Renormalization of electronic properties of semiconductors from electron-phonon interaction

A Dissertation presented

by

Juan Pablo Nery

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Physics

Stony Brook University

May 2018

Stony Brook University

The Graduate School

Juan Pablo Nery

We, the dissertation committe for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation

Philip B. Allen - Dissertation Advisor
Professor Emeritus, Department of Physics and Astronomy

Matthew Dawber - Chairperson of Defense
Associate Professor, Department of Physics and Astronomy

Michael Zingale
Associate Professor, Department of Physics and Astronomy

Sergey Suchalkin
Assistant Professor, Department of Electrical and Computer Engineering

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

Renormalization of electronic properties of semiconductors from electron phonon interaction

by

Juan Pablo Nery

Doctor of Philosophy

in

Physics

Stony Brook University

2018

Electron-phonon effects in crystals have been studied for over a century. They comprise a fundamental aspect of Solid State Physics, including superconductivity, temperature-dependence of optical properties, and transport. Systematic first principles calculations were established the end of the 1980s, with the development of Density Functional Perturbation Theory, which made calculations much more efficient than supercell methods. In this work, we focus mainly on two aspects of the electron-phonon interaction (EPI): (i) How the Fröhlich polaron affects the spectral function of polar semiconductors, and in particular, the quasiparticle (QP) renormalization (Chapters 2 and 3). (ii) The temperature dependence of the (222) forbidden X-ray reflection peak in silicon (Chapter 4).

In the Introduction, we define the retarded Green's function, a fundamental tool to study properties of semiconductors at finite temperature. We also introduce the spectral function and show how it relates to the QP peak. The Born effective charge, closely related to the Fröhlich polaron, is then defined as well. The Fröhlich polaron model is described, and expressions for the electron-phonon matrix element and corresponding zero-point renormalization (ZPR) are derived. A generalized Fröhlich model is subsequently presented, and used to interpolate the electron-phonon matrix elements with a Wannier scheme. Finally, basic descriptions of Density Functional Theory (DFT) and Density Functional Perturbation Theory (DFPT) are given.

The next Chapter deals with a simple method to calculate the temperature dependence of the conduction band minimum or valence band maximum in polar semiconductors. The non-analyticity in long-range modes makes calculations very hard to converge, requiring very dense phonon grids which are computationally very demanding. Our method treats the

problematic term analytically in the region of small phonon wavevector, where it is dominant, while the other terms are treated numerically with the usual adiabatic approximation. We apply the method to zinc-blende GaN.

In the third and longest Chapter, a cumulant representation of the Green's function is used to calculate the spectral function of two polar semiconductors, MgO and LiF. We discuss in detail properties of the method and provide an extensive numerical analysis. It is shown that this approach gives the right distance, approximately the longitudinal-optical phonon frequency, between the QP peak and a prominent feature known as the satellite. On the other hand, the distance in the more usual approach using the Dyson equation is much larger. The position and weight of the QP are also better predicted with the cumulant approach, according to diagrammatic quantum Monte Carlo calculations in the Fröhlich model, while the Dyson approach to the same order considerably underestimates the QP peak. The exact spectral functions of the conduction bands turn out to be surprisingly similar to those of the Fröhlich model. In particular, the ZPR from the Fröhlich polaron comprises a significant portion of the total ZPR.

In the last Chapter, we first obtain a perturbative expression for the temperature dependence of the charge density from a many-body approach, which includes both Debye-Waller (DW) and Fan type terms. Previous models, on the other hand, assumed rigidity of the valence charge, which led to expressions with a DW factor only. By using an acoustic sum rule, we express the second derivatives (DW type-term) as a function of first derivatives (Fan-type terms), simplifying the calculations. Then, we study the convergence of the charge density and of the structure factor, and obtain how the structure factor changes with temperature. The non-adiabatic version of our expression for the change of the charge density with temperature should be useful to study other phenomena such as pyroelectricity.

Finally, an Appendix includes the derivations of several equations used in this work.

Contents

List of Figures

List of Tables

List of Abbreviations

EPI: electron-phonon interaction
QP: quasiparticle
ZPR: zero-point renormalization
CBM: conduction band minimum
VBM: valence band maximum
DW: Debye-Waller
DFT: Density functional theory
DFPT: Density functional perturbation theory
LDA: Local-density approximation
ARPES: Angle-resolved photoemission spectroscopy

Acknowledgements

First and foremost, I would like to thank my advisor Phil Allen for his guidance during the PhD, and for suggesting very interesting problems. Phil was always available for discussion and had useful insights regardless of the question at hand. He helped me improve my presentation and writing skills, always paying close attention to detail. His assistance enabled me to attend conferences in the US and abroad, and to visit another university for several weeks to do research, a rare and rewarding opportunity. Phil also helped me meet excellent researchers around the world. In brief, Phil has been a great advisor and with him I have gained invaluable research experience.

Besides Phil, I want to specially thank Marivi Fernández-Serra for her useful advice and helping me make some important decisions. Although we did not interact frequently, her words were of utmost importance. I also want to thank her for her questions and insights during the group meetings, and for having suggested an interesting problem that I hope to have more time to spend with in the near future.

Additionally, I want to thank our collaborators. Xavier Gonze, for writing part of the cumulant paper, his hard work and investment on the project, his attention to detail, and numerous remarks that clarified the understanding and interpretation of the results. Olle Hellman, for getting me started with the Fortran code we are now using in the forbidden reflection problem, and for approaching problems from a more numerical perspective, to complement analytical techniques. And Lucia Reining, Anna Miglio and Gabriel Antonius for their work and discussions related to the cumulant project, which led to the more important results of the thesis.

Coming to Stony Brook wouldn't have been possible with all the support and encouragement from my parents and family, who gave me the freedom to pursue my own interests. Although it was hard to visit often, they were always there for me.

Another fundamental factor was the financial support from Elena Hirsch during my first year as a Masters student. I sometimes wonder what would I have done without it, and I will always be indebted to her.

Finally, I want to thank Karen for ger generosity and emotional support in the last months of the PhD, and for reading and suggesting improvements to the Introduction and the Conclusion. I also want to thank Adrián and Simon for their help when I started and didn't know anything about Linux, the physics discussions, and their questions and participation during the group meetings. I also want to thank Víctor and Marcelo for all the interesting conversations. Guadalupe, for her hospitality. Kelly, for her company this last year, and Ricardo, Artur and the friends that have been present since I came to Stony Brook.

Chapter 1

Introduction

The study of electron-phonon interactions (EPI) is indispensable for understanding several phenomena of Solid State Physics, including the temperature dependence of optical properties, electron mass renormalization and its relation to the critical temperature in superconductivity, and transport properties such as resistivity and thermal conductivity [1]. Here we focus on analyzing how the Fröhlich polaron [2] affects the electronic spectral function in polar semiconductors. Surprisingly, this is not well understood, and the usual Dyson equation to lowest order gives incorrect results. The theory of temperature dependent properties was not well established until the 1970s [3, 4], and there is still some work that needs revision [5]. In this light, the other focus of this work is to understand how the forbidden reflection changes with temperature, through a perturbative expression for the change of the charge density with temperature.

The concept of the polaron was introduced by Landau as early as 1933. The Fröhlich polaron can be understood as the interacting system between electrons and longitudinal optical modes that generate a macroscopic electric field. Extensive work has been done in this area. There is increasing evidence that polarons are related to high-temperature superconductors, colossal-magnetoresistance oxides, and conducting polymers (see review [6]). They have been tackled analytically and numerically in several ways, including path-integrals, diagrammatic expansions, variational methods and Monte Carlo techniques. Some properties, like the optical absorption measured in [7], have been adequately explained by polaron theory [8]. Although polarons dominate certain properties, it is common in the polaron literature to exclude other terms that enter the electron-phonon interaction. All phonon modes and electronic bands should be considered for a full and accurate description of the properties of crystals.

This has started to change with recent *ab-initio* polaron related calculations [9–12], together with new interpolation schemes [13,14] that are able to calculate non-analytic electron-phonon matrix elements in very dense grids. In the same spirit, in Chapter 2, we develop a simple and computationally non-expensive method to account for the non-analyticity [15]. Improved resolution in ARPES experiments has also been able to resolve sidebands [16–19], which indicate polarons should play an important role in novel electronic devices. Sidebands are separated from the quasiparticle (QP) peak by energies of ∼0.1 eV, the order of the longitudinal phonon frequencies. More details are included in Sec. 3.1. The satellites observed in [16] agree with the *ab-initio* calculations of [10]. In our recent work, Chapter 3, we also calculate the spectral function of polar semiconductors. In Chapter 4, we grapple with a property that has not been adequately addressed in the literature: the change of the charge density with temperature. Another way to summarize this work is through the following questions:

Chapter 2: How can non-adiabatic effects in polar semiconductors be treated adequately?
Chapter 3: What is a correct way to calculate spectral functions in polar semiconductors?
Chapter 4: How does the valence charge change with temperature?

The first question originated with the work of Ref. [20], where it was noticed that the adiabatic approximation diverges in polar semiconductors. It casts doubt on earlier research, like the large ZPR obtained for molecular crystals in [21], and calls for accurate non-adiabatic methods. The conclusions in Chapter 4 indicate that the results of [22, 23], which deal with organic systems, would also have to be revisited. Other active areas of research will be mentioned in the Conclusion.

In the following subsections, we introduce the concepts and mathematical tools that will be used throughout the thesis. The self-energy plays a central role, together with the Fröhlich polaron model, closely related to the Born effective charge and non-analyticity of the electron-phonon matrix elements. Density Funcional Theory (DFT) and Density Functional Perturbation Theory (DFPT) are briefly described, and help explain how macroscopic electric fields can be incorporated. Here we follow existing literature, and our contributions are contained in Chapters 2 to 4. This introduction should be useful to gain some familiarity with the Fröhlich polaron, and contains the derivation of the corresponding coupling constant, which is not easily found in the literature. By including existing expressions for the matrix elements of some Green's function, we also show how to extend a formula for the imaginary part of the Green's function to off-diagonal matrix elements, and use it in Chapter 4.

Unless otherwise specified, we use Hartree atomic units, so Planck's constant, the mass of the electron and the elementary charge are one: $\hbar = 1$,$m_e = 1$ and $e = 1$. In some sections we explicitly include them for clarity, and also in Chapter 2. When considering sums over the electronic or phonon wavevectors $\mathbf{k}$ or $\mathbf{q}$, a factor $1/N_{\mathbf{k}}$ or $1/N_{\mathbf{q}}$ (the number of terms in the sum) is sometimes implied.

Chapters 2 and 3 are previously published works, Ref. [15] and [24], respectively. Chapter 2 was written primarily by the author, while Chapter 3 includes significant contributions from P. B. Allen and X. Gonze. P. B. Allen has also done some of the writing of Chapter 4.

1.1 Green's functions

Green's functions are a very useful tool to calculate observables. In this work, we will be concerned with the renormalization of electronic energies, spectral functions and the charge density, and how these properties change with temperature T. Some derivations are included, but other results are only mentioned. Detailed discussions and proofs can be found in the books of Mahan [25], Fetter and Walecka [26], Bruus and Flensberg [27], and others.

First, we define the thermal average of an operator O,

$$\langle O\rangle = \mathrm{Tr}[e^{-\beta H}O] = \frac{1}{Z}\left(\sum_l e^{-\beta E_l}\langle l|O|l\rangle\right) \tag{1.1}$$

where $\{|l\rangle\}$ is a basis of eigenstates of the Hamiltonian H, $\beta = 1/kT$ with k the Boltzmann constant, and $Z = \mathrm{Tr}[e^{-\beta H}]$ is the partition function. The retarded electron Green's function, useful to calculate propeties at finite temperature, is defined by

$$G_R(kk', t) = -i\theta(t)\langle\{c_k(t), c^\dagger_{k'}(0)\}\rangle \tag{1.2}$$

where k will represent the quantum numbers used to label states. We will not be converned with spin, so $k = \mathbf{k}n$, with $\mathbf{k}$ the electronic wavevector and n the band index. Because of momentum conservation, $\mathbf{k} = \mathbf{k}'$, so $kk' = \mathbf{k}n\mathbf{k}n'$ in this work. $\{,\}$ is the anticommutator, and $c^\dagger_k$ and c_k are the creation and destruction operators of the electron, respectively. They are in the Heisenberg picture, which means their time-evolution is given by

$$O(t) := e^{iHt}Oe^{-iHt} \tag{1.3}$$

where O is the original operator (in the Schrödinger picture) and H is the time-independent Hamiltonian of the system. The spectral function is given by

$$A(k,\omega) = -\frac{1}{\pi}\Im m G_R(kk,\omega), \tag{1.4}$$

experimentally accessible through Angle Resolved Photoemission Spectroscopy (ARPES). The connection of the Green's function to the renormalization of energies and charge density is less straightforward and will be established in Chapter 4. In Chapter 3 we consider only diagonal elements, $k' = k$, and use the notation $G_R(k,\omega)$ instead of $G_R(kk,\omega)$. Considering off-diagonal elements is less common in the literature, but they appear naturally in Chapter 4.

The retarded Green's function satisfies the Dyson equation

$$G(kk',\omega) = G_0(kk',\omega) + \sum_{k''k'''} G_0(kk'',\omega)\Sigma(k''k''',\omega)G(k'''k',\omega), \tag{1.5}$$

where Σ is known as the self-energy, and $G_0(k,\omega)$ is the non-interacting Green's function

$$G_0(kk',\omega) = \frac{\delta_{kk'}}{\omega - \varepsilon_k + i\eta}. \tag{1.6}$$

The parameter $\eta > 0$ is infinitesimal and necessary to properly define the analytical properties of Green's functions, and ϵ_k is the unrenormalized electronic energy. Equation (1.5) is also commonly written in matrix form as $G = G_0 + G_0\Sigma G$. The self-energy of the time-ordered Green's function (see Eq. (1.10) and next section) can in principle be calculated through diagrammatic perturbation theory (Feynman diagrams). In practice, only the lowest order terms are calculated, which is valid when the interaction is weak. Since the self-energy determines the Green's function, it plays a fundamental role in the theory.

To lowest order in the atomic displacements, $\Sigma(kk',\omega) = \Sigma^{\mathrm{Fan}}(kk',\omega) + \Sigma^{\mathrm{DW}}(kk')$, where $\Sigma^{\mathrm{Fan}}(kk',\omega)$ contains squares of first-derivatives, and $\Sigma^{\mathrm{DW}}(kk')$ second-derivatives. They were worked out by Fan [28] and Antončik [29], respectively, and treated together in Yu's unpublished thesis [30]. It was only over a decade later that this was realized in published work [3, 31–33].

Writing the indices more explicitly, $k = \mathbf{k}n$ and $k' = \mathbf{k}n'$, the Fan self-energy at temperature T is given by (same expression as Eq. (4.18), and Eq. (3.6) with $n = n'$)

$$\begin{aligned}\Sigma^{\mathrm{Fan}}(\mathbf{k}nn',\omega) =& \frac{1}{N_{\mathbf{q}}}\sum_{\mathbf{q}s}^{\mathrm{BZ}}\sum_{m}\langle\mathbf{k}n|H^{(1)}_{-\mathbf{q}s}|\mathbf{k}+\mathbf{q}m\rangle\langle\mathbf{k}+\mathbf{q}m|H^{(1)}_{\mathbf{q}s}|\mathbf{k}n'\rangle\times\\ &\times\left[\frac{n_{\mathbf{q}s}(T)+1-f_{\mathbf{k}+\mathbf{q}m}(T)}{\omega-\varepsilon_{\mathbf{k}+\mathbf{q}m}-\omega_{\mathbf{q}s}+i\eta}+\frac{n_{\mathbf{q}s}(T)+f_{\mathbf{k}+\mathbf{q}m}(T)}{\omega-\varepsilon_{\mathbf{k}+\mathbf{q}m}+\omega_{\mathbf{q}s}+i\eta}\right].\end{aligned} \tag{1.7}$$

where $\mathbf{q}$ is the phonon wavevector, s the phonon mode, $\omega_{\mathbf{q}j}$ are the phonon frequencies, $N_{\mathbf{q}}$ is the number of phonon wavevectors, $|\mathbf{k}n\rangle$ an electronic Bloch state, $n_{\mathbf{q}j}(T)$ and $f_{\mathbf{k}n}(T)$ are the Bose-Einstein and Fermi-Dirac distributions, respectively, at temperature T, and $H^{(1)}_{\mathbf{q}s}$ is the change of the Hamiltonian due to phonon $\mathbf{q}s$. It can be written as

$$H^{(1)}_{\mathbf{q}s} = \sum_{i\alpha}\frac{\partial H}{\partial u_{i\alpha}(\mathbf{q})}u_{i\alpha}(\mathbf{q}s) = \sum_{li\alpha}e^{i\mathbf{q}\cdot\mathbf{R}_l}\frac{\partial H}{\partial u_{li\alpha}}u_{i\alpha}(\mathbf{q}s), \tag{1.8}$$

where $u_{li\alpha}$ is the displacement of the ion of cell index l and type i in the Cartesian direction α, $\mathbf{R}_l$ is the position of cell l, and $u_{i\alpha}(\mathbf{q}s) = e_{i\alpha}(\mathbf{q}s)/\sqrt{2M_i\omega_{\mathbf{q}s}}$ (see Sec. 1.3; M_i is the mass of atom i and $e_{i\alpha}(\mathbf{q}s)$ the polarization vector). The electron-phonon matix element $\langle\mathbf{k}+\mathbf{q}m|H^{(1)}_{\mathbf{q}s}|\mathbf{k}n\rangle$ is commonly written as $g_{mns}(\mathbf{k},\mathbf{q})$. In Chapter 4, we also use the notation $V^{(1,\mathbf{q}s)}_{\mathbf{k}+\mathbf{q}m,\mathbf{k}n}$.

A usual approximation consists of omitting the phonon frequencies $\omega_{\mathbf{q}s}$. However, we will see in Sec. 3.1 that this adiabatic approximation does not work in polar semiconductors. Equation (1.7) is used in Sec. 1.4.2 to calculate the ZPR of the Fröhlich polaron, and in Chapter 3 and 4. More details are included in Chapter 3. The derivation can be found in [25], Sec. 3.5.

A useful identity, that we use multiple times, is

$$\frac{1}{x+i\eta} = P\frac{1}{x} - i\pi\delta(x) \tag{1.9}$$

where P denotes the principal value. Together with Eq. (1.4) and Eq. (1.6), the result for the non-interacting case is $A(k,\omega) = \delta(\omega-\varepsilon_k)$. When interactions are added, the position of the quasiparticle peak in general is shifted from ε_k (energy renormalization) and broadened. Additional features can appear as well. See for example Fig. 9 in Chapter 3.

1.1.1 Matsubara formalism and relation to the retarded Green's function

At $T = 0$, the time-ordered Green's function is normally used, which for fermionic operators is defined as

$$\begin{aligned} G_{t-O}(kk', t) &= -i\langle T c_k(t) c^\dagger_{k'}(0)\rangle \\ &= -i\theta(t)\langle c_k(t) c^\dagger_{k'}(0)\rangle + i\theta(-t)\langle c^\dagger_{k'}(0) c_k(t)\rangle, \end{aligned} \tag{1.10}$$

where T in the first line is the time-ordered operator, not to be confused with the temperature; the operators are ordered from right to left in increasing time[1]. The reason for using G_{t-O} is that time-ordered operators satisfy Wick's theorem, which conveniently leads to diagrammatic perturbation theory. On the other hand, when using the Matsubara formalism on the time-ordered Green's function to calculate properties at $T \neq 0$, it turns out that the answer is expressed in terms of G_R when returning to the real axis. Let us see how this works.

The idea of the Matsubara formalism consists of switching to imaginary time $t = i\tau$, and to come back to real time at the end of the calculations through analytic continuation. The Matsubara Green's function, also known as the imaginary-time Green's function, is defined by

$$G(kk', \tau) = -\langle T_\tau c_k(\tau) c^\dagger_{k'}(0)\rangle, \tag{1.11}$$

where T_τ is just the time-ordering symbol (now in imaginary time). It can be shown that $G(kk', \tau + \beta) = -G(kk', \tau)$ for $-\beta < \tau < 0$, and similarly for $0 < \tau < \beta$. Since $G(kk', \tau)$ is an antisymmetric function in a finite interval, $-\beta < \tau < \beta$, it can be written as the Fourier series

$$\begin{aligned} G(kk', i\omega_n) &= \int_0^\beta d\tau e^{i\omega_n\tau} G(kk', \tau) \\ G(kk', \tau) &= \frac{1}{\beta}\sum_{-\infty}^{\infty} e^{-i\omega_n\tau} G(kk', i\omega_n), \end{aligned} \tag{1.12}$$

where $\omega_n = (2n + 1)\pi/\beta$. We adopt the common notation of using $i\omega_n$ as an argument, instead of a subindex n, because of the later analytic continuation to the real axis. First, we obtain the Lehmann representation (which is useful to prove formal results) of the Matsubara and retarded Green's functions.

---

[1] For bosonic operators, the time-ordered product is defined with the same sign on both terms. Also, we use $t - O$ to indicate time-ordered, to differentiate it from the imaginary time-ordered Green's function defined in the next paragraph. It is common not to write a subindex for the time-ordered Green's function, as opposed to the other versions.

Inserting the identity $\sum_{l'} |l'\rangle\langle l'|$ in Eq. (1.2) (see [27], Sec. 8.3.3),

$$\begin{aligned} G_R(kk',t) &= -i\theta(t)\langle\{c_k(t), c^\dagger_{k'}(0)\}\rangle = -i\frac{1}{Z}\sum_l \langle l|e^{-\beta H}\left(c_k(t)c^\dagger_{k'}(0) + c^\dagger_{k'}(0)c_k(t)\right)|l\rangle \\ &= -i\frac{1}{Z}\sum_{ll'} e^{-\beta E_l}\left(\langle l|c_k|l'\rangle\langle l'|c^\dagger_{k'}|l\rangle e^{i(E_l-E_{l'})t} + \langle l|c^\dagger_{k'}|l'\rangle\langle l'|c_k|l\rangle e^{-i(E_l-E_{l'})t}\right) \end{aligned} \tag{1.13}$$

Transforming to frequency space,

$$\begin{aligned} G_R(kk',\omega) &= -i\int_0^\infty dt e^{i(\omega+i\eta)t}\frac{1}{Z}\sum_{l,l'} e^{-\beta E_l}\Big(\langle l|c_k|l'\rangle\langle l'|c^\dagger_{k'}|l\rangle e^{i(E_l-E_{l'})t} \\ &\qquad\qquad + \langle l|c^\dagger_{k'}|l'\rangle\langle l'|c_k|l\rangle e^{-i(E_l-E'_l)t}\Big) \\ &= \frac{1}{Z}\sum_{ll'} e^{-\beta E_l}\left(\frac{\langle l|c_k|l'\rangle\langle l'|c^\dagger_{k'}|l\rangle}{\omega+E_l-E_{l'}+i\eta} + \frac{\langle l|c^\dagger_{k'}|l'\rangle\langle l'|c_k|l\rangle}{\omega-E_l+E'_l+i\eta}\right) \\ &= \frac{1}{Z}\sum_{ll'}\frac{\langle l|c_k|l'\rangle\langle l'|c^\dagger_{k'}|l\rangle}{\omega+E_l-E_{l'}+i\eta}\left(e^{-\beta E_l}+e^{-\beta E_{l'}}\right) \end{aligned} \tag{1.14}$$

This the Lehmann representation of the retarded Green's function. Similarly, for the Matsubara Green's function on Eq. (1.11),

$$\begin{aligned} G(kk',\tau) &= -\frac{1}{Z}\mathrm{Tr}\left[e^{-\beta H}e^{\tau H}c_k e^{-\tau H}c^\dagger_{k'}\right] \\ &= -\frac{1}{Z}\sum_{ll'} e^{-\beta E_l}\langle l|c_k|l'\rangle\langle l'|c^\dagger_{k'}|l\rangle e^{\tau(E_l-E_{l'})} \end{aligned} \tag{1.15}$$

and thus

$$\begin{aligned} G(kk',i\omega_n) &= \int_0^\beta d\tau e^{i\omega_n\tau}\frac{-1}{Z}\sum_{ll'} e^{-\beta E_l}\langle l|c_k|l'\rangle\langle l'|c^\dagger_{k'}|l\rangle e^{\tau(E_l-E_{l'})} \\ &= -\frac{1}{Z}\sum_{ll'} e^{-\beta E_l}\frac{\langle l|c_k|l'\rangle\langle l'|c^\dagger_{k'}|l\rangle}{i\omega_n+E_l-E_{l'}}\left(e^{i\omega_n\beta}e^{\beta(E_l-E_{l'})}-1\right) \\ &= \frac{1}{Z}\sum_{ll'}\frac{\langle l|c_k|l'\rangle\langle l'|c^\dagger_{k'}|l\rangle}{i\omega_n+E_l-E_{l'}}\left(e^{-\beta E_l}-e^{-\beta E_{l'}}\right) \end{aligned} \tag{1.16}$$

We can analytically continue Eq. (1.16) to Eq. (1.14) by the substitution $i\omega_n \to \omega + i\eta$. That is, $G(kk', i\omega_n \to \omega + i\eta) = G_R(kk', \omega)$ [2]. This is the usual procedure to obtain quantities at

---

[2] A known property of uniqueness of analytic functions says that if two holomorphic functions f and g in D coincide on a set of points that has a limit point in D, the functions coincide. In the Matsubara formalism there is a discrete set of points, without an accumulation point. However, Ref. [34] shows that the analytic continuation can still be done in a unique way. Writing the Green's function in a rational way, as we do here, accomplishes this [27].

finite temperature.

1.1.2 Relations between different Green's functions

The greater and lesser Green's function are defined by

$$\begin{aligned} G_>(kk',t) &= -i\langle c_k(t)c^\dagger_{k'}(0)\rangle \\ G_<(kk',t) &= i\langle c^\dagger_{k'}(0)c_k(t)\rangle \end{aligned} \tag{1.17}$$

An important identity, which we use in Chapter 4 (see Eq. (4.13)), is an off-diagonal version of

$$G_<(kk,\omega) = -2if(\omega)\Im m G_R(kk,\omega) \tag{1.18}$$

where $f(\omega)$ is the Fermi-Dirac distribution. To prove this, we follow steps similar to those in Eqs. (1.13) and (1.14) , and get

$$G_>(kk',\omega) = -\frac{2\pi i}{Z}\sum_{ll'} e^{-\beta E_l}\langle l|c_k|l'\rangle\langle l'|c^\dagger_{k'}|l\rangle\delta(E_l - E_{l'} + \omega). \tag{1.19}$$

Also

$$\begin{aligned} G_<(kk',\omega) &= \frac{2\pi i}{Z}\sum_{ll'} e^{-\beta E_l}\langle l|c^\dagger_{k'}|l'\rangle\langle l'|c_k|l\rangle\delta(-E_l + E_{l'} + \omega) \\ &= \frac{2\pi i}{Z}\sum_{ll'} e^{-\beta E_{l'}}\langle l'|c^\dagger_{k'}|l\rangle\langle l|c_k|l'\rangle\delta(-E_{l'} + E_l + \omega) \\ &= \frac{2\pi i}{Z}\sum_{ll'} e^{-\beta(E_l+\omega)}\langle l'|c^\dagger_{k'}|l\rangle\langle l|c_k|l'\rangle\delta(-E_{l'} + E_l + \omega) \\ &= -e^{-\beta\omega}G_>(kk',\omega). \end{aligned} \tag{1.20}$$

The absence of the $\theta(t)$ function gives factors of the type $\delta(E_l - E_{l'} + \omega)$, as opposed to $1/(E_l - E_{l'} + \omega)$ in Eq. (1.14).

Taking $k = k'$ and using Eqs. (1.9) and (1.14),

$$\Im m G_R(kk,\omega) = -\frac{1}{2}i(1 + e^{-\beta\omega})G_>(kk,\omega) \tag{1.21}$$

and combining it with Eq. (1.20), we obtain Eq. (1.18). In Chapter 4, we encounter the off-diagonal expression $\Im m G_R(kk',\omega)$. If $k \neq k'$, then $\langle l|c^\dagger_{k'}|l'\rangle\langle l'|c_k|l\rangle$ in Eq. (1.14) is not necessarily real, and the imaginary part of $G_R(kk',\omega)$ contains another term. However, in Chapter 4, $G_<$ appears in Eq. (4.7) through Eq. (4.8). Since the functions $\psi_{\mathbf{k}n}$ have an arbitrary phase, which cancel out in Eq. (4.4), this implies that the phase of $G_<$ is also arbitrary. Because of the form of Eq. (1.20), this implies that the phase of $\langle l|c_k|l'\rangle\langle l'|c^\dagger_{k'}|l\rangle$ is arbitrary (so we can take it to be real), and thus of the numerator of Eq. (1.14). In this way, we recover Eq. (1.18) for the off-diagonal case. See also the comments after Eq. (4.13).

