K}'}^{\ell}(\mathbf{k}) \int_{1Bz} d^3k' \mathcal{S}^{\ell}(\mathbf{k}-\mathbf{k}'; \mu, \epsilon) \Upsilon_{0,\mathbf{K}'}^{(1)}(\mathbf{k}'; \epsilon - \hat{\mathbf{E}}_{\ell}(\mathbf{k}-\mathbf{k}')), \quad (5.33)$$

$$\mathcal{M}_{G,G'}^{\ell(15)}(\mathbf{k}; \epsilon) = -h_{G,G'}^{\ell}(\mathbf{k}) \int_{1Bz} d^3k' \mathcal{S}^{\ell}(\mathbf{k}-\mathbf{k}'; \mu, \epsilon) v_{0,0}(\mathbf{k}'), \quad (5.34)$$

$$\mathcal{M}_{G,G'}^{\ell(16)}(\mathbf{k}; \epsilon) = [\nabla_{\mathbf{k}} h_{G,G'}^{\ell}(\mathbf{k})] \cdot \int_{1Bz} d^3k' \mathcal{S}^{\ell}(\mathbf{k}-\mathbf{k}'; \mu, \epsilon) \mathbf{k}' v_{0,0}(\mathbf{k}'). \quad (5.35)$$

In the above expressions \sum' denotes summation over all \mathbf{K}, \mathbf{K}' except $\mathbf{K}=0$ and $\mathbf{K}'=0$, while \sum'' stands for summation over all \mathbf{K} and \mathbf{K}' except $\mathbf{K}=0$ or $\mathbf{K}'=0$. The functions occurring in the 32 expressions (5.4)-(5.35) are given by

$$v_{\mathbf{K},\mathbf{K}}(\mathbf{k}) = \frac{e^2}{\epsilon_0 |\mathbf{k}+\mathbf{K}|^2}, \quad (5.36)$$

$$f_{G-\mathbf{K},G'-\mathbf{K}'}^{\ell}(\mathbf{k}') \equiv \sum_{\ell} \Theta(\mu - \text{Re} \hat{\mathbf{E}}_{\ell}(\mathbf{k}')) h_{G-\mathbf{K},G'-\mathbf{K}'}^{\ell}(\mathbf{k}'), \quad (5.37)$$

$$h_{\mathbf{K},\mathbf{K}'}^{\ell}(k') = \tilde{g}_{\ell,\mathbf{k}-\mathbf{k}'} \hat{d}_{\ell,\mathbf{k}}(\mathbf{K}) \hat{d}_{\ell,-\mathbf{k}}(-\mathbf{K}'), \quad (5.38)$$

$$f_{\mathbf{G},\mathbf{G}'}^{\ell}(k',k) = f_{\mathbf{G},\mathbf{G}'}(k') - f_{\mathbf{G},\mathbf{G}'}(k) - [\nabla_{\mathbf{k}} f_{\mathbf{G},\mathbf{G}'}(k)] \cdot (k'-k), \quad (5.39)$$

$$\begin{aligned} \mathcal{H}_{\mathbf{G}-\mathbf{K},\mathbf{G}'-\mathbf{K}'}^{\Delta}(k-k';\epsilon) &= \int_0^{\Delta} d\epsilon' \mathcal{H}_{\mathbf{G}-\mathbf{K},\mathbf{G}'-\mathbf{K}'}(k-k';\epsilon,i\epsilon') \\ &\equiv 2 \sum_{\ell} h_{\mathbf{G}-\mathbf{K},\mathbf{G}'-\mathbf{K}'}^{\ell}(k-k') \arctan\left\{ \frac{\Delta}{\epsilon - \hat{E}_{\ell}(k-k')} \right\}, \end{aligned} \quad (5.40)$$

$$\begin{aligned} \mathcal{H}_{\mathbf{G}-\mathbf{K},\mathbf{G}'-\mathbf{K}'}(k-k';\epsilon,i\epsilon') &\equiv \sum_{\ell} h_{\mathbf{G}-\mathbf{K},\mathbf{G}'-\mathbf{K}'}^{\ell}(k-k') \\ &\times \left\{ \frac{1}{\epsilon + i\epsilon' - \hat{E}_{\ell}(k-k')} + \frac{1}{\epsilon - i\epsilon' - \hat{E}_{\ell}(k-k')} \right\}, \end{aligned} \quad (5.41)$$

$$\arctan(z) = \frac{1}{2} \arctan\left(\frac{2x}{1-x^2-y^2}\right) + \frac{i}{4} \ln\left(\frac{x^2+(y+1)^2}{x^2+(y-1)^2}\right),$$

$$\text{with } z=x+iy, x,y \in \mathbb{R}, z^2 \neq -1 \text{ and } -\pi/2 < \arctan\left(\frac{2x}{1-x^2-y^2}\right) \leq \pi/2, \quad (5.42)$$

$$\mathcal{H}_{\mathbf{G},\mathbf{G}'}^{\Delta}(k,k';\epsilon) = \mathcal{H}_{\mathbf{G},\mathbf{G}'}^{\Delta}(k-k';\epsilon) - \mathcal{H}_{\mathbf{G},\mathbf{G}'}^{\Delta}(k;\epsilon) + [\nabla_{\mathbf{k}} \mathcal{H}_{\mathbf{G},\mathbf{G}'}^{\Delta}(k;\epsilon)] \cdot \mathbf{k}', \quad (5.43)$$

$$\Upsilon_{\mathbf{K},\mathbf{K}'}^{(\alpha)}(k';i\epsilon') = \frac{1}{\mathbf{k}' \cdot \underline{\mathbf{D}}(i\epsilon') \cdot \mathbf{k}'}$$

$$\times \left\{ \frac{\delta_{\mathbf{K},0} \delta_{0,\mathbf{K}'}}{|\mathbf{k}'|^2} + \alpha \frac{\hat{\mathbf{k}}' \cdot \mathbf{F}(\mathbf{K},\mathbf{K}';i\epsilon')}{|\mathbf{k}'|} (1 - \delta_{\mathbf{K},0} \delta_{0,\mathbf{K}'} \right\}, \quad (5.44)$$

$$\hat{W}_{\mathbf{K},\mathbf{K}'}^{(\alpha)}(k';i\epsilon') = \hat{W}_{\mathbf{K},\mathbf{K}'}(k';i\epsilon') + \delta_{\mathbf{K},\mathbf{K}'} v_{\mathbf{K}',\mathbf{K}'}(k') - \Upsilon_{\mathbf{K},\mathbf{K}'}^{(\alpha)}(k';i\epsilon'), \quad (5.45)$$

$$\begin{aligned} \mathcal{H}_{\mathbf{K},\mathbf{K}'}^{\ell}(\mathbf{k},\mathbf{k}';\epsilon,i\epsilon') &= \mathcal{H}_{\mathbf{K},\mathbf{K}'}^{\ell}(\mathbf{k}-\mathbf{k}';\epsilon,i\epsilon') - \mathcal{H}_{\mathbf{K},\mathbf{K}'}^{\ell}(\mathbf{k};\epsilon,i\epsilon') \\ &+ [\nabla_{\mathbf{k}} \mathcal{H}_{\mathbf{K},\mathbf{K}'}^{\ell}(\mathbf{k};\epsilon,i\epsilon')] \cdot \mathbf{k}', \end{aligned} \quad (5.46)$$

$$\mathcal{J}^{\ell}(\mathbf{k}-\mathbf{k}';\mu,\epsilon) = \Theta(\mu - \text{Re} \mathcal{E}_{\mathbf{K}}^{\ell}(\mathbf{k}-\mathbf{k}')) - \Theta(\text{Re}(\epsilon - \mathcal{E}_{\mathbf{K}}^{\ell}(\mathbf{k}-\mathbf{k}'))), \quad (5.47)$$

$$\begin{aligned} \tilde{h}_{\mathbf{K},\mathbf{K}'}^{\ell}(\mathbf{k},\mathbf{k}') &= h_{\mathbf{K},\mathbf{K}'}^{\ell}(\mathbf{k}-\mathbf{k}') - h_{\mathbf{K},\mathbf{K}'}^{\ell}(\mathbf{k}) + [\nabla_{\mathbf{k}} h_{\mathbf{K},\mathbf{K}'}^{\ell}(\mathbf{k})] \cdot \mathbf{k}'. \end{aligned} \quad (5.48)$$

This concludes our complete listing of all expressions for the 32 contributions to M^{GW} . In the next section we will discuss matters concerning the required computation time and storage capacity.

5.2 Computational Aspects

All expressions (5.4)-(5.35) contain a \mathbf{k}' integration over 1Bz. Apart from the bold-faced integrals, which can best be performed by means of a special-point method [112-114], there are also 1Bz integrals with singular integrands. In these latter cases we propose, in accordance with our analysis in section 4.6, to exclude a small sphere or (regular) polyhedron around the singularity point and integrate analytically in this region. The integration over the remaining part of 1Bz can then be carried out by means of a simple integration method. Namely, in all these cases matters have been organized in such a way that the evaluation of the integrands is relatively inexpensive. An exception is the case of the singular-type integrals (5.30)-(5.35) in which the evaluation of the integrands involves the determination of band energies at many points in 1Bz. This happens to be rather expensive, but can be avoided by using a band-structure-interpolation procedure such as that of Slater and Koster [170,171], or a procedure based on the expansion of energy bands in terms of symmetrized plane waves [114,172]. In the latter case the expansion coefficients may be obtained either by fitting to calculated band energies in a restricted number of points in 1Bz or by utilizing the orthonormality property of the symmetrized plane waves so that each

expansion coefficient may be written as a Bz integral of the product of the energy band with the corresponding symmetrized plane wave. Such integrals can be evaluated by means of a special-point method. We mention, however, that although the latter method is advocated by Monkhorst and Pack [114], Schuurmans *et al* [172] are less positive as to the usefulness of symmetrized plane waves for expanding energy bands in general.

Quite generally, the evaluation of (5.4)-(5.35) anyhow requires the determination of wave-function coefficients $\tilde{d}_{L, \mathbf{k}'}(\mathbf{K})$ and energy eigenvalues $\tilde{E}_L(\mathbf{k}'')$ in quite a number of \mathbf{k}'' points in 1Bz. Furthermore, a number of summations over \mathbf{K} and \mathbf{K}' has to be performed. The first requirement, however, is the evaluation of the integrands (summands) in the respective expressions (5.4)-(5.35). Some of these integrands contain functions which do not depend on the variables \mathbf{k} and ϵ occurring in $M_{\mathbf{G}, \mathbf{G}'}^{\text{GW}}(\mathbf{k}; \epsilon)$. As an example, consider the function $\tilde{W}_{\mathbf{K}, \mathbf{K}'}(\mathbf{k}'; i\epsilon')$ in (5.14), which can be determined once and for all in a number of (special) \mathbf{k}' points and a number of $i\epsilon'$ points on the imaginary energy axis and subsequently stored. This will result in a substantial reduction of the computation time when calculating quasi-particle band structures as this implies the evaluation of the self-energy function $M_{\mathbf{G}, \mathbf{G}'}^{\text{GW}}(\mathbf{k}; \epsilon)$ in various \mathbf{k} and ϵ points. As will be shown later on, the determination of \tilde{W} is rather costly and therefore a repeated evaluation of this function should be avoided. In this respect the function \tilde{W} occurring in (5.20) deserves special attention. It depends on \mathbf{k}' , but also on \mathbf{k} and ϵ via the *energy* argument $\epsilon - \tilde{E}_L(\mathbf{k} - \mathbf{k}')$. We propose in this case to calculate $\tilde{W}_{\mathbf{K}, \mathbf{K}'}(\mathbf{k}'; E_i)$ values at a limited number of fixed mesh points $\{E_i\}$ and to use interpolated \tilde{W} quantities between these E_i 's in the actual calculation of (5.20).

Examination of the integrands (summands) in (5.4)-(5.35) reveals that quite a number of functions occurring in the integrands *do* depend on \mathbf{k} and ϵ . Varying these quantities generally requires repeated evaluations of these functions. Taking account of all the above considerations it is now in principle possible to estimate the feasibility of actually calculating M^{GW} in terms of the computation time required for the first iteration cycle.

Let us assume that all required $\tilde{d}_{\underline{\ell}, \underline{\alpha}_j \mathbf{k}_s}(\mathbf{K})$, $\tilde{d}_{\underline{\ell}, \mathbf{k} - \underline{\alpha}_j \mathbf{k}_s}(\mathbf{K})$, $\tilde{E}_{\underline{\ell}}(\underline{\alpha}_j \mathbf{k}_s)$ and $\tilde{E}_{\underline{\ell}}(\mathbf{k} - \underline{\alpha}_j \mathbf{k}_s)$ values have been evaluated and stored, where \mathbf{k}_s , $s=1,2, \dots, N_{sp}$, are special points while $\underline{\alpha}_j$, $j=1,2, \dots, N_\alpha$, are operations of the underlying lattice point group. In general, if \mathbf{k} is a non-symmetric point, the number of $\mathbf{k} - \underline{\alpha}_j \mathbf{k}_s$ points will grow rapidly at increasing number of special points \mathbf{k}_s . For instance, if $\underline{\alpha}_j$ denotes the elements of the lattice point group of a cubic lattice, and the number of Monkhorst-Pack (MP) special points is equal to 2, we have 32 points $\mathbf{k} - \underline{\alpha}_j \mathbf{k}_s$. At 10 MP special points this number is even 256, etc. Furthermore, it will be assumed that functions not depending on \mathbf{k} and ϵ (we consider $\tilde{W}_{\mathbf{K}, \mathbf{K}'}(\underline{\alpha}_j \mathbf{k}_s; \epsilon - \tilde{E}_{\underline{\ell}}(\mathbf{k} - \underline{\alpha}_j \mathbf{k}_s))$ as belonging to this category because of the above-proposed interpolation procedure) have been calculated and stored. In this respect we assume in fact to have sufficient storage capacity at our disposal. Without this facility, computation times will certainly grow beyond practical limits but on the new-generation type computing systems the storage capacity should not pose any principle problem.

With these assumptions we will now estimate the time required for the determination of $M_{\mathbf{G}, \mathbf{G}'}^{\text{GW}}(\mathbf{k}; \epsilon)$ for one combination of \mathbf{k} , ϵ , \mathbf{G} and \mathbf{G}' . Note in this connection that determination of the whole M^{GW} matrix at \mathbf{k} and ϵ requires the knowledge of N_{pw}^2 such matrix elements, where N_{pw} is the number of plane waves taken into account. If, however, \mathbf{k} has some symmetry, the determination of the whole M^{GW} matrix may require appreciably less than N_{pw}^2 times the computation time for one $M_{\mathbf{G}, \mathbf{G}'}^{\text{GW}}(\mathbf{k}; \epsilon)$ matrix element. This remark is based on expression (4.69). Subsequently, we have of course to pay attention to the computation time required for the above-discussed stored quantities \tilde{d} , \tilde{E} , $\tilde{W}_{\mathbf{K}, \mathbf{K}'}(\underline{\alpha}_j \mathbf{k}_s; i\epsilon')$ and $\tilde{W}_{\mathbf{K}, \mathbf{K}'}(\underline{\alpha}_j \mathbf{k}_s; E_i)$.

It is not difficult to see that the most time-consuming contributions to $M_{\mathbf{G}, \mathbf{G}'}^{\text{GW}}(\mathbf{k}; \epsilon)$ originate from (5.8), (5.14) and (5.20). The reason is that these terms contain a summation both over \mathbf{K} and \mathbf{K}' , which is at least one summation more than involved in the remaining contributions. In a calculation with the number of plane waves of the order of hundred, the computation time will be almost completely determined by these three contributions. Though (5.4)

Table I

Estimate performance of the Burroughs B7900 system

name	symbol	required time $\times 10^6$ s
assignment	τ_1	1.0
summation/subtraction	τ_2	1.5
multiplication	τ_3	2.0
division	τ_4	3.5
standard functions	τ_5	45.0
comparison	τ_6	1.0

contains only one summation over \mathbf{K} , we will nevertheless deal with its computation time, as it may be of interest for a Hartree-Fock self-consistent-field calculation; (5.5)-(5.7) are not considered in this respect as they contain no \mathbf{K} summation at all.

In Table I times are given for elementary operations such as approximately valid for a Burroughs B7900 computing system, which we will take as our reference. Our aim is to express the evaluation time for $M_{\mathbf{G},\mathbf{G}'}^{\text{GW}}(\mathbf{k};\epsilon)$ quite generally in terms of these elementary times τ_1 - τ_6 . Let us denote the minimum computation time required to evaluate a function f at fixed values of its arguments, subscripts and superscripts by $T\{f\}$. Let furthermore N_α be the number of operations in the lattice point group; L_v (L_c) the number of valence (conduction) bands; L the total number of bands; N_{sp} the number of special points and $N_{1\text{Bz}}$ the number of distinct points in 1Bz equivalent to the points in a given special point set. According to our best estimates we can write:

$$T\{\mathcal{M}_{\mathbf{G},\mathbf{G}'}^{(1)}(\mathbf{k})\} \lesssim N_{\text{sp}} N_{\text{pw}}^2 T\{f_{\mathbf{K},\mathbf{K}'}(\mathbf{k})\} + N_\alpha N_{\text{sp}} N_{\text{pw}} (7\tau_1 + 7\tau_2 + 2\tau_3), \quad (5.49)$$

$$T\{\mathcal{M}_{\mathbf{G},\mathbf{G}'}^{(1)}(\mathbf{k};\epsilon,\Delta)\} \sim N_{1\text{Bz}} N_{\text{pw}}^2 T\{\mathcal{K}_{\mathbf{K},\mathbf{K}'}^\Delta(\mathbf{k};\epsilon)\} + N_\alpha N_{\text{sp}} N_{\text{pw}}^2 (13\tau_1 + 2\tau_2 + 8\tau_3), \quad (5.50)$$

$$\begin{aligned}
T\{\mathcal{M}_{\mathbf{G},\mathbf{G}}^{(1)}(\mathbf{k};\epsilon,\epsilon')\} &\sim N_{1Bz} N_{pw}^2 T\{\mathcal{H}_{\mathbf{K},\mathbf{K}'}(\mathbf{k},\epsilon,i\epsilon')\} \\
&+ N_{\alpha} N_{sp} N_{pw}^2 (13\tau_1 + 2\tau_2 + 8\tau_3), \tag{5.51}
\end{aligned}$$

$$\begin{aligned}
T\{\mathcal{M}_{\mathbf{G},\mathbf{G}}^{\ell(1)}(\mathbf{k};\epsilon)\} &\sim N_{1Bz} \left\{ T\{\mathcal{S}^{\ell}(\mathbf{k};\mu,\epsilon)\} + N_{pw}^2 T\{h_{\mathbf{K},\mathbf{K}'}^{\ell}(\mathbf{k})\} \right\} \\
&+ N_{sp} N_{\alpha} N_{pw}^2 (14\tau_1 + 14\tau_2 + 12\tau_3) \\
&\sim N_{1Bz} N_{pw}^2 T\{h_{\mathbf{K},\mathbf{K}'}^{\ell}(\mathbf{k})\} \\
&+ N_{sp} N_{\alpha} N_{pw}^2 (14\tau_1 + 14\tau_2 + 12\tau_3). \tag{5.52}
\end{aligned}$$

in which

$$T\{f_{\mathbf{K},\mathbf{K}'}(\mathbf{k})\} \simeq L_v (10\tau_1 + 5\tau_2 + 8\tau_3), \tag{5.53}$$

$$T\{\mathcal{H}_{\mathbf{K},\mathbf{K}'}^{\Delta}(\mathbf{k};\epsilon)\} \simeq L (19\tau_1 + 11\tau_2 + 11\tau_3 + 2\tau_4 + 2\tau_5), \tag{5.54}$$

$$T\{\mathcal{H}_{\mathbf{K},\mathbf{K}'}(\mathbf{k},\epsilon,i\epsilon')\} \simeq L (15\tau_1 + 12\tau_2 + 8\tau_3 + 4\tau_4), \tag{5.55}$$

$$T\{\mathcal{S}^{\ell}(\mathbf{k};\mu,\epsilon)\} \simeq 4\tau_1 + 3\tau_6, \tag{5.56}$$

$$T\{h_{\mathbf{K},\mathbf{K}'}^{\ell}(\mathbf{k})\} \simeq 10\tau_1 + 4\tau_2 + 8\tau_3. \tag{5.57}$$

In (5.49)-(5.52) by the symbol " \sim " we mean that the relations are asymptotic ones, in the sense that they are to be used (only) for N_{pw} of the order of ten or larger (in (5.49) \lesssim means asymptotically smaller). Assuming N_e energy mesh points $i\epsilon'$ needed for the evaluation of the integral running from Δ to ∞ and assuming N_{ℓ} bands in the ℓ summation for $M_{\mathbf{G},\mathbf{G}}^{\text{Res}}(\mathbf{k};\epsilon)$, we come to the estimate

$$\begin{aligned}
T\{M_{G,G}^{GW}(\mathbf{k};\epsilon)\} &\sim T\{\mathcal{M}'_{G,G}(\mathbf{k};\epsilon,\Delta)\} + N_e T\{\mathcal{M}''_{G,G}(\mathbf{k};\epsilon,\epsilon')\} \\
&+ N_\ell T\{\mathcal{M}^\ell_{G,G}(\mathbf{k};\epsilon)\}.
\end{aligned}
\tag{5.58}$$

For determining the whole $M^{GW}(\mathbf{k};\epsilon)$ matrix, requiring the determination of $N_{pw} \times N_{pw}$ matrix elements, the following estimate holds (note that by suppressing the subscripts of a matrix in $T\{\cdot\}$, the computation time of the whole matrix is meant):

$$\begin{aligned}
T\{M^{GW}(\mathbf{k};\epsilon)\} &\sim \alpha_{\mathbf{k}} N_{pw}^2 T\{M_{G,G}^{GW}(\mathbf{k};\epsilon)\} \\
&- \alpha_{\mathbf{k}} N_{1Bz} N_{pw}^4 \left[T\{\mathcal{K}_{K,K}^\Delta(\mathbf{k};\epsilon)\} + T\{\mathcal{K}_{K,K}(\mathbf{k};\epsilon,i\epsilon')\} + T\{h_{K,K}^\ell(\mathbf{k})\} \right],
\end{aligned}
\tag{5.59}$$

in which the quantity $\alpha_{\mathbf{k}}$, $0 < \alpha_{\mathbf{k}} \leq 1$, takes account of the fact that for symmetric \mathbf{k} vectors, according to (4.69), only a part of the self-energy matrix elements need to be evaluated *directly*. In case \mathbf{k} is a non-symmetric point, $\alpha_{\mathbf{k}}$ equals unity. Note that the second term on the right-hand side of (5.59) accounts for the fact that the matrices \mathcal{K}^Δ , \mathcal{K} and h^ℓ are to be evaluated only once (see (5.50)-(5.52)). Using, as an example, $N_\alpha=48$, $N_{sp}=2$, $N_{1Bz}=32$, $N_{pw}=15$, $L=10$, $L_v=4$, $N_e=10$ and $N_\ell=3$ we obtain

$$T\{M_{G,G}^{GW}(\mathbf{k};\epsilon)\} \sim 57.8 \text{ s.}
\tag{5.60}$$

Determination of the whole $M^{GW}(\mathbf{k};\epsilon)$ matrix then requires for a non-symmetric wave vector \mathbf{k} about 9.02×10^3 s, which is about 2.5 hours. If, however, \mathbf{k} is chosen to be a symmetric point, this time will be enormously reduced.

Let us now consider the computation times involved in the evaluation of $\mathfrak{d}_{\ell,\underline{\alpha}_j\mathbf{k}_s}(\mathbf{K})$, $\mathfrak{d}_{\ell,\mathbf{k}-\underline{\alpha}_j\mathbf{k}_s}(\mathbf{K})$, $\mathfrak{E}_{\ell}(\underline{\alpha}_j\mathbf{k}_s)$, $\mathfrak{E}_{\ell}(\mathbf{k}-\underline{\alpha}_j\mathbf{k}_s)$, $\mathfrak{W}_{K,K'}(\underline{\alpha}_j\mathbf{k}_s;i\epsilon')$ and $\mathfrak{W}_{K,K'}(\underline{\alpha}_j\mathbf{k}_s;E_i)$, again as above in the first iteration step. First we deal with

the times involved in the evaluation of $\tilde{d}_{\ell,q}(K)$ and $\tilde{E}_{\ell}(q)$, with q some point in 1Bz such as $\underline{\alpha}_j k_s$ or $k - \underline{\alpha}_j k_s$. These quantities are obtained by the solution of an eigenvalue problem in which use has been made of *ab initio* norm conserving pseudopotentials such as given in Ref. [173]. As is well-known, the latter procedure for an $N_{pw} \times N_{pw}$ system of equations, takes approximately $\tau_7 N_{pw}^3$ s per q point [174]. The parameter τ_7 which, among other things, depends on the computing capacity of a system, is taken equal to 0.4×10^{-4} s in our reference system. In view of the relatively large number of $k - \underline{\alpha}_j k_s$ points involved, it is necessary to do these evaluations in many points in 1Bz which is very time consuming. Incidentally, the calculation of \tilde{d} and \tilde{E} quantities in a higher-iteration step will also require the computation of $M^{GW}(k - \underline{\alpha}_j k_s; \epsilon)$ at all these $k - \underline{\alpha}_j k_s$ points. As according to the above estimates the determination of the M^{GW} matrix takes at least 2.5 hours (see below (5.60)), it is clear that it will be problematic to reach self-consistency.

As a last estimate let us consider the computation time involved in the evaluation of an $N_{pw} \times N_{pw}$ screened-interaction matrix at an imaginary energy, i.e. $T\{W(k''; i\epsilon')\}$. This time will be representative for $T\{W(k''; E_i)\}$ as well. We start by dealing with $T\{P_{K,K'}(k''; i\epsilon')\}$. Making use of a special-point method and assuming the precalculation of all required terms of the kind (see (4.37)):

$$a_{K'}(-\underline{\alpha}_j k_s, \underline{\alpha}_j k_s + k''; \ell_v, \ell_c) a_K^*(-\underline{\alpha}_j k_s, \underline{\alpha}_j k_s + k''; \ell_v, \ell_c),$$

$$\tilde{E}_{\ell_v}(\underline{\alpha}_j k_s) - \tilde{E}_{\ell_c}(\underline{\alpha}_j k_s + k''),$$

we obtain the following result:

$$T\{P_{K,K'}(k''; i\epsilon')\} \sim N_{sp} L_v L_c N_{\alpha} (17\tau_1 + 7\tau_2 + 18\tau_3 + \tau_4). \quad (5.61)$$

If the wave vector k'' is a general point in 1Bz, all elements of the polarization matrix have to be calculated independently, so that the computation time of an $N_{pw} \times N_{pw}$ polarization matrix amounts to (note that in the first iteration cycle

the polarization function along the imaginary axis is Hermitian):

$$T\{P(\mathbf{k}''; i\epsilon')\} \sim \frac{1}{4} N_{pw}^2 T\{P_{\mathbf{K}, \mathbf{K}'}(\mathbf{k}''; i\epsilon')\}. \quad (5.62)$$

According to the relation $W=(1-vP)^{-1}v$ and the fact that the involved computation time in a matrix inversion procedure amounts asymptotically to $\tau_8 \times N_{pw}^3$, [175], with $\tau_8 \approx 0.4 \times 10^{-4}$, we obtain

$$T\{W(\mathbf{k}''; i\epsilon')\} \sim T\{P(\mathbf{k}''; i\epsilon')\} + N_{pw}^2 (34\tau_1 + 2\tau_2 + 8\tau_3 + 2\tau_6) + \tau_8 N_{pw}^3. \quad (5.63)$$

If we use the parameters employed in the calculation of (5.60) and choose $N_{pw}=15$, we obtain

$$T\{W(\mathbf{k}''; i\epsilon')\} \sim 17.4 \text{ s}. \quad (5.64)$$

As it is necessary to compute \tilde{W} at N_{sp} special points \mathbf{k}'' and at N_e energy points along the imaginary energy axis, the latter computation time will be enhanced by a factor $N_e \times N_{sp}$. Assuming the screened interaction to be a smooth function along the imaginary energy axis we expect that $N_e \approx 10$ will be sufficient, in general.

In conclusion, we have shown that a first-principles calculation of the three most time consuming contributions to the $15 \times 15 M^{GW}(\mathbf{k}; \epsilon)$ matrix in the *first iteration cycle* of a self-consistency procedure at one *non-symmetric* \mathbf{k} point and at one ϵ value takes about 2.5 hours. Taking into account the time required to calculate the remaining 29 contributions, we estimate the total evaluation time of $M^{GW}(\mathbf{k}; \epsilon)$ to be about 5 hours. There is however additional computing time to be taken into account because of the necessary determinations of \tilde{d} , \tilde{E} and \tilde{W} functions. In view of this we estimate the total evaluation time of $M^{GW}(\mathbf{k}; \epsilon)$ at a non-symmetric \mathbf{k} point to be about 5.25 hours. In giving this estimate, we took $N_{sp}=2$, however, which might be too optimistic in view of required accuracy. On the other hand, if we consider *symmetric* \mathbf{k} points, for

instance in the (1,0,0) or (1,1,1) direction, the computation time will significantly reduce. Let this reduction be in the order of ten. As we will need $M^{\text{GW}}(\mathbf{k};\epsilon)$ at a given \mathbf{k} in about twenty to thirty ϵ values, in order to be able to solve the quasi-particle equations [62], we estimate that the evaluation of one set of energy values $E_{\mathbf{k}}^{\text{QP}}(\mathbf{k})$ at one such symmetric \mathbf{k} point will take about 14 hours.

This indicates that a first-iteration-cycle determination of the band structure is within reach for a system in which we restrict ourselves to fifteen plane waves. It indicates also that extension to more plane waves is only possible at this moment if further refinements in our calculation procedure can be carried through, among which, for instance, parallel computing. Extension to a self-consistent way of determining energy band structures is *not* yet within reach, unless the above-used τ_1 - τ_2 values reduce dramatically, or unless actual first-iteration-cycle calculation indicate that the obtained $M_{\mathbf{G},\mathbf{G}'}^{\text{GW}}(\mathbf{k};\epsilon)$ matrix elements appear to show little structure in \mathbf{k} . We repeat, however, that pursuing self-consistency may not be necessary at all [64,65].

We find ourselves in the unfavorable position of not being able to produce numerical results for the self-energy function or the quasi-particle band structure, not even in a first iteration step within the GW approximation. However, several partial results have been obtained and progress is still being made. Moreover, valuable preparatory work has been done and presented, while a fairly complete exposition has been given in the present chapter of the way in which the GW self-energy function is to be calculated. Without the claim of having foreseen all practical problems or presenting the ultimate strategy for finishing the work, we are convinced that all basic difficulties have been recognized and strategies for tackling them have been proposed.

APPENDIX A

DIAGRAMMATIC APPROACH TO THE CALCULATION OF THE ONE-PARTICLE GREEN AND SELF-ENERGY FUNCTION

This appendix summarizes the main features of the representation of the one-particle Green function G and the related self-energy function M for an interacting many-particle system in terms of diagrams [176-179]. The method accounts for all perturbation-expansion terms when expanding the full Green function G in terms of the interaction part of the Hamiltonian. The diagrammatic technique, originally due to Feynman, enables one to obtain the whole perturbation series merely on the basis of topological properties of diagrams. The new aspects in this appendix are the inclusion of the $z_{\mathcal{L}}$ [57,180] and $z_{n\mathcal{L}}$ interaction terms and the diagrammatic derivation of Hedin's equations [56,181]. Note that, although we are interested in the behavior of particles with spin, the material in this appendix mainly deals with spinless particles. We have chosen to do so because the results for particles without spin can be straightforwardly modified to cover the case of particles with spin. Thus, unnecessary complicated notational work is avoided. The above-mentioned modifications will be discussed at the end of this appendix.