1.1.3 Quasiparticle peak

As mentioned earlier, the position of the QP is generally broadened and shifted when interactions are taken into account, and additional features can appear. Using the Lehmann representation, it can be shown that (see Eq. (A.4) in the Appendix)

$$\int d\omega A(k,\omega) = 1. \tag{1.22}$$

If the peak is indeed broadened and shifted, and most of the weight of the spectral function is under such peak, then we refer to it as the QP peak. If other prominent features appear, but carry a smaller weight than the QP peak, they are referred to as sidebands or satellites.

Using Eq. (1.4), (1.5) and (1.6), one gets (see also Eq. (3.8))

$$A(\mathbf{k}n,\omega) = \frac{(-1/\pi)\Im m\Sigma(\mathbf{k}n,\omega)}{(\omega - \varepsilon_{\mathbf{k}n} - \Re e\Sigma(\mathbf{k}n,\omega))^2 + (\Im m\Sigma(\mathbf{k}n,\omega))^2}. \tag{1.23}$$

Usually, the imaginary part of the self-energy does not vary appreciably around the position of the QP, and the position of the QP is given by

$$E^{QP}_{\mathbf{k}n} = \varepsilon_{\mathbf{k}n} + \Sigma(\mathbf{k}n, E^{QP}_{\mathbf{k}n}) \tag{1.24}$$

This equation can be solved self-consistently and coincides with the Brillouin-Wiger approach [25].

Another approach, which a priori appears to be a coarser approximation, is to replace $E^{QP}_{\mathbf{k}n}$ on the right-hand side by the unrenormalized energy $\varepsilon_{\mathbf{k}n}$ (*on the mass shell* approximation), giving

$$E^{QP}_{\mathbf{k}n} = \varepsilon_{\mathbf{k}n} + \Sigma(\mathbf{k}n, \varepsilon_{\mathbf{k}n}) \tag{1.25}$$

This coincides with Rayleigh-Schrödinger perturbation theory when considering Σ to lowest order. However, we will see in Chapter 3 that Eq. (1.24) significantly underestimates the position of the QP for the Fröhlich polaron, while Eq. (1.25) is much more accurate.

1.2 Born effective charge

The Born effective charge, also known as effective or dynamical charge, is defined as

$$Z^*_{i,\alpha\beta} = \Omega_0 \frac{\partial P_{\mathrm{mac},\beta}}{\partial u_{i\alpha}(\mathbf{q}=0)}\bigg|_{E_\alpha} = \frac{\partial F_{i,\alpha}}{\partial E_\beta}\bigg|_{u_{i\alpha}=0}. \tag{1.26}$$

where $P_{\mathrm{mac},\alpha}$ is the macroscopic polarization, E_α the macroscopic electric field, $F_{i,\alpha}$ the force on atom i, and $u_{i\alpha}$ is the displacement from equilibrium of atom i in direction α. In the first derivative, $\mathbf{q} = 0$ indicates that the whole sublattice of atoms i are displaced together, and $|_{E_\alpha}$ that the derivative is at constant electric field. Thus it can be expressed as the derivative of the macroscopic polarization with respect to the atomic displacements, or as the derivative of the force induced on an ion by a change of a macroscopic electric field. A brief and clear discussion of Born effective charges and Wannier functions (see Sec. 1.4.4) can be found in [35].

With the Born effective charge, the change of polarization can be expressed in terms of these *effective* charges times their displacement,

$$\delta P_{i,\beta} = \frac{1}{\Omega_0} Z^*_{i,\alpha\beta} u_{i\alpha}, \tag{1.27}$$

as opposed to the formal charges. The Born effective charge is a dynamical and measurable quantity, while the formal charge of an ion is arbitrary and depends on how the charge of the bonds is divided. While the formal charge is related to the amount of charge that participates in the bonds, polarization or ferroelectricity are given by how that charge changes with ionic displacements.

The macroscopic electric field or polarization associated with the Born effective charge (and with the Fröhlich coupling, see Eqs. (1.61), (1.62), and (1.64)) then arises from the (small) displacements of the nuclei. The macroscopic polarization coupled to the displacements induces additional dipoles that vary on the microscopic scale, and give rise to a macroscopic quadrupole potential [36]. In other words, the displacements of the nuclei give rise to higher order interactions, like quadrupole-quadrupole or octupole-octupole, that decay as $1/d^5$ and $1/d^7$, respectively [37]. The Fröhlich coupling is associated with the dipole interaction between atomic displacements, and in real space the force decays as $1/d^3$ (with d the distance between atoms). So, in principle, a polar material can have zero Born effective charges, since other non-analytic terms are present. In practice, however, the Fröhlich coupling dominates over higher dipoles, and we refer to polar materials as those that have non-zero Born effective charges. For example, in [38] we studied the energy renormalization of the band-gap from piezoacoustic modes. We observed an interesting low temperature non-adiabatic behavior, but at higher temperatures, the piezoacoustic effects can be treated adiabatically and are not particularly important.

The Born effective charge also determines the long-range part of the interatomic force constant (Sec. 1.3.1), the dielectric function, and as we will see in Section 1.4, the long-range part of the electron-phonon interaction associated to the Fröhlich polaron.

1.3 Interatomic force constant matrix

Here, we define the interatomic force constant (IFC) matrix in a periodic crystal, and show how it can be separated into analytical and non-analytical parts [39]. It is defined by

$$C_{i\alpha,j\beta}(l,m) = \frac{\partial^2 E_{\text{tot}}}{\partial u_{li\alpha} \partial u_{mj\beta}}. \tag{1.28}$$

where E_{tot} is the total energy, and $u_{li\alpha}$ is the displacement of ion i in cell l in direction α. The Fourier transform is

$$C_{i\alpha,j\beta}(\mathbf{q}) = \frac{1}{N} \sum_{l} C_{i\alpha,j\beta}(0,m) e^{i\mathbf{q}\cdot\mathbf{R}_l} \tag{1.29}$$

where $\mathbf{R}_l$ are the lattice vectors, and the dynamical matrix is defined by

$$D_{i\alpha,j\beta}(\mathbf{q}) = \frac{C_{i\alpha,j\beta}(\mathbf{q})}{\sqrt{M_i M_j}}, \tag{1.30}$$

where M_i is the mass of atom i. The equation of motion is a generalized eigenvalue problem

$$M_i \omega^2_{\mathbf{q}s} U_{i\alpha}(\mathbf{q}s) = \sum_{j\beta} C_{i\alpha,j\beta}(\mathbf{q}) U_{j\beta}(\mathbf{q}s) \tag{1.31}$$

or equivalently

$$\omega^2_{\mathbf{q}s} e_{i\alpha}(\mathbf{q}s) = \sum_{j\beta} D_{i\alpha,j\beta}(\mathbf{q}) e_{j\beta}(\mathbf{q}s) \tag{1.32}$$

where $e_{i\alpha}(\mathbf{q}s) = \sqrt{M_i} U_{i\alpha}(\mathbf{q}s)$ are the polarization vectors, and s is the phonon mode. The normalization condition is

$$\sum_{i\beta} e^*_{i\beta}(\mathbf{q}s) e_{i\beta}(\mathbf{q}s') = \delta_{ss'} \tag{1.33}$$

Equation (1.32) determines the phonon frequencies $\omega_{\mathbf{q}s}$ in Eq. (1.7)(in particular, the phonon frequency ω_{LO} of the Fröhlich polaron) and the polarizaation vectors $e_{i\alpha}(\mathbf{q}s)$, which enter in the self-energy Eq. (1.7) through Eq. (1.8).

1.3.1 Non-analyticity

Let us consider a crystal in the presence of a long range electric field $E_\alpha \exp[i(\mathbf{q}.\mathbf{r} - \omega t)]$, with displacements $u_{i\alpha}(\mathbf{q}s)\exp[i(\mathbf{q}.\mathbf{r} - \omega t)]$, and assume an energy density E_{tot}/Ω_0 can be defined at every point. Following Sec. 7 in p.265 of [40] (see also [39] and Sec. A.3 of [20]), and omitting indices $\mathbf{q}s$, the energy can be written as

$$E_{\text{tot}} = \frac{1}{2} \sum_{ij\alpha\beta} u^*_{i\alpha} C^{\text{AN}}_{ij\alpha\beta} u_{j\beta} - \sum_{i\alpha\beta} u^*_{i\alpha} Z^*_{i,\alpha\beta} E_\beta - \frac{\Omega_0}{8\pi} \sum_{\alpha\beta} E_\alpha \chi_{\alpha\beta} E_\beta \tag{1.34}$$

where $C^{\text{AN}}_{i\alpha,j\beta}$ is the analytic part of the IFC, Eq. (1.28), $\chi_{\alpha\beta}$ is the susceptibility, and $Z^*_{i,\alpha\beta}$ the Born effective charge. It is instructive to see from this expression that the IFC, Born effective charge and susceptibility can be written as second derivatives of the total energy.

From the Euler-Lagrange equations, the equation of motion is

$$\omega^2 M_i u_{i\alpha} = \sum_{j\beta} C^{\text{AN}}_{i\alpha,j\beta} u_{j\beta} - \sum_{\beta} Z^*_{i,\alpha\beta} E_\beta \tag{1.35}$$

Thus, in addition to the analytic term, there is also a term resulting from the macroscopic electric field. Let us see that the second term is non-analytic in q, justifying the notation. The dielectric displacement is defined by

$$\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P} = \boldsymbol{\epsilon} \cdot \mathbf{D} \tag{1.36}$$

In the absence of magnetic field, $\nabla \times \mathbf{E} = 0$, so $\mathbf{q} \times \mathbf{E} = 0$: $\mathbf{E}$ is parallel to $\mathbf{q}$. If there are no free charges,

$$\nabla.\mathbf{D} = 0, \tag{1.37}$$

so $\mathbf{q}.\mathbf{D} = 0$: $\mathbf{D}$ is perpendicular to $\mathbf{q}$. Thus, the parallel component of the polarization has to cancel the electric field. The polarization is

$$\mathbf{P} = -\frac{1}{\Omega_0}\frac{\partial E_{\text{tot}}}{\partial \mathbf{E}}, \tag{1.38}$$

so

$$P_\beta = \sum_{i\alpha} \frac{1}{\Omega_0} u_{i\alpha} Z^*_{i,\alpha\beta} + \sum_{\alpha} \frac{1}{4\pi} E_\alpha \chi_{\alpha\beta} \tag{1.39}$$

The first term corresponds to Eq. (1.27), the contribution from the ionic displacements at constant electric field.

Replacing in the definition of $\mathbf{D}$ and multiplying by $\mathbf{q}$,

$$0 = E_\alpha q_\alpha + 4\pi \left(\sum_{i\alpha} \frac{1}{\Omega_0} u_{i\alpha} Z^*_{i,\alpha\beta} q_\beta + \sum_{\alpha} \frac{1}{4\pi} E_\alpha \chi_{\alpha\beta} q_\beta \right). \tag{1.40}$$

Now se use that

$$\epsilon^\infty_{\alpha\beta} = \delta_{\alpha\beta} + \chi_{\alpha\beta} \tag{1.41}$$

and $E_\alpha = |\mathbf{E}| q_\alpha / q$. Therefore

$$\mathbf{E} = -\frac{4\pi}{\Omega_0} \sum_i \frac{\mathbf{u}_i \cdot \mathbf{Z}^*_i \cdot \mathbf{q}}{\mathbf{q} \cdot \boldsymbol{\epsilon}_\infty \cdot \mathbf{q}} \mathbf{q} \tag{1.42}$$

Substituting in Eq. (1.35), we can write [37, 39]

$$C_{i\alpha,j\beta}(\mathbf{q} \to 0) = C^{\text{AN}}_{i\alpha,j\beta}(\mathbf{q} = 0) + C^{\text{NA}}_{i\alpha,j\beta}(\mathbf{q} \to 0) \tag{1.43}$$

with

$$C^{\text{NA}}_{i\alpha,j\beta} = \frac{4\pi}{\Omega_0} \frac{(\mathbf{q} \cdot \mathbf{Z}^*_i)_\alpha (\mathbf{q} \cdot \mathbf{Z}^*_j)_\beta}{\mathbf{q} \cdot \boldsymbol{\epsilon}_\infty \cdot \mathbf{q}}. \tag{1.44}$$

So, as we mentioned earlier, the Born effective charge determines the non-analytic part of the IFC. If the LO eigendisplacements of IFC at $\mathbf{q} = 0$ are the same as those of $\mathbf{q} \to 0$, then in the limit $\mathbf{q} \to 0$ (reintroducing indices $\mathbf{q}s$),

$$\omega_{\mathbf{q}s}^2 = \omega_{\mathbf{q}=0,s}^2 + \frac{4\pi}{\Omega_0}\frac{|\sum_i \mathbf{q}\cdot\mathbf{Z}_i^*\cdot\mathbf{U}_i(\mathbf{q}=0,s)|^2}{\mathbf{q}\cdot\boldsymbol{\epsilon}_\infty\cdot\mathbf{q}} \tag{1.45}$$

where we used the normalization equation, Eq. (1.33), and $U_{i\alpha}(\mathbf{q}s)$ are the eigenvectors of the generalized eigenvalue problem, Eq. (1.31).

Equation (1.45) manifestly shows what is known as LO-TO splitting. In Sec. 1.4.1, for the case of two atoms in the primitive cell, we derive Eq. (1.45) and an explicit expression for the Born effective charge.

1.4 Fröhlich polaron

The Fröhlich polaron is a model developed in 1954 [2]. It corresponds to the interacting system between an electron in the bottom of the conduction band and a longitudinal optical phonon. The Hamiltonian of the Fröhlich polaron is given by

$$\begin{gathered}
H = \sum_{\mathbf{k}} \frac{q^2}{2m^*} c_{\mathbf{k}}^\dagger c_{\mathbf{k}} + \omega_{LO} \sum_{\mathbf{q}} a_{\mathbf{q}}^\dagger a_{\mathbf{q}} + \sum_{\mathbf{kq}} M_{\mathbf{k},\mathbf{q}} c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}} (a_{\mathbf{q}} + a_{-\mathbf{q}}^\dagger), \\
M_{\mathbf{k},\mathbf{q}} = \frac{i}{q}\left[\frac{4\pi}{\Omega_0}\frac{\omega_{LO}}{2}\left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0}\right)\right]^{1/2}
\end{gathered} \tag{1.46}$$

This implies: (i) An isotropic effective mass approximation for the energy of the electron. (ii) A dispersionless phonon (Einstein model). (iii) Isotropic Born effective charge Z^*, isotropic high-frequency electronic dielectric constant ϵ_∞, and consequently, isotropic dielectric constant ϵ_0 (compare with the general form of the Hamiltonian in the Appendix, Eq. (A.1)). The $1/q$ in the electron-phonon matrix element is associated to a macroscopic electric field, as we will now see.

1.4.1 Electron-phonon matrix element

To derive the electron-phonon matrix element, we follow closely Ref. [41], Sec. 6.4 and Sec. 3.3.5, and Ref. [39]. In this way, the derivation of Ref. [41] is contained in one section, and Ref. [39] helps to establish the result more rigurously.

The derivation is based on the equation of motion of a harmonic oscillator in the presence of a electric field, and the Gauss equation in a system with no free charges. As a result of this analysis, the Lydanne-Sachs-Teller (LST) relation is also derived[3], together with the expression that relates the transverse and longitudinal phonon frequencies.

We consider a crystal with two atoms in the unit cell, interacting with a long wavelength (macroscopic) electric field

$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{q}\cdot\mathbf{r}-\omega t)} \tag{1.47}$$

---

[3] Although not a result we explicitly use in our work, it is directly related to the non-analyticity in Eq. (1.46).

The atoms can be described as a single harmonic oscillator with a relative displacement vector $\mathbf{u}$, reduced mass $\mu = (M_1 + M_2)/(M_1 M_2)$, effective charge Z^* (we will obtain an expression for it in terms of other quantities of the problem later) and natural vibrational frequency ω_{TO}. The equation of motion is

$$\mu \frac{d^2\mathbf{u}}{dt^2} = -\mu\omega_{TO}^2 \mathbf{u} + Z^*\mathbf{E} \tag{1.48}$$

Using (1.47), solutions to (1.48) can be written as

$$\mathbf{u} = \mathbf{u}_0 e^{i(\mathbf{q}.\mathbf{r}-\omega t)}. \tag{1.49}$$

which corresponds to the long-wavelength limit of the lattice vibrations. Substituting in (1.48), we obtain

$$\mathbf{u}_0 = \frac{Z^*\mathbf{E}_0}{\mu(\omega_{TO}^2 - \omega^2)} \tag{1.50}$$

The expression for the polarization is given in Eq. (1.39). Combining it with Eqs. (1.36) and (1.41), we obtain

$$\epsilon_{\alpha\beta}^{\infty} E_\alpha + \sum_{i\alpha} \frac{4\pi}{\Omega_0} u_{i\alpha} Z_{i,\alpha\beta}^* = \epsilon_{\alpha\beta} E_\beta, \tag{1.51}$$

In the isotropic case we are considering here, together with Eq. (1.50), we get

$$\epsilon(\omega) = \epsilon_\infty + \frac{4\pi(Z^*)^2}{\Omega_0\mu(\omega_{TO}^2 - \omega^2)} \tag{1.52}$$

This can also be written as $\epsilon(\omega) = \epsilon_\infty + (\epsilon_0 - \epsilon_\infty)/(1 - (\omega/\omega_{TO})^2)$, and the constants in this equation can be directly measured. The infrared-red dispersion frequency ω_{TO} can be measured as the absorption frequency of a thin film or crystal; the static or low-frequency dielectric constant ϵ_0 is the dielectric constant measured in a static field, or with a frequency very low compared to ω_{TO}; and the high-frequency dielectric constant ϵ_∞ can be obtained from the refraction of waves with frequency high compared to ω_{TO}.

In Ref. [41], the authors follow a more intuitive approach to derive Eq. (1.52), that we reproduce here. First, they consider the harmonic oscillator produces a macroscopic polarization

$$\tilde{\mathbf{P}} = \frac{1}{\Omega_0} Z^* \mathbf{u}. \tag{1.53}$$

Comparing with Eq. (1.39), this corresponds to taking a zero susceptibility $\chi = \epsilon^\infty - 1$. Using $\mathbf{E} + 4\pi\tilde{\mathbf{P}} = \epsilon\mathbf{E}$ and Eq. (1.50), they obtain

$$\epsilon = 1 + \frac{4\pi(Z^*)^2}{\Omega_0\mu(\omega_{TO}^2 - \omega^2)} \tag{1.54}$$

In this equation, the contribution to the dielectric constant coming from the valence electrons $\epsilon_e(\omega)$ is ignored. If the frequency ω of the radiation field $\mathbf{E}$ is much smaller than the band gap, $\omega << E_g$, then the radiation can be approximated as static for the electrons, and $\epsilon_e(0)$ can be used instead of $\epsilon_e(\omega)$. On the other hand, if $\omega >> \omega_{TO}$, the atoms cannot follow the electric field and do not contribute to the dielectric constant. When both of this limits are satisfied, the dielectric function is $\epsilon \sim \epsilon_e(0)$, which is usually designated as ϵ_∞, the high frequency dielectric constant. This corresponds to a frequency much higher than the phonon frequencies, but much smaller than the electronic excitation energies. Including the contribution from the electrons, Eq. (1.54) becomes Eq. (1.52).

In a crystal with no free charges, $\mathbf{D}$ satisfies Eq. (1.37), which is equivalent to

$$\epsilon \mathbf{q}.\mathbf{E}_0 = 0 \tag{1.55}$$

There are two solutions to this equation: $\mathbf{q}.\mathbf{E}_0 = 0$ and $\epsilon = 0$.

If $\mathbf{q}.\mathbf{E}_0 = 0$, which corresponds to a transverse field, there is no restriction on ω in Eq. (1.52). So when the direction of propagation is perpendicular to the electric field, Eq. (1.52) shows that ω_{TO} corresponds to a resonant frequency.

When the electric field has a longitudinal component (so that $\mathbf{q}.\mathbf{E}_0 \neq 0$), a second solution is obtained with $\epsilon = 0$, and Eq. (1.52) fixes the frequency at

$$\omega_{LO}^2 = \omega_{TO}^2 + \frac{4\pi (Z^*)^2}{\Omega_0 \mu \epsilon_\infty} \tag{1.56}$$

Also, $\epsilon = 0$ and Eq. (1.51) imply

$$\mathbf{E}_{LO} = -\frac{4\pi}{\epsilon_\infty \Omega_0} Z^* \mathbf{u}_{LO}. \tag{1.57}$$

Consequently, the electric field, proportional to the displacement, can be non-zero even if there are no external charges. This longitudinal electric field is produced by the polarization resulting from the displacement of the charges. It has the opposite sign of the displacement, so it gives an additional restoring force in Eq. (1.48). This explains why ω_{LO} is larger than ω_{TO}.

Setting $\omega = 0$ in Eq. (1.52), we have

$$\epsilon_0 = \epsilon_\infty + \frac{4\pi (Z^*)^2}{\Omega_0 \mu \omega_{TO}^2} \tag{1.58}$$

Using Eq. (1.56), we obtain

$$\frac{\epsilon_0}{\epsilon_\infty} = \frac{\omega_{LO}^2}{\omega_{TO}^2}, \tag{1.59}$$

the well known Lydanne-Sachs-Teller (LST) relation [42]. This assumes two atoms in the unit cell, but the same result holds under much more general considerations (see for example Ref. [39]). Although the derivation here implies an ordered crystal, the LO-TO splitting has been observed in first-principle calculations in water [43]. A detailed discussion can be found in [44].

From Eq. (1.56) and the LST relation Eq. (1.59), the effective charge can be written as

$$Z^* = \omega_{LO}\epsilon_\infty \sqrt{\left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0}\right)\frac{\Omega_0 \mu}{4\pi}} \tag{1.60}$$

Combining this with Eq. (1.57), we get

$$\mathbf{E}_{LO} = -\omega_{LO}\sqrt{\left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0}\right)\frac{4\pi\mu}{\Omega_0}}\mathbf{u}_{LO}, \tag{1.61}$$

getting an expression $\mathbf{E}_{LO} = -F\mathbf{u}_{LO}$ (see Ref. [41], Sec. 3.3.5). The longitudinal electric field can be written as $\mathbf{E}_{LO} = -\nabla\phi_{LO}$, where

$$\phi_{LO} = \frac{F}{iq}u_{LO} \tag{1.62}$$

and $u_{LO} = |\mathbf{u}_{LO}|$. Thus, the gradient is proportional to $\mathbf{q}$ and $\mathbf{E}_{LO}$ is longitudinal. Now we need to quantize the displacement Eq. (1.49):

$$u_{LO} = \frac{1}{\sqrt{2\mu\omega_{LO}}} e^{i(\mathbf{q}\cdot\mathbf{r} - \omega_{LO}t)}(a^\dagger_{\mathbf{q}} + a_{-\mathbf{q}}) \tag{1.63}$$

The term in the Hamiltonian is (explicitly including the elementary charge e for clarity),

$$H_{\mathrm{Fr}} = -e\phi_{LO} = \frac{ieF}{q}u_{LO}. \tag{1.64}$$

This finally gives us the electron-phonon matrix element written earlier in Eq. (1.46).

Born and Huang [40] (Sec. 7, p.82) consider the problem from a slightly different perspective, by solving the linear equations

$$\begin{aligned} \frac{\partial^2 \mathbf{u}}{\partial t^2} &= b_{11}\mathbf{u} + b_{12}\mathbf{E} \qquad (1.65) \\ \mathbf{P} &= b_{21}\mathbf{u} + b_{22}\mathbf{E}, \qquad (1.66) \end{aligned}$$

where the coefficients are scalars if isotropy is assumed. Equation (1.66) has two contributions to the polarization: one comes from the atomic displacements, and the one proportional to $\mathbf{E}$ corresponds to the ions being polarizable. Otherwise, $\mathbf{P}$ would be fully determined by the relative displacement of the ions.

The displacement $\mathbf{u}$ is separated into a solenoidal and irrotational part, which for plane waves correspond to the transverse and longitudinal parts, and they obtain the same results. That is, for the transverse wave, $\mathbf{E}$ vanishes and the vibration frequency is determined only by the local restoring force. For the longitudinal waves, there is a macroscopic electric field that contributes with an additional restoring force, and the ions oscillate with a higher frequency ω_{LO}. Since a general vibration can be expressed as a linear superposition of planewaves, all the transverse waves vibrate with the infra-red dispersion frequency ω_{TO}, and the longitudinal waves with ω_{LO}.

A more rigorous treatment would include retardation effects, as opposed to the instantaneous interaction implicit in the previous approach, and the time derivative of the magnetic field has to be included in the equations. However, the longitudinal vibrations are not affected ([40], Sec. 8, p.89).

We are interested in several aspects of the Fröhlich model, related to the energy renormalization of the electronic energies. As we will see In Chapters 2 and 3, a significant portion of the total zero-point-renormalization (ZPR) is given by the Fröhlich polaron (acoustic modes and transverse optical modes do not contribute as much, or partially cancel each other). It also gives the satellites in Chapter 3.

1.4.2 Zero-point renormalization of the conduction band minimum

The energy renormalization is given by Eq. (1.25). Using Eq. (1.7) with $k = k'$ at $T = 0$ for an electron at the bottom of the conduction band, and evaluating at $\omega = \varepsilon_{\mathbf{k}n}$, we have (denoting by $\tilde{M}$ the part of M in Eq. (1.46) that does not depend on q)

$$\begin{aligned}
\Delta\epsilon_{\mathbf{k}n} &= \sum_{\mathbf{q}} \frac{|\tilde{M}|^2}{q^2} \frac{1}{\varepsilon_{\mathbf{k}n} - \varepsilon_{\mathbf{k+q}n} - \omega_{LO}} \\
&= |\tilde{M}|^2 \frac{\Omega_0}{(2\pi)^3} \int_0^{q_F} d\theta \cos\theta d\phi q^2 dq \frac{1}{q^2} \frac{1}{-q^2/2m^* - \omega_{LO}} \\
&\sim -|\tilde{M}|^2 \frac{\Omega_0}{(2\pi)^3} 4\pi \int_0^{\infty} dq 2m^* \frac{1}{q^2 + 2m^*\omega_{LO}} \\
&= -|\tilde{M}|^2 \frac{\Omega_0}{(2\pi)^3} 4\pi 2m^* \frac{1}{\sqrt{2m^*\omega_{LO}}} \arctan\left(\frac{q}{\sqrt{2m^*\omega_{LO}}}\right)\Bigg|_0^{\infty} \\
&= -\alpha\omega_{LO}
\end{aligned} \tag{1.67}$$

where

$$\alpha = \left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0}\right)\sqrt{\frac{m^*}{2\omega_{LO}}}, \tag{1.68}$$

and m^* is the effective mass of the electron. In the second line, the limit of integration q_F would be that of a sphere whose volume is equal to the Brillouin zone volume, $(2\pi)^3/\Omega_0$. It is common practice to extend it to ∞ as we do in line 3. Because the denominator becomes smaller, this introduces typically an error of the order of 10%. Additional comments related to the validity of the effective mass approximation can be found in Sec. 2.3. Values of α between 0 and 1 are referred to as weak coupling; between 1 and 6 as intermediate coupling, and α greater than 6 is considered strong coupling. Several techniques have been employed to study the Fröhich polaron, like the all-coupling variational path-integral formalism developed by Feynman [45], or the Lang-Firsov canonical transformation [46]. Diagrammatic quantum Monte Carlo is used in [47], obtaining in principle an exact ZPR.

The Fröhlich polaron is extended to degenerate and anisotropic bands using the **k**.**p** method in [48]. This is useful to treat the valence bands of many polar semiconductors, as we do later in Sec. 2.8.1.2 (spin-orbit coupling tends to lift degeneracies, as seen in the calculations of [49], so it might need to be taken into account to obtain accurate results in certain materials). To analyze GaN-zincblende in Chapter 2, these methods seemed sufficient. However, more complex materials, with stronger anisotropy and additional modes, in principle need a more general formulation like the Vogl model, which we describe in the following section.

1.4.3 Generalization of the Fröhlich model

The work by Vogl [36] consists of a generalization of the Fröhlich model, by using a many-body approach and considering analytical properties of the dielectric function [50]. The Vogl model does not assume isotropy in the Born effective charges nor in the dielectric constant, as opposed to the Fröhlich polaron model. The electron-phonon matrix element, including periodicity, is given by [13, 14]

$$g_{mns}(\mathbf{k},\mathbf{q}) = i\frac{4\pi}{\Omega_0}\sum_{i,\mathbf{G}\neq s}\frac{1}{\sqrt{2M_i\omega_{\mathbf{q}s}}}\frac{(\mathbf{q}+\mathbf{G}).\mathbf{Z}_i^*.\mathbf{e}_i(\mathbf{q}s)}{(\mathbf{q}+\mathbf{G}).\boldsymbol{\epsilon}_\infty.(\mathbf{q}+\mathbf{G})}\langle\psi_{\mathbf{k+q}m}|e^{i(\mathbf{q}+\mathbf{G}).\mathbf{r}}|\psi_{\mathbf{k}n}\rangle \tag{1.69}$$

A clear advantage of this formulation is that all modes are automatically incorporated; it is not necessary to identify the divergent modes beforehand. In crystals with enough symmetry like GaN-zincblende, the Born effective charge matrix Z_i^* is diagonal for each atom i. That is, it can be treated like a scalar, and then the component of the eigenvector that contributes is the one parallel to **q**. Thus, we can consider the eigenvector to be aligned with **q**, as in Sec. 1.4.1. But in general, Z_i^* is not diagonal, and **q** and the eigendisplacements have different orientations.