We first note that the solution of Dyson's equation (2.10) can symbolically be written

$$\begin{aligned} G &= (1 - G^{\circ}M)^{-1}G^{\circ} \\ &= G^{\circ} + G^{\circ}MG^{\circ} + G^{\circ}MG^{\circ}MG^{\circ} + \dots \end{aligned} \quad (\text{A.1})$$

By introducing the function \hat{M} defined by

$$\hat{M} = M + MG^{\circ}M + MG^{\circ}MG^{\circ}M + \dots = MGG^{\circ-1}, \quad (\text{A.2})$$

we can, making use of the right-hand side of (A.1), express the Green function as

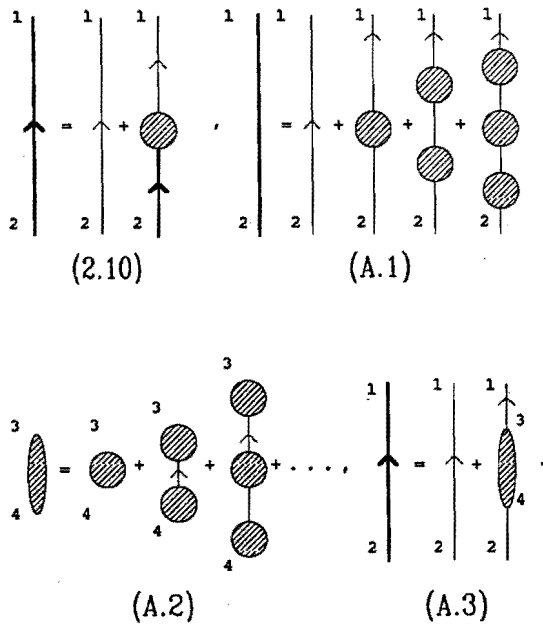


Fig. A1. The diagrammatic notations of equations (2.10), (A.1), (A.2) and (A.3).

$$G = G^0 + G^0 \hat{M} G^0 \tag{A.3}$$

The function \hat{M} is usually called the *improper* self-energy, contrary to M which often is referred to as the *proper* self-energy.

In Fig. A1 a diagrammatic notation of equations (2.10), (A.1) (A.2) and (A.3) are given. It should be realized that these diagrammatic equations are nothing but formal visualizations of the respective equations and do not at this stage contribute to solving G and M in terms of G^0 , v , z_{ℓ} and $z_{n\ell}$.




The important point is now that the general theory [176-179] shows that G can be written in the form of a perturbation series in the interaction, each term being a multiple space-time integral of products of unperturbed Green functions G^0 , interaction functions $v(1,2)=v(r_1-r_2)\delta(t_1-t_2)$, $z_{\ell}(r_1)$, $z_{n\ell}(1,2)=z_{n\ell}(r_1,r_2)\delta(t_1-t_2)$, and a numerical factor to be specified below. Among all

possible products of the functions G^0 , v , z_ℓ and $z_{n\ell}$ only those products are allowed which do not fall apart into factors depending on disjoint subsets of the space-time variables; these allowed terms are called *linked* or *connected*. Of most practical importance is the fact that each term of this perturbation series can uniquely be represented by diagrams which are commonly referred to as *Feynman diagrams*.

In a Feynman diagram, different types of functions are represented by different types of lines, whereas the arguments of the functions, being space-time points, are symbolized by points accompanied by a unique number, j say, representing the space-time point $r_j t_j$.

A direct consequence of the specific kind of terms, allowed in the perturbation series of G , is the restriction of the diagrams to linked or connected ones. A connected or linked diagram is a diagram which does not consist of separate parts. A diagram representing a term with n interactions, either of the kind v , z_ℓ or $z_{n\ell}$ is called a *diagram of n th order*.

The prescription of drawing an n th-order diagram involving l v -interactions, m $z_{n\ell}$ -interactions and $(n-l-m)$ z_ℓ -interactions is as follows:

- (i) Mark two points 1 and 2 on the paper. These two points specify the space-time points in the function $G(r_1 t_1, r_2 t_2)$. We shall refer to these points as *external* points, as opposed to the other points in the diagram which are referred to as *internal* points.
- (ii) Mark $l+m$ pairs of internal points (vertices) on the paper and label them $3, 3'; 4, 4'; \dots, (l+m+2), (l+m+2)'$. Join the pairs of points (j, j') , $j=3, 4, \dots, l+2$, $j'=3', 4', \dots, (l+2)'$ by v -interaction lines (dotted lines ) and the pairs of points (j, j') , $j=l+3, l+4, \dots, (l+m+2)$, $j'=(l+3)', (l+4)', \dots, (l+m+2)'$ by $z_{n\ell}$ -interaction lines (thick wiggled lines ) . Mark $(n-l-m)$ additional points $l+m+3, l+m+4, \dots, n+2$ and join to each one a z_ℓ -interaction line (thin wiggled line ) .
- (iii) Draw *directed* lines, the so-called *particle-lines*, (thin lines $j \longrightarrow i$), representing unperturbed Green functions $G^0(i, j)$, such that each of the internal points connected with a v - or z_ℓ -interaction line has precisely one particle line entering and one leaving. To the points connected with a $z_{n\ell}$ -interaction line only *one* directed line should be attached; it is either entering or leaving.

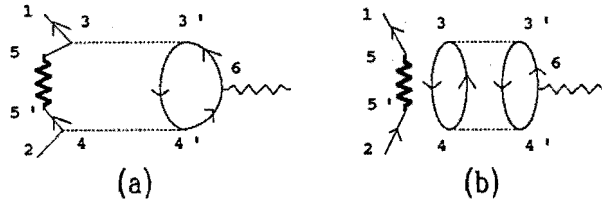


Fig. A2. Examples of two 4th-order Feynman diagrams. (a) A linked or connected diagram. (b) An unlinked or disconnected diagram.

As an example, consider Fig.A2 representing two 4th-order diagrams containing all three kinds of interaction v , z_l and z_{nl} .

In considering the diagrams, we may restrict ourselves to the collection of *topologically inequivalent* diagrams. *Topologically equivalent* diagrams, which can readily be shown to have the same contribution [176], are accounted for by just calculating the contribution of one of them and multiplying this contribution by the number of equivalent diagrams. Two diagrams are said to be topologically equivalent if they can be transformed into one another, irrespective of the indices of the *internal* vertices, by continuous deformations. The latter consist of all kinds of rotations, either of the whole or a part of the diagram, stretchings, shortenings, etc., provided that none of the lines is cut. For instance, the diagrams in Fig.A3 are all topologically equivalent.

We call a representative of a class of topologically equivalent diagrams, the *topological structure* of the corresponding class. The linked topological structures contributing to G , up to the second order, are given in Fig. A4.

The prescription of calculating the contribution of a given topological structure of n th order, with l v -interaction lines, m z_{nl} -lines and $(n-l-m)$ z_l -lines is given below:

- (i) Assign to each v line connecting j and j' , $v(j,j')$; to each z_{nl} line between j and j' , $-z_{nl}(j,j')$ and to each z_l line attached to j , $-z_l(r_j)$. Each particle line directed from j' to j is to be identified with $G^0(j,j')$; in the case $j \equiv j'$, it has to be assumed that $t_j' = t_j + \eta$ with η an infinitesimal small positive number. The latter is meant to guarantee the correct ordering of the field operators in the defining relation of G^0 .

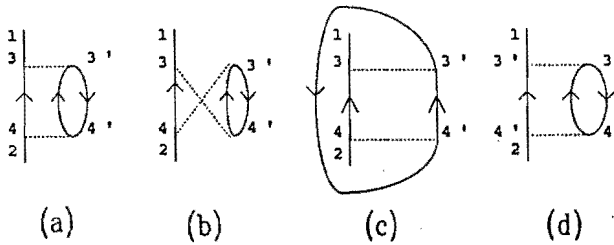


Fig. A3. Four 2nd-order diagrams which are topologically equivalent with each other.

- (ii) Multiply the contribution of all lines in the diagram, and subsequently integrate over all *internal* variables j, j' .
- (iii) Multiply the result obtained in (ii) by a factor $(-i)(-i/\hbar)^n \times (i)^{n+1}(-1)^F = (-1)^F i^1/\hbar^n$. Here F is the number of *closed loops* in the diagram. For example, F 's in the diagrams $(b_3), (c_3)$ and (c_4) of Fig. A4 are 0, 1 and 2, respectively.

As an example, we write down the contribution to G of the topological structures (b_1) and (b_2)

$$\begin{array}{c} 1 \\ 3 \\ 2 \end{array} \left| \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \right. = -\hbar^{-1} \int d(3) G^0(1,3) z_{\ell}(r_3) G^0(3,2). \quad (\text{A.4})$$

$$\begin{array}{c} 1 \\ 3 \\ 2 \end{array} \left| \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \right. \text{---} \text{---} \text{---} = (-1) i \hbar^{-1} \int d(3) d(4) G^0(1,3) v(3,4) G^0(4,4^+) G^0(3,2) \\ = -i \hbar^{-1} \int d(3) d^3 r_4 G^0(1,3) v(r_3 - r_4) G^0(r_4 t_3, r_4 t_3^+) G^0(3,2). \quad (\text{A.5})$$

We are now able to obtain G in terms of G^0 , v , z_{ℓ} and $z_{n\ell}$ in the form of a diagrammatic expression,

$$\begin{array}{c} 1 \\ 2 \end{array} \left| \begin{array}{c} \text{---} \\ \text{---} \end{array} \right. = \begin{array}{c} 1 \\ 2 \end{array} \left| \begin{array}{c} \text{---} \\ \text{---} \end{array} \right. + \begin{array}{c} 1 \\ 2 \end{array} \left| \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \right. + \begin{array}{c} 1 \\ 2 \end{array} \left| \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \right. \text{---} \text{---} \text{---} + \begin{array}{c} 1 \\ 2 \end{array} \left| \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \right. \text{---} \text{---} \text{---} + \begin{array}{c} 1 \\ 2 \end{array} \left| \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \right. \text{---} \text{---} \text{---} + \dots \quad (\text{A.6})$$

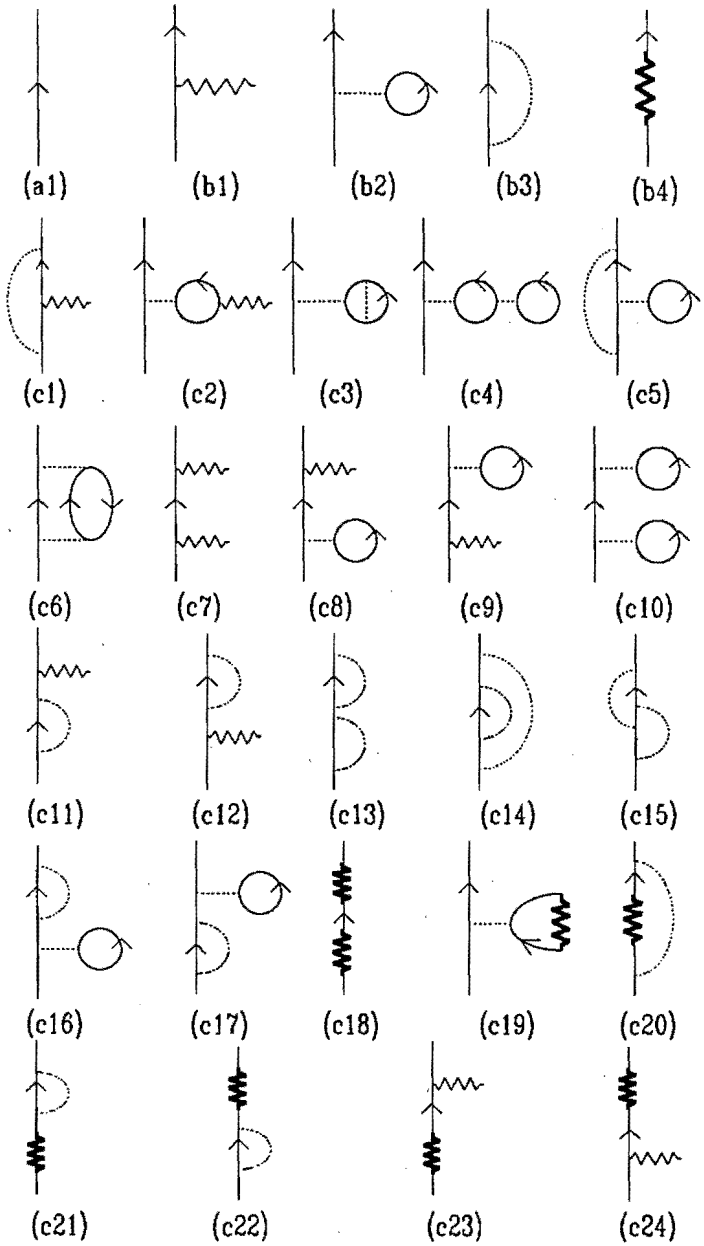
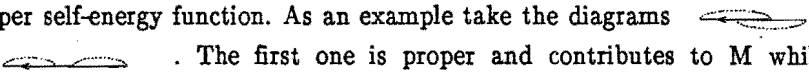







Fig. A4. All linked topological structures contributing to G , up to the second order; (a₁) is a zeroth-order structure; (b₁)-(b₄) are 1st-order structures; (c₁)-(c₂₄) are 2nd-order structures.

By identifying this expression for G with the expression (A.1), one can extract an expansion for M in terms of functions G^0 , v , z_ℓ and $z_{n\ell}$. To this end, let us consider a diagram contributing to G with external lines left out, i.e. without the particle lines entering vertex 1 and leaving vertex 2. From (A.3) we deduce that this diagram is an \tilde{M} diagram; only if it cannot be split up into separate diagrams by cutting a single particle-line, it is an M diagram as well (cf. (A.1)). In the former case we call this diagram an *improper* diagram whereas in the latter case a *proper* diagram. The collection of all proper diagrams to be obtained in this way, i.e. by leaving out the external lines, gives the complete proper self-energy function. As an example take the diagrams . The first one is proper and contributes to M while the second one is improper; both diagrams contribute to \tilde{M} .

Let us now introduce the concept of a *skeleton* M diagram [182]. Such a diagram is defined by demanding that it is an M diagram with the restriction that it does not contain any internal M diagram. For example  is a skeleton diagram, whereas  is not, for this contains . It will be clear that all diagrams for M may be obtained by drawing all skeleton M diagrams and then inserting all possible M parts. This is equivalent with [182,183]:

$$M = \{ \text{Contribution of all possible skeleton } M \text{ diagrams} \\ \text{with } G^0 \text{ replaced by } G \} . \quad (\text{A.7})$$

Equation (A.7) is actually an implicit equation for M since G also contains M .

In considering the skeleton M diagrams we notice that the two first-order diagrams  and  are special ones, as they represent the only local contributions to M . As the self-energy function, introduced in (2:10), is by definition a two-point function, it turns out that in these two cases the contributions are to be expressed with the help of δ functions [cf. (A.4) and (A.5)]:

$$\frac{1}{2} \text{wavy line} = -\hbar^{-1} \delta(1,2) z_\ell(r_1), \quad (\text{A.8})$$

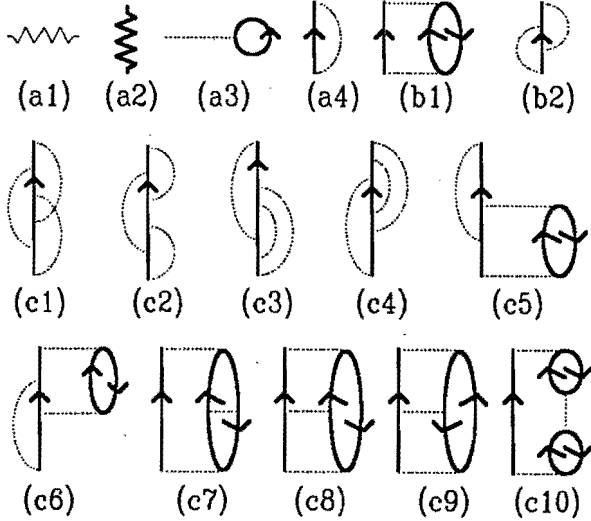


Fig. A5. All skeleton M structures up to the third order in interactions. (a_1)-(a_4) are 1st-order structures. (b_1), (b_2) are 2nd-order and (c_1)-(c_{10}) are 3rd-order structures.

$$\frac{1}{2} \text{---} \text{---} \bigcirc \text{---} = -i\hbar \delta(1,2) \int d^3r_1' v(r_1-r_1') G(r_1't_1', r_1't_1'^+). \quad (\text{A.9})$$

Also the first order skeleton diagrams involving the non-local interaction z_{nl} is special and leads to the M contribution

$$\frac{1}{2} \text{---} \text{---} \text{---} = -\hbar^{-1} z_{nl}(r_1, r_2) \delta(t_1-t_2). \quad (\text{A.10})$$

Note that a thick line in (A.9) represents a G function.