We follow here the derivation of Eq. (1.69) in [14] and its Supplemental Material. In electrostatics, $\mathbf{E} = -\nabla\phi$. Using Eq. (1.36) in its anisotropic version (where $\boldsymbol{\epsilon}$ is a tensor) and $\nabla.\mathbf{D} = -\rho$, it becomes $\nabla.\boldsymbol{\epsilon}.\nabla\phi = -\rho$. Considering a point charge Q at position **d**, and a compensating background $-Q/N\Omega_0$, the Poisson equation is

$$\nabla.\boldsymbol{\epsilon}.\nabla\phi(\mathbf{r},\mathbf{d}) = -\sum_{\mathbf{R}}\left(Q\delta(\mathbf{r}-\mathbf{d}-\mathbf{R}) - \frac{Q}{N\Omega_0}\right). \tag{1.70}$$

where **R** are the lattice vectors. Since ϕ is periodic, it can be written as

$$\phi(\mathbf{r},\mathbf{d}) = \sum_{\mathbf{qG}} e^{i(\mathbf{q}+\mathbf{G}).\mathbf{r}}\phi(\mathbf{q}+\mathbf{G},\mathbf{d}). \tag{1.71}$$

Replacing this expansion in Eq. (1.70), and using the usual Fourier representation of the δ function, the result is

$$\phi(\mathbf{r},\mathbf{d}) = \frac{4\pi}{\Omega_0}\frac{Q}{N}\sum_{\mathbf{q},\mathbf{G}\neq-\mathbf{q}} \frac{e^{i(\mathbf{q}+\mathbf{G}).(\mathbf{r}-\mathbf{d})}}{(\mathbf{q}+\mathbf{G}).\boldsymbol{\epsilon}.(\mathbf{q}+\mathbf{G})} \tag{1.72}$$

The electrostatic potential of a dipole $\mathbf{p} = Q\mathbf{d}$ is given by $\phi_{\text{dip}}(\mathbf{r}) = \phi(\mathbf{r},\mathbf{d}) - \phi(\mathbf{r},0)$ in the limit where $\mathbf{d} \to 0$ and $\mathbf{p}$ is constant:

$$\phi_{\text{dip}} = -i\frac{4\pi}{\Omega_0}\frac{1}{N}\sum_{\mathbf{q},\mathbf{G}\neq-\mathbf{q}} \mathbf{p}.\frac{(\mathbf{q}+\mathbf{G})e^{i(\mathbf{q}+\mathbf{G}).\mathbf{r}}}{(\mathbf{q}+\mathbf{G}).\boldsymbol{\epsilon}.(\mathbf{q}+\mathbf{G})} \tag{1.73}$$

The corresponding Hamiltonian piece is $V = -e\phi_{\text{dip}}$ (just as Eq. (1.64)). Then, replacing $\mathbf{p}$ with $\mathbf{Z}_i^*.\mathbf{u}_{li}(\mathbf{q}s)$ (as in Eq. (1.27), without Ω_0), using $\sum_l e^{i(\mathbf{q}-\mathbf{q}')\mathbf{R}_l} = N\delta_{\mathbf{q}\mathbf{q}'}$ and that the matrix element $g_{mns}(\mathbf{k},\mathbf{q})$ is given by $\langle\psi_{\mathbf{k}+\mathbf{q}m}|V_{\mathbf{q}s}(\mathbf{r})|\psi_{\mathbf{k}n}\rangle$, Eq. (1.69) can be obtained. See [14] for a more detailed discussion.

1.4.4 Interpolation scheme

The electron-phonon matrix elements can be separated into a long-range or non-analytic part, and a short-range or analytic part, just as for the IFC in Sec. 1.3.1. The basic idea is that the long-range part is subtracted from the electron-phonon matrix element, then the standard Wannier interpolation can be used [51], and then the long-range part is added at the end. We describe the interpolation procedure following [13].

The electron-phonon matrix element can be written

$$g_{mns}(\mathbf{k},\mathbf{q}) = \sum_i \frac{\mathbf{e}_i(\mathbf{q}s)}{\sqrt{2M_i\omega_{\mathbf{q}s}}}.\mathbf{d}_{mni}(\mathbf{k},\mathbf{q}) \tag{1.74}$$

where $\mathbf{d}_{mni}(\mathbf{k},\mathbf{q}) = \langle\mathbf{k}+\mathbf{q}m|\delta V_{\text{SCF}}/\delta\mathbf{u}_i(\mathbf{q})|\mathbf{k}n\rangle$ (see Sec. 1.5 for the definition of V_{SCF}). First, a set of Wannier functions is defined by

$$|\mathbf{R}m\rangle = \frac{1}{\sqrt{N_\mathbf{k}}}\sum_{\mathbf{k}n} e^{-i\mathbf{k}.\mathbf{R}}U_{nm}(\mathbf{k})|\mathbf{k}n\rangle, \tag{1.75}$$

where $U_{nm}(\mathbf{k})$ is a unitary matrix determined by the Wannierization procedure [52], and $N_\mathbf{k}$ is the number of points in the k-grid. Let

$$\tilde{\mathbf{d}}_{m'n'i}(\mathbf{k},\mathbf{q}) = U^*_{mm'}(\mathbf{q})\mathbf{d}_{m'n'i}(\mathbf{k},\mathbf{q})U_{n'n}(\mathbf{k}) \tag{1.76}$$

Fourier transforming to direct space, we get

$$\mathbf{d}_{mni}(\mathbf{R},\mathbf{R}') = \frac{1}{N_\mathbf{k}}\sum_{\mathbf{k},\mathbf{q}}\sum_{m'n'} e^{-i\mathbf{k}.\mathbf{R}+i\mathbf{q}.\mathbf{R}'}\tilde{\mathbf{d}}_{m'n'i}(\mathbf{k},\mathbf{q}) \tag{1.77}$$

Then, one transforms back to momentum space by

$$\mathbf{d}_{mni}(\mathbf{k},\mathbf{q}) = \frac{1}{N_{\mathbf{k}}^2} \sum_{\mathbf{R}\mathbf{R}'m'n'} e^{i\mathbf{k}.\mathbf{R}+i\mathbf{q}.\mathbf{R}'} U_{m'm}(\mathbf{k}+\mathbf{q}) \mathbf{d}_{m'n'i}(\mathbf{R},\mathbf{R}') U^*_{nn'}(\mathbf{k}) \tag{1.78}$$

and now $\mathbf{k}$ and $\mathbf{q}$ can take any values. In the case of polar materials, the authors of [13] define the non-analytic deformation potential (compare with Eq. (1.69))

$$\mathbf{d}_i^{\mathrm{NA}}(\mathbf{q}) = 4\pi i e \sum_{\mathbf{G}} \frac{(\mathbf{q}+\mathbf{G}) \cdot \mathbf{Z}_i^*}{(\mathbf{q}+\mathbf{G}) \cdot \boldsymbol{\epsilon}_\infty \cdot (\mathbf{q}+\mathbf{G})} e^{-(\mathbf{q}+\mathbf{G})^2/4\gamma} \tag{1.79}$$

which corresponds to the long-range contribution (γ is just a convergence parameter). Then they subtract it from $\tilde{\mathbf{d}}_{m'n'i}$, giving the short-range contribution $\mathbf{d}^{\mathrm{AN}}_{m'n'i}$,

$$\mathbf{d}^{\mathrm{AN}}_{m'n'i}(\mathbf{k},\mathbf{q}) = \tilde{\mathbf{d}}_{m'n'i}(\mathbf{k},\mathbf{q}) - \delta_{m'n'} \mathbf{d}_i^{\mathrm{NA}}(\mathbf{q}), \tag{1.80}$$

and AN stands for analytic. Then the transformation of Eq. (1.77) is applied to $\mathbf{d}^{\mathrm{AN}}_{m'n'i}(\mathbf{k},\mathbf{q})$, giving $\mathbf{d}^{\mathrm{AN}}_{m'n'i}(\mathbf{R},\mathbf{R}')$, to which one applies the transformation of Eq. (1.78), giving $\mathbf{d}^{\mathrm{AN}}_{mni}(\mathbf{k},\mathbf{q})$. Finally, the long-range part can be added back,

$$\mathbf{d}_{mni}(\mathbf{k},\mathbf{q}) = \mathbf{d}^{\mathrm{AN}}_{mni}(\mathbf{k},\mathbf{q}) + \mathbf{d}_i^{\mathrm{NA}}(\mathbf{q}) \delta_{mn} \tag{1.81}$$

obtaining the matrix element for any $\mathbf{q}$, and thus of the electron-phonon matrix element through Eq. (1.74). The deformation potential is defined as

$$D_{nms}(\mathbf{k},\mathbf{q}) = \frac{\sqrt{2\rho\Omega_0\omega_{\mathbf{q}s}}}{\hbar} |g_{mns}(\mathbf{k},\mathbf{q})|. \tag{1.82}$$

In [13] the authors consider LO and TO modes for GaAs, a polar material, and compare the deformation potential calculations resulting from their interpolation method with DFPT results, obtaining an excellent agreement. The same type of interpolation is applied in [14], and analogous results are obtained when comparing to DFPT calculations.

1.5 Density Functional Theory

For completeness, we include a brief discussion of Density Functional Theory (DFT). DFT is a formalism to calculate properties of condensed matter systems that has found a widespread use in the last decades. It is based on the work of Hohenberg and Kohn [53] and Kohn and Sham [54]. Solving the Schrödinger equation for N particles in principle involves $3N$ degrees of freedom. In DFT, this is reduced to the 3 degrees of freedom of the charge density $n(\mathbf{r})$.

The first theorem states that different external potentials give rise to different ground-state electronic-densities. So the ground-state density determines uniquely the external potential. Thus the Hamiltonian is fully determined, and the many-body wavefunctions as well. Therefore, all the ground-state properties of the system are determined by the ground-state charge density [55]. The second theorem states that there is a universal functional of the density, $F[n]$, such that the functional

$$E[n] = F[n] + \int n(\mathbf{r})V(\mathbf{r})d\mathbf{r} \tag{1.83}$$

is minimized by the ground-state density corresponding to the external potential $V(\mathbf{r})$, with a value equal to the ground-state energy. This provides the first fundamental result for DFT. However, the form of the functional $F[n]$ is not known. This was addressed by Kohn and Sham [54].

In this work, the original many-body problem of interacting electrons is mapped into an equivalent non-interacting problem with an effective potential V_{SCF}. That is, one has to solve the one-electron Schrödinger equation

$$\left(-\frac{1}{2}\frac{\partial^2}{\partial\mathbf{r}^2} + V_{\mathrm{SCF}}(\mathbf{r})\right)\psi_n(\mathbf{r}) = \varepsilon_n\psi_n(\mathbf{r}) \tag{1.84}$$

where

$$V_{SCF}(\mathbf{r}) = V(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + v_{xc}(\mathbf{r}),\ \ v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}}{\delta n(\mathbf{r})}. \tag{1.85}$$

We can write

$$n(\mathbf{r}) = 2\sum_{n=1}^{N_e/2}|\psi_n(\mathbf{r})|^2 \tag{1.86}$$

$$T_0[n] = -2\frac{1}{2}\sum_{n=1}^{N_e/2}\int \psi_n^*(\mathbf{r})\frac{\partial^2\psi_n(\mathbf{r})}{\partial\mathbf{r}^2}d\mathbf{r}, \tag{1.87}$$

where N_e is the total number of electrons, and the exchange correlation energy is defined by

$$F[n] = T_0[n] + \frac{1}{2}\int\frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' + E_{xc}[n]. \tag{1.88}$$

The difficulty of DFT is to find good approximations to E_{xc}. In the Local Density Approximation (LDA), the exchange correlation energy depends only on the charge density evaluated at each point in space (and not, for example, on derivatives of the charge density), and it is written as

$$E_{xc}[n] = \int \epsilon_{xc}(n)|_{n=n(\mathbf{r})}n(\mathbf{r})d\mathbf{r}. \tag{1.89}$$

It is a reasonable approximation, since the long-range Hartree effects are contained in the Hartree term [55]. This approach has been more successful than originally anticipated, and it acccurately describes properties of weakly correlated materials, such as semiconductors and simple metals.

1.5.1 Density functional perturbation theory

The most direct or intuitive approach to calculate the variations appearing in Eq. (1.7) is by displacing the ions a small and finite amount in a supercell. However, it is well-known that the supercell method is very time-consuming and only wave vectors that are commensurate with the supercell can be calculated. In particular, it is inconvenient for $\mathbf{q} \to 0$, which is necessary when calculating long-range properties of polar semiconductors, like the electron-phonon matrix element of LO modes. Here we follow [56].

The Hellman-Feynman theorem states (see for example the derivation in [55])

$$\frac{\partial E}{\partial \lambda_i} = \int \frac{\partial V_\lambda(\mathbf{r})}{\partial \lambda} n_\lambda(\mathbf{r}) d\mathbf{r}, \tag{1.90}$$

which results from the wavefunction being stationary in the equilibrium position. Then the second derivative (which is useful to determine for example the IFC when $\lambda = u_{li\alpha}$) is

$$\frac{\partial^2 E}{\partial \lambda_i \lambda_j} = \int \frac{\partial^2 V_\lambda(\mathbf{r})}{\partial \lambda_i \lambda_j} n_\lambda(\mathbf{r}) d\mathbf{r} + \int \frac{\partial n_\lambda(\mathbf{r})}{\partial \lambda_i} \frac{\partial V_\lambda(\mathbf{r})}{\partial \lambda_j} d\mathbf{r} \tag{1.91}$$

The derivative $\partial n_\lambda(\mathbf{r})/\partial \lambda_i$ needs to be determined. From Eq. (1.86) one has

$$\Delta n(\mathbf{r}) = 4\Re e \sum_{n=1}^{N_e/2} \psi_n^*(\mathbf{r}) \Delta \psi_n(\mathbf{r}) \tag{1.92}$$

and in the following calculations, $\Delta\cdot = \sum_i \partial \cdot /\partial \lambda_i \Delta \lambda_i$. From standard perturbation theory [57], the change $\Delta\psi_n$ of the Kohn-Sham orbitals to first-order is

$$(H_{SCF} - \varepsilon_n)|\Delta\psi_n\rangle = -(\Delta V_{SCF} - \Delta\varepsilon_n)|\psi_n\rangle \tag{1.93}$$

where

$$H_{SCF} = -\frac{1}{2}\frac{\partial^2}{\partial \mathbf{r}^2} + V_{SCF}(\mathbf{r}) \tag{1.94}$$

is the unperturbed Kohn-Sham Hamiltonian

$$\Delta V_{SCF}(\mathbf{r}) = \Delta V(\mathbf{r}) + \int \frac{\Delta n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \left.\frac{dv_{xc}(n)}{dn}\right|_{n=n(\mathbf{r})} \Delta n(\mathbf{r}) \tag{1.95}$$

and $\Delta\varepsilon_n = \langle\psi_n|\Delta V_{SCF}|\psi_n\rangle$. The equations can be solved self-consistently: $\Delta V_{SCF}(\mathbf{r})$ is a linear functional of $\Delta n(\mathbf{r})$, which depends linearly on the $\Delta\psi_n$, which in turns depends on $\Delta V_{SCF}(\mathbf{r})$ through Eq. (1.93). In this way, quantities like $\partial V_{SCF}/\partial u_{li\alpha}$, which enters into the electron-phonon matrix elements of Eq. (1.7), and the IFC matrix ([56], p. 526), can be determined. To see how different wavevectors $\mathbf{q}$ decouple and that they can be arbitrary (as opposed to the case of supercells), see the discussion on Monochromatic Perturbations on p.520 of [56].

An equation analogous to (1.93) is known as the Sternheimer equation. In the context of the renormalization of electronic energies, it is more efficient to solve this equation iteratively [58], than to invert the operator on the left and write a sum over states as in Eq. (1.96). A few more comments are included in Sec. 3.10.

1.5.2 Homogeneous electric field

In Sec. 1.4, we saw how the Fröhlich polaron is associated with a macroscopic electric field. The corresponding electrostatic potential is $V_{\mathbf{E}}(\mathbf{r}) = e\mathbf{E}.\mathbf{r}$, but this it not well defined because it is not bounded from below in an infinite crystal, and it is incompatible with periodic boundary conditions. We will now see how it can be properly defined [56].

The change in the charge density Eq. (1.92) depends on $\Delta\psi_n(\mathbf{r})$, given by

$$\Delta\psi_n(\mathbf{r}) = \sum_{m\neq n} \psi_m(\mathbf{r}) \frac{\langle\psi_m|\Delta V_{SCF}|\psi_n\rangle}{\varepsilon_n - \varepsilon_m}, \tag{1.96}$$

and the matrix element for $\mathbf{r}$, entering in $V_{\mathbf{E}}(\mathbf{r})$, can be written as

$$\langle\psi_m|\mathbf{r}|\psi_m\rangle = \frac{\langle\psi_m|[H_{SCF},\mathbf{r}]|\psi_n\rangle}{\varepsilon_m - \varepsilon_n} \tag{1.97}$$

for all $m \neq n$. If the self-consistent potential is local,

$$[H_{SCF},\mathbf{r}] = -\frac{\partial}{\partial\mathbf{r}} \tag{1.98}$$

which is well defined. In this way, DFPT can handle macroscopic electric fields, and thus include the long-range Fröhlich type interaction. Similarly, the integral for the polarization operator in all space involves $\mathbf{r}$, which is written as a wave vector derivative $i\partial/\partial\mathbf{k}$ [35, 59], which is well defined in a unit cell.

Chapter 2

Influence of Fröhlich polaron coupling on renormalized electron bands in polar semiconductors. Results for zincblende GaN.

We develop[4] a simple method to study the zero-point and thermally renormalized electron energy $\varepsilon_{\mathbf{k}n}(T)$ for $\mathbf{k}n$ the conduction band minimum or valence maximum in polar semiconductors. We use the adiabatic approximation, including an imaginary broadening parameter $i\delta$ to suppress noise in the density-functional integrations. The finite δ also eliminates the polar divergence which is an artifact of the adiabatic approximation. Non-adiabatic Fröhlich polaron methods then provide analytic expressions for the missing part of the contribution of the problematic optical phonon mode. We use this to correct the renormalization obtained from the adiabatic approximation. Test calculations are done for zincblende GaN for an 18×18×18 integration grid. The Fröhlich correction is of order -0.02 eV for the zero-point energy shift of the conduction band minimum, and +0.03 eV for the valence band maximum; the correction to renormalization of the 3.28 eV gap is -0.05 eV, a significant fraction of the total zero point renormalization of -0.15 eV.

2.1 Introduction

Electron quasiparticles in crystals form energy bands $\varepsilon_{\mathbf{k}n}$. Computations normally use the Born-Oppenheimer approximation, that atoms are fixed rigidly at crystalline coordinates. Vibrations around these fixed coordinates (phonon quasiparticles) are the main cause of temperature-dependent shifts of the electron bands. At temperature T the shift is typically $2-4k_BT$, which can have noticeable effects on electron behavior in semiconductors. There is also a zero-point shift, caused by phonon zero-point fluctuations, which is comparable in size to the thermal shift at room temperature. Cardona and collaborators [60, 61] have given brief reviews of the vibrational renormalization of semiconductor bands. Since energy band calculations omit these effects, a correction should be made when comparing with experiment.

---

[4] This work was done in collaboration with P. B. Allen. Only minor modifications were introduced to Ref. [15].

These effects have an analog in electronic energy levels in molecules. When an electron is excited, interatomic separations and vibrational spectra are altered compared to the ground state. To compute the correct electron excitation energy, Born-Oppenheimer energies are not enough. This topic is usually described as "Franck-Condon effects" [62–65]. Zero-point vibrational contributions to a molecular excited state energy are different from the zero-point vibrational contributions to the ground state energy. In the molecule, one generally thinks of the change in vibrational energies caused by electronic excitation, whereas in the crystal one generally thinks of the change in electronic energies caused by vibrational excitation. These two points of view are united by what is known [66] as "Brooks' Theorem" [67]: the shift in an electron energy $\varepsilon_{\mathbf{k}n}$ caused by a unit increase in phonon occupancy of mode $\omega_{\mathbf{q}s}$ equals the shift of the phonon energy $\hbar\omega_{\mathbf{q}s}$ caused by a unit increase in electron occupancy $\varepsilon_{\mathbf{k}n}$.

Computation by density functional theory (DFT) of the temperature dependence of electronic properties of semiconductors and insulators, and also metals, has grown recently [20,68–74]. Fits to experimental data with different models have also been done [75–77]. For arbitrary bands, the electron-phonon contribution to the renormalization $(E_{\mathbf{k}n} - \varepsilon_{\mathbf{k}n})$ of the electronic bands, to second order in the ion's displacement, is

$$\begin{aligned} E_{\mathbf{k}n} - \varepsilon_{\mathbf{k}n} = \frac{1}{N}\sum_{\mathbf{q}sn'}^{\mathrm{BZ}} |\langle \mathbf{k}n|H_s^{(1)}|\mathbf{k}+\mathbf{q}n'\rangle|^2 \Bigg[&\frac{n_{\mathbf{q}s} + 1 - f_{\mathbf{k}+\mathbf{q}n'}}{\varepsilon_{\mathbf{k}n} - \varepsilon_{\mathbf{k}+\mathbf{q}n'} - \hbar\omega_{\mathbf{q}s} + i\eta} + \\ &\frac{n_{\mathbf{q}s} + f_{\mathbf{k}+\mathbf{q}n'}}{\varepsilon_{\mathbf{k}n} - \varepsilon_{\mathbf{k}+\mathbf{q}n'} + \hbar\omega_{\mathbf{q}s} + i\eta} \Bigg] \\ + \frac{1}{N}\sum_{\mathbf{q}s}^{\mathrm{BZ}} \langle \mathbf{k}n|H_{ss}^{(2)}|\mathbf{k}n\rangle [2n_{\mathbf{q}s} + 1] \end{aligned} \tag{2.1}$$

Here, $\langle \mathbf{k}n|H_s^{(1)}|\mathbf{k}+\mathbf{q}n'\rangle$ is the matrix element for scattering an electron $\mathbf{k}$ by a phonon $\mathbf{q}$; it has units of energy and a typical size of roughly the geometric mean of electron and phonon energies (the notation for the electron-phonon matrix element varies slighlty with respect to the one in the Introduction). The Debye-Waller term $\langle \mathbf{k}n|H_{ss}^{(2)}|\mathbf{k}n\rangle$ is the second order interaction energy involving two phonons $\mathbf{q}s$ and $\mathbf{q}'s'$, but only $\mathbf{q}s = -\mathbf{q}'s'$ enters in lowest order. The Fermi-Dirac and Bose-Einstein equilibrium occupation factors are denoted f and n. The infinitesimal parameter $i\eta$ ensures the real and imaginary parts are well defined. Only the real part is discussed here. We omit the smaller thermal expansion contribution in this work.

The formulas used by Allen, Heine and Cardona [3, 4] intentionally drop the phonon energy $\pm\hbar\omega_{\mathbf{q}s}$ from the denominators in comparison with the electron energy difference $\varepsilon_{\mathbf{k}n} - \varepsilon_{\mathbf{k}+\mathbf{q}n'}$. This is an adiabatic approximation. The justification is that, in semiconductors, typical energy denominators are much larger than $\hbar\omega_{\mathbf{q}s}$. However, it was pointed out by Poncé *et al.* [20] that for polar materials, it is necessary to keep the $\pm\hbar\omega_{LO}$ for longitudinal optic (LO) modes to avoid an unphysical divergence in the intraband $(n' = n)$ term at band extrema, caused by the adiabatic treatment of the long-range Fröhlich-type electron-phonon interaction.

A converged non-adiabatic evaluation of Eq. (2.1), summed on a fine enough mesh to accurately get the Fröhlich part of the renormalization, requires a very fine and very expensive mesh. Our aim in this Chapter is to explore a simplified method that works adequately on a coarser mesh. We test our method by computations for zincblende (cubic) GaN, abbreviated c-GaN. Our corrections use the effective mass approximation ($\varepsilon_{\mathbf{k}} = \hbar^2 k^2/2m^*$ for band edges near $\mathbf{k} = 0$. This works well for the conduction band where m^* is small, $\approx 0.16 m_e$. The top of the valence band is triply-degenerate (because we ignore the spin-orbit interaction) and involves higher effective masses, which work a bit less well.

One reason for choosing Gallium nitride is its useful properties, including a high thermal conductivity [78], and a high melting point that allows it to operate at high temperatures. Its wide and direct band gap make it efficient for lasers [79], and for high-power and high-frequency electronic devices [80–82]. It is used in white LED's. Alloying with InN and AlN allows engineering of optical and electrical properties [83]. For simplicity we study c-GaN rather than the more stable wurtzite (hexagonal) GaN, or h-GaN. Although h-GaN has been more thoroughly studied, c-GaN has several advantages: it has better n and p doping properties [84, 85], higher saturated electron drift mobilities [83, 86], and it is convenient to work in the 510 nm region.

2.2 "Adiabatic plus $i\delta$" approximation corrected using effective-mass theory

For convenience, we assume (correctly for c-GaN) that band extrema are at $\mathbf{k} = 0$. The Fröhlich part of the integral in Eq. (2.1) involves $\int d^3q$ and a factor $1/q^2$ from the long-range polar electron-phonon matrix element. If the $\pm\hbar\omega_{\mathrm{LO}}$ is omitted, then in the small-$\mathbf{q}$ Fröhlich region, denominators in Eq. (2.1) behave as q^2. The integral then involves $\int dq/q^2$ which diverges at $q = 0$. When $\pm\hbar\omega_{\mathrm{LO}}$ is kept, the divergence is removed from the first denominator in Eq. (2.1), and the singularity in the second denominator is integrable. When $\mathbf{k}$ is not chosen to be 0, there are (integrable) singular denominators $\varepsilon_{\mathbf{k}n} - \varepsilon_{\mathbf{k}+\mathbf{q}n'} \pm \hbar\omega_{\mathbf{q}s} \to 0$ on extended surfaces in $\mathbf{q}$-space. All these cases create problems if integrated numerically by summing points on a simple mesh. Regardless of how dense the $\mathbf{q}$-mesh is, singular integrals of this type do not converge (as already noted in [4]) except with a carefully tempered mesh, designed to give the correct principal-value treatment in three dimensions. A useful procedure is to change the $i\eta$ in the denominator to a finite imaginary energy $i\delta$. Convergence in this parameter was studied by Poncé *et al.* [20]. Since the true result around this type of singularity integrates to a small contribution when done correctly, it is safe to add a finite imaginary energy consistent with the mesh size. Unlike Poncé *et al.*, we do not need δ to be particularly small or less than $\hbar\omega_{\mathrm{LO}}$. Specifically, δ should not be smaller than the typical energy jump $\Delta_s\varepsilon = \varepsilon_{\mathbf{k}+\mathbf{q}+\Delta\mathbf{q}} - \varepsilon_{\mathbf{k}+\mathbf{q}}$ associated with the mesh size $\Delta\mathbf{q}$ when $\varepsilon_{\mathbf{k}+\mathbf{q}}$ lies near the singularity surface. The singular part of the integrand $1/\Delta_s\varepsilon$ is then replaced by $\Delta_s\varepsilon/(\Delta_s\varepsilon^2 + \delta^2)$. The subscript "s" indicates "singularity." Errors associated with the random location of mesh points relative to the singularity surface are then reduced from $N_s^{1/2}/\Delta_s\varepsilon$ to $N_s^{1/2}\Delta_s\varepsilon/\delta^2$, where N_s is the number of mesh points neighboring the singularity surface. When the singularity is at $\mathbf{k} = 0$, $N_s \approx 1$, but for an extended singularity, the value of N_s is likely to be of order $N_{\text{mesh}}^{2/3}$. Therefore the value of δ should be greater than $\Delta_s\epsilon$ or $N_{\text{mesh}}^{1/6}\Delta_s\epsilon$, depending on whether the singularity is at a point or on an extended area in $\mathbf{k}$-space. At a minimum or maximum (local or absolute) of $\varepsilon_{\mathbf{k}n}$, there is a singular point which requires a non-adiabatic treatment in polar materials. When $\varepsilon_{\mathbf{k}n}$ is not at an absolute band maximum or minimum, the extended singularity surface can be safely approximated by replacing $\pm\omega_{\mathbf{q}s}$ by $i\delta$. The reason is, if the surface is redefined by $\Delta\varepsilon = 0$ instead of $\Delta\varepsilon \pm \omega_{\mathbf{q}s} = 0$, it causes only a small shift of the surface in $\mathbf{k}$-space. This should do little to change the small remainder after principal-parts cancellation of the singularity. The replacement of $\pm\omega_{\mathbf{q}s}$ by $i\delta$ is what we call the "adiabatic + $i\delta$" method.

When the state of interest $\mathbf{k}n$ is a (local or absolute) band extremum (taken here to be $\mathbf{k} = 0$), replacement of $\pm\hbar\omega_{\text{LO}}$ by a finite $i\delta$ does not correctly treat the Fröhlich intraband renormalization effect. This is especially true in the first denominator of Eq. (2.1). This "emission term" with $n_{\mathbf{q}s}+1$ in the numerator, integrates only over one side of the singularity, and thus has no principal-parts cancellation. The long-range polar interaction, when treated correctly (non-adiabatically), makes an additional renormalization. Our aim is to use a mesh fine enough to capture all the less singular contributions, but coarse enough for rapid computation (for example 20×20×20). Then to include the Fröhlich effect, we want to focus on a small $\mathbf{q}$ "central region" and treat it by an analytic integration using effective mass theory. For this purpose we need a central region large enough that outside it, $\pm\hbar\omega_{\text{LO}}$ can be safely replaced by $i\delta$, but small enough that inside, the energy $\varepsilon_{\mathbf{k},n}$ can be replaced by $\varepsilon_{0n} + \hbar^2k^2/2m^*$. The mesh should be fine enough that the "adiabatic plus $i\delta$" calculation (by mesh summation) is reasonably converged in the central region, and therefore adequately approximated by an analytic effective-mass integration of the "adiabatic plus $i\delta$" intraband central region sum. If these conditions can be satisfied, then we can subtract the analytic effective-mass version of the "adiabatic plus $i\delta$" and add the analytic effective-mass version of the Fröhlich renormalization to get a good computation of the full non-adiabatic theory.

For the direct $\mathbf{k} = 0$ gaps of c-GaN (the case we study in detail), the relevant energy jump is $\Delta_s\varepsilon = (\hbar^2/2m^*)(\Delta q)^2$, where Δq is the size of the $\mathbf{q}$-grid. The value of m^* for the conduction band is $0.16m_e$, and $\hbar\omega_{\text{LO}}$ is 0.089 eV. A desirable value of δ is 0.1eV, which requires $\Delta q = 0.065\mathring{A}$ to make $\Delta_s\varepsilon < \delta$. However, we find that an 18×18×18 mesh is sufficient. This corresponds to $\Delta q = 0.155\mathring{A}$. The reason why this works is because the grid and the singular point are both centered at $\mathbf{k} = 0$. The integrand is then sampled at symmetric points, an appropriate "tempered mesh" that converges with far less noise to the correct principal value integral. A confirmation that this works comes from the plots of Poncé *et al.* [20]. See for example the middle graph of Fig.6(a), which shows very good convergence for a 20×20×20 grid.

2.3 Correction formulas

The full theory is contained in the perturbative expressions worked out by Vogl [36]. The singular part corresponds to the Fröhlich polaron [2]. In [13, 14] a Vogl expression is studied from an *ab-initio* perspective, and is shown to coincide for small $\mathbf{q}$ with DFPT calculations. All agree that the polaron is the dominant contribution in the small $\mathbf{q}$ region, and needs to be treated carefully.

A polaron describes the coupled system of an electron and phonons. Most often, only zero temperature is considered, but the concept works also at $T > 0$. The most famous case is the Fröhlich, or "large" polaron, present in ionic crystals and polar semiconductors [2]. Fröhlich theory is designed for the bottom of the conduction band where an effective mass approximation $\varepsilon_{\mathbf{k}} = \hbar^2 k^2/2m^*$ is accurate, and for intraband ($n = n'$) coupling only to the polar LO mode, where ω_{LO} has negligible $\mathbf{q}$-dependence. It can also be used for the valence band, which will be discussed later. In the conduction band case, the matrix element $|M|^2$ is $4\pi\alpha(\hbar\omega_{\mathrm{LO}})^2(a_{\mathrm{LO}}/\Omega_0 q^2)$. It is the factor q^{-2} which comes from long range polarization. The distance $a_{\mathrm{LO}} = \sqrt{\hbar/2m^*\omega_{\mathrm{LO}}}$ is of order $10\mathring{A}$, larger than the zero-point root mean square vibrational displacement $u_{\mathrm{LO}} = \sqrt{\hbar/2M_{\mathrm{red}}\omega_{\mathrm{LO}}}$ by the large factor $\sqrt{M_{\mathrm{red}}/m^*}$, where M_{red} is the appropriate ionic reduced mass. The Fröhlich coupling constant is $\alpha = V_c/\hbar\omega_{\mathrm{LO}}$, where V_c is a Coulomb interaction strength $V_c = e^2/(8\pi\tilde{\epsilon}_0\epsilon^* a_{\mathrm{LO}})$. The $\tilde{\epsilon}_0$ is the permittivity of free space, and the ϵ^* is defined in terms of the low and high frequency dielectric constants as $1/\epsilon^* = 1/\epsilon_\infty - 1/\epsilon_0$. Since we are interested in the renormalization of the band gap, we focus on the band extrema at $\mathbf{k} = 0$. For a non-degenerate band (*e.g.* the conduction band), the Fröhlich contribution to the renormalization at temperature T is (see [87] for the $T = 0$ result)

$$\begin{aligned}
[E_{\mathrm{kc}} - \varepsilon_{\mathrm{kc}}]_{\mathrm{Fr},\mathbf{k}=0} =& -\frac{\alpha\hbar\omega_{LO}}{2\pi^2 a_{LO}} \int_0^{q_F} \frac{4\pi q^2 dq}{q^2} \left[\frac{n_B(T)+1}{q^2 + a_{LO}^{-2}} + \frac{n_B(T)}{q^2 - a_{LO}^{-2}} \right] \\
=& -\alpha\hbar\omega_{LO} \left\{ \frac{\tan^{-1}(q_F a_{LO})}{\pi/2}[n_B(T)+1] + \frac{1}{\pi}\ln\left|\frac{q_F - a_{LO}^{-1}}{q_F + a_{LO}^{-1}}\right| [n_B(T)] \right\}
\end{aligned} \tag{2.2}$$

where ω_{LO} is the longitudinal optical frequency, and $n_B(T) = 1/[\exp(\hbar\omega_{LO}/kT) - 1]$ is the Bose- Einstein distribution. The radius of integration is q_F. This and other radii in reciprocal space used in this work, together with their approximate values, are included in Table 1. In most polaron studies, the approximation $q_F \to \infty$ is used. One might instead use the radius q_D of the Debye sphere whose volume is the BZ volume. However, the integrand becomes inaccurate if q_F is larger than the radius q_m where the effective mass treatment works well. The first term of Eq.(2.2) corresponds to phonon emission. It is included in Fröhlich's treatment at $T = 0$. The second term is only present at non-zero temperature and it corresponds to phonon absorption. At $T = 0$, extending the sum over the Brillouin Zone to infinity, the famous result [87] is $E_{\mathbf{k}=0,c} - \varepsilon_{\mathbf{k}=0,c} = -\alpha\hbar\omega_{\mathrm{LO}}$.

In the adiabatic approximation, the term in brackets [ ] in Eq. (2.2) is replaced by

$(2n_B+1)/(q^2 - i2m^*\delta/\hbar^2)$, and then the real part is taken:

$$\begin{aligned}[E_{\mathrm{kc}} - \varepsilon_{\mathrm{kc}}]_{\mathrm{Ad,k=0}} &= -\frac{\alpha\hbar\omega_{LO}}{2\pi^2 a_{LO}}\Re\int_0^{q_F}\frac{4\pi q^2 dq}{q^2}\left[\frac{2n_B(T)+1}{q^2 - i2m^*\delta/\hbar^2}\right] \\ &= -\frac{\alpha\hbar\omega_{LO}}{a_{LO}}\Re\frac{1}{\pi z}\ln\left(-\frac{q_F - z}{q_F + z}\right)[2n_B(T)+1]\end{aligned} \tag{2.3}$$

where $z = \sqrt{2m^*\delta/\hbar^2}\exp(i\pi/4)$. Subtracting this term from the *ab-initio* calculation and adding the correct Fröhlich contribution (2.2), with an appropriate radius of integration q_c, we obtain in principle our desired correction.

In the adiabatic approximation, the denominator (for $\mathbf{k}=\mathbf{0}$ and $\mathbf{q}\to 0$) is $i\delta$, i.e., pure imaginary. Because the energy renormalization is given by the real part, the central mesh-cell contribution is 0 in the adiabatic approximation. This mis-represents a converged adiabatic calculation (like Eq.(2.3)). We should not subtract the part of Eq.(2.3) that represents the missing contribution from the central grid cell.

To determine the optimal integation radius q_c to use for the correction, we calculate the difference between the Fröhlich-polaron contribution Eq. (2.2) and the adiabatic approximation Eq. (2.3) (replacing ω_{LO} with $i0.1$ eV) for different radii q_F of integration. We denote q_c the radius for which the curves differ by less than 1 meV for all temperatures, and we refer to it as the convergence radius. The adiabatic expression is a good approximation to the Fröhlich polaron for radii greater than q_c. The analysis can be separated in two cases:

(i) $q_{\mathrm{mesh}} < q_c$. This is the case in our calculation, for both the conduction and valence band. It is discussed in the Appendix.

(ii) $q_{\mathrm{mesh}} \geqslant q_c$. We will illustrate this case with our c-GaN calculation, although the expression in the Appendix is required for a more precise result. Since the adiabatic DFT calculation has no contribution from the central cell, it does not have to be subtracted. Therefore, the correction is just given by (2.2) with $q_F = q_{\mathrm{mesh}}$, which is a good enough radius of integration since $q_{\mathrm{mesh}} \geqslant q_c$. Case (i) is similarly simple, but the correction involves an extra term.

As long as the effective mass approximation is accurate, both Eq.(2.2) and Eq.(2.3), and thus q_c, will be accurate. If there are worries about the applicability of the effective mass approximation, then one coud use the small $\mathbf{q}$, intraband, LO phonon part of Eq. (2.1) with $\omega_{\mathrm{LO}} \to i\delta$ to subtract the adiabatic contribution, without using the effective mass approximation. However, this is not necessary for the accuracy of a few meV we are interested in.

We first study how the correction changes with the integration radius. Then we calculate the temperature dependence of the minimum of the conduction band, including the polaron correction. Finally, we study the valence band. We use ABINIT [88, 89] to carry out the *ab-initio* calculations.

Symbol	Definition	Approximate choices
q_F	Upper limit of Fröhlich integral Eq. (2.2)	$\frac{4\pi}{3}q_F^3 = \Omega_{BZ}$
q_c	Convergence radius beyond which $\Delta\varepsilon_{\mathrm{Fr}} \approx \Delta\varepsilon_{\mathrm{Ad}}$	Conduction $1.2q_{\mathrm{mesh}}$ Valence $2.5q_{\mathrm{mesh}}(T = 0)$
q_m	Wavevector limit for effective mass approximation	Conduction $1.2q_{\mathrm{mesh}}$ Valence $2.5q_{\mathrm{mesh}}$
q_{mesh}	$\frac{4\pi}{3}q_{\mathrm{mesh}}^3 = \frac{\Omega_{BZ}}{N}$	$N = 18 \times 18 \times 18$

Table 1: Definitions and approximate values of the different radii in momentum space used in this work. The convergence radius q_c determines the region in which the correction has to be applied. Note the similarity between q_c and q_m^* both for the valence and conduction bands. However, $q_c = 6.3q_{\mathrm{mesh}}$ at $T = 1000$K because of the absorption term in the valence band. See the discussion in the Appendix.

2.4 Results and Discussion

We use Troullier-Martins pseudopotentials for both Gallium and Nitrogen, in the Perdew-Wang [90] parameterization of local density approximation (LDA), generated using the fhi98PP code [91]. The Ga-3d electrons are included as valence electrons. We use a 6×6×6 Monkhorst-Pack (MP) [92] centered **k**-point grid in our calculations, and a high energy cutoff of 1700 eV in order to converge the total energy to less than 0.018 meV per atom (h-GaN converges well with 1400 eV). The resulting lattice constant is $a = 4.499\mathring{A}$. Experimental values are 4.507 $\mathring{A}$ [93] and 4.52 $\mathring{A}$ [94]. The phonons and electron-phonon interaction matrix elements use DFPT in the rigid-ion approximation, as specified by ABINIT, to speed up calculations. It has been shown reliable for simple crystals [71]. We use an 18×18×18 MP **q**-point grid, and the adiabatic + iδ approximation, with $\delta = 0.1$ eV.

2.5 Conduction band

The conduction band is very isotropic, with an effective mass $m^* = 0.16m_e$. The differences between the Fröhlich contribution and the adiabatic approximation are shown in Fig. 1. Going beyond a radius of integration of $q_F = 0.068\ 2\pi/a$, the curves differ by less than 1 meV for all temperatures. Therefore, $q_c = 0.068\ 2\pi/a$. To obtain an accurate result, the analytic integration in (2.2) should be restricted to a small radius q_c close to $\mathbf{q} = 0$, because the effective mass approximation is only valid close to $\mathbf{q} = 0$. From Table 1, we see that for the conduction band $q_m = q_c$, so the method is indeed accurate.

Using the method of case (ii), the corrections at $T = 0$ K and $T = 1000$ K are -19 meV and -22 meV, respectively. For the more precise method (i) described in the Appendix, since actually $q_{\mathrm{mesh}} < q_c$, the corrections are -17 meV at $T = 0$ K and -17 meV at T = 1000 K. Fig. 2 shows the adiabatic calculation of $E_{\Gamma c} - \varepsilon_{\Gamma c}$ done with ABINIT, the corrected result, and the total Fröhlich contribution at finite temperature (taking as q_F the radius of the BZ).

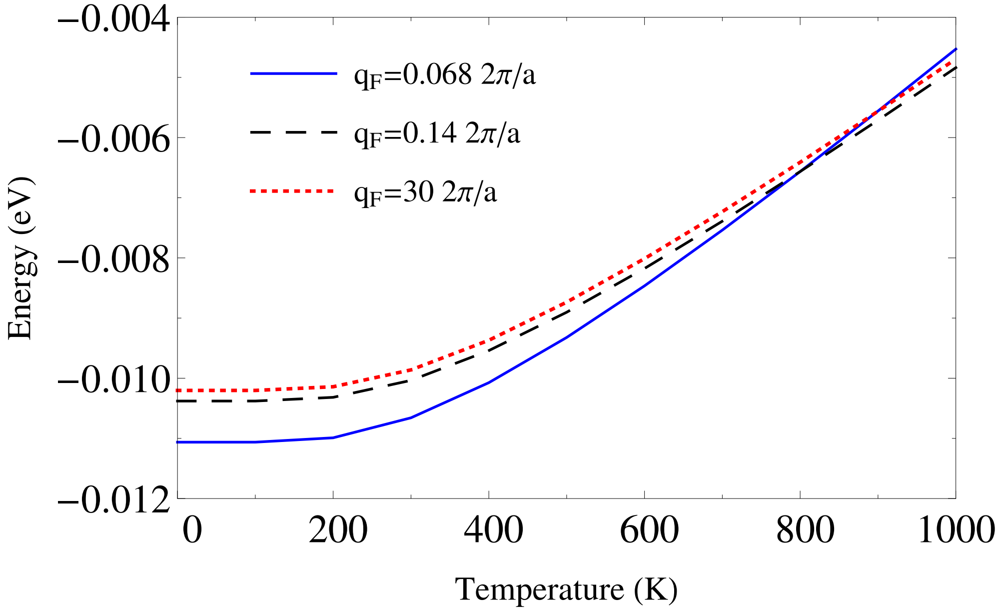


Figure 1: Difference between the Fröhlich contribution and the corresponding adiabatic approximation with $i\delta = i0.1$eV for the conduction band, for different radii q_F of integration.

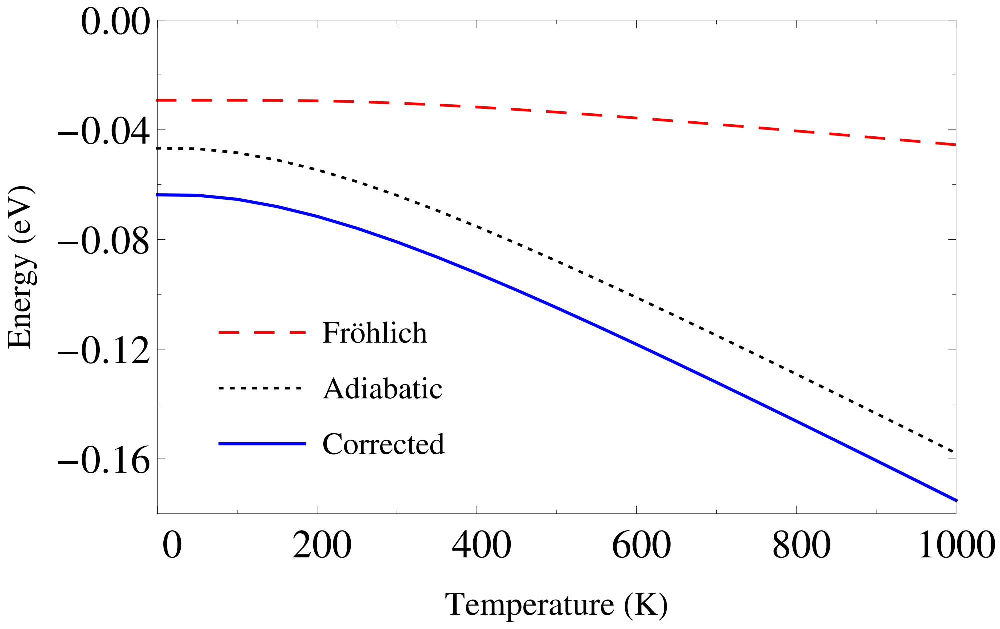


Figure 2: Temperature dependence of the conduction band: direct adiabatic calculation (including interband parts in Eq. (2.1) in a 18×18×18 MP grid with $\delta = 0.1$eV (dotted), corrected calculation (full), and the pure Fröhlich term at finite temperature (dashed).

2.6 Valence band

For the valence band, the correction is more complicated because of two factors: bands are degenerate and they are not isotropic. Since we are not considering spin-orbit coupling, the top of the valence band is triply degenerate. The $\mathbf{k}\cdot\mathbf{p}$ method fixes the valence band energy dispersion to be the eigenvalues of [95]

$$D = \begin{bmatrix} Ak_z^2 + B(k_y^2 + k_z^2) & Ck_xk_y & Ck_xk_z \\ Ck_xk_y & Ak_y^2 + B(k_x^2 + k_z^2) & Ck_yk_z \\ Ck_xk_z & Ck_yk_z & Ak_z^2 + B(k_x^2 + k_y^2) \end{bmatrix} \tag{2.4}$$

Comparing with the *ab-initio* calculation, we obtain $A = -3.14\hbar^2/m_e$, $B = -0.61\hbar^2/m_e$ and $C = -3.49\hbar^2/m_e$. These correspond, for example, to effective masses $m^* = 0.16m_e$ and $0.82m_e$ in the (100) direction.

Renormalization does not lift the triple degeneracy of the top of the valence band. For degenerate and isotropic bands, Trebin and Rössler [48] use the $\mathbf{k}\cdot\mathbf{p}$ method to generalize Fröhlich's result (giving analytic expressions). Following their procedure, we write the band renormalization in the case, without requiring band isotropy:

$$\begin{aligned}[E_{\mathbf{k}v} - \varepsilon_{\mathbf{k}v}]_{\mathrm{Fr},\mathbf{k}=0} = \frac{e^2}{4\pi\tilde{\epsilon_0}N\Omega_0}\frac{2\pi\hbar\omega_{LO}}{\epsilon^*}\sum_q^{q<q_F}\sum_{s=1}^{3}\frac{1}{q^2}|\langle n_s(\mathbf{q})|n\rangle|^2\times \\ \times\Re\left[\frac{n_B(T)+1}{\varepsilon_{\Gamma v} - \varepsilon_{\mathbf{q}n_s} + \hbar\omega_{LO}} + \frac{n_B(T)}{\varepsilon_{\Gamma v} - \varepsilon_{\mathbf{q}n_s} - \hbar\omega_{LO} + i\Delta}\right]\end{aligned} \tag{2.5}$$

where n_s indicates the degenerate bands and $|n_s(\mathbf{q})\rangle$ are the eigenstates of Eq.((2.4)) at $\mathbf{q}$. The initial state $|n\rangle$ can be any of the $\mathbf{k}=0$ degenerate eigenstates; all give the same answer. We include a small $i\Delta = i0.001$eV only in the second denominator to allow a good numerical evaluation of the principal part. Now, $\varepsilon_\Gamma - \varepsilon_\mathbf{q} > 0$ and the factors $n_B + 1 - f$ and $n_B + f$ have become n_B and $n_B + 1$ instead of $n_B + 1$ and n_B, respectively (because $f_v = 1$). As a result, we get an extra minus sign with respect to the conduction band; the band renormalization is now positive.

The adiabatic *ab-initio* calculation gives a band renormalization of 62 meV at T=0 K and 185 meV at T=1000 K. The valence band has in case (i) ($q_{\mathrm{mesh}} < q_c$) and the method is described in the Appendix. The correction is 28 meV and 11 meV at $T = 0$ K and $T = 1000$ K, respectively. Fig. 3 shows the results for the valence band analogous to Fig. 2.

Therefore, we see that the Fröhlich correction provides approximately a constant shift of the renormalization by -20 meV for all temperatures in the conduction band. For the valence band, it is +28 meV at $T = 0$ and it decreases to 11 meV at $T = 1000$ K. The Fröhlich correction is about 30% of the total ZPR of both the conduction and valence band. At 1000 K, the corrections are between 6-9% of the total renormalization.

2.7 Conclusions

Our procedure allows a calculation of the whole electronic renormalization of a polar material, using the adiabatic approximation with an $i\delta$, and an affordable mesh. The $i\delta$ cures

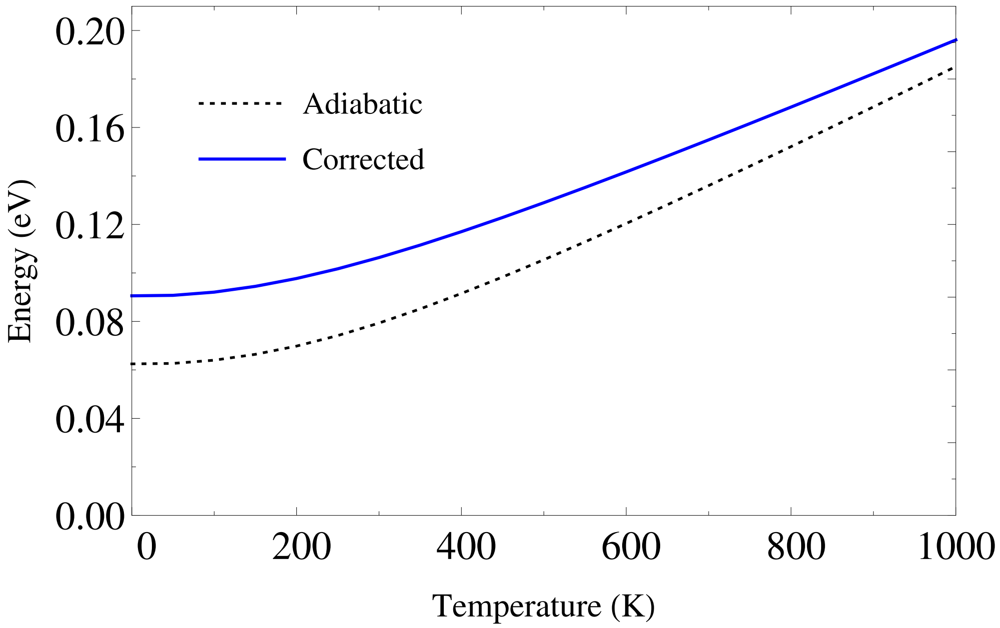


Figure 3: Temperature dependence of the valence band: direct adiabatic calculation with an 18×18×18 MP grid with $\delta = 0.1$eV (dotted) and the corrected calculation (full).

the divergence of the adiabatic approximation at the extrema of polar materials, but does not correctly include Fröhlich polaron corrections. Then we add the Fröhlich polaron contribution in the central mesh-cell, omitted in the DFT adiabatic calculation due to the pure imaginary denominator $i\delta$. Finally, we add the difference between the Fröhlich and adiabatic expressions if $q_c > q_{\text{mesh}}$. This method is then a combination of the adiabatic and non-adiabatic approximations. We avoid using a very dense **q**-grid by treating the Fröhlich polaron analytically. By this method, we calculated for c-GaN the electron-phonon temperature dependence of the minimum of the conduction band and the maximum of the valence mand. At high temperatures, the method is approximate for the valence band. The correction is a significant fraction of the total electron-phonon renormalization, although it decreases as the temperature increases.

2.8 Appendix

2.8.1 Method, case (i)

Here we describe the method for case (i), where $q_{\text{mesh}} < q_c$. The difference between the Fröhlich contribution and the adiabatic expression is small beyond a radius q_c. Because of the $i\delta$ there is no adiabatic contribution from the central cell (in c-GaN, a truncated octahedron), which can be approximated by a sphere of radius q_{mesh}. Instead of subtracting the adiabatic contribution from 0 to q_c, we have to subtract it from q_{mesh} to q_c. Therefore we need to determine what is the adiabatic contribution in this region.

2.8.1.1 Conduction band

The correction is given by

$$\begin{aligned}\Delta(E_{\mathbf{k}c} - \varepsilon_{\mathbf{k}vc})_{\mathbf{k}=0} = &-\alpha\hbar\omega_{LO}\left\{\frac{\tan^{-1}(q_c a_{LO})}{\pi/2}[n_B(T)+1] + \frac{1}{\pi}\ln\left|\frac{q_c - a_{LO}^{-1}}{q_c + a_{LO}^{-1}}\right|[n_B(T)]\right\} \\ &+ \frac{\alpha\hbar\omega_{LO}}{a_{LO}}\Re\frac{1}{\pi z}\ln\left(\frac{q_c - z}{q_c + z}\frac{q_{\text{mesh}} + z}{q_{\text{mesh}} - z}\right)[2n_B(T)+1]\end{aligned} \tag{2.6}$$

Note that this is just the difference between equations (2.2), evaluated between 0 and q_c, and (2.3), evaluated between q_{mesh} and q_c. The plot of the adiabatic calculation with $\delta = 0.1\text{eV}$ and the correction is included in Fig. 2 in the main text.

The effective mass varies between $0.157m_e$ and $0.175m_e$ when taking different radii up to $0.067\ 2\pi/a$. The difference of the Fröhlich contribution for these two effective masses in the correction is only 0.6 meV or less for all temperatures. So the change of the effective mass with k causes negligible errors in our method.

2.8.1.2 Valence band

The expression we use for the correction is

$$\begin{aligned}\Delta(E_{\mathbf{k}v} - \varepsilon_{\mathbf{k}v})_{\mathbf{k}=0} &= +\frac{e^2}{4\pi\tilde{\epsilon_0}\Omega_0}\frac{2\pi\hbar\omega_{LO}}{\epsilon^*}(I_{\text{Fr}} - I_{\text{Ad}}), \text{where} \\ I_{\text{Fr}} &= \frac{\Omega_0}{(2\pi)^3}\sum_{s=1}^{3}\int_0^{q_c}\frac{d^3q}{q^2}|\langle n_s(\mathbf{q})|n\rangle|^2\Re\left[\frac{n_B(T)+1}{\varepsilon_{\Gamma v} - \varepsilon_{\mathbf{q}n_s} + \hbar\omega_{LO}} + \frac{n_B(T)}{\varepsilon_{\Gamma v} - \varepsilon_{\mathbf{q}n_s} - \hbar\omega_{LO} + i\Delta}\right] \\ I_{\text{Ad}} &= \frac{\Omega_0}{(2\pi)^3}\sum_{s=1}^{3}\int_{q_{\text{mesh}}}^{q_c}\frac{d^3q}{q^2}|\langle n_s(\mathbf{q})|n\rangle|^2\Re\left[\frac{2n_B(T)+1}{\varepsilon_{\Gamma v} - \varepsilon_{\mathbf{q}n_s} + i\delta}\right]\end{aligned} \tag{2.7}$$

Here, I_{Fr} corresponds to Eq. (2.5) and I_{Ad} is the corresponding adiabatic equation (these expressions are analogous to Eq. (2.2) and (2.3) for the conduction band). As a reminder, we use $\delta = 0.1$ eV, the value that was used in the DFT calculation, and $\Delta = 0.001$ eV to calculate the principal value adequately. The values of $\varepsilon_{\mathbf{q}n_s}$ and $n_s(\mathbf{q})$ come from diagonalizing the matrix in Eq. (2.4), so Eq. (2.5) can be readily calculated, albeit not having an analytic expression as for the conduction band.