In Fig.A5 the set of skeleton M structures, with full Green functions, up to the third order are presented.

The prescription of calculating the contributions of the skeleton

diagrams of order $n > 1$ is:

- (i) Assign to each v line connecting j and j' , $v(j, j')$, and to each thick line $j' \longrightarrow j$ directed from j' to j , $G(j, j')$.
- (ii) Multiply the contributions of all lines in the diagram, and subsequently integrate over all internal space-time variables, except the ones connected with only one particle-line.
- (iii) Multiply the results obtained in (ii) by the factor $(-i)(-i/\hbar)^n \times (i)^{2n+1}(-1)^F = (-1)^F(i/\hbar)^n$; F being, as before, the number of closed particle loops in the diagram.

In this way we have a prescription at our disposal giving the self-energy function M in terms of the Green function G and the known functions v , z_ℓ and $z_{n\ell}$

We now want to show (diagrammatically) that the contribution to M due to all diagrams (but the first three diagrams in Fig.A5) can also be expressed in a closed analytical form, in terms of G and v . The resulting expression, however, has to be supplemented with three additional equations, as it turns out to be helpful to introduce three additional functions, i.e. a screened interaction function W , a polarization function P and a vertex function Γ . The resulting equations are known as Hedin's equations [56]. A direct derivation of the above-mentioned equations, making use of a non-perturbative method, will be given in Appendix B.

We first introduce the dynamically screened interaction function $W(1,2)$, (dashed interaction line) by means of the equation [56,57,181,184]

$$\begin{array}{c} 1 \\ \vdots \\ 2 \end{array} = \begin{array}{c} 1 \\ \vdots \\ 2 \end{array} - \begin{array}{c} 1 \\ \vdots \\ 2 \end{array} + \begin{array}{c} 1 \\ \vdots \\ \text{diamond} \\ \vdots \\ 2 \end{array} \quad (A.11)$$

where

$$\begin{aligned}
P(3,4) = & \begin{array}{c} 3 \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ 4 \end{array} - \text{---} \quad \text{(zeroth order in } W) \\
& + \text{---} \quad \text{(first order in } W) \\
& + \text{---} + \text{---} + \text{---} \\
& + \text{---} + \text{---} + \text{---} \quad \text{(second order in } W) \\
& + \dots \quad \text{(higher order in } W). \quad (A.12)
\end{aligned}$$

In analogy with Dyson's equation for the one-particle Green function, we may now write this as

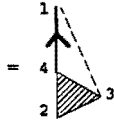
$$W(1,2) = v(1,2) + \int d(3)d(4) v(1,3)P(3,4)W(4,2), \quad (A.13)$$

or symbolically $W=v+vPW$. The function $P(i,j)$ is called polarization function. An alternative way of writing (A.13) is

$$W(1,2) = \int d(3) \epsilon^{-1}(1,3)v(3,2), \quad (A.14)$$

thus introducing the inverse dielectric screening function ϵ^{-1} .


It is now possible to express the self-energy function M completely in terms of the contributions of $z_{\ell} z_{n\ell}$, Hartree potential and a series of skeleton diagrams in which the dotted v lines have been replaced by dashed W lines. Up to the third order in W we now have [61]



$$= \frac{i}{\hbar} \int d(3)d(4) W(1^+, 3)G(1,4)\Gamma(4,2;3), \quad (\text{A.16})$$

where the vertex part, according to (A.15), is formally represented by

$$\begin{aligned}
 \Gamma(4,2;3) &= \begin{array}{c} 4 \\ \triangleleft \\ 2 \end{array} 3 \\
 &= \begin{array}{c} 4 \\ \bullet \\ 2 \end{array} 3 \quad (\text{zeroth order in } W) \\
 &+ \begin{array}{c} 4 \\ \triangleleft \\ 2 \end{array} 3 \quad (\text{first order in } W) \\
 &+ \begin{array}{c} 4 \\ \triangleleft \\ 2 \end{array} 3 + \begin{array}{c} 4 \\ \triangleleft \\ 2 \end{array} 3 + \begin{array}{c} 4 \\ \triangleleft \\ 2 \end{array} 3 \\
 &+ \begin{array}{c} 4 \\ \triangleleft \\ 2 \end{array} 3 + \begin{array}{c} 4 \\ \triangleleft \\ 2 \end{array} 3 + \begin{array}{c} 4 \\ \triangleleft \\ 2 \end{array} 3 \\
 &+ \dots \quad (\text{higher order in } W).
 \end{aligned} \quad (\text{A.17})$$

The plus sign of the argument of W on the right-hand side of (A.16) has its origin in the fact that otherwise the contribution of the first-order diagram  due to the instantaneous time behavior of the bare Coulomb interaction, contained in W , would not be specified.

From the diagrammatic structure of the vertex part Γ in (A.17) it is straightforward to obtain the corresponding analytical expression for the vertex function: The zeroth-order contribution to $\Gamma(4,2;3)$ is clearly equal to

$\delta(4,2)\delta(4,3)$ as can be deduced from the M diagram in (A.15) contributing to first order in W. As far as the higher-order contributions to $\Gamma(4,2;3)$ are concerned, it should be realized that the structure of the vertex diagrams in (A.17), with the exception of the zeroth-order diagrams in W, is such that each Γ diagram can be obtained by starting from any skeleton diagram contributing to M' , in which, however, one internal line representing $G(k,l)$ is replaced by two particle lines representing $G(k,i)$, $G(j,l)$, and a diagram contributing to $\Gamma(i,j;3)$. Note that this replacement has to take place in the polarization parts, contributing to dashed interaction lines, as well. The elimination from an $M'(4,2)$ diagram of subsequently *all* internal particle lines representing the functions $G(k,l)$ is in fact equivalent with taking the *functional derivative* [54,185] $\delta M'(4,2)/\delta G(k,l)$. Therefore, taking the above structural property of the vertex part Γ into account, we conclude to the analytical expression

$$\Gamma(4,2;3) = \delta(4,2) \delta(4,3) + \int d(i)d(j)d(k)d(l) \frac{\delta M'(4,2)}{\delta G(k,l)} \times G(k,i)G(j,l)\Gamma(i,j;3). \quad (\text{A.18})$$

Until now, the polarization function $P(1,2)$ is the only function which has not yet been given in analytical form. In view of our above definition (A.17) of Γ it is straightforward, however, to conclude to

$$P(1,2) = -\frac{i}{\hbar} \int d(i)d(j) G(1,i)G(j,1^+)\Gamma(i,j;2). \quad (\text{A.19})$$

This completes our attempt to write M' , and therefore M , in closed analytical form. We succeeded in doing so by means of equations (A.13), (A.16), (A.18) and (A.19) which are Hedin's equations.

Thus far, we have dealt with spinless particles. We will now discuss the required modifications of the above results, if particles with spin are dealt with. As stated earlier, the above-presented prescriptions are easily carried over to the case of particles with spin. In fact the scalar functions we have been dealing with until now, such as the one-particle Green functions, the self-energy function, etc., become operators in the spinor space. As a consequence, the ordinary multiplications of the functions, as described in the prescriptions for evaluating

the one-particle Green and the self-energy function, become matrix multiplications. Therefore, in addition to the space-time integrations over the coordinates of the internal vertices of diagrams, a summation should be carried out over all of the internal spin indices. In the simplest case, in which the interparticle interaction is spin independent, all of the above-mentioned matrices become multiples of a unit tensor of appropriate rank. For instance, for the one-particle Green function we have $G_{\alpha,\beta}(1,2)=\delta_{\alpha,\beta}G(1,2)$, in which α and β are spin indices and $G(1,2)\equiv G_{\alpha,\alpha}(1,2)$, for all α . Note, that for electrons α assumes the values $\pm\frac{1}{2}$. The spin summations can be carried out directly, resulting in an enhancement of the contributions of diagrams containing *closed loops*, as compared to the case of spinless particles, while the contributions of other kind of diagrams remain unchanged. For particles with spin s , the above-mentioned enhancement factor is $(2s+1)^F$, in which F denotes the number of closed loops in the diagram under consideration. For electrons, the latter factor is equal to 2^F , which has been carried through in the main text.

APPENDIX B

LINEAR-RESPONSE APPROACH TO THE DETERMINATION OF THE ONE-PARTICLE GREEN FUNCTION AND THE CORRESPONDING SELF-ENERGY FUNCTION

In this appendix we reformulate Hedin's derivation [56] of the general relationship between the one-particle Green function G and the corresponding self-energy function M for an interacting system of electrons such as met in, e.g., a metal or a semiconductor. This relationship is commonly expressed in terms of the four equations (A.13), (A.16), (A.18) and (A.19) which in fact give M as a functional of G . We call these four equations *Hedin's equations*. In order to fix both G and M , these four equations are to be supplemented by either the Green function equation (2.7) or Dyson's equation (2.10). Hedin's derivation is based on *linear response* theory and falls back on an earlier treatment by Martin and Schwinger [55]. The advantage over the diagrammatic derivation given in appendix A is its *non-perturbative* character, in the sense that no use is made of series expansions in terms of the bare Coulomb interaction function v . This advantage is indeed important, as it is well-known that the *individual* terms in a series expansion in v turn out to suffer from divergencies [186]. Such divergencies are believed not to show up if G and M are expressed in terms of a *screened* interaction function W which appears the natural function to choose if linear response theory is applied.

Hedin's equations together with Dyson's equation (2.10) directly lead to the diagrammatic representation of M given in (A.15) involving (dashed) screened interaction lines. As mentioned above there are no indications that the *individual* terms in this series expansion in W suffer from divergencies, which is therefore an obvious advantage over the similar diagrammatic expansion of M given in Fig. A5. Still there is no *a priori* proof that the expansion (A.15) in terms of W converges [91]. In this thesis, however, we strongly suggest this to be the case; we even approximate the expression by its very first term (the GW term). Though Hedin's derivation as given in his paper [56] is certainly correct, a number of intermediate steps are not explicitly commented on. In view of this, we considered it worthwhile to present our (rather lengthy) reformulation of

Hedin's derivation of the expressions relating G and M.

We start by introducing an auxiliary time-dependent scalar and local potential $\phi(\mathbf{r},t)$ acting on the system under consideration. By studying the response of the systems on variations of ϕ , we will be able to establish the relation between G and M. We may think of ϕ as being the potential associated with an external (longitudinal) electric field which polarizes the system. The potential $\phi(\mathbf{r},t)$ will be assumed to vanish outside the time interval $(-T_0, T_0)$ where T_0 may have any fixed positive value.

In accordance with (2.5), we write the Hamiltonian of the system as

$$\begin{aligned} \hat{H}_\phi(t) = & \int d^3\mathbf{r} \hat{\psi}^\dagger(\mathbf{r}) \left\{ \frac{-\hbar^2}{2m} \nabla^2 + u(\mathbf{r}) + \phi(\mathbf{r},t) \right. \\ & \left. + 1/2 \int d^3\mathbf{r}' \hat{\psi}^\dagger(\mathbf{r}') v(\mathbf{r}-\mathbf{r}') \hat{\psi}(\mathbf{r}') \right\} \hat{\psi}(\mathbf{r}), \end{aligned} \quad (\text{B.1})$$

where $\hat{\psi}^\dagger$ and $\hat{\psi}$ are Schrödinger creation and annihilation field operators fulfilling the anticommutation relations (2.2) and where the time dependence of \hat{H}_ϕ is entirely through the additional time dependent $\phi(\mathbf{r},t)$. Note that, due to the presence of $\phi(\mathbf{r},t)$, the Schrödinger and Heisenberg representations of the Hamiltonian will be different, unlike the situation with $\phi \equiv 0$, in which case \hat{H} represents a constant of motion. In (B.1), the Hamiltonian is given in the Schrödinger representation.

The time evolution of state vectors is given by the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\Psi_\phi(t)\rangle_S = \hat{H}_\phi(t) |\Psi_\phi(t)\rangle_S, \quad (\text{B.2})$$

which, after introducing the unitary time-evolution operator $\hat{\Lambda}_\phi(t,t')$ through

$$|\Psi_\phi(t)\rangle_S = \hat{\Lambda}_\phi(t,t') |\Psi_\phi(t')\rangle_S, \quad (\text{B.3})$$

can also be expressed as

$$i\hbar \frac{\partial}{\partial t} \hat{\Lambda}_\phi(t,t') = \hat{H}_\phi(t) \hat{\Lambda}_\phi(t,t'), \quad (\text{B.4a})$$

and

$$i\hbar \frac{\partial}{\partial t'} \hat{\Lambda}_\phi(t, t') = -\hat{\Lambda}_\phi(t, t') \hat{H}_\phi(t). \quad (\text{B.4b})$$

It can be verified that $\hat{\Lambda}_\phi$ satisfies the following relations:

$$\hat{\Lambda}_\phi^\dagger(t, t') = \hat{\Lambda}_\phi(t', t) = \hat{\Lambda}_\phi^{-1}(t, t'), \quad (\text{B.5a})$$

$$\hat{\Lambda}_\phi(t, t') \hat{\Lambda}_\phi(t', t'') = \hat{\Lambda}_\phi(t, t''), \quad (\text{B.5b})$$

$$\hat{\Lambda}_\phi(t, t) = 1. \quad (\text{B.5c})$$

Heisenberg (i.e., time independent) state vectors will be introduced by means of the relation

$$|\Psi_\phi\rangle_{\text{H}} = \hat{\Lambda}_\phi(-T_0, t) |\Psi_\phi(t)\rangle_{\text{S}}, \quad (\text{B.6})$$

implying that in view of (B.5c) the Heisenberg and Schrödinger states coincide at $t=-T_0$.

Let \hat{O} be a time-independent Schrödinger operator. The Heisenberg representation of \hat{O} is given by

$$\hat{O}_\phi(t) = \hat{\Lambda}_\phi(-T_0, t) \hat{O} \hat{\Lambda}_\phi(t, -T_0). \quad (\text{B.7})$$

By using (B.4) we readily find that operator $\hat{O}_\phi(t)$ satisfies the Heisenberg equation of motion

$$i\hbar \frac{\partial}{\partial t} \hat{O}_\phi(t) = [\hat{O}_\phi(t), \hat{H}_\phi^{\text{H}}(t)], \quad (\text{B.8})$$

where $\hat{H}_\phi^{\text{H}}(t)$ is obtained from (B.1) by replacing $\hat{\psi}^\dagger(\mathbf{r})$ and $\hat{\psi}(\mathbf{r})$ by their

Heisenberg versions $\hat{\psi}_\phi^\dagger(\mathbf{r}t)$ and $\hat{\psi}_\phi(\mathbf{r}t)$.