We study the convergence radius q_c in the same way as we did for the conduction band. At $T = 0$, we obtain $q_c = 2.5q_{\text{mesh}}$ (for larger radii, the correction changes by less than 1 meV). At $T = 1000K$, however, we observe a difference of 6 meV between the correction at $q = 2.5q_{\text{mesh}}$ and the convergence radius $q_c = 6.3q_{\text{mesh}}$. What occurs is that the absorption term in the Fröhlich integral changes more with the radius of integration than the emission

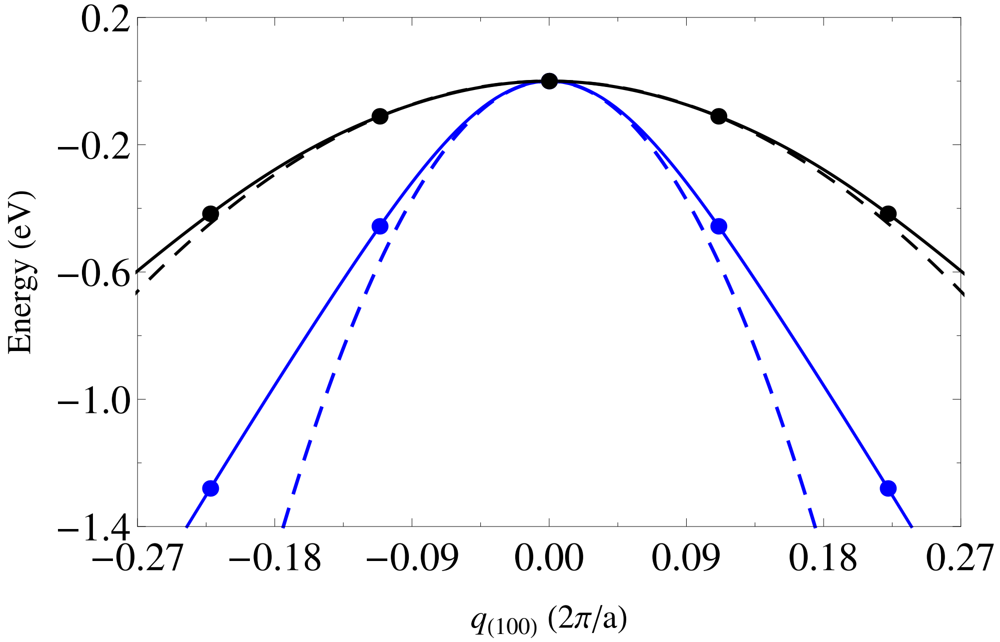


Figure 4: Plot of the heavy and light holes in the (100) direction (full lines) with their corresponding effective mass fit (dashed lines). The dots indicate the sampling points in the *ab-initio* calculation.

term in I_{Fr}, and I_{Ad}. While the absorption term does not contribute at $T = 0$ because it is suppressed by $n_B(T)$, it does at higher temperatures.

From Fig. 4 we see that the effective mass approximation is accurate up to around $q_{m^*} = 2.5 q_{\mathrm{mesh}} = 0.13\ 2\pi/a$ for the heavier masses, the same value we found for q_c. For the lighter mass, the effective mass approximation breaks down for a smaller q, but the convergence radius is much smaller (as for the conduction band). Varying the effective mass of the light hole, we can see that the error introduced is less than 1 meV (assuming a contribution of one third for each band; see the following paragraph). Therefore, our method is accurate for the valence band for temperatures below $500K$, and less accurate for higher temperatures.

In the isotropic case, it is shown in [48] how the renormalization is an average of the light and heavy holes at **k**=0. We can average the effective mass of each band over a sphere using Eq. (2.4). We obtain $m^*_{1,\mathrm{av}} = 0.14 m_e$, $m^*_{2,\mathrm{av}} = 0.94 m_e$ and $m^*_{3,\mathrm{av}} = 1.72 m_e$. Assuming each band is isotropic, we can calculate the renormalization by using the standard Fröhlich result Eq. (2.2) for each band and then averaging over the bands. Integrating from 0 to q_F, with $0 < q_F < 6.3 q_{\mathrm{mesh}}$, the renormalization differs from Eq. (2.5) by less than 1 meV at $T = 0$. At $T = 1000$ K, they differ by 5 meV or less, depending on the value of q_F. Therefore, at $T = 0$ the renormalization can be just calculated by averaging over the Fröhlich contribution of the average effective masses. At higher temperatures, using averaged isotropic masses is less accurate.

2.8.2 Piezopolaron

In a piezoelectric material, a strain induces a macroscopic electric field. If the strain is produced by a long-wavelength acoustic phonon, the coupled system of an electron and the acoustic phonon is known as the piezopolaron [96]. It turns out that, like the Fröhlich polaron effect, the piezopolaron also causes a divergent intraband term at band extrema in the adiabatic approximation. Adding an artificial $i\delta$ removes the divergence, but does not correctly include the true non-adiabatic behavior, namely, part of the acoustic contribution to zero-point renormalization, and a new low T contribution scaling as T^2 with a positive coefficient (increasing the gap at low T). This topic is covered in [38]. There we show that both the zero point contribution and the high T contribution are quite small, and the T^2 term only plays a dominant role at very low T. Therefore there is no need to add an analytic correction for piezo-effects to the result obtained from adiabatic + $i\delta$ approximation. To clarify a little, the formula for band renormalization from intraband acoustic phonon processes, at a band extremum, is

$$[E_{\mathrm{k}v} - \varepsilon_{\mathrm{k}v}]_{\mathrm{acoustic},\mathbf{k}=0} = \frac{1}{N}\sum_{\mathbf{q}s} |\langle \mathbf{q}|V_1(\mathbf{q}s)|\mathbf{0}\rangle|^2 \left[\frac{1+n_{\mathbf{q}s}}{-\hbar^2q^2/2m^* - \hbar v_s q} + \frac{n_{\mathbf{q}s}}{-\hbar^2q^2/2m^* + \hbar v_s q}\right], \tag{2.8}$$

where v_j is the velocity of sound. This keeps the small q part of the theory only. The piezoelectric matrix element is

$$\langle \mathbf{q}|V_1(\mathbf{q}s)|\mathbf{0}\rangle = -\frac{e}{4\pi\tilde{\varepsilon}_0}\frac{\mathbf{q}.\mathbf{e}_m.(\mathbf{q}\delta\mathbf{R})}{q^2\epsilon_\infty} \tag{2.9}$$

where $\delta\mathbf{R}$ is the acoustic vibration or acoustic phonon amplitude and $\mathbf{e}_m$ is the electromechanical or piezoelectric tensor (see [41] for a derivation). The acoustic phonon displacement factor δR is $\sqrt{\hbar/M_{\mathrm{tot}}v_s q}$. Therefore the squared matrix element $|\langle \mathbf{q}|V_1(\mathbf{q}s)|\mathbf{0}\rangle|^2$ behaves as $1/q$. The adiabatic approximation replaces the factor [ ] by the approximation $-[(1+2n_{\mathbf{q}s})/(\hbar^2q^2/2m^*)]$. Therefore, at low T, the sum over $\mathbf{q}$ becomes, at small q, $-\int dq q^2(1/q)[(1+2k_BT/\hbar v_s q)/(\hbar^2q^2/2m^*)]$. This is valid for when the acoustic phonon energy $\hbar v_s q$ is smaller than k_BT. The zero-point part diverges logarithmically, and the thermal part as $1/q$. The correct non-adiabatic version of this is $\int dq q^2(1/q)[-1/\hbar v_s q + (2k_BT/\hbar v_s q)(\hbar^2q^2/2m^*)/(\hbar v_s q)^2]$. Both zero-point and thermal parts converge as $\int dq$. It turns out that the difference between the true non-adiabatic contribution and the artificially converged adiabatic part (adding $+i\delta$ in the denominator) is small, except for a small (but interesting) non-adiabatic T^2 term at very low T which has little effect at higher T.

Chapter 3

Quasiparticles and phonon satellites in spectral functions of semiconductors and insulators: Cumulants applied to full first principles theory and Fröhlich polaron.

The electron-phonon[5] interaction causes thermal and zero-point motion shifts of electron quasiparticle (QP) energies $\epsilon_k(T)$. Other consequences of interactions, visible in angle-resolved photoemission spectroscopy (ARPES) experiments, are broadening of QP peaks and appearance of sidebands, contained in the electron spectral function $A(k,\omega) = \frac{-1}{\pi}\text{Im}G_R(k,\omega)$, where G_R is the retarded Green's function. Electronic structure codes (*e.g.* using density-functional theory) are now available that compute the shifts and start to address broadening and sidebands. Here we consider MgO and LiF, and determine their non-adiabatic Migdal self energy. The spectral function obtained from the Dyson equation makes errors in the weight and energy of the QP peak and the position and weight of the phonon-induced sidebands. Only one phonon satellite appears, with an unphysically large energy difference (larger than the highest phonon energy) with respect to the QP peak. By contrast, the spectral function from a cumulant treatment of the same self energy is physically better, giving a quite accurate QP energy and several satellites approximately spaced by the LO phonon energy. In particular, the positions of the QP peak and first satellite agree closely with those found for the Fröhlich Hamiltonian by Mishchenko *et al.* (2000) using diagrammatic Monte Carlo. We provide a detailed comparison between the first-principles MgO and LiF results and those of the Fröhlich Hamiltonian. Such an analysis applies widely to materials with infrared(IR)-active phonons.

3.1 Introduction

The notion of a single-particle (quasiparticle, or QP) spectrum ϵ_k for electrons (k is short for all needed quantum numbers - wavevector, band, spin - $\mathbf{k}, n, \sigma$) underlies much of solid

---

[5] This work was done in collaboration with P. B. Allen, G. Antonius, L. Reining, A. Miglio and X. Gonze. Only minor modifications were introduced to Ref. [24].

state physics. Evidence for the existence of such quasiparticle spectra [97, 98] relies on experiment. Optical experiments, combined with theoretical guidance, *e.g.* the "empirical pseudopotential method" [99], have been used for decades, and allowed extraction of accurate ϵ_k from reflectivity data for simple semiconductors. Excitonic effects cause deviation from an independent-particle interpretation, but theory can determine their consequences and help to extract one-electron properties. Angle-resolved photoemission spectroscopy (ARPES) provides a more direct QP spectrum [100, 101]. The data can be approximately related to the rigorously defined one-particle spectral function, obtained from the retarded Green's function $G(k, \omega)$ as [98, 102]

$$A(k, \omega) = -\frac{1}{\pi}\Im m G_R(k, \omega), \tag{3.1}$$

where $G_R(k, \omega)$ is the Fourier transform

$$\int dt \exp(i\omega t) G_R(k, t) \tag{3.2}$$

of the retarded Green's function

$$G_R(k, t) = -i\langle\{c_k(t), c_k^+(0)\}\rangle\theta(t), \tag{3.3}$$

and $\{a, b\}$ is an anticommutator.

When the spectral function $A(k, \omega)$ exhibits a strong peak that correlates with the corresponding single-particle theory, this defines a QP energy. The total spectral weight $\int d\omega A(k, \omega)$ is 1, but the QP peak has reduced weight $Z_k < 1$. It is broadened, and accompanied by features at other energies. When clearly differing from a structureless background, these features are called satellites. This has been seen in many photoemission experiments, for example, Ref. [16] for polaron satellites, and Ref. [103] for plasmon satellites. Thus $A(k, \omega)$ contains more information than just the QP energy ϵ_k. Full interpretation is a challenge to theory; progress on plasmon [104, 105] and polaron [10] satellites in real semiconductors is occurring.

On the basis of a one-band analysis for metals, Migdal [106] showed that the electron self energy due to phonons has important low temperature dynamical effects, which can be accurately described by a lowest-order self energy diagram Σ_M. Vertex corrections can be omitted because of the small ratio $\omega_{\rm ph}/E_F$ of phonon energies to the Fermi energy. Even though this argument that allows to neglect vertex corrections does not strictly apply in semiconductors, we will use the term "Migdal" to indicate the Migdal formula, Σ_M, given later as Eq.(3.6), where it is labeled $\Sigma^{\rm Fan}$, referring to earlier works on insulators by Fan in the fifties [28, 107]. Using Dyson's equation, the corresponding Green's function is

$$G^D(k, \omega) = G_0(k, \omega) + G_0(k, \omega)\Sigma_M(k, \omega)G^D(k, \omega). \tag{3.4}$$

When inserted into Eq. (3.1), it gives a spectral function which we will call "Dyson-Migdal", $A^{\rm DM}$. Given the success of Migdal theory in metals and the expected success of low order perturbation theory for most electron-phonon problems, it is embarrassing to realize [72,108, 109] that $A^{\mathrm{DM}}(k,\omega)$ often agrees poorly with measured $A(k,\omega)$. Typically, only one distinct satellite is found in the side-band spectrum, while in reality several satellites are possible [16], corresponding to emission of several phonons. Even worse, the QP peak is misplaced [110].

An alternative approach, also approximate, involves a cumulant treatment (see e.g. Refs. [25, 111–113] and pp. 410–415 of Ref. [98]). This will be denoted as G^{C} for the Green's function, and A^{C} for the spectral function. The cumulant consists of an exponential representation of the Green's function, $G^{C}(k,t) = G_0(k,t)e^{C(k,t)}$. By equating this expansion to a perturbation associated with the Dyson's equation (see the Appendix, derivation of Eqs. (3.18-3.22)), an expression for the cumulant $C(k,t)$, and thus for the Green's function, can be obtained. As we will see later, the cumulant gives better results than the Dyson approach for the spectral funcion of polar semiconductors. This seems to be related to the fact that, through the exponential, higher order terms (although approximately) are included as well. The cumulant approach and equations are discussed in detail in Sec. 3.4.

The cumulant was advocated for describing plasmon satellite effects in metals by Steiner *et al.* [103] and Hedin [114]. For electron-phonon effects, there is early work by Mahan [115] who used cumulant methods for the Urbach tail on optical $\sigma(\omega)$, and Dunn [108] for the polaron spectral function. Gunnarsson *et al.* [109] introduced the cumulant method for electron-phonon effects in metals. More recently, Verdi *et al.* [10] applied the cumulant method to doped semiconductors, where both plasmon and phonon effects occur. Here we apply the cumulant method to undoped semiconductors with electron-phonon renormalization but no dynamical electron-electron coupling.

To further motivate the interest to work with the cumulant approach instead of the Dyson-Migdal one, consider Fig. 5, which illustrates where the QP peak appears in various treatments of the Fröhlich polaron problem [2,25]. Lowest-order Rayleigh-Schrödinger perturbation theory puts the renormalized polaron energy below the conduction band minimum by $\alpha\omega_{LO}$, where α is the Fröhlich coupling constant. An accurate high-order treatment [47] gives a surprisingly similar answer, shown as the blue squares on the graph. This is where the QP peak of the spectral function should be located. If the Migdal self energy Σ_M is evaluated at $\omega = \epsilon_k$, the unperturbed energy, the corresponding energy shift $\Delta\epsilon_{k=0} = \Sigma_M(0,\epsilon_k)$ agrees with the Rayleigh-Schrödinger result $-\alpha\omega_{LO}$. However, the actual QP peak of the spectral function A^{DM} is located at a fairly different energy, shown by the red circles on the graph [25,110]. The cumulant method used in the present Chapter puts the QP peak exactly on the Rayleigh-Schrödinger line, close to the correct polaron answer. The cumulant method also greatly improves the position of the first satellite peak, as will be shown later. It also gives additional peaks, corresponding to multiphonon excitations. However, they differ in location and strength from the other peaks in the accurate Fröhlich spectral function [47].

Another improvement given by the cumulant method is a much more reasonable value of the QP spectral weight Z, defined later in Eq.(3.11) and shown in Fig. 6.

Since the characteristic energy of plasmons is much bigger than that of phonons, the resolution needed to see plasmon satellites in photoemission experiments has been available for many years. The valence photoemission spectra of alkali metals, that exhibit multiple

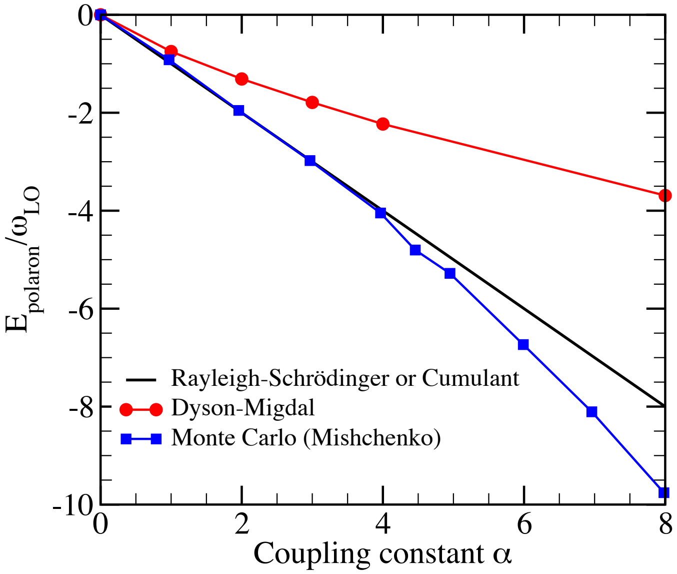


Figure 5: Quasiparticle energy of the Fröhlich polaron, as a function of the Fröhlich coupling constant, in units of ω_{LO}: accurate results from Ref. [47] (blue squares), from the cumulant approach (which agrees with lowest-order Rayleigh-Schrödinger theory) (black line), and from Dyson-Migdal spectral function [25, 110] (red circles).

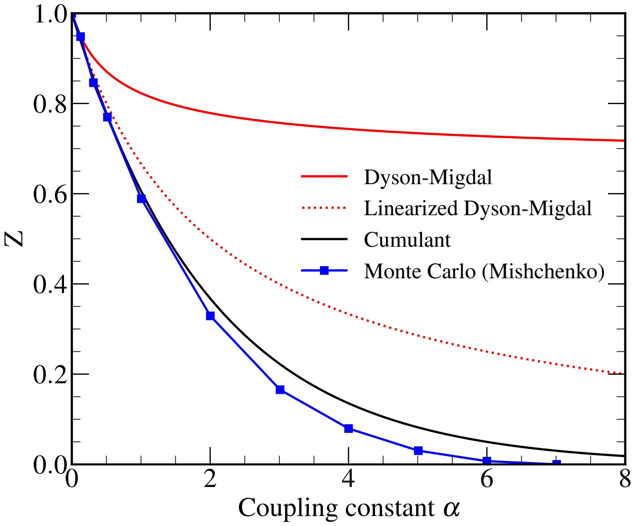


Figure 6: Quasiparticle spectral weight $Z_{k=0}$ of the Fröhlich polaron, as a function of the Fröhlich coupling constant. The blue squares are accurate Monte Carlo results from Ref. [47]. The solid red line comes from the full Dyson-Migdal spectral function, Eq. (3.11). The dotted red line comes from the linearized approach, Eq.(3.14). The black line is the cumulant result using Eq. (3.26) with the retarded Green's function.

plasmon satellite structures caused by the Coulomb interaction, was modeled from first principles by Aryasetiawan *et al.* [104] using a cumulant treatment. More recently, Kheifets *et al.* [116] and Guzzo *et al.* [105] did ARPES for valence electrons in Si, and observed plasmon satellites. Their calculated spectrum based on a cumulant approach G^{C}, agrees much better with their data than a theory based on the Dyson G^{D}. Similar observations were made for graphene, doped graphene and graphite [117, 118], and for several other materials [103, 119]. The homogeneous electron gas has also been studied using the cumulant approach [120–122].

Electron-phonon interaction (EPI) effects are ubiquitous in solids [1], but were not seen in ARPES until improved resolution became available. In metals, temperature shifts of photoemission linewidths were seen for surface states in Cu [123] and Ag [124]. Phonon-induced "kinks" in the quasiparticle dispersion were seen by photoemission from surface states [125] of Be(0001) [126, 127] and Mo(110) [128]. Bulk electron-phonon effects were seen by ARPES in superconducting and normal Pb [129]. These experiments on metals did not resolve EPI-induced satellite structures. Recent ARPES experiments on monolayer metallic (and superconducting) FeSe, grown on both $SrTiO_3$ [130] and rutile TiO_2 [131], have seen replica bands displaced by the energy of a phonon of the substrate. ARPES studies of EPI effects on electrons near the band gaps of semiconductors are now available, with resolved EPI-induced satellites for electrons and holes doped into non-metals, for bulk [16], surface [17, 18], and interface [19] doped regions.

First-principle studies of the EPI effects on the electronic structure of insulators (zero-point motion as well as temperature dependence) have mostly used perturbation theory to second order in displacement (Fan and Debye-Waller terms) [3]. The focus has mostly been on the quasiparticle shift in simple solids, and occasionally on the quasiparticle broadening [1, 15, 20–23, 68, 70, 71, 132–143]. Polarons, by contrast, have been studied for a long time using high-order perturbative treatment of the singular Fröhlich first-order matrix element, with no second or higher-order matrix elements, and simplified model band structures [25, 144]. When the Fröhlich coupling constant α exceeds ≈ 5, nonperturbative corrections start to be needed, and eventually small polarons may form [145] in real materials where the polaron radius in the continuum approach would be comparable to or smaller than the interatomic distance. Also, defects or higher than harmonic lattice displacement terms cannot be ignored in real materials. Only recently [20] was it noticed that first-principles studies of semiconductors had incorrectly ignored nonadiabatic aspects familiar in polaron literature for the quasiparticle energy shift (also sometimes referred to as renormalization) due to LO phonons in IR-active semiconductors. Perturbation theory had previously been simplified by omission of $\pm\omega_Q$ from denominators. Because of the $1/q$ divergence of the Fröhlich coupling, LO phonon contributions cause unphysical divergences (3D-integral of $1/q^4$ at small q phonon wavevectors) if the $\pm\omega_Q$ pieces are omitted [20].

The literature on large polarons has focused on models such as the Fröhlich Hamiltonian, generally ignoring the existence of multiple phonon branches, non-parabolic electronic bands, Debye-Waller, and interband effects. We advocate unification of the separate skills of polaron and energy-band communities. This has started, with the above-mentioned understanding of the LO-phonon role in first-principles calculations [12, 15, 20], as well as the first-principles approach to the Fröhlich vertex developed by Sjakste *et al.* [13] and Verdi and Giustino [14].

Beyond the computation of quasiparticle shifts, first-principles studies of spectral functions, side bands and satellites have also appeared. Cannuccia and Marini [22, 23] showed that optical data for diamond and polyethylene contain subgap EPI effects also visible in the single-electron theoretical spectral function $A(k,\omega)$. Spectral functions were computed by electron-phonon perturbation theory for the full band structure of C, BN, MgO and LiF by Antonius *et al.* [138]. A satellite is distinctly seen in the spectral function at the top of the valence band of LiF. However, these computations used G^D with a Migdal self-energy. Cumulant studies of EPI effects [108, 115] have recently been revived by Verdi *et al.* [10] who discussed the doped $\mathrm{TiO_2}$ data of Ref. [16], explaining the evolution from polaron to metallic-type EPI-renormalization as the doping level increases. The current Chapter reconsiders the results of Antonius *et al.* [138] using G^{C}, and shows, on the basis of the comparison with the Fröhlich Hamiltonian, how a cumulant calculation improves such results.

Cumulants [146], and their relatives, have been discussed a lot in older [104, 108, 109, 112, 114, 147, 148] as well as more recent [72, 98, 105, 113, 117, 122, 149–152] literature. The improvement they give is incomplete [113, 152, 153]. Despite the abundant literature, quantitative tests for EPI effects in non-metals treated with a cumulant approach are almost nonexistent. We are only aware of the above-mentioned model calculations for the Fröhlich Hamiltonian by Mahan [115] and Dunn [108], and first-principles study of $\mathrm{TiO_2}$ by Verdi *et al.* [10].

The current Chapter deals with the zero-temperature spectral function $A(k,\omega,T=0)$, for the top of the valence band or bottom of the conduction band of two real insulators, MgO and LiF, arising from EPI effects. We first summarize the theoretical background for (a) computation of self energies, and (b) spectral functions. Then we consider the Fröhlich Hamiltonian, and discuss results for several values of the coupling constant α. Subsequently we calculate the self energy and the spectral function of MgO and LiF, using *ab initio* density functional perturbation theory (DFPT) calculations with the code ABINIT [88, 89] to determine phonons and their coupling with electrons. We consider, for the mentioned states, the Dyson-Migdal approach to the spectral function ($A^{\mathrm{DM}}(k,\omega)$), as well as the cumulant approach ($A^{\mathrm{C}}(k,\omega)$), and show that the cumulant results are physically more sensible. The DM approach gives qualitatively very different results, which calls for a reconsideration of the results given in Refs. [22, 23, 138].

3.2 Self-energy

The Hartree atomic unit system is used throughout ($\hbar = m_e = e = 1$). Starting from now, we will use the more explicit notations $\mathbf{k}$ for wavevectors and n for bands, instead of k. Spin is irrelevant in this article. The self energy of an unperturbed state, labeled by wavevector and band, includes two contributions at the lowest order of perturbation theory (quadratic in the strength of the EPI), namely, the Fan self energy and the Debye-Waller self energy [1]:

$$\Sigma(\mathbf{k}n,\omega) = \Sigma^{\mathrm{Fan}}(\mathbf{k}n,\omega) + \Sigma^{\mathrm{DW}}(\mathbf{k}n). \tag{3.5}$$

The Debye-Waller self energy (see its expression in Refs. [1,137], a first-order effect from the second-order matrix elements) is static and real. On the contrary, the Fan self energy is dynamical (but often treated adiabatically). In matrix notation, the Fan self energy is given by $\Sigma = iG\Gamma D$, where G is the electron propagator, Γ is the vertex, and D is the phonon propagator. As often done [106], the vertex is approximated as $\Gamma = 1$. Approximating the electronic Green's function by its non-interacting counterpart, $G = G^{(0)}$, e.g. Kohn-Sham Green's function without electron-phonon corrections, corresponds to a non-self-consistent treatment, and gives the standard result for the retarded Fan self energy [1] :

$$\Sigma^{\text{Fan}}(\mathbf{k}n,\omega) = \frac{1}{N_{\mathbf{q}}}\sum_{\mathbf{q}s}^{\text{BZ}}\sum_{n'} |\langle \mathbf{k}n|H^{(1)}_{\mathbf{q}s}|\mathbf{k}+\mathbf{q}n'\rangle|^2 \times$$
$$\left[\frac{n_{\mathbf{q}s}+1-f_{\mathbf{k}+\mathbf{q}n'}}{\omega-\varepsilon_{\mathbf{k}+\mathbf{q}n'}-\omega_{\mathbf{q}s}+i\eta}+\frac{n_{\mathbf{q}s}+f_{\mathbf{k}+\mathbf{q}n'}}{\omega-\varepsilon_{\mathbf{k}+\mathbf{q}n'}+\omega_{\mathbf{q}s}+i\eta}\right]. \tag{3.6}$$

In this expression, contributions from phonon modes with harmonic phonon energy $\omega_{\mathbf{q}j}$ and occupation number $n_{\mathbf{q}s}$, are summed for all phonon branches, labeled s, and wavevectors, labelled $\mathbf{q}$, in the entire Brillouin zone. The latter is discretized, with $N_{\mathbf{q}}$ indicating the number of wavevectors in the sum. The limit for infinite number of wavevectors is implied. Contributions from intermediate electronic states $|\mathbf{k}+\mathbf{q}n'\rangle$ with electron energy $\varepsilon_{\mathbf{k}+\mathbf{q}n'}$ (not renormalized by phonons) and occupation number $f_{\mathbf{k}+\mathbf{q}n'}$ are summed for all bands n' (valence and conduction). $H^{(1)}_{\mathbf{q}s}$ is the self-consistent change of potential due to the phonon labelled $\mathbf{q}s$ [1,154]. Eq.(3.6) is also the Migdal result Σ_M.

We will work with semiconductors at zero temperature, in which case the phonon occupation numbers vanish, and the electron occupation numbers $f_{\mathbf{k}+\mathbf{q}n'}$ are either one, for the valence states, or zero, for the conduction states. At zero temperature, the phonon occupation numbers vanish. Since we will work with semiconductors, the electron occupation numbers $f_{\mathbf{k}+\mathbf{q}n'}$ are either one, for the valence states, or zero, for the conduction states. For semiconductors with IR-active phonons, the intraband ($n' = n$) contribution with small-q LO phonons (the Fröhlich problem) gives the most important dynamical features for the frequency range near the bare electronic energy ($\omega \approx \epsilon_{\mathbf{k}n}$), due to the combination of small-q diverging matrix element $\langle \mathbf{k}n|H^{(1)}_{\mathbf{q}s}|\mathbf{k}+\mathbf{q}n'\rangle \rightarrow C_n\delta_{nn'}/q$, see Ref. [20], and small denominator ($\omega \approx \epsilon_{\mathbf{k}n} \approx \epsilon_{\mathbf{k}+\mathbf{q}n}$) in Eq.(3.6).

3.3 Quasiparticles and spectral function

The second-order self energy $\Sigma(\mathbf{k}n,\omega)$ is the basis of different approximations for the quasiparticle shift and the spectral function. In the Rayleigh-Schrödinger (RS) approximation, the new quasiparticle energy $E_{\mathbf{k}n}$ is shifted from its initial value $\epsilon_{\mathbf{k}n}$ by the real part of the self energy evaluated at $\epsilon_{\mathbf{k}n}$:

$$E^{\text{RS}}_{\mathbf{k}n} = \epsilon_{\mathbf{k}n} + \Re e\Sigma(\mathbf{k}n,\omega=\epsilon_{\mathbf{k}n}). \tag{3.7}$$

The spectral function $-\Im m G_R(\mathbf{k}n,\omega)/\pi$ has dynamical effects beyond Rayleigh-Schrödinger. In the Dyson-Migdal (DM) approach [1], the spectral function obtained from the self energy is

$$\begin{aligned} A^{\mathrm{DM}}(\mathbf{k}n,\omega) &= -\frac{1}{\pi}\Im m G_R^{\mathrm{DM}}(\mathbf{k}n,\omega) \\ &= \frac{(-1/\pi)\Im m\Sigma(\mathbf{k}n,\omega)}{(\omega-\varepsilon_{\mathbf{k}n}-\Re e\Sigma(\mathbf{k}n,\omega))^2+(\Im m\Sigma(\mathbf{k}n,\omega))^2}. \end{aligned} \tag{3.8}$$

There is typically a QP peak at $\omega = E_{\mathbf{k}n}^{\mathrm{DM}}$ where $\omega - \epsilon_{\mathbf{k}n} + \Re e\Sigma(\mathbf{k}n,\omega) = 0$ (*i.e.* where the first term in the denominator of Eq.(3.8) vanishes). This assumes small values of $\Im m\Sigma$ near $\omega = E_{\mathbf{k}n}^{\mathrm{DM}}$. If $\Im m\Sigma$ is not small at $E_{\mathbf{k}n}^{\mathrm{DM}}$, the QP peak can be strongly broadened. The value of $E_{\mathbf{k}n}^{\mathrm{DM}}$ is usually shifted from $E_{\mathbf{k}n}^{\mathrm{RS}}$ (Eq.(3.7)) by a non-negligible ω-dependence of $\Re e\Sigma(\mathbf{k}n,\omega)$. An additional possible slight shift of the QP peak in Eq.(3.8) can occur if the ω-dependence of $\Im m\Sigma$ is not negligible at $E_{\mathbf{k}n}^{\mathrm{DM}}$. Eq.(3.8) usually also gives one broad satellite, above the quasi-particle peak (for the conduction states), or below it (for the valence states).

A QP part of the spectral function, in Dyson-Migdal theory, can be separated out by Taylor expanding Σ^{Fan} around $\omega = E_{\mathbf{k}n}^{\mathrm{DM}}$, keeping only the constant term in the imaginary part and both constant and linear terms in the real part. The answer is

$$A_{\mathrm{QP}}^{\mathrm{DM}}(\mathbf{k}n,\omega) = \frac{Z_{\mathbf{k}n}^{\mathrm{DM}}}{\pi}\frac{\Gamma_{\mathbf{k}n}^{\mathrm{DM}}}{(\omega-E_{\mathbf{k}n}^{\mathrm{DM}})^2+(\Gamma_{\mathbf{k}n}^{\mathrm{DM}})^2} \tag{3.9}$$

where the QP energy is

$$E_{\mathbf{k}n}^{\mathrm{DM}} = \epsilon_{\mathbf{k}n} + \Re e\Sigma(\mathbf{k}n,\omega = E_{\mathbf{k}n}^{\mathrm{DM}}), \tag{3.10}$$

the quasiparticle weight Z^{DM} is

$$Z_{\mathbf{k}n}^{\mathrm{DM}} = \left(1 - \Re e\frac{\partial\Sigma(\mathbf{k}n,\omega)}{\partial\omega}\bigg|_{\omega=E_{\mathbf{k}n}^{\mathrm{DM}}}\right)^{-1}, \tag{3.11}$$

and the QP broadening is

$$\Gamma_{\mathbf{k}n}^{\mathrm{DM}} = Z_{\mathbf{k}n}^{\mathrm{DM}}|\Im m\Sigma(\mathbf{k}n,\omega = E_{\mathbf{k}n}^{\mathrm{DM}})|. \tag{3.12}$$

If $T = 0$ and the imaginary part of the self-energy vanishes at the QP energy, Eqs.(3.10) and (3.11) give the exact location and weight of a scaled Dirac delta peak $Z_{\mathbf{k}n}^{\mathrm{DM}}\delta(\omega - E_{\mathbf{k}n}^{\mathrm{DM}})$.

Often [98, 112], Eq.(3.10) is linearized to find an approximate quasiparticle energy and the related weight from quantities defined at the bare eigenenergy $\epsilon_{\mathbf{k}n}$:

$$E_{\mathbf{k}n}^{\mathrm{DM0}} = \epsilon_{\mathbf{k}n} + Z_{\mathbf{k}n}^{\mathrm{DM0}}\Re e\Sigma(\mathbf{k}n,\omega = \epsilon_{\mathbf{k}n}), \tag{3.13}$$

with

$$Z_{\mathbf{k}n}^{\mathrm{DM0}} = \left(1 - \Re e\frac{\partial\Sigma(\mathbf{k}n,\omega)}{\partial\omega}\bigg|_{\omega=\epsilon_{\mathbf{k}n}}\right)^{-1}. \tag{3.14}$$

The Lehmann representation of the spectral function is derived from Eqs.(3.1,3.3) using the exact eigenstates $|m\rangle$, of energy E_m, of the full Hamiltonian,

$$\begin{aligned} A(\mathbf{k}n,\omega) &= \sum_{m',m} \frac{e^{-\beta E_m}}{Z} \left[|\langle m'|c^+_{\mathbf{k}n}|m\rangle|^2 \delta(\omega - E_{m'} + E_m) \right. \\ &+ \left. |\langle m'|c_{\mathbf{k}n}|m\rangle|^2 \delta(\omega + E_{m'} - E_m) \right], \end{aligned} \tag{3.15}$$

where $Z = \Sigma_m \exp(-\beta E_m)$ is the partition function and β is $1/k_B T$. Only $T = 0$ is directly relevant to the computations of this Chapter, but the $T > 0$ properties are also important and interesting. From Eq.(3.15) it is clear that $A(\mathbf{k}n,\omega) \geq 0$ at all ω, and the integrated spectral weight ($\int d\omega A(\mathbf{k}n,\omega)$) is 1. The QP part (Eq.(3.9)) has total weight $Z^D_{\mathbf{k}n}$ which must therefore be less than 1. The non-QP part $A - A_{\mathrm{QP}}$ has the rest of the spectral weight.

There is an interesting property of the first moment [152],

$$\int_{-\infty}^{\infty} d\omega \omega A(\mathbf{k}n,\omega) = \langle \{[c_{\mathbf{k}n}, H], c^+_{\mathbf{k}n}\}\rangle. \tag{3.16}$$

For noninteracting electrons, the right-hand side of Eq.(3.16) is the band energy $\epsilon_{\mathbf{k}n}$. If the only interaction is with phonons, the right-hand side has an extra piece, $\langle \{[c_{\mathbf{k}n}, H_{e-p}], c^+_{\mathbf{k}n}\}\rangle$. Terms in the electron-phonon interaction H_{e-p} which have odd powers of lattice displacement do not contribute. The even powers do, however, and the total right hand side of Eq.(3.16) is $\langle \mathbf{k}n|\langle H\rangle|\mathbf{k}n\rangle$ where $\langle H\rangle$ is the thermal average of the vibrating one-electron Hamiltonian. This is exactly a Debye-Waller shifted single particle energy, both zero-point and thermal if $T > 0$. If only the leading second-order term in a vibrational Taylor expansion is kept, the answer is $\epsilon_{\mathbf{k}n} + \Sigma^{\mathrm{DW}}(\mathbf{k}n)$, with only the Debye-Waller part of Eq.(3.5). It is interesting that an observable (in principle) property can separate the Debye-Waller from the Fan effects, given that translational invariance forces a partial cancellation [3] and indicates an underlying connection between these terms.

Although the quasiparticle energy from the DM spectral function for the Fröhlich problem does not occur at the right place and the DM quasiparticle spectral weight is quite bad, as shown by Figs. 1 and 2, the integral of the DM spectral function is correctly 1, and its first moment is correctly $\epsilon_{\mathbf{k}n}$ for the Fröhlich problem which has no 2-phonon matrix element and therefore no D-W term.

3.4 Cumulant

The cumulant expansion is an alternative to the usual Dyson diagrammatic perturbation theory. It derives from an exponential representation of the retarded Green's function $G(\mathbf{k}n,t)$ in the time domain:

$$G^{\mathrm{C}}(\mathbf{k}n,t) = G_0(\mathbf{k}n,t)e^{C(\mathbf{k}n,t)}. \tag{3.17}$$

To lowest order, it treats the Fan diagram, Fig. 7 (a), exactly, and higher-order diagrams such as (b) and (c), approximately.

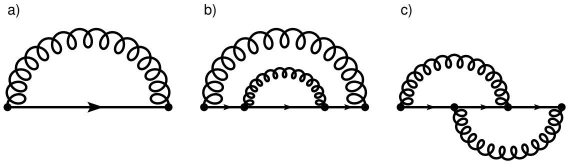


Figure 7: Some of the lowest order Feynman diagrams. (a) Usual Fan self energy diagram. (b) Diagram that contributes to the self-consistency of the electron propagator. (c) Diagram that contains a vertex correction.

Different methods are used to derive the cumulant $C(\mathbf{k}n,t)$ from the self energy, including identification of the terms of the same power of the interaction in a diagrammatic expansion of the left- and right-hand side of Eq.(3.17). In the case of one isolated electronic level (without dispersion), the cumulant approach gives the exact result using only a second-order self energy [147]. Among others, Langreth [147], Hedin [114], Gunnarsson *et al.* [109], Aryasetiawan *et al.* [104], Guzzo *et al.* [105] and Kas *et al.* [149] examined, developed and applied the cumulant expansion.

In a recent paper, Kas *et al.* [149] considered a cumulant expansion of the retarded Green's function. See Sec. 3.11 Appendix II for a discussion of the alternative time-ordered Green's function approach. Their equations, applied to the EPI self energy (including Fan and Debye-Waller contributions), are

$$A^{\mathrm{C}}(\mathbf{k}n,\omega) = -\frac{1}{\pi}\Im m G^{\mathrm{C}}(\mathbf{k}n,\omega), \tag{3.18}$$

$$G^{\mathrm{C}}(\mathbf{k}n,\omega) = \int_{-\infty}^{\infty} e^{i\omega t} G^{\mathrm{C}}(\mathbf{k}n,t)dt, \tag{3.19}$$

$$G^{\mathrm{C}}(\mathbf{k}n,t) = -i\theta(t)e^{-i(\varepsilon_{\mathbf{k}n}+\Sigma^{\mathrm{DW}}_{\mathbf{k}n})t}e^{C(\mathbf{k}n,t)}, \tag{3.20}$$

$$C(\mathbf{k}n,t) = \int_{-\infty}^{\infty} \beta(\mathbf{k}n,\omega)\frac{e^{-i\omega t}+i\omega t-1}{\omega^2}d\omega, \tag{3.21}$$

$$\beta(\mathbf{k}n,\omega) = \frac{1}{\pi}|\Im m\,\Sigma^{\mathrm{Fan}}(\mathbf{k}n,\omega+\varepsilon_{\mathbf{k}n})|. \tag{3.22}$$

Eq.(3.20) shows that the static Debye-Waller self energy $\Sigma^{\mathrm{DW}}_{\mathbf{k}n}$ shifts (in frequency) the whole spectral function with respect to a quasiparticle situated at $\varepsilon_{\mathbf{k}n}$.

Eq.(3.21) also shows that the cumulant and its first time derivative vanish at $t=0$, from which one deduces that the integral of the spectral function $A^{\mathrm{C}}(\mathbf{k}n,\omega)$ over all frequencies is 1, while its first moment is equal to $\varepsilon_{\mathbf{k}n}+\Sigma^{\mathrm{DW}}_{\mathbf{k}n}$. These properties agree perfectly with the exact results of the previous section. Because the phonon density of states (DOS) from acoustic modes vanishes quadratically at zero frequency, for the top of the valence band and the bottom of the conduction band, $\beta(\mathbf{k}n,\omega)$ vanishes quadratically around $\omega=0$. Then following the method of Ref. [104], three separate effects can be attributed to the three pieces $(e^{-i\omega t}+i\omega t-1)$ of Eq.(3.21). Specifically, $e^{-i\omega t}$ generates the satellites, $+i\omega t$ shifts the quasiparticle peak, and -1 generates the renormalization of the quasiparticle weight. The latter two effects make use of Kramers-Kronig relations (satisfied by the retarded self-energy),

$$\Re e\Sigma^{\mathrm{Fan}}(\mathbf{k}n,\varepsilon_{\mathbf{k}n}) = -P\int_{-\infty}^{\infty}\frac{\beta(\mathbf{k}n,\omega)}{\omega}d\omega, \tag{3.23}$$

$$\Re e\frac{\partial\Sigma^{\mathrm{Fan}}(\mathbf{k}n,\omega)}{\partial\omega}|_{\omega=\varepsilon_{\mathbf{k}n}} = -\Re e\int_{-\infty}^{\infty}\frac{\beta(\mathbf{k}n,\omega)}{(\omega+i\delta)^2}d\omega, \tag{3.24}$$

where P denotes the principal part of the integral. The first of these gives the quasiparticle peak shift,

$$\begin{aligned} E_{\mathbf{k}n} &= \epsilon_{\mathbf{k}n} + \Re e\Sigma^{\mathrm{Fan}}(\mathbf{k}n,\omega=\epsilon_{\mathbf{k}n}) + \Re e\Sigma^{\mathrm{DW}}(\mathbf{k}n) \\ &= \epsilon_{\mathbf{k}n} + \Re e\Sigma(\mathbf{k}n,\omega=\epsilon_{\mathbf{k}n}). \end{aligned} \tag{3.25}$$

This is identical to the shift $E^{\mathrm{RS}}_{\mathbf{k}n}$ from Rayleigh-Schrödinger perturbation theory. The second yields the quasiparticle weight,

$$Z_{\mathbf{k}n} = \exp\left(\Re e\frac{\partial\Sigma^{\mathrm{Fan}}(\mathbf{k}n,\omega)}{\partial\omega}|_{\omega=\varepsilon_{\mathbf{k}n}}\right). \tag{3.26}$$

In Eq.(3.17), or equivalently Eq.(3.20), the exponential can be Taylor-expanded: the Green's function and spectral function in the time domain are the sum of the product of the independent-electron contribution multiplied by powers of the cumulant, the latter being a linear functional of the imaginary part of the self energy. In the frequency domain, this creates a satellite series [104, 113], coming from repeated convolution in frequency of the undressed particle spectral function (a Dirac delta) by the "satellite spectral function" A^{S} (See Eqs. 8, 13, and 14 of Ref. [104] or Eqs. 3, 4, and 6 of the supplemental materials of Ref. [10]). Supposing A^{S} to be also a Dirac delta (as in the Lundqvist model Hamiltonian), one obtains for the spectral function a Poisson distribution of Dirac function satellites, each being spaced by the characteristic bosonic satellite energy. More generally, supposing the imaginary part of the self-energy to be non-zero only for positive frequencies beyond a given threshold, the threshold for each satellite contribution will be determined by the self-energy threshold multiplied by the order of the satellite. In the next section, such an effect can be seen in the case of the Fröhlich Hamiltonian.

When the imaginary part of the self energy has contributions at both negative and positive frequencies, this simple picture is lost, unless the cumulant is clearly dominated by one of these. In the case of the first-principles EPI for wide-gap insulators, the self energy has indeed both negative and positive contributions. However, only one of these contributions will dominate for wide-gap insulators, as will be seen later.

3.5 Fröhlich Hamiltonian self-energy

The relationship between the Fröhlich coupling and first-principles calculations has been established in Refs. [13, 14], in the general case of several electronic bands, several phonon branches, as well as anisotropic Born effective charge tensor and dielectric tensor. Here, we consider the simple original Fröhlich Hamiltonian, corresponding to the following hypotheses: (1) Only one isotropic electronic parabolic band (we will first treat the conduction band), with a minimum at $\mathbf{k} = 0$. We choose $\varepsilon_{\mathbf{k}=0,c} = 0$, and use parabolic dispersion governed by the effective mass m^*. (2) Only one LO-phonon branch with constant phonon frequency ω_{LO}. (3) Isotropic Born effective charge Z^*, isotropic electronic (optical) dielectric constant ϵ_∞, and thus, isotropic low-frequency dielectric constant ϵ_0 [37].

Only intraband terms $n' = n$ are present, thus, the general $g_{nn'j}(\mathbf{k},\mathbf{q}) = \langle \mathbf{k}n | H^{(1)}_{\mathbf{q}j} | \mathbf{k}+\mathbf{q}n' \rangle$ reduces to $g_{nn'j}(\mathbf{k},\mathbf{q}) = g_{\mathbf{q}} \delta_{nn'}$ for $j = LO$, with

$$g_{\mathbf{q}} = \frac{i}{q} \left[\frac{4\pi}{\Omega_0} \frac{\omega_{LO}}{2} \left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0} \right) \right]^{1/2}, \tag{3.27}$$

where Ω_0 is the volume of the primitive cell.

In computing the self energy, the constant Debye-Waller shift is neglected. This might seem a strong approximation. However, hypothesis (2) implies the neglect of the Fan term from the acoustic modes as well, and it is known that the acoustic mode Fan contributions and the Debye-Waller contributions cancel each other in the vanishing-q limit [3]. Furthermore, the LO-phonon Fan term dominates, due to the integrable divergence mentioned at the end of Sec. II.

The expression for the Fröhlich self energy is easily found, see *e.g.* Ref. [25], but is presented here as well, for convenience and comparison with the first-principle results. The zero-temperature formula for the Fan self energy of the (Fröhlich) electron state at the bottom of the conduction band ($\mathbf{k} = 0$) comes from Eq.(3.6):

$$\Sigma^{\mathrm{Fr}}_e(\mathbf{k}=0,\omega) = \frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}} \frac{|g_{\mathbf{q}}|^2}{\omega - \epsilon_{\mathbf{q}} - \omega_{LO} + i\eta}. \tag{3.28}$$

The intermediate electron energy $\epsilon_{\mathbf{q}}$ becomes $q^2/2m^*$ in the parabolic band with effective mass approximation. Using the Debye sphere for the Brillouin zone, the equation becomes

$$\Sigma^{\mathrm{Fr}}_e(0,\omega) = \int_0^{q_D} dq \frac{\Omega_0}{(2\pi)^3} \frac{4\pi q^2 |g_q|^2}{\omega - \frac{q^2}{2m^*} - \omega_{LO} + i\eta}. \tag{3.29}$$

Provided the electronic energy on the boundary of the Debye sphere is much bigger than ω_{LO}, the upper limit q_D can be safely extended to infinity [25]. For $q_D \to \infty$, this gives

$$\Re e\, \Sigma^{\mathrm{Fr}}_e(\omega) = -\frac{\alpha\omega_{LO}}{\sqrt{1 - \omega/\omega_{LO}}} \theta(\omega_{LO} - \omega), \tag{3.30}$$

$$\Im m\, \Sigma^{\mathrm{Fr}}_e(\omega) = -\frac{\alpha\omega_{LO}}{\sqrt{\omega/\omega_{LO} - 1}} \theta(\omega - \omega_{LO}), \tag{3.31}$$

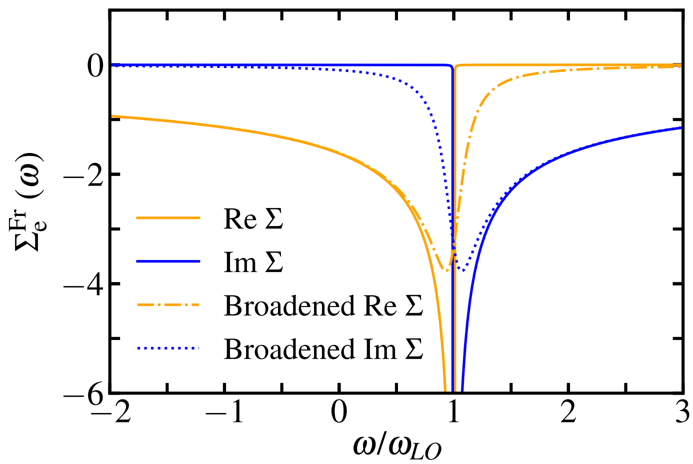


Figure 8: Fröhlich Hamiltonian self-energy. Real part in orange, imaginary part in blue. The functions with a negligible broadening, $\delta = 0.001\omega_{LO}$, are represented by continuous lines, while functions with a broadening $\delta = 0.12\omega_{LO}$, similar to the one used in first-principles calculations, are represented by dashed lines.

where the Fröhlich coupling constant α is

$$\alpha = \left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0}\right)\left(\frac{m^*}{2\omega_{LO}}\right)^{1/2}. \tag{3.32}$$

Above ω_{LO}, the self energy is purely imaginary, while below ω_{LO}, it is purely real. Both are negative and diverge like an inverse square root of the frequency around ω_{LO}.

For the valence band, with the eigenenergy of the top of the valence band now taken as zero, the corresponding retarded self energy is

$$\Re e\, \Sigma_h^{\rm Fr}(\omega) = +\frac{\alpha\omega_{LO}}{\sqrt{1+\omega/\omega_{LO}}}\theta(\omega+\omega_{LO}), \tag{3.33}$$

$$\Im m\, \Sigma_h^{\rm Fr}(\omega) = -\frac{\alpha\omega_{LO}}{\sqrt{-\omega/\omega_{LO}-1}}\theta(-\omega-\omega_{LO}). \tag{3.34}$$

For a given material with well-defined dielectric constant and LO frequency, the coupling constant α from Eq.(3.32) has different values for electron and hole polarons, due to differing effective masses.

In Secs. VII and VIII dealing with first-principles calculations, we will maintain a small finite broadening factor δ [155], of order $0.12\omega_{LO}$, for numerical reasons. Thus the self-energy functions, Eqs.(3.30) and (3.34), will not retain their inverse-square-root shape. Eqs.(3.30) and (3.31) and their broadened versions are shown in Fig. 8.

The self energies in Eqs.(3.30)-(3.34) were derived in lowest order perturbation theory. They include only the Fan diagram, without vertex corrections. Calculations of self energies

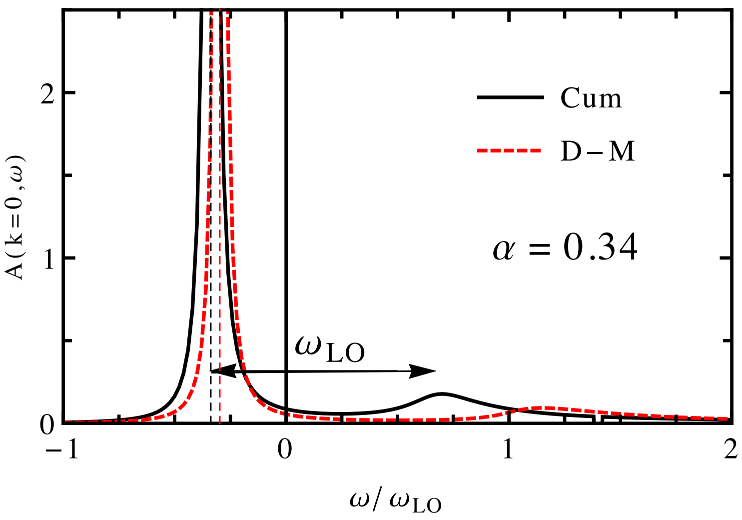


Figure 9: Fröhlich Hamiltonian spectral function using the cumulant approach $G^{\rm C}$ (solid, black) and the Dyson-Migdal approach $G^{\rm DM}$ (dashed, red) for α=0.34. The position of the quasiparticle peak slightly differs between the two. The cumulant version deviates from the Fröhlich value $-\alpha\omega_{\rm LO}$ only because a non-zero broadening $\delta \approx 0.12\omega_{LO}$ [155] is used in numerical evaluation of Eq.(3.29), for consistency with later calculations. In the DM case, the onset of the phonon-emission “satellite” is higher by $\omega_{\rm LO}$ than the bare band energy $\omega = \varepsilon^{(0)}_{\mathbf{k}=0,c} = 0$ [110]. By contrast, it is higher by $\omega_{\rm LO}$ than the quasiparticle peak in the cumulant method, corresponding to states that combine the dressed quasiparticle with one LO phonon.

at higher orders have been performed for the Fröhlich Hamiltonian, see *e.g.* Refs. [25, 108]. For values of α in the range considered in this Chapter, those higher-order corrections to the quasiparticle shift are small, as argued by Migdal [106]. Calculations using these formulas for $\Sigma^{\rm Fan}$ and the corresponding spectral functions $A^{\rm DM}$ and $A^{\rm C}$, are discussed in the next section, and plotted in Figs. 9-12.

3.6 Fröhlich Hamiltonian quasiparticle energy and spectral function

The spectral function $-\Im m G_R(\mathbf{k}n,\omega)/\pi$ has “dynamical effects” beyond Rayleigh-Schrödinger. In Fig. 9 we show the spectral function for $\alpha = 0.34$ (a small value, typical of many semiconductors, e.g. electrons at the conduction band minimum of GaN), for the Fröhlich Hamiltonian using two approximations. The black curve corresponds to the cumulant expansion Eqs.(3.18-3.22), while the dash-red curve was obtained with the DM approach Eq.(3.8). For consistency with later first-principles calculations, we used a small broadening factor $\delta \approx 0.12\omega_{LO}$ for the self energy, which causes a small artificial shift and broadening of the quasiparticle peak.

There are two effects. (i) In the cumulant case, the quasiparticle peak agrees with the value $-\alpha\omega_{\mathrm{LO}}$ predicted by Rayleigh-Schrödinger perturbation theory. (ii) A phonon-emission side band appears, with one clearly visible satellite, at different energies in the DM and cumulant cases. The separation between the quasiparticle and satellite is about ω_{LO} in the cumulant expansion, but slightly higher than $(1+\alpha)\omega_{LO}$ in the DM approach. Physically, this satellite ought to start at the quasiparticle energy E_{QP} plus ω_{LO}. The DM shift of E_{QP} away from $-\alpha\omega_{\mathrm{LO}}$ does not agree with the results obtained for many years by the polaron community [25, 144].

Increasing the value of α to values typical for the valence and conduction band extrema of MgO and LiF, 1.62, 4.01 and 8.00 (see later) gives the three next figures 10, 11 and 12. Note that the position E_{QP} of the quasiparticle peak differs more and more between the two, with the cumulant version staying at the RS answer $-\alpha\omega_{\mathrm{LO}}$, as expected. A^{DM} becomes increasingly unphysical for stronger couplings: the side band has one broad satellite, setting in at ω_{LO}, with a maximum at a frequency that increases with α. By contrast, A^{C} has the satellite onset at $E_{\mathrm{QP}} + \hbar\omega_{\mathrm{LO}}$. Several satellites are clearly visible in A^{C} of Fig. 11, spaced approximately by ω_{LO}. The side bands become broader and less well-defined as α increases, with a long tail extending to higher energies. The numerical value of the broadening $\delta \approx 0.12\omega_{LO}$, albeit small, has an impact on the threshold at which it becomes impossible to distinguish the satellites from the overall smooth behavior. This broadening factor is multiplied by the order of the satellite in the repeated convolution of the undressed particle spectral function mentioned in Sec. IV.

These cumulant results globally agree with the previous cumulant-based study by Dunn [108], for $\alpha = 2, 4$, and 6. He worked, however, at finite temperature, and also included the next order of perturbation theory in his calculations of the self energy. Higher orders of perturbation theory tend to sharpen features of the spectral function. In the case of a model core-electron spectrum, for which the exact solution is known [147], the next order of perturbation theory improves significantly the position of the peak, and sharpens it with respect to a first-order self-consistent treatment. On the other hand, a first-order non-self-consistent treatment also gives a sharper plasmon satellite than the first-order self-consistent treatment, albeit located at nearly the same too low energy [148].

The first satellite shape and position, in our cumulant calculations, resemble reasonably well those of the diagrammatic Monte Carlo (MC) calculations of Mischenko *et al.* [47], apparently the best reference results available at present. However, the MC results do not show the second and third satellites and instead develop a satellite in the range from 3.5 to 4.0 ω_{LO} if α is larger than one. For values of α larger than 4, another satellite appears in the range 8.0 ... 9.0 ω_{LO}. As MC results for optical $\sigma(\omega)$ compare well with other approaches [156], we believe that the cumulant approach for $A(k,\omega)$ has some errors for ω beyond the first satellite. The physical reason for the disappearance of the second and third multiphonon peaks, and the appearance of other peaks, has been discussed in Ref. [157]. The new peaks might be related to so-called “relaxed excited states”, not treated by the lowest order cumulant approach, that dominate the spectral function in the energy range beyond the first phonon threshhold.