We now introduce a one-particle Green function which we want to coincide with $G(\mathbf{r}t, \mathbf{r}'t')$ of (2.4) if $\phi \equiv 0$. The function

$$G_\phi(\mathbf{r}t, \mathbf{r}'t') = -i \frac{\text{H}\langle \Psi_N | \hat{\Lambda}(T_0, -T_0) \mathcal{T} \{ \hat{\psi}_\phi(\mathbf{r}t) \hat{\psi}_\phi^\dagger(\mathbf{r}'t') \} | \Psi_N \rangle \text{H}}{\text{H}\langle \Psi_N | \hat{\Lambda}_\phi(T_0, -T_0) | \Psi_N \rangle \text{H}} \quad (\text{B.9})$$

obviously reduces to (2.4) for $\phi \equiv 0$. As we will show, it also satisfies a Green-function type of equation of precisely the same form as $G(\mathbf{r}t, \mathbf{r}'t')$ in case of vanishing ϕ . In order to derive the equation of motion for G_ϕ , we depart from (B.8) in which $\hat{O}_\phi(t)$ is replaced by $\hat{\psi}_\phi(\mathbf{r}t)$, use the anticommutation rules (2.2), which are valid for the Heisenberg field operators at equal times, and multiply the members of the obtained equation on the right by $\hat{\psi}_\phi^\dagger(\mathbf{r}'t')$. Subsequent applications of the time-ordering operator then leads to

$$\begin{aligned} & -i \left[i\hbar \frac{\partial}{\partial t} + \frac{\hbar^2}{2m} \nabla^2 - u(\mathbf{r}) - \phi(\mathbf{r}, t) \right] \mathcal{T} \{ \hat{\psi}_\phi(\mathbf{r}t) \hat{\psi}_\phi^\dagger(\mathbf{r}'t') \} \\ & + i \int d^3r'' dt'' v(\mathbf{r}t'', \mathbf{r}'t'') \mathcal{T} \{ \hat{\psi}_\phi^\dagger(\mathbf{r}''t'') \hat{\psi}_\phi(\mathbf{r}''t'') \hat{\psi}_\phi(\mathbf{r}t) \hat{\psi}_\phi^\dagger(\mathbf{r}'t') \} \\ & = \hbar \delta(\mathbf{r}-\mathbf{r}') \delta(t-t'). \end{aligned} \quad (\text{B.10})$$

In obtaining (B.10) we have also used $\partial\Theta(t)/\partial t = \delta(t)$; the superscript "+" occurring in (B.10) over the time symbols are introduced to guarantee the correct order of operators.

In order to obtain the equation of motion for G_ϕ , we will first prove the important identity:

$$\begin{aligned} & \frac{\delta}{\delta\phi(\mathbf{3})} \left[\hat{\Lambda}_\phi(T_0, -T_0) \mathcal{T} \{ \hat{\psi}_\phi(1) \hat{\psi}_\phi^\dagger(2) \} \right] \\ & = -\frac{i}{\hbar} \hat{\Lambda}_\phi(T_0, -T_0) \mathcal{T} \{ \hat{\psi}_\phi^\dagger(3') \hat{\psi}_\phi(3) \hat{\psi}_\phi(1) \hat{\psi}_\phi^\dagger(2) \}, \end{aligned} \quad (\text{B.11})$$

where the symbol $\delta/\delta\phi(3)$ stands for the *functional derivative* with respect to $\phi(3)$. The functional derivation $\delta F_\phi/\delta\phi(1)$ of a functional F_ϕ is defined by means of the implicit relation

$$\int d(1) \frac{\delta F_\phi}{\delta\phi(1)} f(1) = \lim_{\eta \rightarrow 0} \frac{1}{\eta} \{F_{\phi+\eta f} - F_\phi\}, \quad (\text{B.12})$$

where $f(1)$ is an arbitrary, but smooth and integrable function. For a short introduction to the theory of functional derivative as well as for a compendium of useful differentiation rules the reader is referred to the literature [54,185,186]. Let us, however, consider a simple example in which $F_\phi = \int d(1) \mathcal{A}(1)\phi(1)$. By applying (B.12) we obtain

$$\begin{aligned} & \int d(1) \frac{\delta F_\phi}{\delta\phi(1)} f(1) \\ &= \lim_{\eta \rightarrow 0} \frac{1}{\eta} \left\{ \int d(1) \mathcal{A}(1)[\phi(1)+\eta f(1)] - \int d(1) \mathcal{A}(1)\phi(1) \right\} \\ &= \int d(1) \mathcal{A}(1)f(1). \end{aligned}$$

As this relation holds for every suitable f , we conclude that $\delta F_\phi/\delta\phi(1) = \mathcal{A}(1)$. In the same manner one obtains $\delta\phi(1)/\delta\phi(2) = \delta(1,2)$.

The equation which results if in (B.4a) ϕ is varied by an infinitesimal amount $\delta\phi$ reads

$$i\hbar \frac{\partial}{\partial t} \hat{\Delta}_\phi(t,t') = \hat{\delta H}_\phi(t) \hat{\Delta}_\phi(t,t') + \hat{H}_\phi(t) \delta \hat{\Delta}_\phi(t,t'). \quad (\text{B.13})$$

It can easily be verified by substitution that the solution to (B.13) is given by

$$\delta \hat{\Delta}_\phi(t,t') = \frac{-i}{\hbar} \int_{t'}^t dt_3 \hat{\Delta}_\phi(t,t_3) \delta \hat{H}_\phi(t_3) \hat{\Delta}_\phi(t_3,t'), \quad (\text{B.14})$$

which by using (B.1) can be written

$$\begin{aligned}
\delta \hat{\Lambda}_\phi(t, t') &= -\frac{i}{\hbar} \int_{t'}^t dt_3 \hat{\Lambda}_\phi(t, t_3) \\
&\times \left\{ \int d^3 r_3 \hat{\psi}^\dagger(\mathbf{r}_3) [\delta \phi(\mathbf{r}_3, t_3)] \hat{\psi}(\mathbf{r}_3) \right\} \hat{\Lambda}_\phi(t_3, t') \\
&\equiv -\frac{i}{\hbar} \int d(3) \{ \Theta(t-t_3) \Theta(t_3-t') - \Theta(t'-t_3) \Theta(t_3-t) \} \\
&\times \hat{\Lambda}_\phi(t, t_3) \hat{\psi}^\dagger(\mathbf{r}_3) \hat{\psi}(\mathbf{r}_3) \hat{\Lambda}_\phi(t_3, t') \delta \phi(3), \tag{B.15}
\end{aligned}$$

so that using (B.5) and (B.7), we obtain for the functional derivative (note that we change variables: $t-t_1, t'-t_2$)

$$\begin{aligned}
\frac{\delta \hat{\Lambda}_\phi(t_1, t_2)}{\delta \phi(3)} &= \frac{-i}{\hbar} \{ \Theta(t_1-t_3) \Theta(t_3-t_2) - \Theta(t_2-t_3) \Theta(t_3-t_1) \} \\
&\times \hat{\Lambda}_\phi(t_1, -T_0) \hat{\psi}^\dagger(3) \hat{\psi}(3) \hat{\Lambda}_\phi(-T_0, t_2). \tag{B.16}
\end{aligned}$$

This general relation is used, together with the general expression (B.7) for Heisenberg operators, to establish the relation

$$\frac{\delta}{\delta \phi(3)} \hat{\psi}_\phi(1) = \frac{i}{\hbar} \Theta(t_1-t_3) \{ \hat{\psi}_\phi^\dagger(3) \hat{\psi}_\phi(3) \hat{\psi}_\phi(1) - \hat{\psi}_\phi(1) \hat{\psi}_\phi^\dagger(3) \hat{\psi}_\phi(3) \}. \tag{B.17}$$

A similar relation for $\hat{\psi}_\phi^\dagger(1)$ can be obtained by taking the Hermitian conjugation of (B.17). If we now consider the left-hand side of (B.11) in which $\hat{\Lambda}_\phi(T_0, -T_0) \hat{\psi}_\phi(1) \hat{\psi}_\phi^\dagger(2)$ has to be differentiated with respect to $\phi(3)$, it is, using the chain rule of differentiation [54,187] together with (B.16) and (B.17), an easy task to obtain the identity (B.11).

We proceed by multiplying (B.10) on the left by $\hat{\Lambda}_\phi(T_0, -T_0)$ and by taking the expectation value of the equation in the ground state $|\Psi_N\rangle_H$. Using the definition (B.9), we straightforwardly arrive at

$$\begin{aligned}
& {}_H\langle\Psi_N|\hat{\Lambda}_\phi(T_o,-T_o)|\Psi_N\rangle_H [i\hbar\frac{\partial}{\partial t_1} + \frac{\hbar^2}{2m}\nabla_1^2 - u(r_1) - \phi(1)] G_\phi(1,2) \\
& -i\hbar\int d(3) v(1^+,3) \frac{\delta}{\delta\phi(3)} \{ {}_H\langle\Psi_N|\hat{\Lambda}_\phi(T_o,-T_o)|\Psi_N\rangle_H G_\phi(1,2) \} \\
& = \hbar {}_H\langle\Psi_N|\hat{\Lambda}_\phi(T_o,-T_o)|\Psi_N\rangle_H \delta(1,2). \tag{B.18}
\end{aligned}$$

The functional derivative occurring in (B.18) can be shown, again using (B.16), to equal

$$\begin{aligned}
& \frac{\delta}{\delta\phi(3)} \{ {}_H\langle\Psi_N|\hat{\Lambda}_\phi(T_o,-T_o)|\Psi_N\rangle_H G_\phi(1,2) \} \\
& = \frac{1}{\hbar} G_\phi(3,3^+) G_\phi(1,2) + {}_H\langle\Psi_N|\hat{\Lambda}_\phi(T_o,-T_o)|\Psi_N\rangle_H \frac{\delta G_\phi(1,2)}{\delta\phi(3)}, \tag{B.19}
\end{aligned}$$

where the first term on the right-hand side explicitly takes apart the contribution which factorizes with $G_\phi(1,2)$. As the explicit external-potential term $\phi(1)G_\phi(1,2)$ on the left-hand side of (B.18) is precisely of the same form, it is natural to introduce a one-electron response potential function in the following way:

$$\nu_\phi(1) \equiv \phi(1) - i \int d(3) v(1^+,3) G_\phi(3,3^+), \tag{B.20}$$

such that the equation of motion for $G_\phi(1,2)$ reduces to the form

$$\begin{aligned}
& [i\hbar\frac{\partial}{\partial t_1} + \frac{\hbar^2}{2m}\nabla_1^2 - u(r_1) - \nu_\phi(1)] G_\phi(1,2) \\
& -i\hbar\int d(3) v(1^+,3) \frac{\delta}{\delta\phi(3)} G_\phi(1,2) = \hbar \delta(1,2). \tag{B.21}
\end{aligned}$$

Note that, $\nu_\phi(1)$ reduces to the well-known Hartree potential if $\phi \equiv 0$. This implies that the local (Hartree) part of the interaction is now accounted for in $\nu_\phi(1)$. In accordance with the procedure followed in (2.8), we finally define a self-energy function M'_ϕ through the relation

$$\int d(3) M'_{\phi}(1,3) G_{\phi}(3,2) = i \int d(3) v(1^+,3) \frac{\delta}{\delta \phi(3)} G_{\phi}(1,2), \quad (\text{B.22})$$

such that (B.21) is alternatively written

$$\begin{aligned} & [i\hbar \frac{\partial}{\partial t_1} + \frac{\hbar^2}{2m} \nabla_1^2 - u(r_1) - \nu_{\phi}(1)] G_{\phi}(1,2) \\ & - \hbar \int d(3) M'_{\phi}(1,3) G_{\phi}(3,2) = \hbar \delta(1,2). \end{aligned} \quad (\text{B.23})$$

Note that for $\phi \equiv 0$, the above M'_{ϕ} coincides with the M' of (A.16), being equal to the M of (A.15) without the first three diagrams. In the main text we reserved the notation M , without prime, to generally indicate the self-energy including z_p , $z_{n\ell}$ and Hartree contributions.

In accordance with the procedure followed in section 2.1 we may also introduce the unperturbed Green function G_{ϕ}^0 , fulfilling the equation

$$[i\hbar \frac{\partial}{\partial t_1} + \frac{\hbar^2}{2m} \nabla_1^2 - u(r_1) - \nu_{\phi}(1)] G_{\phi}^0(1,2) = \hbar \delta(1,2). \quad (\text{B.24})$$

It can then again be shown (see the main text below (2.9)), that G_{ϕ} fulfills an inhomogeneous integral equation of the second kind (Dyson's equation)

$$G_{\phi}(1,2) = G_{\phi}^0(1,2) + \int d(3)d(4) G_{\phi}^0(1,3) M'_{\phi}(3,4) G_{\phi}(4,2), \quad (\text{B.25})$$

which is symbolically written

$$G_{\phi} = G_{\phi}^0 + G_{\phi}^0 M'_{\phi} G_{\phi}. \quad (\text{B.26})$$

The functional derivative occurring in (B.22), can alternatively be written

$$\frac{\delta G_{\phi}(1,2)}{\delta \phi(3)} = - \int d(4)d(5) G_{\phi}(1,4) \frac{\delta G_{\phi}^{-1}(4,5)}{\delta \phi(3)} G_{\phi}(5,2), \quad (\text{B.27})$$

which straightforwardly follows from the definition of the *inverse* of a two-point functional, being given by

$$\int d(3) F_{\phi}(1,3) F_{\phi}^{-1}(3,2) = \int d(3) F_{\phi}^{-1}(1,3) F_{\phi}(3,2) = \delta(1,2), \quad (\text{B.28})$$

or symbolically $F_{\phi} F_{\phi}^{-1} = F_{\phi}^{-1} F_{\phi} = 1$. Namely, taking the functional derivative of (B.28) with respect to ϕ directly leads to $\delta F_{\phi} / \delta \phi = -F_{\phi} \{ \delta F_{\phi}^{-1} / \delta \phi \} F_{\phi}$ which proves (B.27). Substitution of (B.27) in (B.22) and subsequent "multiplication" of the resulting equation on the right by G_{ϕ}^{-1} yields

$$M'_{\phi}(1,2) = -i \int d(3) d(4) v(1^*,3) G_{\phi}(1,4) \frac{\delta G_{\phi}^{-1}(4,2)}{\delta \phi(3)}. \quad (\text{B.29})$$

The final step in deriving Hedin's equations will now be to eliminate the functional derivative of G_{ϕ}^{-1} with respect to ϕ and to replace it by an expression involving a functional derivative of M'_{ϕ} with respect to G_{ϕ} . To this end, we first introduce the inverse of a *dielectric response* function [188,189]:

$$\epsilon_{\phi}^{-1}(1,2) \equiv \frac{\delta \nu_{\phi}(1)}{\delta \phi(2)}, \quad (\text{B.30})$$

being the "ratio" of the variation of the *response* function $\nu_{\phi}(1)$ and that of the *stimulating* function $\phi(2)$. In view of the definition of ν_{ϕ} in (B.20), we can express ϵ^{-1} as

$$\begin{aligned} \epsilon_{\phi}^{-1}(1,2) &= \delta(1,2) - i \int d(3) v(1^*,3) \frac{\delta G_{\phi}(3,3^*)}{\delta \phi(2)} \\ &= \delta(1,2) + i \int d(3) d(4) d(5) v(1^*,3) G_{\phi}(3,4) \frac{\delta G_{\phi}^{-1}(4,5^*)}{\delta \phi(2)} G_{\phi}(5,3^*), \end{aligned} \quad (\text{B.31})$$

where the last equality is due to (B.27). The leading idea in all this is that two electrons in the system will *not* interact with each other via the bare Coulomb interaction v but rather via the modified (screened) interaction:

$$\begin{aligned} W_{\phi}(1,2) &= \int d(3) \epsilon_{\phi}^{-1}(1,3) v(3,2) \\ &= v(1,2) + i \int d(3) d(4) d(5) d(6) v(1^*,4) G_{\phi}(4,5) \end{aligned}$$

$$\times \frac{\delta G_{\phi}^{-1}(5,6^*)}{\delta \phi(3)} G_{\phi}(6,4^*) v(3,2). \quad (\text{B.32})$$

Using the chain rule of differentiation

$$\frac{\delta}{\delta \phi(1)} = \int d(2) \frac{\delta \nu_{\phi}(2)}{\delta \phi(1)} \frac{\delta}{\delta \nu_{\phi}(2)}, \quad (\text{B.33})$$

and introducing a polarization function P_{ϕ} by means of

$$\begin{aligned} P_{\phi}(4,3) &= i \int d(5)d(6) G_{\phi}(4,5) \frac{\delta G_{\phi}^{-1}(5,6^*)}{\delta \nu_{\phi}(3)} G_{\phi}(6,4^*) \\ &\equiv -i \frac{\delta G_{\phi}(4,4^*)}{\delta \nu_{\phi}(3)}, \end{aligned} \quad (\text{B.34})$$

we straightforwardly rewrite (B.32) in the form

$$W_{\phi}(1,2) = v(1,2) + \int d(3)d(4) v(1^*,3) P_{\phi}(3,4) W_{\phi}(4,2), \quad (\text{B.35})$$

which is one of Hedin's equations (originally due to Hubbard [57]). If we define the so-called vertex function Γ_{ϕ} by means of

$$\Gamma_{\phi}(1,2;3) = -\hbar \frac{\delta G_{\phi}^{-1}(1,2)}{\delta \nu_{\phi}(3)}, \quad (\text{B.36})$$

equation (B.34) can also be written in the form (we change names of variables)

$$P_{\phi}(1,2) = \frac{-i}{\hbar} \int d(3)d(4) G_{\phi}(1,3) G_{\phi}(4,1^*) \Gamma_{\phi}(3,4;2), \quad (\text{B.37})$$

which is another equation of Hedin. Application of the chain rule (B.33) and use of (B.30), (B.32) and (B.36), when applied to (B.29) gives

$$M'_{\phi}(1,2) = \frac{i}{\hbar} \int d(3)d(4) G_{\phi}(1,3) W_{\phi}(4,1^*) \Gamma_{\phi}(3,2;4), \quad (\text{B.38})$$

what is again one of Hedin's equations.