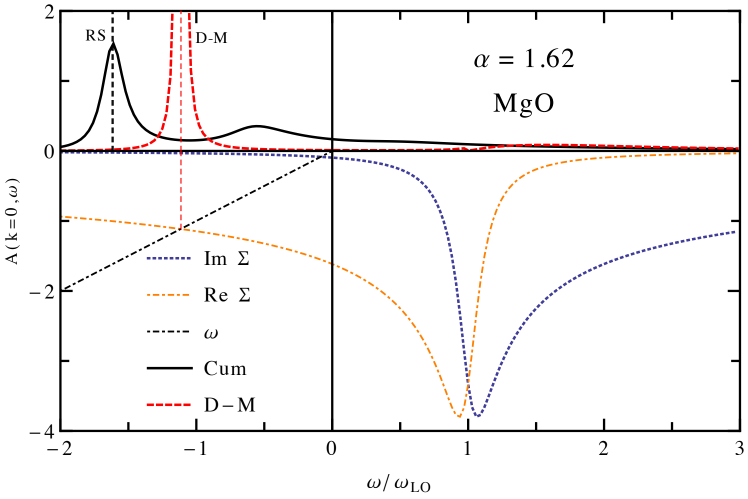


Figure 10: The lower part shows the DM self energy (in units of ω_{LO}) for a Fröhlich electron with $\alpha = 1.62$ (typical of the conduction band minimum of MgO), using Eq.(3.30) except broadened with $\delta = 0.12\omega_{LO}$ in Eq.(3.28). The position of the DM QP peak is at the crossing between the real part and the line $\Re e\Sigma = \omega$. The upper part shows both the resulting DM spectral function and the cumulant version from G^{C}. The satellite setting in at $\omega/\omega_{LO} = 1$ in the DM case is barely visible in this picture.

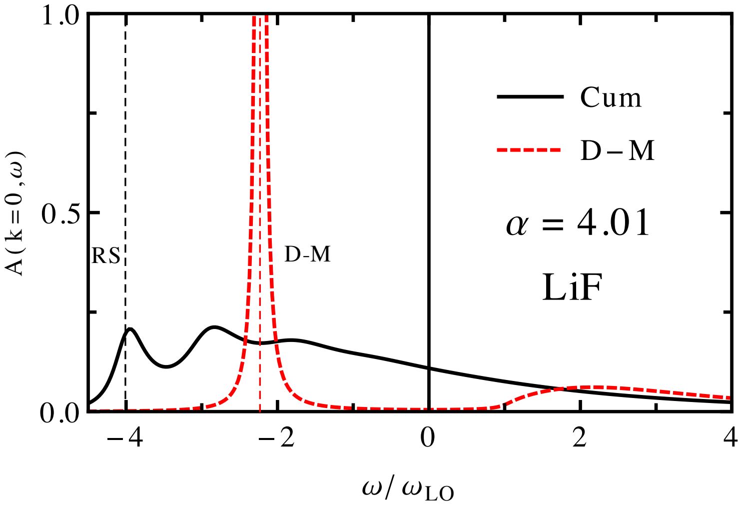


Figure 11: Fröhlich Hamiltonian spectral function using the cumulant G^{C} (in black) and the Dyson-Migdal approach G^{DM} (dashed, in red) for α=4.01 (typical of the conduction band minimum of LiF), from the Migdal self energy broadened by $\delta = 0.12\omega_{LO}$.

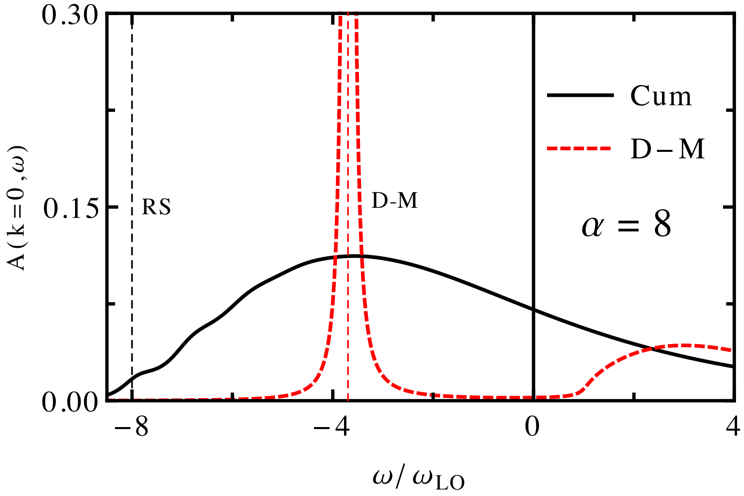


Figure 12: Fröhlich Hamiltonian spectral function using the cumulant G^{C} (in black) and the Dyson-Migdal approach G^{DM} (dashed, in red) for α=8, from the Migdal self energy broadened by $\delta = 0.12\omega_{LO}$.

Let us analyze the behavior of the quasiparticle peak in the DM case in more detail, in the case without any numerical broadening. The quasiparticle energy is found from

$$E_{QP}^{\mathrm{DM}} = \Re e\, \Sigma(E_{QP}^{\mathrm{DM}}). \tag{3.35}$$

Using Eq.(3.30) for electron states gives a cubic equation,

$$\alpha^2 = \left(\frac{E_{QP}^{\mathrm{DM}}}{\omega_{LO}}\right)^2 - \left(\frac{E_{QP}^{\mathrm{DM}}}{\omega_{LO}}\right)^3 \;;\;\; (E_{QP}^{D} < 0). \tag{3.36}$$

For all α, this has one negative real root; $E_{QP}^{\mathrm{DM}} \sim -\alpha\omega_{LO}$ for small α and $E_{QP}^{\mathrm{DM}} \sim -\alpha^{2/3}\omega_{LO}$ for large α. This agrees with Fig. 5 (red circles). For small α, to second order, the Dyson-Migdal QP energy is

$$E_{QP}^{\mathrm{DM}} \approx -(\alpha - \alpha^2/2)\omega_{LO} \tag{3.37}$$

The DM quasiparticle energy Eq.(3.35) corresponds to the intersection of the $\Re e\,\Sigma(\omega)$ function with the straight ω line, as shown in Fig. 10. As shown in Fig. 5, the second-order Rayleigh-Schrödinger answer $E_{QP}^{\mathrm{RS}} = -\alpha\omega_{LO}$, although not perfect, is better than the Dyson-Migdal answer (with $\Sigma_M = G_0 D$, not the self-consistent $\Sigma_M = GD$, see Ref. [110]).

The Fröhlich spectral function $A = -\Im m\, G_R(\omega)/\pi$, in DM approximation, has two parts. The quasiparticle part is $Z^{\mathrm{DM}}\delta(\omega - E_{QP}^{\mathrm{DM}})$, where $Z^{\mathrm{DM}} = 1/(1 - d\Sigma/d\omega)$ is evaluated at $\omega = E_{QP}^{\mathrm{DM}}$. To lowest order, the quasiparticle weight $Z^{\mathrm{DM}} \approx 1 - \alpha/2$. At large α, Z^{DM} tends asymptotically to 2/3, but the linearized weight, Eq.(3.14) is $Z^{\mathrm{DMlin}} \approx 1/(1 + \alpha/2)$ for all values of α, which tends to zero asymptotically. The DM spectral function (without numerical broadening) is

$$A(x) = \frac{Z^{\mathrm{DM}}}{\omega_{LO}}\delta(x - x_{QP}) + \frac{1}{\pi\omega_{LO}}\frac{\alpha\sqrt{x-1}}{x^2(x-1)+\alpha^2}\theta(x-1), \tag{3.38}$$

where $x = \omega/\omega_{LO}$ and $x_{QP} = E_{QP}^{\mathrm{DM}}/\omega_{LO}$. The DM side-band always starts at $\omega = \omega_{LO}$, rather than at the intuitively correct value of $E_{QP} + \omega_{LO}$. The Monte Carlo spectral functions [47] show sidebands starting close to the intuitive energy.

3.7 Full Migdal self-energy from first-principles

We present now first-principles results for the full self energies (real and imaginary parts, from all phonon modes, including interband and Debye-Waller effects) of MgO and LiF, at the valence band maximum (VBM) and conduction band minimum (CBM). These results will be used in the next section to find DM and cumulant spectral functions. This section also gives the related first-principles parameters to be fed into the Fröhlich model. For MgO and LiF band extrema, the Fröhlich coupling α ranges from about 1.5 to 15. We also tabulate the magnitudes of the separate Debye-Waller and Fan terms, as well as their contributions from unoccupied and occupied states.

Technical details of the first-principles calculations are in Appendix I. The most delicate issue concerns the sampling of phonon wavevectors in the Brillouin zone, and the numerical broadening factor needed to treat Eq.(3.6). To obtain well-converged self energies, the eigenenergy differences between sampled wavevectors as well as the numerical broadening factor must be significantly smaller than the LO phonon frequency. This is especially important at the unperturbed quasiparticle eigenenergy, where the real and imaginary parts and their derivatives govern the asymptotics of the cumulant, and hence the quasiparticle peak characteristics. We choose a broadening of 0.01 eV, approximately $\omega_{LO}/8$ (see Table 2), and wavevector grids up to 96 × 96 × 96 points for MgO and 48 × 48 × 48 points for LiF. This is considerably better than in Ref. [138] for the same materials (diamond and BN were also studied in that work). In Ref. [138], the broadening factors ranged between 0.1eV and 0.4eV, and phonon wavevector grids had at most 32 × 32 × 32 points.

MgO and LiF both crystallize in the (cubic) rocksalt structure, with one formula unit per primitive cell. Density-Functional Theory (DFT-GGA) Kohn-Sham electronic structure of both materials can be found elsewhere [138], and will not be reproduced here. In both materials, the CBM is not degenerate. It is parabolic in a large region around Γ, so we expect the effective mass parabolic approximation to be adequate. The VBM is triply degenerate at Γ. One light hole band rapidly separates from two heavy hole bands away from Γ, the latter being degenerate along the Γ-X and Γ-L directions. The deviation with respect to parabolicity is faster than for the conduction band. The DFT eigenenergy is set to zero at the relevant band extremum. In both materials, there are three acoustic and three optic phonon branches. At Γ, the LO branch is separated from the doubly-degenerate TO branches.

Table 2 presents the computed geometric, electronic, dielectric, and dynamical properties of MgO and LiF, that determine the corresponding Fröhlich parameter α, also reported in this table. The primitive cell parameter is $a_0 = a/\sqrt{2}$, where a is the size of the conventional cube. Different effective masses are mentioned for the valence bands, corresponding to the heavy hole (hh) and to the light holes (lh), and also to different directions of the non-spherical electronic structure [158]. Note that the ω_{LO} and ω_{TO} of the two materials are rather similar, while their dielectric properties and effective masses differ significantly. The Fröhlich parameter α provides a rough estimate of the phonon-induced zero-point renormalization of the quasiparticle energy. For the conduction band minimum, the estimated shifts ($-\alpha_e\omega_{LO}$) are -0.137eV for MgO and -0.332eV for LiF. For valence bands, we do not attempt to integrate over all effective mass directions and hole types, but simply provide the corresponding α values deduced from Eq.(3.32). The values presented in Table I are in reasonable agreement with those recently computed in Ref. [12] for the same materials. However, in Ref. [12], the Fröhlich polaron binding energy is defined as half the value from the usual theoretical approach (that we adopt), because the authors cut off the q-integral at π over the polaron radius instead of infinity. We find on the contrary that the Fröhlich values underestimate first-principles values, as will appear later.

Table 2: Computed basic characteristics of MgO and LiF. See text for the different symbols.

	Unit	MgO	LiF
a_0	[$\AA$]	3.01	2.88
Ω_0	[$\AA^3$]	19.2	16.9
ε_g (DFT-GGA)	[eV]	4.49	8.54
ω_{LO}/ω_{TO}	[eV]	0.0844/0.0454	0.0828/0.0466
$\epsilon_\infty/\epsilon_0$		3.23/11.14	2.04/6.44
m_e^*		0.340	0.873
m_{hh}^* (Γ-X/Γ-L)		2.164/3.822	3.622/11.955
m_{lh}^* (Γ-X/Γ-L)		0.387/0.335	1.346/0.887
α_e		1.624	4.009
α_{hh} (Γ-X/Γ-L)		4.101/5.450	8.165/14.834
α_{lh} (Γ-X/Γ-L)		1.734/1.610	4.977/4.040

The self energy $\Sigma_1+i\Sigma_2$ for the CBM of MgO, in a 2eV window around the bare electronic energy ϵ^0, is in Fig. 13. Fröhlich-type real and imaginary peaks, both negative, occur at $\omega_{LO} = 0.0844eV$, just as in Fig. 10. Despite a very fine 96^3 q-point grid sampling, numerical noise is still visible for the small 0.01 eV broadening of the denominators (see Appendix I). The Debye-Waller contribution, and the Fan contributions from bands other than lowest conduction, shift Σ_1 downwards in Fig. 13 compared to Fig. 10, and give it a slight linear slope in the vicinity of the conduction band minimum. Fig. 14 shows the same data as Fig. 13 in a wider energy window, also with the electronic density of states (DOS). Additional structures are present, in the valence band region (below -4 eV), with the same van Hove singularities as the electronic DOS. Small structures in the conduction band region are seen as well.

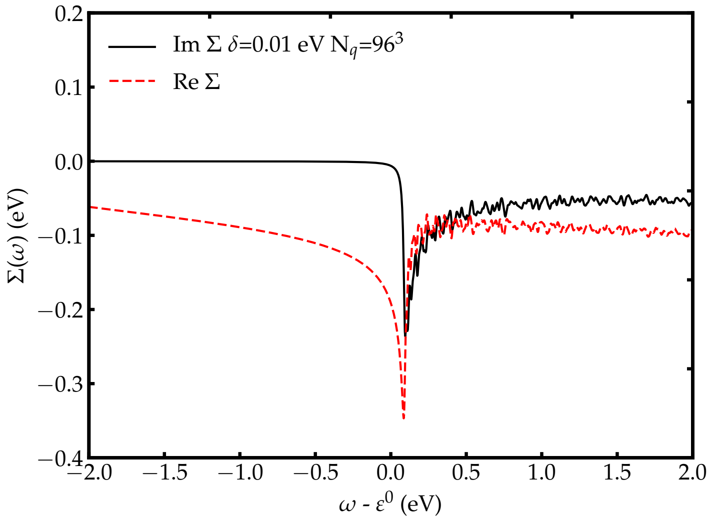


Figure 13: The MgO conduction band minimum retarded self energy with $\delta = 0.01$ eV and a 96^3 q-grid. Full black line: imaginary part; dotted red line: real part. The reference energy ε^0 is the unrenormalized conduction band energy minimum. This figure is surprisingly similar to Fig. 6,considering that it includes all phonons and interband effects rather than just the analytic Fröhlich result.

Similarly, Σ_1 and Σ_2 for the VBM of MgO are shown in Fig. 15. With the same sampling and broadening as the CBM, the noise can hardly be seen. The curvature of the hole band is less pronounced for the VBM than for the CBM, making the numerical work less difficult. Close to $\omega - \epsilon^0 = 0$, the structure of the self energy is close to the Fröhlich self energy, with the appropriate sign change for a hole polaron. Additional valence band characteristic features are seen, clearly related to the electronic DOS. By contrast, the imaginary part of the self energy in the conduction region is very small, and the real part is nearly structureless.

Unlike the electronic dispersion, the role of the phonon dispersion is apparently minor. For the phonon frequencies to have an impact on the self energy, the difference between $\omega \approx \epsilon_{\mathbf{k}n}$ and the electronic eigenenergies $\epsilon_{\mathbf{k+q}n'}$ must be comparable to phonon frequencies (see Eqs.(3.42) and (3.43)). This happens only in a small Brillouin zone region around Γ, in which the phonon branches are practically constant.

The characteristics of the self energy, evaluated at the bare eigenenergy $\varepsilon_{\mathbf{k}n}$, are reported in Table 3, including the decomposition into Debye-Waller and Fan (and unoccupied/occupied) components. Also, the frequency derivative of the self energy and components are given. A $48 \times 48 \times 48$ q-wavevector grid has been used by default for this table, except for the real part of the self energies and their decomposition, which is also given using the more converged $96 \times 96 \times 96$ q-wavevector grid.

The convergence of the Allen-Heine-Cardona zero-point renormalization (ZPR) with respect to the wavevector sampling has been thoroughly analyzed in Sec. IV.B.2 of Ref. [20]. In particular, for IR-active materials treated in the non-adiabatic approximation, at the band

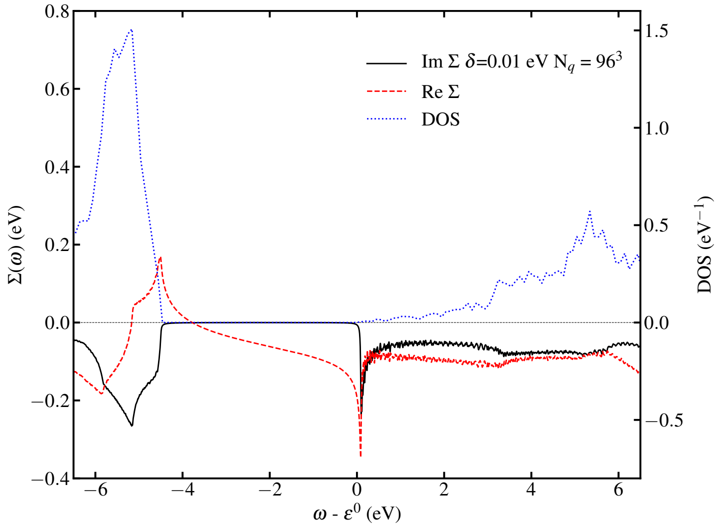


Figure 14: Retarded self energy for the bottom of the conduction band of MgO in a wider range of energy than in Fig. 13: imaginary part in black, real part in dashed red. The electronic DOS is also shown (dotted blue), for comparison.

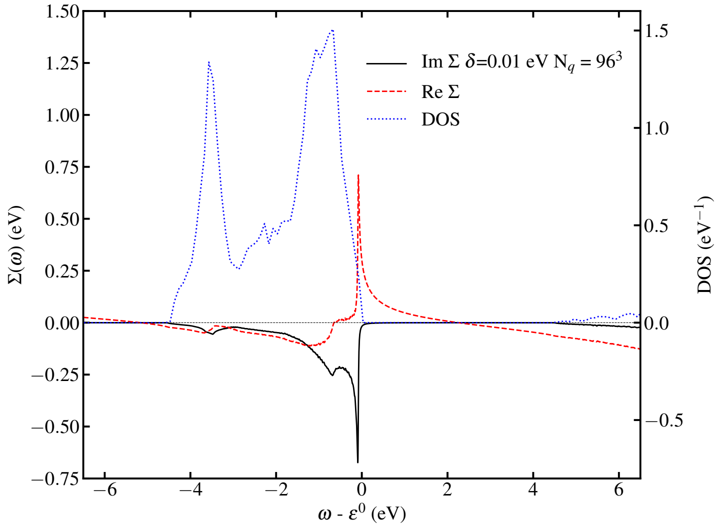


Figure 15: Retarded self energy for the top of the valence band of MgO: imaginary part in black, real part in dashed red. The electronic DOS is also shown (dotted blue), for comparison. The VBM eigenenergy is the reference energy ε^0, which explains the horizontal shift of the DOS with respect to Fig. 14.

Table 3: MgO and LiF first-principles self energy (eV) and frequency derivative (dimensionless) at $\omega = \varepsilon_{\mathbf{k}n}$, and their components, for the conduction band minimum and valence band maximum. The Debye-Waller self energy is static (frequency-independent) and real. The quasiparticle weights, from linearized DM and retarded cumulant approaches are also given, as well as their occupied and unoccupied bands factors. For the real part of the self energy, results are reported with two different phonon wavevector grids (96^3 and 48^3), while for the imaginary part, the derivatives and the quasiparticle weights, only the results obtained with the 48^3 grid are reported. Both Σ_2 and $\partial\Sigma_2/\partial\omega$ vanish; their non-zero values arise from artificial broadening by $\delta = 0.12\omega_{LO}$. See Appendix II for explicit expressions of $\Sigma^{\mathrm{Fan}}_{\mathrm{oc}}$ and $\Sigma^{\mathrm{Fan}}_{\mathrm{un}}$.

	MgO CBM	MgO VBM	LiF CBM	LiF VBM
96^3 grid				
Σ_1	-0.191	0.302	-0.370	0.723
Σ_1^{DW}	-0.056	4.271	0.078	6.785
$\Sigma_{1,\mathrm{un}}^{\mathrm{Fan}}$	-0.371	-4.336	-0.524	-6.911
$\Sigma_{1,\mathrm{oc}}^{\mathrm{Fan}}$	0.235	0.367	0.077	0.850
48^3 grid				
Σ_1	-0.175	0.285	-0.342	0.695
Σ_1^{DW}	-0.054	4.263	0.078	6.772
$\Sigma_{1,\mathrm{un}}^{\mathrm{Fan}}$	-0.354	-4.327	-0.497	-6.898
$\Sigma_{1,\mathrm{oc}}^{\mathrm{Fan}}$	0.233	0.349	0.077	0.821
Σ_2	-0.005	0.016	-0.014	0.053
$\Sigma_{2,\mathrm{un}}^{\mathrm{Fan}}$	-0.005	0.000	-0.014	0.000
$\Sigma_{2,\mathrm{oc}}^{\mathrm{Fan}}$	-0.000	0.016	0.000	0.053
$\partial\Sigma_1/\partial\omega$	-0.455	-1.594	-1.353	-4.780
$\partial\Sigma_{1,\mathrm{un}}^{\mathrm{Fan}}/\partial\omega$	-0.446	-0.007	-1.345	-0.012
$\partial\Sigma_{1,\mathrm{oc}}^{\mathrm{Fan}}/\partial\omega$	-0.009	-1.587	-0.008	-4.768
$\partial\Sigma_2/\partial\omega$	-0.058	-0.252	-0.206	-1.248
$\partial\Sigma_{2,\mathrm{un}}^{\mathrm{Fan}}/\partial\omega$	-0.058	0.000	-0.206	0.000
$\partial\Sigma_{2,\mathrm{oc}}^{\mathrm{Fan}}/\partial\omega$	0.000	-0.252	0.000	-1.248
$Z_{\mathbf{k}n}^{Dlin}$	0.687	0.386	0.425	0.173
$Z_{\mathbf{k}n}$	0.634	0.204	0.258	0.008
$Z_{\mathbf{k}n}^{\mathrm{un}}$	0.640	0.993	0.260	0.988
$Z_{\mathbf{k}n}^{\mathrm{oc}}$	0.991	0.205	0.992	0.008

structure extrema, a $1/N_q$ behavior is obtained, where N_q is the linear density of q-points of the three-dimensional sampling. As shown in Appendix I, Table III, such a trend matches well numerical results. Thus, the 48^3 and 96^3 grids Σ_1 results can be extrapolated to infinity, giving for the CBM and the VBM MgO, respectively, a ZPR (or polaron binding energy) of -207meV and 319meV, and for the CBM and VBM of LiF, respectively, a ZPR of -398meV and 751meV. The total band gap ZPR for MgO is 526meV while for LiF it is 1149meV.

The Fröhlich estimated CBM shifts ($-\alpha_e \omega_{LO}$), *i.e.* -137meV for MgO and -332meV for LiF, are in qualitative agreement with first-principles results, but underestimate their absolute value by about 50-70meV. The first-principles ZPR of the CBM in these materials is thus largely dominated by the Fröhlich part. The analysis of the VBM shift is more complex due to the band warping, and will not be given here. Still, the range of α for holes mentioned in Table I and the zero-point renormalization for the VBM in Table II are quite consistent. A similar dominance of the Fröhlich part of the electron-phonon interaction in other IR-active materials with large LO-TO splitting is expected, and would be consistent with the widespread use of the Fröhlich Hamiltonian for the interpretation of many experimental results.

Our self energy values compare favorably with those of Table I, column $\Sigma^{dyn}(\varepsilon^0)$ of Antonius et al. [138]. Remember however, that in the latter study, the broadening factor δ was much larger and the sampling of phonon wavevectors much coarser than in the present study, see Appendix I. Actually, the quantities reported in the column $\Sigma^{stat}(\varepsilon^0)$ of Table I of Ref. [138] should diverge for vanishing broadening factor and perfect Brillouin Zone sampling, for the IR-active materials BN, MgO and LiF, as shown in Ref. [20]. The similarity of $\Sigma^{stat}(\varepsilon^0)$ and $\Sigma^{dyn}(\varepsilon^0)$ is thus an artifact, simply due to the similarity of the chosen broadening factor value ($\delta = 0.1$ eV) with the LO phonon frequency in these materials (see *e.g.* Table 2).

The decomposition of Σ_1 into its Debye-Waller and Fan components highlights the dramatic cancellation between the Debye-Waller component and the unoccupied bands Fan components, for the VBM of the two materials. As a consequence, the occupied band Fan component has the same magnitude as the total zero-point renormalization value. By contrast, the CBM zero-point renormalization comes from contributions with different signs, without noticeable cancellation. Thus it is surprising that the total shift is given quite accurately by the Fröhlich part. The sum rule for acoustic modes, presented in Ref. [3], contributes to the final dominance of the Fröhlich estimation.

In Table 3, we also report the quasi-particle weights, from linearized DM and retarded cumulant approaches, that are directly obtained from the derivative of the real part of the self energy with respect to the frequency at the bare electronic energy, see Eqs. (3.14) and (3.26). The weights can be decomposed in their hole and electron factors, as discussed in Appendix II.

$Z^{\mathrm{DM0}}_{\mathbf{k}n}$ and $Z_{\mathbf{k}n}$ spectral weights differ the most in the VBM case; for LiF, the ratio exceeds one order of magnitude. It is slightly less than two for the MgO VBM. Taking into account the results from large polaron studies of the Fröhlich Hamiltonian (clearly favoring the Rayleigh or cumulant shifts, Eq.(3.7) and (3.25)), the values from Table I of

Ref. [138], column $\Sigma^{dyn}(\varepsilon^0)$ have also to be preferred over the values in columns $Z\Sigma^{dyn}(\varepsilon^0)$ or $\Sigma^{dyn}(\varepsilon)$. Similarly the values published in Ref. [20], Table VII, column "ZPR Non-adiabatic" correspond to the preferred expressions Eq.(3.7) and (3.25).

For completeness, Table 3 also gives the imaginary parts of the self energy and their derivatives, which are, respectively, linked physically to the broadening of the quasiparticle peak and its asymmetry. Note that these values are actually artificial effects of numerical broadening. Our computations only include band extrema at zero temperature, for which the imaginary part of Eqs. 3.42 and 3.43 vanish exactly if there is no artificial broadening. Nonzero values of $\Im m\Sigma$ at the QP energy thus indicate the magnitude of the broadening parameter.

3.8 First-principles spectral functions

Figs. 16 to 19 present DM and retarded cumulant spectral functions, for the CBM and VBM of MgO and LiF, and also show the Fröhlich spectral function obtained with estimated α (Table II) for the CBM case. The LiF VBM spectral function A^{DM} from G^{DM} was previously given in Fig. 1 of Ref. [138], but our numerical treatment of G^{DM} is significantly improved. The Brillouin zone has been sampled by a 96 × 96 × 96 grid for the CBM of MgO, and a 48 × 48 × 48 grid for the other cases. In all cases, a 0.01 eV $\approx 0.12\omega_{LO}$ broadening of the self energy has been used.

We only use the retarded cumulant approach due to the discussion in the previous section. It also appears to be the preferred method to obtain the quasiparticle spectral functions for insulators as well as metals.

For the MgO CBM, with Fröhlich $\alpha = 1.62$, Fig.16, the position of the quasi-particle peak in the first-principles DM case is lower than from the first-principles cumulant case. Also, the Fröhlich peak position is closer to the first-principles DM position than with the first-principles cumulant one. This agreement between the DM position and the Fröhlich peak position is accidental: the Fröhlich constant is too small to reproduce the band gap shift from first-principles (cf. the above mentioned $50-70$ meV underestimation), while the DM shift is also too small, but this is due to the incorrect underestimation highlighted in Fig. 1. The position of the satellite follows the same pattern as observed for the Fröhlich Hamiltonian: the DM satellite is separated from the quasiparticle peak by much more than the LO phonon frequency, while the distance between the satellite and the quasiparticle peak in the cumulant case is close to the LO phonon frequency value. Hence, we conclude that for the CBM of MgO, the spectral function shape is dominated by the LO phonon. First-principles and Fröhlich Hamiltonian approaches yield very similar shapes, although the Fröhlich approach underestimates the QP energy shift. This is an important result of the present work. The same conclusion will be obtained for the other band extrema, for both MgO and LiF.