Finally, we derive an equation for Γ_ϕ in terms of a functional derivative of M' with respect to G . This is achieved by first multiplying (B.23) on the right by $G_\phi^{-1}(2,3)$ and integrating with respect to 2. This gives

$$\hbar G_\phi^{-1}(1,3) = [i\hbar \frac{\partial}{\partial t_1} + \frac{\hbar^2}{2m} \nabla_1^2 - u(r_1) - \nu_\phi(1)] \delta(1,3) - \hbar M'_\phi(1,3). \quad (\text{B.39})$$

Differentiating this equation with respect to $\nu_\phi(2)$ gives (use (B.36))

$$\Gamma_\phi(1,3;2) = \delta(1,3)\delta(1,2) + \hbar \frac{\delta M'_\phi(1,3)}{\delta \nu_\phi(2)}. \quad (\text{B.40})$$

By applying the chain rule of differentiation for a two-point function:

$$\hbar \frac{\delta M'_\phi(1,3)}{\delta \nu_\phi(2)} = \int d(4)d(5) \frac{\delta G_\phi(4,5)}{\delta \nu_\phi(2)} \frac{\delta M'_\phi(1,3)}{\delta G_\phi(4,5)}, \quad (\text{B.41})$$

and by writing in accordance with (B.27) and (B.36):

$$\begin{aligned} \hbar \frac{\delta G_\phi(4,5)}{\delta \nu_\phi(2)} &= - \int d(6)d(7) G_\phi(4,6) \hbar \frac{\delta G_\phi^{-1}(6,7)}{\delta \nu_\phi(2)} G_\phi(7,5) \\ &= \int d(6)d(7) G_\phi(4,6) \Gamma_\phi(6,7;2) G_\phi(7,5), \end{aligned} \quad (\text{B.42})$$

we may rewrite (B.40) in the form

$$\begin{aligned} \Gamma_\phi(1,3;2) &= \delta(1,3)\delta(1,2) + \int d(4)d(5)d(7) \frac{\delta M'_\phi(1,3)}{\delta G_\phi(4,5)} \\ &\quad \times G_\phi(4,6) G_\phi(7,5) \Gamma_\phi(6,7;2), \end{aligned} \quad (\text{B.43})$$

which is the last one of the Hedin equations.

The above-derived Hedin equations, (B.35), (B.37), (B.38) and (B.43)

indeed coincide for $\phi \equiv 0$ with equations (A.13), (A.19), (A.16) and (A.18), respectively; they can as such be taken as the expressions generating the diagrammatic expansions for M' , Γ and P in terms of W , as given in (A.15), (A.17) and (A.12). If subsequently M' , Γ and P are all approximated by taking only the lowest-order expansion term in W , i.e., the zeroth-order term for P and Γ , and the first-order term for M' , we obtain the so-called bubble GW scheme, adopted in this thesis. It remains to be shown, however, whether this simplifying procedure, adopted by many authors [37,62-68], is a valid procedure indeed. Until now its "justification" is merely based on a few preliminary apparent successes in predicting energy band structures of some semiconducting crystalline materials [37,62-68].

To conclude this appendix, it is useful, in view of its application in chapter 3, to introduce the time-ordered density-density correlation function [153], (cf. (B.9)):

$$D_{\phi}(1,2) \equiv \frac{\text{H}\langle \Psi_{\text{N}} | \hat{\Lambda}_{\phi}(T_0, -T_0) \mathcal{T} \{ \hat{\rho}'_{\phi}(1) \hat{\rho}'_{\phi}(2) \} | \Psi_{\text{N}} \rangle_{\text{H}}}{\text{H}\langle \Psi_{\text{N}} | \hat{\Lambda}_{\phi}(T_0, -T_0) | \Psi_{\text{N}} \rangle_{\text{H}}}, \quad (\text{B.44})$$

where $\hat{\rho}'_{\phi}(1)$ stands for the density-deviation (Heisenberg) operator

$$\hat{\rho}'_{\phi}(1) \equiv \hat{\psi}_{\phi}^{\dagger}(1) \hat{\psi}_{\phi}(1) - \frac{\text{H}\langle \Psi_{\text{N}} | \hat{\Lambda}_{\phi}(T_0, -T_0) \hat{\psi}_{\phi}^{\dagger}(1) \hat{\psi}_{\phi}(1) | \Psi_{\text{N}} \rangle_{\text{H}}}{\text{H}\langle \Psi_{\text{N}} | \hat{\Lambda}_{\phi}(T_0, -T_0) | \Psi_{\text{N}} \rangle_{\text{H}}}. \quad (\text{B.45})$$

It can be verified, using (B.11) and (B.17) and differentiating by parts, that

$$\frac{\delta G_{\phi}(1,1^*)}{\delta \phi(2)} = \frac{1}{\hbar} D_{\phi}(1,2). \quad (\text{B.46})$$

This enables us to rewrite the first line of (B.31) in the form

$$\epsilon_{\phi}^{-1}(1,2) = \delta(1,2) - \frac{i}{\hbar} \int d(3) v(1^*,3) D_{\phi}(3,2). \quad (\text{B.47})$$

Equation (B.47) expresses the well-known simple relationship between the inverse dielectric function and the density-density correlation function. This relationship is used for $\phi \equiv 0$, in chapter 3.

REFERENCES

- [1] D.R. Hartree, Cambridge Phil. Soc., 24, 89 (1928).
- [2] V. Fock, Z.Physik, 61, 126 (1930).
- [3] J.C. Slater, Phys. Rev. 35, 210 (1930).
- [4] F. Seitz, *The Modern Theory of Solids*, McGraw-Hill, New York, 1940, pp. 227-246.
- [5] N.H. March, W.H. Young and S. Sampanthar, *The Many-body Problem in Quantum Mechanics*, Cambridge, 1967, pp. 20-22.
- [6] E. Wigner and F. Seitz, Phys. Rev. 43, 804 (1933).
- [7] J.Slater, Phys. Rev. 81, 385 (1951).
- [8] R.N. Euwema, D.L. Wilhite and G.T. Surratt, Phys. Rev. B7, 818 (1973): application to diamond.
- [9] A. Mauger and M. Lannoo, Phys. Rev. B 15, 2324 (1977): application to diamond.
- [10] R. Dovesi, C. Pisani, C. Roetti and P. Dellarole, Phys. Rev. B24, 4170 (1981): application to cubic boron nitrid.
- [11] R. Dovesi, M.Causà and G. Angonoa, Phys. Rev. B24, 4177 (1981): application to silicon.
- [12] W. von der Linden, P. Fulde and K.P. Bohnen, Phys. Rev. B34, 1063 (1986): application to silicon.
- [13] A. Svane, Phys. Rev. B35, 5496 (1987): application to diamond, silicon, germanium and alfa tin.
- [14] P. Hohenberg and W. Kohn, Phys. Rev. B136, 864 (1964).
- [15] W. Kohn and L.J. Sham, Phys. Rev. 140, A1133 (1965).
- [16] O. Gunnarsson and R.O. Jones, Phys. Scr. 21, 394 (1980).
- [17] V.L. Moruzzi, J.F. Janak and A.R. Williams, *Calculated Electronic Properties of Metals*, Pergamon Press, Oxford, 1978.
- [18] A. Zunger and A.J. Freeman, Phys. Rev. B15, 5049 (1977).
- [19] M.T. Yin and M.L. Cohen, Phys. Rev. B26, 5668 (1982).
- [20] P. Denteneer and W. van Haeringen, J. Phys. C18, 4127 (1985).
- [21] N.D. Lang, *Solid State Physics*, Academic Press, New York, 1973, Vol. 28, pp. 225-300.
- [22] G.A. Baraff and M. Schlüter, Phys. Rev. B28, 2296 (1983).

- [23] R.N. Euwema, D.L. Wilhite and G.T. Surratt, Phys. Rev. B7, 818 (1973).
- [24] A.B. Kunz, Phys. Rev. B12, 5890 (1975).
- [25] A. Mauger and M. Lannoo, Phys. Rev. B15, 2324 (1977).
- [26] R. Dovesi, C. Pisani, C. Roetti and P. Dellarole, Phys. Rev. B24, 4170 (1981).
- [27] J.P. Perdew and A. Zunger, Phys. Rev. B23, 5048 (1981).
- [28] R.A. Heaten, J.G. Harrison and C.C. Lin, Solid State Commun. 41, 827 (1982).
- [29] J.P. Perdew and M.R. Norman, Phys. Rev. B26, 5445 (1982).
- [30] J.A. Alonso and L.A. Girifalco, Phys. Rev. B17, 3735 (1978).
- [31] O. Gunnarsson, M. Jonson and B.I. Lundqvist, Phys. Rev. B20, 3136 (1979).
- [32] G.P. Kerker, Phys. Rev. B24, 3468 (1981).
- [33] C.S. Wang and W.E. Pickett, Phys. Rev. Lett., 51, 597 (1983).
- [34] J. P. Perdew and M. Levy, Phys. Rev. Lett. 51, 1884 (1983).
- [35] L.J. Sham and M. Schlüter, Phys. Rev. Lett. 51, 1888 (1983).
- [36] L.J. Sham, Phys. Rev. B32, 3876 (1985).
- [37] W. von der Linden and P. Horsch, Phys. Rev. B37, 8351 (1988).
- [38] O. Gunnarsson and K. Schönhammer, Phys. Rev. Lett. 56, 1968 (1986).
- [39] O. Gunnarsson and K. Schönhammer, Phys. Rev. Lett. 60, 1582 (1988).
- [40] F. Manghi, G. Riegler, C.M. Baroni, C. Calandra and G.B. Bachelet, Phys. Rev. B28, 6157 (1983).
- [41] H. Lehmann, Nuvo Cimento 11, 342 (1954).
- [42] G. Green, *An Essay on the Application of Mathematical Analysis to the Theories of Electricity and Magnetism*, Nottingham, 1828.
- [43] G.C. Wick, Phys. Rev. 80, 268 (1950).
- [44] R.P. Feynman, Phys. Rev. 76, 749 (1949); 76, 769 (1949).
- [45] F.J. Dyson, Phys. Rev. 75, 486 (1949); 75, 1736 (1949).
- [46] R.P. Feynman, Rev. Modern Phys. 20, 367 (1948).
- [47] P.A.M. Dirac, Physik. Zeits. Sowjetunion 3, 64 (1933).
- [48] J. Schwinger, Phys. Rev. 82, 914 (1951).
- [49] W. Yourgrau and S. Mandelstam, *Variational Principles in Dynamics and Quantum Theory*, Dover Publication, Inc., New York, 1968, pp. 127-141.

- [50] J. Schwinger, Proc. Natl. Acad. Sci. 37, 452 (1951).
- [51] J. Schwinger, Proc. Natl. Acad. Sci. 37, 455 (1951).
- [52] T. Matsubara, Progr. Theor. Phys. 14, 351 (1955).
- [53] J.A. Anderson, Phys. Rev. 94, 703 (1954).
- [54] T. Kato, T. Kobayashi and M. Namiki, Progr. Theor. Phys. Suppl. 15, 3 (1960).
- [55] P.C. Martin and J. Schwinger, Phys. Rev. 115, 1342 (1959).
- [56] L. Hedin, Phys. Rev. 139, A796 (1965).
- [57] J. Hubbard, Proc. Roy. Soc. A240, 539 (1957).
- [58] P.M. Morse and H. Feshbach, *Methods of Theoretical Physics* McGraw-Hill, New York 1953, p. 884-866.
- [59] A.J. Layzer, Phys. Rev. 129, 897 (1963).
- [60] L.J. Sham and W. Kohn, Phys. Rev. 145, 561 (1966).
- [61] W. van Haeringen, B. Farid and D. Lenstra, Physica Scr. T19, 282 (1987).
- [62] G. Strinati, H.J. Mattausch and W. Hanke, Phys. Rev. B25, 2867 (1982).
- [63] C.S. Wang and W.E. Pickett, Phys. Rev. B30, 4719 (1984).
- [64] M.S. Hybertsen and S.G. Louie, Phys. Rev. Lett. 55, 1418 (1985).
- [65] M.S. Hybertsen and S.G. Louie, Phys. Rev. B34, 5390 (1986).
- [66] R.W. Godby, M. Schlüter and L.J. Sham, Phys. Rev. Lett. 56, 2415 (1986).
- [67] R.W. Godby, M. Schlüter and L.J. Sham, Phys. Rev. B35, 4170 (1987).
- [68] R.W. Godby, M. Schlüter and L.J. Sham, Phys. Rev. B37, 10159 (1988).
- [69] B. Farid, R. Daling, D. Lenstra and W. van Haeringen, Phys. Rev. B38, 7530 (1988).
- [70] B. Farid, D. Lenstra and W. van Haeringen, Solid State Commun. 67, 7 (1988); also section 4.5 of this thesis.
- [71] A.L. Fetter and J.D. Walecka, *Quantum Theory of Many-Particle Systems*, McGraw-Hill, New York (1971), pp. 19-21, 53-59.
- [72] A.L. Fetter and J.D. Walecka, Ref. 71, Ch. 3.
- [73] D.A. Kirzhnits, *Field Theoretical Methods in Many-Body Systems*, Pergamon Press, London, 1967, pp. 184-192.
- [74] F. Seitz, Ref. 4, p. 235.
- [75] F. Seitz, Ref. 4, p. 335.

- [76] L. Hedin and S. Lundqvist, *Solid State Physics*, Academic Press, New York, 1969, Vol. 23, p. 37.
- [77] L. Hedin and S. Lundqvist, Ref. 76, pp. 25-26.
- [78] A.L. Fetter and J.D. Walecka, Ref. 71, pp. 72-79.
- [79] L.J. Sham and M. Schlüter, *Phys. Rev.* B32, 3883 (1985).
- [80] G. Goertzel and N. Tralli, *Some Mathematical Methods of Physics*, McGraw-Hill, New York, 1960, Chs. 4 and 5.
- [81] E. Kreyszig, *Introductory Functional Analysis with Applications*, John Wiley & Sons, New York, 1978, p. 444.
- [82] A. Messiah, *Quantum Mechanics*, vol. II, North Holland Publishing Company, Amsterdam, 1965, pp. 665-672.
- [83] J. Callaway, *Energy Band Theory*, Academic Press, New York, 1964, pp. 52-54.
- [84] G.W. Pratt, Jr., *Phys. Rev.* 118, 462 (1960).
- [85] W. Kohn and P. Vashishta, *Theory of Inhomogeneous Electron Gas*, Plenum Press, New York, 1983, pp. 79-147.
- [86] D.R. Hamann, *Phys. Rev. Lett.* 42, 662 (1979).
- [87] F. Herman, R.L. Kortum, C.D. Kuglin and R.A. Short, *Quantum Theory of Atoms, Molecules, and the Solid State*, Academic Press, New York, 1966, p. 381.
- [88] R. Heaten and E. Lafon, *Phys. Rev.* B17, 1958 (1978).
- [89] R.C. Chaney, C.C. Lin and E.E. Lafon, *Phys. Rev.* B3, 459 (1971).
- [90] D.J. Stukel and R.N. Euwema, *Phys. Rev.* B1, 1635 (1970).
- [91] L. Hedin and S. Lundqvist, Ref. 76, pp. 80-82.
- [92] S. Baroni and R. Resta, *Phys. Rev.* B33, 7017 (1986).
- [93] W. von der Linden, P. Horsch and W.D. Lukas, *Solid State Commun.* 59, 485 (1986).
- [94] D. Pines and D. Bohm, *Phys. Rev.* 85, 338 (1952).
- [95] H. Ehrenreich and M. H. Cohen, *Phys. Rev.* 115, 786 (1959).
- [96] J.M. Ziman, *Elements of Advanced Quantum Theory*, Cambridge University Press, London, 1980, pp. 158-160.
- [97] A. J. Glick and R.A. Ferrell, *Annals of Physics*: 11, 359 (1960).
- [98] K.S. Singwi, M.P. Tosi, R.H. Land and A. Sjölander, *Phys. Rev.* 176, 589 (1968).

- [99] K.S. Singwi, A. Sjölander, M.P. Tosi and R.H. Land, Phys. Rev. B1, 1044 (1970).
- [100] F. Brosens and J.T. Devreese, Phys. Stat. Sol.(b) 147, 173 (1988).
- [101] N. Wiser, Phys. Rev. 129, 62 (1963).
- [102] H.R. Philipp and H. Ehrenreich, Phys. Rev. 129, 1550 (1963).
- [103] R.A. Roberts and W.C. Walker, Phys. Rev. 161, 730 (1967).
- [104] J.A. van Vechten and R.M. Martin, Phys. Rev. Lett. 28, 446 (1972).
- [105] J.P. Walter and M.L. Cohen, Phys. Rev. B5, 3101 (1972).
- [106] S. Louie, J.R. Chelikowsky and M.L. Cohen, Phys. Rev. Lett. 34, 155 (1975).
- [107] L. van Hove, Phys. Rev. 84, 1189 (1953).
- [108] J.C. Phillips, Phys. Rev. 104, 1263 (1956).
- [109] A.A. Maradudin, E.W. Montroll and G.H. Weiss, Solid State Physics, Supplement 3, Academic Press, New York, 1971, pp. 150-158.
- [110] G. Lehmann and M. Taut, Phys. Stat. Sol.(b) 54, 469 (1972).
- [111] J. Rath and A.J. Freeman, Phys. Rev. B11, 2109 (1975).
- [112] A. Baldereschi, Phys. Rev. B7, 5212 (1973).
- [113] D.J. Chadi and M.L. Cohen, Phys. Rev. B8, 5747 (1973).
- [114] H.J. Monkhorst and J.D. Pack, Phys. Rev. B13, 5188 (1976).
- [115] A. Baldereschi and E. Tosatti, Phys. Rev. B17, 4710 (1978).
- [116] M.S. Methfessel, M.H. Boon and F.M. Muller, J. Phys. C16, L949 (1983).
- [117] M.S. Methfessel, M.H. Boon and F.M. Muller, J. Phys. C20, 1069 (1987).
- [118] M.H. Boon, M.S. Methfessel, F.M. Muller, To be published in J. Phys. C.
- [119] Also, the analytic linear tetrahedron method shows poor convergence, in case of the here considered two-dimensional integral. This can be attributed to the fact that only the first two terms, i.e zero- and first-order terms, of the Taylor series expansions of $\epsilon_{\lambda}(\mathbf{k}')$ around the vertices of the tetrahedrons are taken into account. Therefore, the method is not capable of describing the behavior of the integrand in the direct vicinity of the van Hove singularities. It has been observed [116-118] that a better convergence can be achieved if one takes also into account the quadratic term in the above-mentioned Taylor series expansion. We will

not, however, deal with such a refined method. We note incidentally that in all methods based on Taylor series expansions of the energy eigenvalues, one systematically neglects the possibility of occurrence of non-analytic behavior [107-109] in certain points in 1Bz. For instance, at all points where one or more components of $\nabla_{\mathbf{k}'}\epsilon_{\mathbf{k}'}(\mathbf{k}')$ change sign discontinuously, the Taylor expansion of the energy eigenvalue $\epsilon_{\mathbf{k}'}(\mathbf{k}')$ is prohibited.

- [120] R.M. Pick, M.H. Cohen and R.M. Martin, Phys. Rev. B1, 910 (1970).
- [121] J.J.Quinn and R.A Ferrell, Phys. Rev. 112, 812 (1958).
- [122] K. Knopp, *The Theory of Functions*, Part I, Dover Publications, New York, 1945, pp. 92-111.
- [123] K. Knopp, *The Theory of Functions*, Part II, Dover Publications, New York, 1947, pp. 93-118.
- [124] L. Hedin and S. Lundqvist, Ref. 76, p. 33.
- [125] L. Hedin and S. Lundqvist, Ref. 76, p.180.
- [126] G. Rickayzen, *Green's Functions and Condensed Matter*, Academic Press, London, 1980, pp. 15-18.
- [127] N.H. March, W.H. Young and S. Sampanthar, Ref. 5, pp. 367-385.
- [128] W.M. Saslow and G.F. Reiter, Phys. Rev. B7, 2995 (1973).
- [129] J.M. Ziman, Ref. 96, pp. 112-114, 150-154.
- [130] J.M. Ziman, Ref. 96, P. 107.
- [131] K. Knopp, Ref. 122, pp. 129-131.
- [132] The necessary conditions to be satisfied by \tilde{W} are the so-called Hölder condition [133,134],

$$|\tilde{W}(\mathbf{r}_1, \mathbf{r}_2; i\epsilon_1) - \tilde{W}(\mathbf{r}_1, \mathbf{r}_2; i\epsilon_2)| < A |\epsilon_1 - \epsilon_2|^\gamma, \quad A > 0, \quad \gamma > 0,$$

and the asymptotic relation

$$\tilde{W}(\mathbf{r}_1, \mathbf{r}_2; \epsilon') = \mathcal{O}(|\epsilon'|^{-\xi}), \quad \xi > 0 \text{ for } |\epsilon'| \rightarrow \infty.$$

Since \tilde{W} is analytic along the imaginary axis, Hölder condition is automatically fulfilled.

- [133] B. Davies, *Integral Transforms and their Applications*, Springer-Verlag, New York, 1985, pp. 313-321, 340.
- [134] B. Noble, *Methods based on the Wiener-Hopf Technique for the Solution of Partial Differential Equations*, Pergamon Press, London 1958, pp. 11-13.

- [135] B. Noble, Ref. 134, pp. 141-147.
- [136] E.C. Titchmarsh, *The Theory of Functions*, Oxford University Press, London, Second Edition (reprinted and corrected), 1985, p. 157.
- [137] It can be shown that the probability of transfer of momentum $\hbar(\mathbf{k}+\mathbf{G})$ and energy ϵ by an electron to the system per unit time is proportional to $\tilde{W}_{\mathbf{G},\mathbf{G}}^*(\mathbf{k};\epsilon)$, [128,138]; therefore owing to the above considerations we conclude that in the energy region $0 < \epsilon < E_g$ no energy and momentum transfer to our semiconducting system is possible.
- [138] D. Pines and P. Nozière, *The theory of Quantum Liquids*, Vol. I: *Normal Fermi Liquids*, W.A. Benjamin, Inc., New York, 1966, pp. 85-90.
- [139] J.J. Quinn, Phys. Rev. 126, 1453 (1962).
- [140] E.C. Titchmarsh, Ref. 136, pp. 155-157.
- [141] M.S. Hybertsen and S.G. Louie, Phys. Rev. B37, 2733 (1988).
- [142] W. Brinkman and B. Goodman, Phys. Rev. 149, 597 (1966). In this work no account has been taken of local-field effects. As a result no improvements have been observed with regard to the X_α band structures.
- [143] L. Hedin and S. Lundqvist, Ref. 76, pp. 40-41.
- [144] P. Nozières, *Theory of Interacting Fermi Systems*, W.A. Benjamin, Inc., New York, 1964, pp. 45-50.
- [145] D.L. Johnson, Phys. Rev. B9, 4475 (1974).
- [146] W. Schülke, J. Phys. C: Solid State Phys. 16, 5629 (1983).
- [147] M Taut, J. Phys. C: Solid State Phys. 18, 2677 (1985).
- [148] L.P. Kadanoff and P.C. Martin, Phys. Rev. 124, 670 (1961).
- [149] L. Hedin and S. Lundqvist, Ref. 76, p. 26.
- [150] S.L. Adler, Phys. Rev. 126, 413 (1962).
- [151] D. Pines and P. Nozières, Ref. 138, pp. 101-105.
- [152] E.T. Whittaker and G.N. Watson, *A Course of Modern Analysis*, Cambridge, 1958 (Forth Edition), pp. 211-231.
- [153] L. Hedin and S. Lundqvist, Ref. 76, p. 180.
- [154] J.M Ziman, Ref. 96, p. 161.
- [155] As a general rule we may state that the failure of the series expansion to converge in some energy region may be thought of as the first indication of new types of excitations (in our case these are plasmons), brought

- about by interaction, in that energy region [156].
- [156] J.C. Inkson, *Many-Body Theory of Solids*, Plenum Press, New York, 1984, pp. 36-38, 289-290.
 - [157] D.A. Kirzhnitz, *Usp. Fiz. Nauk* 119, 357 (1976) [*Sov. Phys.-Usp.* 19, 530 (1976)].
 - [158] O.V. Dolgov, D.A. Kirzhnitz, E.G. Maksimov, *Rev. Mod. Phys.* 53, 81 (1981).
 - [159] J.C. Slater, *Quantum Theory of Molecules and Solids*, Vol. 2, McGraw-Hill, New York, 1965, pp. 3-14.
 - [160] P.J. H. Denteneer, Ph.D Thesis, Eindhoven University of Technology, 1987, pp. 71-73.
 - [161] N.W. Ashcroft and N.D. Mermin, *Solid State Physics*, Holt-Sanders, Ltd., Japan, 1976, pp. 111-129.
 - [162] J. Callaway, *Ref.* 83, pp. 11-28.
 - [163] P.J. H. Denteneer, *Ref.* 161, p. 52.
 - [164] G. Gilat, *J. Comput. Phys.* 10, 432 (1972).
 - [165] G. Gilat and L.J. Raubenheimer, *Phys. Rev.* 144, 390 (1966).
 - [166] N.W. Dalton, *Solid State Commun.* 8, 1047 (1970).
 - [167] L.J. Raubenheimer and G. Gilat, *Phys. Rev.* 157, 586 (1967).
 - [168] Z. Kam and G. Gilat, *Phys. Rev.* 175, 1156 (1968).
 - [169] F. Ayres, Jr., *Theory and Problems of Matrices*, Schaum Publishing Co., New York, 1962, pp. 56-58.
 - [170] J.C. Slater and G.F. Koster, *Phys. Rev.* 94, 1498 (1954).
 - [171] C. Salustri, *Computer Physics Commun.* 30, 271 (1983). We mention that the corresponding computer program, named BAPAR with Catalogue number ACKU, obtainable from CPC Program Library, Queen's University of Belfast, N. Ireland, contains a syntactical error, namely that all numbers 0.5 are typed as "1/2" rather than "1/2.0". As may be well-known "1/2" according to the standard FORTRAN protocol is equal to zero.
 - [172] M.F.H. Schuurmans, W. van Haeringen and H.G. Junginger, *Solid State Commun.* 10, 549 (1972).
 - [173] G.B. Bachelet, D.R. Hamann and M. Schlüter, *Phys. Rev.* B26, 4199 (1982).

- [174] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, Springer-Verlag, New York, 1980. Ch. 6.
- [175] J. Stoer and R. Bulirsch, Ref. 175, Ch. 4.
- [176] A.L. Fetter and J.D. Walecka, Ref. 71, pp. 92-111.
- [177] D.A. Kirzhnits, Ref. 73, Ch. 3.
- [178] N. H. March, W.H. Young and S. Sampanthar, Ref. 5, Ch. 4 and Appendix 10A.
- [179] P. Nozières, Ref. 144, Ch. 5.
- [180] J. Goldstone, Proc. Roy. Soc. A239, 267 (1957).
- [181] L. Hedin and S. Lundqvist, Ref. 76, p. 36.
- [182] P. Nozières, Ref. 144, p. 221.
- [183] J.M. Luttinger and J.C. Ward, Phys. Rev. 118, 1417 (1960).
- [184] J. Hubbard, Proc. Roy. Soc. A243, 336 (1958).
- [185] N.N. Bogoliubov and D.V. Shirkov, *Introduction to the Theory of Quantized Fields*, Interscience Publishers, Inc., New York (1959), p. 202.
- [186] R.D. Mattuck, *A Guide to Feynman Diagrams in the Many-Body Problem*, McGraw-Hill, Ltd., New York, 1967, pp. 158-164.
- [187] R.G. Parr and L.J. Bartolotti, J. Phys. Chem. 87, 2810 (1983).
- [188] G. Baym and L.P. Kadanoff, Phys. Rev. 124, 287 (1961).
- [189] P. Nozières, Ref. 144, Ch. 2.
- [190] L. Hedin and S. Lundqvist, Ref. 76, pp. 12-19.

SUMMARY

The central theme in this thesis is the *ab initio* calculation of electron energies in a semiconductor. The fact that the commonly used effective-potential methods, such as the Hartree, Hartree-Fock and Kohn-Sham (local) density functional methods invariably lead to wrong predictions, indicates that there is a serious problem. Indeed, these methods do not seem to take into account the effect of electron-electron interaction in a proper way. Therefore, we return to the rigorous theory and reconsider the incorporation of the mutual electron interaction.

Our framework is the theory of Green functions. In this theory the Fourier transform with respect to time of the one-particle Green function G plays an essential role. It can formally be shown that the excitation energies of a system show up as the singularities of this function in the energy domain. We introduce a representation of G in terms of wave functions which satisfy Schrödinger-like wave equations and thus can be thought of as representing particle-like entities. These entities are referred to as quasi-particles and their "energies", being the simple poles in the above-mentioned representation, are given by the corresponding eigenvalues in the equations.

A crucial role in our considerations is played by the self-energy function M , which relates the one-particle Green function of an unperturbed system to that of the interacting system through the Dyson equation and which acts as an energy-dependent non-local potential in the above-mentioned quasi-particle wave equations. The non-Hermiticity of this function gives rise to complex-valued quasi-particle "energies", that is, to quasi-particles with finite lifetimes.

Each effective-potential method can uniquely be characterized in terms of its corresponding approximated Hermitian self-energy function. In this thesis we concentrate on a more general non-Hermitian approximation of M , which is called the GW self-energy function. In view of its recent successes in energy band calculations of semiconductors, the study of the GW scheme takes a central position in this thesis. We have, among other things, thoroughly investigated the analytic behavior of the GW self-energy function in the complex energy plane, enabling us to justify a Taylor-series expansion of M , which will very likely facilitate its numerical evaluation. Moreover, a number of further

approximations within the framework of GW are discussed which may be of practical relevance in actual calculations.

As the name suggests, the self-energy function in the GW scheme asks for the determination of both the one-particle Green function G of the interacting system and the dynamically screened Coulomb interaction function W . The latter function, in contrast to the bare Coulomb interaction, takes account of the polarization effects in the interacting system. These effects are described by means of the polarization function P whose study in this thesis is restricted to the simplest approximation, namely the so-called bubble approximation. In this approximation one assumes no interaction between the members of excited electron-hole pairs. Besides the presentation of a number of general expressions and properties concerning both the bubble polarization function and the exact screened interaction function, two new methods for the determination of P are introduced and numerically tested. These methods substantially facilitate the numerical evaluation of P and W at real energies.

It is well-known that the bare Coulomb interaction, due to its long range, is responsible for many inconveniences in the treatment of interacting systems. In our case, the bare Coulomb interaction gives rise to singular behavior in the wave-vector dependence of a number of plane-wave matrix elements of the dynamically screened interaction matrix. As the evaluation of M involves a Brillouin zone integration of an expression containing these matrix elements, a number of precautionary measures must be taken in order to make numerical integration possible. Based on a thorough analysis of these matrix elements, we propose a decomposition of the expression for the self-energy function in such a way that the above-mentioned Brillouin zone integration can be performed numerically, either by a special-point method or by some more common numerical method.

Finally, we give an analysis of the numerical feasibility of the proposed calculation strategy. We conclude that an evaluation of M and the related quasi-particle spectrum within a GW scheme, in which no further approximations are carried through, is very time consuming. There will be no problem in this evaluation if we restrict ourselves to the first iteration step of the required self-consistency procedure, at least if the number of plane waves in which the quasi-particle wave functions are expanded is not too large. Subsequent iteration steps can only be carried out if it turns out that M depends smoothly on its wave-

vector argument. This conclusion is based on the computing power of an ordinary computing system without facilities for vector and parallel computing. The prospects are much better if such facilities can be used. It is a lucky circumstance in this connection that recent studies on *simplified* GW schemes seem to indicate that there is no need to go beyond the first iteration step. In view of all this, we conclude that an *ab initio* (GW) calculation of electron energies in a semiconductor is indeed within reach.

SAMENVATTING

Het centrale thema in dit proefschrift is het model-vrij berekenen van elektron-energieën in halfgeleiders. Het feit dat de gangbare effectieve-potentiaal methoden, zoals die volgens Hartree en Hartree-Fock, alsmede de methode van de (locale) dichtheidsfunctionaal volgens Kohn en Sham, zonder uitzondering tot foutieve voorspellingen leiden, laat zien dat hier sprake is van een ernstig probleem. Inderdaad blijkt geen van deze methoden het effect van de wisselwerking tussen elektronen op behoorlijke wijze in rekening te brengen. Derhalve keren wij terug naar de rigoureuze theorie ten einde het in rekening brengen van deze wisselwerking opnieuw te bezien.

Ons raamwerk is de theorie van de Green-functies. In deze theorie speelt de naar de tijd Fourier-getransformeerde ééndeeltje Greense functie G een essentiële rol. Het kan formeel worden aangetoond dat de excitatie-energieën van een systeem zich manifesteren als singulariteiten van deze functie in het energie-domein. Wij introduceren een representatie van G in termen van golf functies die aan vergelijkingen van het Schrödinger-type voldoen en derhalve deeltjesachtige entiteiten geacht kunnen worden te representeren. Naar deze entiteiten wordt verwezen als quasi-deeltjes, en hun "energieën", zijnde de enkelvoudige polen van de bovengenoemde representatie, zijn de betreffende eigenwaarden van de vergelijkingen.

Een cruciale rol in onze beschouwingen wordt gespeeld door de zelfenergie-functie M die de ééndeeltje Green-functie van een ongestoord systeem met die van het wisselwerkend systeem, via de vergelijking van Dyson, aan elkaar relateert en als een energie-afhankelijke, niet-lokale potentiaal in de bovengenoemde quasideeltjes-golfvergelijkingen voorkomt. De niet-Hermiticiteit van deze functie geeft aanleiding tot complexwaardige "energieën", dat wil zeggen, tot quasideeltjes met eindige levensduur.

Elke effectieve-potentiaal methode kan op unieke wijze worden gekarakteriseerd in termen van een daaraan gerelateerde Hermitische zelfenergie-functie. In dit proefschrift concentreren wij ons op een meer algemene, niet-Hermitische benadering van M , die de GW-zelfenergie-functie wordt genoemd. In het licht van recente successen die zijn behaald in de bandstructuurberekeningen van halfgeleiders, krijgt de benadering van het GW-schema een centrale plaats in

dit proefschrift. Wij hebben ondermeer een grondige studie gemaakt van het analytische gedrag van de GW-zelfenergie in het complexe energie-vlak, waardoor het nu mogelijk is om een Taylor-ontwikkeling van M te rechtvaardigen, die wellicht de berekening van M aanzienlijk zal vergemakelijken. Tevens worden een aantal verdere benaderingen binnen het raamwerk van GW besproken die van praktisch nut kunnen zijn in een daadwerkelijke berekening.

Zoals de naam suggereert, vraagt de zelfenergie-functie in het GW-schema om de berekening van zowel de ééndeeltje Green-functie G van het wisselwerkende systeem als van de dynamisch afgeschermd Coulomb-interactie W . De laatstgenoemde functie houdt, in tegenstelling tot de niet-afgeschermd Coulomb-interactie, rekening met de polarisatie-effecten in het wisselwerkende systeem. Deze effecten worden beschreven door middel van de polarisatie-functie P , waarvan de studie in dit proefschrift wordt beperkt tot de meest eenvoudige benadering, de zogenaamde "bubble"-benadering. In deze benadering veronderstelt men dat bij elektron-gat paar-creatie geen wisselwerking optreedt tussen het elektron en het gat. Naast de presentatie van een aantal algemene uitdrukkingen en eigenschappen van zowel de "bubble"-polarisatie-functie als van de exacte afgeschermd interactie-functie, worden twee nieuwe methoden ter berekening van P geïntroduceerd en op numerieke wijze getoetst. Deze methoden vergemakkelijken het berekenen van P en W voor reële energieën.

Het is bekend dat de naakte Coulomb-interactie, vanwege haar lange dracht, verantwoordelijk is voor veel ongemak in de behandeling van wisselwerkende systemen. In ons geval geeft de naakte Coulomb-interactie aanleiding tot singulier gedrag in de golfvector-afhankelijkheid van een aantal vlakke-golf-matrixelementen van de dynamisch afgeschermd interactie. Daar bij de berekening van M een Brillouin-zone-integratie moet worden uitgevoerd van een uitdrukking die deze matrixelementen bevat, dient een aantal voorzorgsmaatregelen genomen te worden om numerieke integratie mogelijk te maken. Uitgaande van een grondige analyse van deze matrixelementen, stellen wij een zodanige decompositie van de zelfenergie-uitdrukking voor, dat de bovengenoemde Brillouin-zone-integratie toch op numerieke wijze uitgevoerd kan worden, met behulp van hetzij een speciale-punten methode, hetzij een meer gewone methode.

Tot slot geven wij een analyse van de numerieke haalbaarheid van de

voorgestelde rekenkundige strategie. Geconcludeerd wordt dat een berekening van M en het daaraan gerelateerde quasideeltjes-spectrum binnen het GW-raamwerk zonder verdere benaderingen zeer tijdrovend is. Er zullen zich geen problemen voordoen zolang we ons beperken tot de eerste iteratiestap van de vereiste zelfconsistente-procedure, ten minste, als het aantal vlakke golven waarin de quasideeltjes-golffuncties worden ontwikkeld niet te groot is. De iteratiestappen die daarna volgen, kunnen alleen worden uitgevoerd als zal blijken dat M op voldoende vloeiende wijze van haar golfvectorargument afhangt. Deze conclusie is gebaseerd op het rekenvermogen van een normale computer die niet geschikt is voor vectoriële of parallele berekeningen. De vooruitzichten worden beter als dergelijke opties wel aanwezig zijn. Het is een gelukkige omstandigheid dat recente studies in het kader van vereenvoudigde GW-schemas erop lijken te wijzen dat het niet nodig is verder te gaan dan de eerste iteratiestap. In het licht hiervan concluderen wij dat een modelvrije GW-berekening van elektron-energieën in een halfgeleider wel degelijk binnen bereik ligt.

CURRICULUM VITAE

- jan. '60 : geboren te Teheran (Iran)
- sep. '65 – juni '71 : lagere school te Teheran
- sep. '71 – juni '77 : middelbare school te Teheran (VWO, exacte richting)
- sep. '77 : ingeschreven als student aan de Universiteit van Teheran, Faculteit der Wis en Natuurkunde; studie onmogelijk wegens politieke situatie
- mei '79 : aankomst in Nederland
- sep. '79 – feb. '85 : studie Elektrotechniek aan de Technische Universiteit Delft, Faculteit der Elektrotechniek
1982 kandidaatsexamen
1985 doctoraalexamen (cum laude)
Afstudeerrichting: Theoretische Elektriciteitsleer
Afstudeerhoogleraren: Prof. dr. ir. P.M. van den Berg en Prof. dr. ir. A.T. De Hoop
Titel afstudeerscriptie: Diffraction of an Electromagnetic Plane Wave by a Perfectly Reflecting (Rough) Surface
- maart '85 – maart '89 : promotie-onderzoek bij Prof. dr. W. van Haeringen, vakgroep Theoretische Natuurkunde, Faculteit der Technische Natuurkunde, Technische Universiteit Eindhoven
Titel proefschrift: Towards *ab initio* Calculation of Electron Energies in Semiconductors

STELLINGEN

behorende bij het proefschrift van

Behnam Farid

Eindhoven, 21 maart 1989

1. Het door Hubbard aangegeven voorschrift om uitgaande van een expressie behorende bij een polarisatie-diagram de bijdrage van het daaraan geassocieerde "exchange conjugate" diagram te bepalen, schendt de wet van behoud van impuls. Voorts is onduidelijk op welke manier hij de "gemiddelde interactie" voor processen rondom het Fermi oppervlak uitrekt.

J. Hubbard, Proc. Roy. Soc. A243, 336 (1958).

2. Semi-empirische clusterberekeningen [1,2] leveren geen wezenlijke bijdrage tot de studie van defecten in halfgeleiders.

[1] P. Deák, L.C. Snyder, R. Singh and J. Corbett, Phys. Rev. B36, 9612 (1987).

[2] P. Deák and L.C. Snyder, Phys. Rev. B36, 9619 (1987).

3. De door Godby, Schlüter en Sham [1] afgeleide uitdrukking voor de zelf-energie-functie is onvolledig [2]. De invloed van de weggelaten term op de bandstructuur van quasi-deeltjes is onduidelijk en dient nader onderzocht te worden.

[1] R.W. Godby, M. Schlüter and L.J. Sham, Phys. Rev. B37, 10159 (1988).

[2] Dit proefschrift, hoofdstuk 3.

4. De door Mattuck aangegeven diagrammatische ontwikkeling ter bepaling van de zelf-energiefunctie behorend bij een systeem wisselwerkende deeltjes bevat een principiële onjuistheid met als gevolg dat er te veel diagrammen in rekening worden gebracht.

R. D. Mattuck, *A Guide to Feynman Diagrams in the Many-Body Problem*, McGraw-Hill, New York, 1967, pp. 170-172.

5. De veelgebruikte benaming "Random Phase Approximation", afgekort RPA, van Pines en Bohm [1,2], die een bepaalde benadering van de polarisatie-functie van uniforme elektron-gassen voorstelt, is misleidend als het om kristallen gaat. Het is dan beter om in plaats van de bovengenoemde benaming de aanduiding "Bubble Approximation" te gebruiken.

[1] D. Pines and D. Bohm, Phys. Rev. 85, 338 (1952).

[2] J.M. Ziman, *Elements of Advanced Quantum Theory*, Cambridge University Press, Cambridge, 1969, pp. 158-162.

6. De door Gygi en Baldereschi [1] voorgestelde methode ter bepaling van de zelf-energiefunctie met behulp van "speciale punten" doet geen recht aan het anisotrope karakter van bepaalde integranden in integralen over de Brillouin-zone. Het is goed mogelijk de methode in dit opzicht aan te passen [2].

[1] E. Gygi and A. Baldereschi, *Phys. Rev.* B34, 4405 (1986).

[2] Dit proefschrift, hoofdstukken 4 en 5.

7. De door Pratt voorgestelde methode om tot de juiste bandstructuur van halfgeleiders te komen met behulp van een effectieve potentiaal, berust op een onjuiste veronderstelling.

G.W. Pratt, Jr., *Phys. Rev.* 118, 462 (1960).

8. De bevindingen van Hawrylak, Eliasson en Quinn aangaande het kwalitatieve gedrag van de zelf-energie, de effectieve massa en de levensduur van quasi-deeltjes in een gelaagde structuur zijn pas dan interessant als kan worden aangetoond dat deze bevindingen niet een gevolg zijn van de door hen veronderstelde, maar niet gerechtvaardigde, wisselwerking tussen de deeltjes.

P. Hawrylak, G. Eliasson and J.J. Quinn, *Phys. Rev.* B37, 10187 (1987).

9. Beschouw een in één richting (y-richting) oneindig uitgestrekte starre plaat ter breedte $2a$ (in de x-richting), trillend met frequentie ω en amplitude v_{p1} op een lokaal reagerende, instantane, tijdinvariante, lineaire, isotrope, homogene vloeistof met massadichtheid ρ en compressiemodulus κ . Het drukprofiel onder de plaat laat zich dan verrassend nauwkeurig beschrijven door $p(x; \omega) = C(ka) i \omega \rho a v_{p1} (1 - x^2/a^2)^{1/2}$ waarin $C(ka)$ een bekende functie is van ka , $k = \omega \sqrt{\rho/\kappa}$ en i de imaginaire eenheid voorstelt, als tenminste voldaan is aan de eis $ka < 1$.

10. Het door W. Oevel gesuggereerde verband tussen hoekvariabelen en zogenaamde "master symmetries" is in de meeste van zijn voorbeelden niet aanwezig.

W. Oevel, *Topics in Soliton Theory and Exactly Solvable Nonlinear Equations*, World Scientific Publ., Singapore, 1987, pp. 108-124.

11. In tegenstelling tot wat algemeen werd aangenomen [1], kan de dynamische dieëlectrische functie van een halfgeleider met success worden berekend door gebruik te maken van de "speciale punten"-methode [2].

[1] A. Baldereschi and E. Tosatti, *Phys. Rev. B* **17**, 4710 (1978).
[2] Dit proefschrift, hoofdstuk 4.

12. De uitdrukking voor het verschil tussen de exacte bandsprong in een halfgeleider en het in de LDA methode daarvoor te verkrijgen resultaat kan eenvoudig worden bepaald door gebruik te maken van Brillouin-Wigner of Rayleigh-Schrödinger storingstheorie [1,2]. Dit maakt (een deel van) een aantal bestaande, dikwijls langdradige, artikelen [3-5] in feite overbodig.

[1] N.H. March, W.H. Young and S. Sampanthar, *The Many-Body Problem in Quantum Mechanics*, Cambridge University Press, Cambridge, 1967, pp.16-19.
[2] W. von der Linden, and P. Horsch, *Phys. Rev. B* **37**, 8351 (1988).
[3] L.J. Sham, *Phys. Rev. B* **32**, 3876 (1985).
[4] W. Hanke, N. Meskini and H. Weiler, *Electronic Structure, Dynamics, and Quantum Structural Properties of Condensed Matter*, Plenum Press, New York, 1985, pp. 113-156.
[5] B. Farid, D. Lenstra and W. van Haeringen, *Derivation of an Expression for the Energy Gap in a Semiconductor*, Technische Universiteit Eindhoven, Faculteit der Technische Natuurkunde, Internal Report 1986-4.

13. Door een deel van de opvoeding van kinderen door grootouders te laten verzorgen, wordt niet alleen geprofiteerd van een opvoedkundig voordeel, maar worden tevens de gelijke kansen op maatschappelijke ontplooiing van man en vrouw bevorderd.
14. Het "sport"-onderdeel boksen dient uit de lijst van officiële sporten verwijderd te worden.