In the LiF CBM case, with Fröhlich $\alpha = 4.01$ (Fig.17), the position of the quasi-particle peak in the DM case is much higher than in the cumulant case. The Fröhlich-only cumulant peak position is close to the full-band cumulant one. Because the value of α is larger than

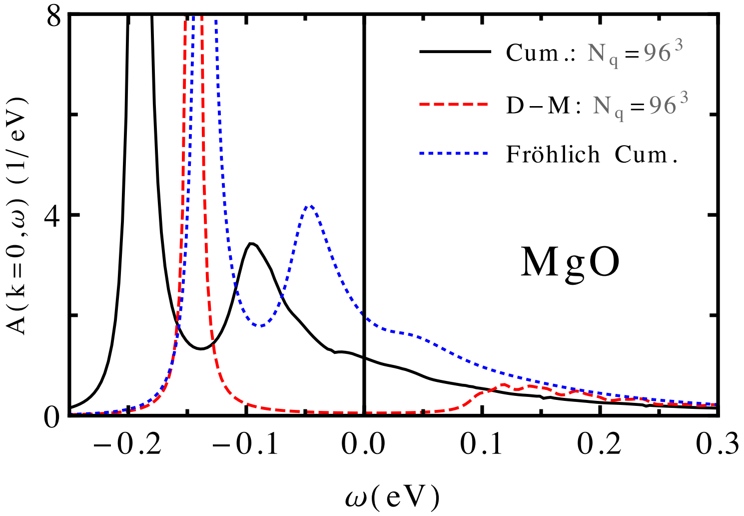


Figure 16: Cumulant (black) and Dyson-Migdal (dashed red) spectral functions for the conduction band minimum of MgO. The Fröhlich spectral function (dotted blue) with $\alpha = 1.62$ is also shown for comparison.

for MgO, LiF has a larger second satellite before the smoothing of the spectral function. The same observations as for the MgO CBM, concerning the shape and position of the peaks, can also be made.

Despite the MgO valence band being three-fold degenerate at Γ, the MgO VBM case, Fig.18, is actually very similar to the LiF CBM case, with positive energy shifts seen instead of negative energy shifts.

Finally, in the case of LiF VBM, with the Fröhlich α being at most 14.8 for the heavy hole effective mass, Fig.19, the cumulant spectral function has become a broad peak (similar to Fig. 12) without any quasiparticle peak or satellite structure, unlike in the DM case. The lack of structure in the cumulant spectral function results from the large value of the Fröhlich α, directly linked to the large hole effective mass, *i.e.* the rather flat LiF valence bands [159].

3.9 Summary and perspectives

In this work, we compute from first principles the electron-phonon self energy for the electronic states at the band extrema of MgO and LiF, at zero temperature, from which we obtain the spectral function using both Dyson-Migdal and the cumulant methodologies. We also analyze the self energy and spectral function of the Fröhlich Hamiltonian using the same methodologies.

For the imaginary part of the self energy, the characteristic inverse square root behavior of the Fan self energy found in the Fröhlich Hamiltonian, starting at the LO phonon frequency threshold, is also clearly present in MgO and LiF. However, additional structures, mirroring

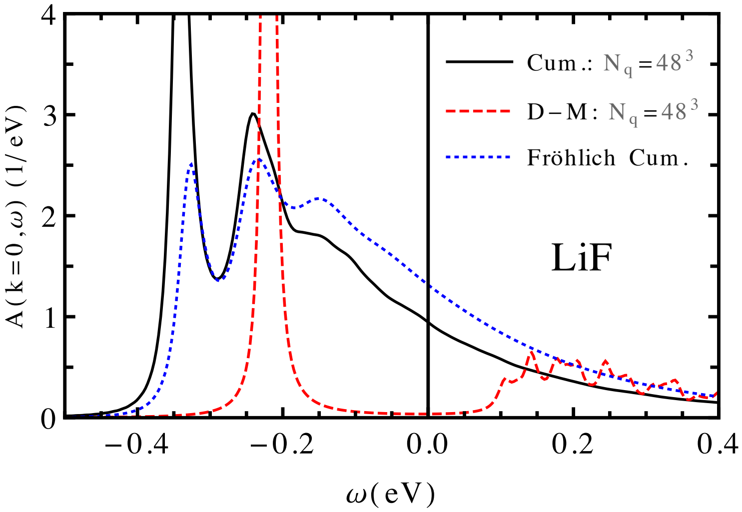


Figure 17: Cumulant (black) and Dyson-Migdal (dashed red) spectral functions for the conduction band minimum of LiF. The Fröhlich spectral function (dotted blue) with $\alpha = 4.01$ is also shown for comparison.

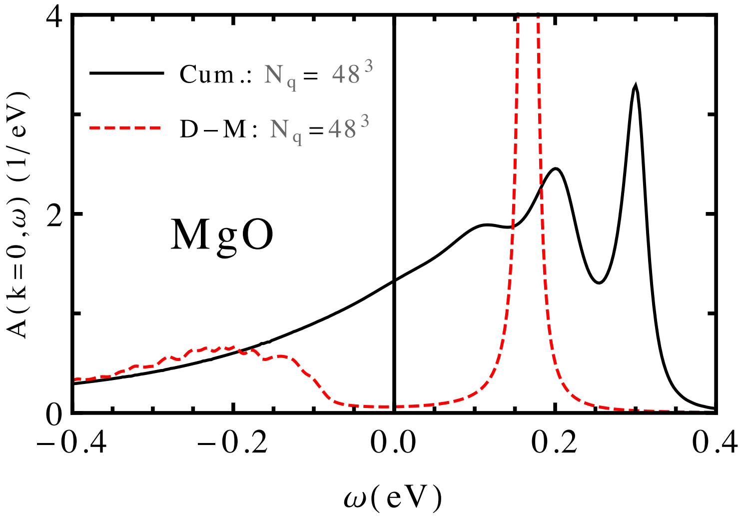


Figure 18: Cumulant (black) and Dyson-Migdal (dashed red) spectral functions for the valence band maximum of MgO.

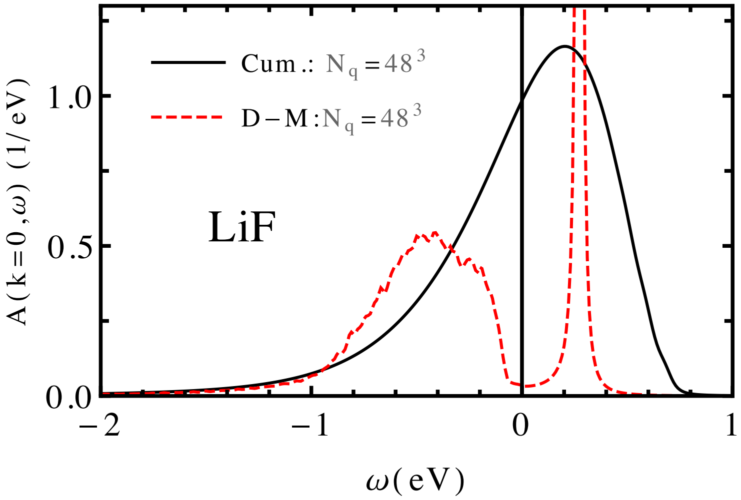


Figure 19: Cumulant (black) and Dyson-Migdal (dashed red) spectral functions for the valence band maximum of LiF.

the electronic DOS, are also present. They originate from occupied as well as unoccupied bands, for both electron and hole self energies. By contrast, the role of the phonon dispersion is minor.

The real part of the Fröhlich Hamiltonian self energy has also an inverse square root behavior, on the other side of the LO phonon frequency threshold, compared to the imaginary part. This feature is also present in first-principles calculations. The Debye-Waller self energy, that complements the Fan self energy, gives a global shift, impacting the zero-point renormalization of eigenenergies.

To correctly obtain the inverse square root behavior in first-principles self energies, the numerical integration over the phonon wavevectors has to be done carefully, and its convergence monitored. Numerical broadening of the denominator present in the Fan self energy helps to reach convergence, but affects the accuracy of the prediction. Although schemes to overcome such convergence problems start to appear [12, 15], we have simply performed the summation over a very fine grid, for a small broadening.

Our first-principles eigenenergy shifts range from 207 meV for the CBM of MgO to 751 meV for the VBM of LiF. Simple Fröhlich Hamiltonian estimates are too low although they account for a large fraction of such shift. There is cancellation between the Debye-Waller self energy (not taken into account in the Fröhlich Hamiltonian) and the Fan self energy. This cancellation is especially strong between the unoccupied part of the Fan self energy and the Debye-Waller self energy, in the case of the VBM for both MgO and LiF.

We analyze spectral functions from both cumulant and Dyson-Migdal approaches, using both the Fröhlich Hamiltonian and full DFT. Since the quasiparticle peak and the satellite location are both badly predicted from Dyson-Migdal theory even at low Fröhlich coupling constants, it is safe to argue that the shape of spectral function from the Dyson-Migdal approach can never be trusted for IR-active materials.

Our conclusions should apply to a large class of IR-active materials. In particular, the polar LO phonon contribution will give the biggest part of the zero-point renormalization and the spectral lineshape. It would be worth to examine more materials, and find a rule for when the Fröhlich Hamiltonian might give a reasonably accurate estimate of the electron-phonon quasiparticle shift. For the conduction band minimum, Fröhlich alone accounts for 4/5 of the CBM shift for LiF, but only 2/3 of the CBM shift for MgO. By contrast, the case of diamond is a remarkable example of a large QP shift without polar phonons. Its band edge zero-point renormalization (−330 meV for the indirect gap, and −416 meV for the direct gap) [70, 71] has a magnitude similar to those of the IR-active materials studied here.

Finally, angle-resolved photoemission (ARPES), although difficult in insulators, could provide a nice test of some of the predicted spectral functions in this Chapter. In particular, consider the MgO VBM shown in Fig. 18. We predict a quasiparticle carrying 20% of the spectral weight, and clear phonon satellites. These may be accessible to experiment. By contrast, consider the VBM of LiF, shown in Fig. 19. The prediction is a completely blurred quasiparticle ($Z = 0.008$), with a broad peak rather than distinct satellites. However, this prediction is outside the trust range of our approach. The α values for LiF are in the range 4 to 15, as opposed to 1.6 to 5.5 for MgO. Experiment [160] shows that a hole at the top of the valence band of LiF self-traps in a local distortion similar to an F_2^- ion. This trapped polaron is known as a V_K center [161–163]. The present simple Hamiltonian (e.g. without anharmonicities) is insufficient to yield such a bound state. The perturbative spectral function correctly tells us that there is negligible weight in the quasiparticle peak. However, for holes at the top of the valence band in materials like LiF and NaCl, perturbation theory is not enough.

3.10 Appendix I. First-principles calculations: technical details

All first-principles calculations are done using ABINITv8.4.1 (main executable abinit as well as postprocessor ElectronPhononCoupling). The computations use density functional theory (DFT, ground state and electronic structure) and density functional perturbation theory (DFPT, phonons and electron-phonon coupling), as described in Refs. [88, 89], with the PBE generalized gradient approximation (PBE-GGA) [164, 165] for MgO, and local density approximation (LDA) for LiF. See Ref. [55] for more information about the basic theories. The effective masses have been computed directly using the scheme from Ref. [166]. The Allen-Heine-Cardona formalism is used for the computation of the Fan and Debye-Waller self energies [3, 4, 137, 138]. The summation over unoccupied bands is handled efficiently thanks to a reformulation in terms of a Sternheimer equation [58]. However, in order to treat correctly the dynamical self energy, the contribution from a number of "active" bands is computed explicitly by a sum over states, as described in Ref. [20], see Eq.(15).

Atomic masses from natural isotopic ratios have been used throughout, namely (in atomic mass units), $M_{Mg} = 24.305$, $M_O = 15.9994$, $M_{Li} = 6.941$ and $M_F = 18.9984032$.

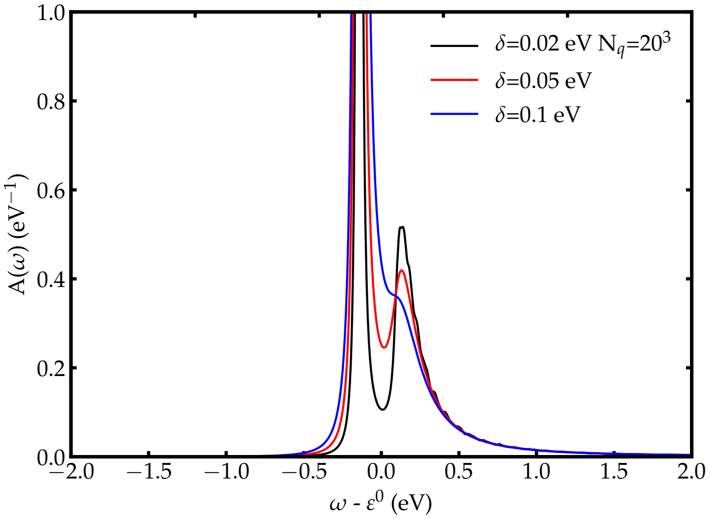


Figure 20: Spectral function (Dyson-Migdal) of the MgO CBM self energy with a 20^3 q-grid, and decreasing $\delta = 0.1, 0.05, 0.02$ eV. The separation between the quasiparticle peak and the satellite is still not complete with the smallest δ value.

Norm-Conserving Pseudopotentials for MgO were taken from the ABINIT web site, generated using ONCVPSP [167], while for LiF Troullier-Martins pseudopotentials have been used. For both MgO and LiF, we used a 8^3 Monkhorst-Pack grid sampling of the Brillouin Zone for the electronic wavefunctions, and a 50 Ha kinetic energy plane-wave cut-off. The summation over unoccupied bands for the dynamical self energy is made over 6 conduction bands for MgO, and 3 conduction bands for LiF, corresponding to a range of more than 10 eV above the CBM. The difference between the dynamical denominator (including the phonon frequency) and the static denominator (with only the electronic eigenenergy differences) is smaller than 1% above these bands, hence the summation over states can be safely replaced by the Sternheimer equation beyond them. Note that the replacement of an explicit sum over states, that includes a physical infinitesimal imaginary η, by a static contribution from the Sternheimer equation [20], destroys the Kramers-Kronig relations Eqs. 3.23 and 3.24. Of course, this has no influence on the real and imaginary parts directly computed in the low-energy region of the conduction bands or in the valence bands.

The phonon wavevector sampling that is needed to get converged self energies and spectral functions is a delicate issue, already mentioned in Ref. [138]. The most difficult case is the CBM of MgO, since this band disperses quite strongly, as can be deduced from its low effective mass. In order to get the imaginary part of the self energy, the wavevector sampling ought to sample electronic eigenenergies such that their differences are not larger than the LO phonon frequency. This is hard to achieve when the effective mass is small. As described in Ref. [20], see in particular Eq. (16), a broadening factor δ is used to smooth the computed functions. Too large a δ value prevents distinguishing characteristic features of the self energy or spectral function, which may typically appear at scales of the LO phonon frequency.

Table 4: MgO self energy (eV) and frequency derivative (adimensional) at $\omega = \varepsilon_{\mathbf{k}n}$, for the conduction band minimum, for different wavevector samplings and broadening factors δ .

Wavevector grid	δ(eV)	Σ	$\partial\Sigma/\partial\omega$
20^3	0.01	-0.1391 - i 0.0008	-0.1603 - i 0.0089
20^3	0.02	-0.1390 - i 0.0032	-0.1589 - i 0.0177
20^3	0.05	-0.1381 - i 0.0078	-0.1498 - i 0.0421
20^3	0.10	-0.1352 - i 0.0147	-0.1232 - i 0.0712
20^3	0.01	-0.1391 - i 0.0016	-0.1604 - i 0.0089
32^3	0.01	-0.1602 - i 0.0031	-0.3086 - i 0.0294
48^3	0.01	-0.1746 - i 0.0050	-0.4975 - i 0.0619
96^3	0.01	-0.1912 - i 0.0062	-0.6264 - i 0.0968

Fig. 20 presents the DM spectral function of the CBM of MgO with a 20^3 sampling, and δ values 0.1, 0.05 eV and 0.02 eV. Larger δ smoothes the function, but also modifies its shape: the clear and physical separation between the quasiparticle peak and the satellite is washed out.

The values of the self energy at the CBM are also strongly affected, as can be seen in Table 4. From the numbers in this table, one can appreciate the convergence of the real part of the self energy as the inverse of N_q, mentioned in Sec. VII, mathematically derived in Ref. [20]. This convergence study also highlights why the ZPR of the MgO CBM obtained in the present work, namely -207meV, is quite different from the same quantity presented in Table I of Ref. [138], $\Sigma^{dyn}(\epsilon^0)$=-143meV, although the same software and pseudopotentials have been used.

A 0.01 eV broadening appears maximum for a meaningful investigation of the different spectral functions and self energies. The self energy is even noisier than the spectral function (see Fig. 21). Going from a 20^3 sampling, that delivers the meaningless Fig. 22, to our best 96^3 grid, Fig. 13, requires increasing by two orders of magnitude the computational effort, introducing possible computer memory problems, and still leaves some noise.

The convergence parameters also impact the cumulant spectral functions, although they are more easily converged than the self energy or the DM spectral function. The cumulant spectral functions corresponding to the VBM of MgO with $\delta = 0.01$ eV and 20^3, 32^3, 48^3, and 64^3 grids are presented in Fig. 23.

3.11 Appendix II. Different flavors of cumulants

The calculations presented in this Chapter have been performed using the retarded cumulant. Another approach that has been used in the literature is the time-ordered (t-O) cumulant. At temperature $T = 0$, the t-O Green's function yields also a simple expression (cf. Eq. (3.1)),

$$A(k,\omega) = |\Im m G_{t-O}(k,\omega)|/\pi, \tag{3.39}$$

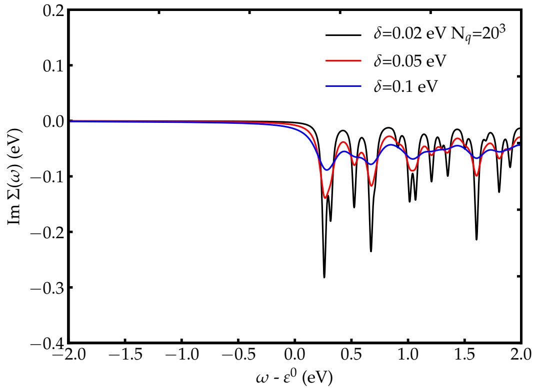


Figure 21: The imaginary part of the MgO CBM self energy with a 20^3 q-grid and decreasing $\delta = 0.1, 0.05, 0.02$ eV.

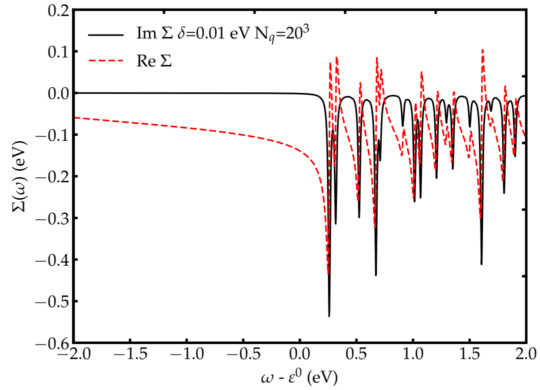


Figure 22: The MgO CBM self energy with $\delta = 0.01$ eV and a 20^3 q-grid. Full black line: imaginary part; dotted red line: real part.

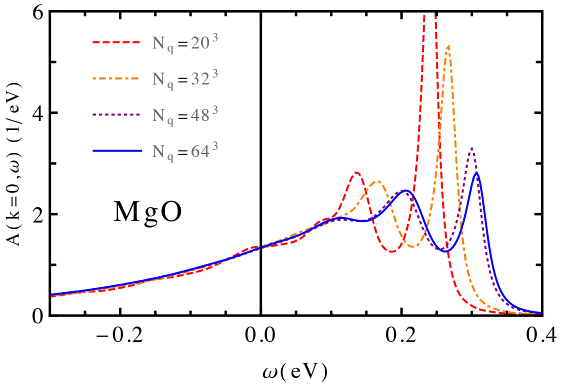


Figure 23: Cumulant spectral function of the MgO VBM with $\delta = 0.01$ eV and 20^3, 32^3, 48^3 and 64^3 q-grids.

since the $T = 0$ t-O function has the same imaginary part as $G_R(k, \omega)$, except for a sign change at $\omega = \mu$. In diagrammatic perturbation theory the t-O Green's function is normally computed at $T = 0$. However there are two main advantages to the retarded representation: (i) The retarded Green's function enters more naturally at $T > 0$, through analytic continuation of Matsubara Green's functions to the real frequency axis [98]. (ii) Previous t-O versions [168] typically make an additional approximation, by considering only one of the terms of the t-O Green's function:

$$\begin{aligned} G_{t-O}(k,t) &= -i\langle c_k(t)c_k^{+}(0)\rangle\theta(t) + i\langle c_k^{+}(0)c_k(t)\rangle\theta(-t) \\ &= G^{>}(k,t)\theta(t) + G^{<}(k,t)\theta(-t) \end{aligned} \tag{3.40}$$

where $G^{<}$ is the lesser Green's function and $G^{>}$ the greater Green's function. For the hole, only the $G^{<}$ term is kept and a cumulant expansion $G(t) = G_0(t)e^{C(t)}$ is written for $t < 0$. An analogous expansion can be written for $t > 0$. These expansions are more accurately described as lesser (hole) or greater (electron) cumulant expansions. This explains discrepancies in the literature between retarded and so called time-ordered representations, despite Eqs.(3.1) and (3.39) indicating the spectral functions should be the same (and second order approximations to the self-energy should yield very similar results). The retarded version, having only one θ prefactor, treats greater and lesser contributions on the same footing and naturally leads to a cumulant Green's function with satellites on both sides of the chemical potential.

As a result of (ii), the spectral function of the t-O cumulant is also not necessarily normalized to 1 (for more details, see Ref. [168]). However, the discrepancies disappear in core level spectra, where the cumulant approach, in its time-ordered version, is justified. Here, we will also see that for large gaps, the retarded cumulant and other versions yield only minor differences.

We first notice that there are separate contributions to the Fan self energy from the intermediate states $|\mathbf{k} + \mathbf{q}n'\rangle$ in the conduction bands (labeled 'un' for unoccupied) and in the valence band (labeled 'oc' for occupied) [169]. No matter whether the initial state $|\mathbf{k}n\rangle$ is from the valence or the conduction band, both contributions occur in Eq.(3.6),

$$\Sigma^{\mathrm{Fan}}(\mathbf{k}n, \omega) = \Sigma^{\mathrm{Fan}}_{\mathrm{un}}(\mathbf{k}n, \omega) + \Sigma^{\mathrm{Fan}}_{\mathrm{oc}}(\mathbf{k}n, \omega). \tag{3.41}$$

Explicit equations are

$$\Sigma^{\mathrm{Fan}}_{\mathrm{un}}(\mathbf{k}n, \omega) = \frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}s}^{\mathrm{BZ}} \sum_{n'}^{unocc} \frac{|\langle \mathbf{k}n|H_j^{(1)}|\mathbf{k}+\mathbf{q}n'\rangle|^2}{\omega - \varepsilon_{\mathbf{k}+\mathbf{q}n'} - \omega_{\mathbf{q}s} + i\eta}, \tag{3.42}$$

for the intermediate unoccupied state contribution to the self energy, and

$$\Sigma^{\mathrm{Fan}}_{\mathrm{oc}}(\mathbf{k}n, \omega) = \frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}s}^{\mathrm{BZ}} \sum_{n'}^{occ} \frac{|\langle \mathbf{k}n|H_s^{(1)}|\mathbf{k}+\mathbf{q}n'\rangle|^2}{\omega - \varepsilon_{\mathbf{k}+\mathbf{q}n'} + \omega_{\mathbf{q}s} + i\eta}, \tag{3.43}$$

for the intermediate occupied state contribution to the retarded self energy.

Thanks to Eqs.(3.23-3.24), the retarded cumulant of Eq.(3.21) can be rewritten as

$$\begin{aligned} C(\mathbf{k}n,t) &= \int_{-\infty}^{\infty} \beta(\mathbf{k}n,\omega)e^{-i\omega t}\Re e\frac{1}{(\omega+i\delta)^2}d\omega \\ &- it\Re e\Sigma^{\text{Fan}}(\mathbf{k}n,\varepsilon_{\mathbf{k}n}) + \frac{\partial\Re e\Sigma^{\text{Fan}}(\mathbf{k}n,\omega)}{\partial\omega}|_{\omega=\varepsilon_{\mathbf{k}n}}. \end{aligned} \tag{3.44}$$

In the t-O cumulant approach of Ref. [104], Eqs.(3.44) and (3.22) change as follows for the electrons:

$$\begin{aligned} C_e^{t-O}(\mathbf{k}n,t) &= \int_{\mu-\varepsilon_{\mathbf{k}n}}^{\infty} \beta_e^{t-O}(\mathbf{k}n,\omega)e^{-i\omega t}\Re e\frac{1}{(\omega+i\delta)^2}d\omega \\ &- it\Sigma^{\text{Fan}}(\mathbf{k}n,\varepsilon_{\mathbf{k}n}) + \frac{\partial\Sigma^{\text{Fan}}(\mathbf{k}n,\omega)}{\partial\omega}|_{\omega=\varepsilon_{\mathbf{k}n}}, \end{aligned} \tag{3.45}$$

$$\beta_e^{t-O}(\mathbf{k}n,\omega) = \frac{1}{\pi}|\Im m\,\Sigma_{\text{un}}^{\text{Fan}}(\mathbf{k}n,\omega+\varepsilon_{\mathbf{k}n})|. \tag{3.46}$$

Note the reduced range of the integral, as well as the selection of part of the self energy in the β factor. By contrast, the contributions that are either constant in time or linear in time are computed from the whole self energy. Unlike the retarded cumulant, this version of the t-O cumulant does not vanish at $t = 0$, nor does its time-derivative, which means that the spectral function is not normalized to 1, and its first moment is changed by the dynamical contribution. Corresponding expressions for the holes are also presented in Ref. [168]:

$$\begin{aligned} C_h^{t-O}(\mathbf{k}n,t) &= \int_{-\infty}^{\mu-\varepsilon_{\mathbf{k}n}} \beta_h^{t-O}(\mathbf{k}n,\omega)e^{-i\omega t}\Re e\frac{1}{(\omega+i\delta)^2}d\omega \\ &- it\Sigma^{\text{Fan}}(\mathbf{k}n,\varepsilon_{\mathbf{k}n}) + \frac{\partial\Sigma^{\text{Fan}}(\mathbf{k}n,\omega)}{\partial\omega}|_{\omega=\varepsilon_{\mathbf{k}n}}, \end{aligned} \tag{3.47}$$

$$\beta_h^{t-O}(\mathbf{k}n,\omega) = \frac{1}{\pi}|\Im m\,\Sigma_{\text{oc}}^{\text{Fan}}(\mathbf{k}n,\omega+\varepsilon_{\mathbf{k}n})|. \tag{3.48}$$

Our computations for MgO and LiF have such large band gaps that the differing limits of integration in Eqs. 3.44, 3.45, and 3.47 have negligible consequences. Similarly, the function $\beta(\mathbf{k}, n = c)$ is almost identical to β_e^{t-O} and $\beta(\mathbf{k}, n = v)$ is almost identical to β_h^{t-O}, in the relevant range of ω-integration where the denominator ω^2 is small. Also, the imaginary part of the self energy and its derivative, evaluated at the unrenormalized energy, are usually small (numerically this might not be the case due to the use of a broadening factor, see for example Fig. 4). This completes the comparison between Eq.(3.44), and Eq.(3.45) and Eq.(3.47).

Different approximations give rise to other representations, described in Ref. [168], where the full Fan self energy and its derivative are replaced in the last two terms of Eqs.(3.45) and (3.47), by their unoccupied and occupied counterparts, respectively, as in Ref. [113] (see Eq.36), where the normalization of the spectral function is again one. While the *retarded* cumulant weights (or the t-O versions introduced here) can be decomposed in their hole and electron factors, following Eqs.(3.41) and (3.26):

$$Z^R_{\mathbf{k}n} = Z^{\mathrm{un}}_{\mathbf{k}n} Z^{\mathrm{oc}}_{\mathbf{k}n}, \tag{3.49}$$

only one of these factors is included in the t-O cumulant weights of Refs. [104], [105], [113], and [168] (Eq. (3.64)). However, as can be judged by the closeness to unity of $Z^e_{\mathbf{k}n}$ in the VBM case and $Z^h_{\mathbf{k}n}$ in the CBM case (see Table II) , in MgO and LiF, the normalization defect is very small. The smallest of these weights is at least 0.988. This lack of impact of the unoccupied states on the VBM self energy, and of the occupied states on the CBM self energy can be traced back to the large ratio between the electronic gap and the largest phonon frequency. This might not be true for small-gap semiconductors. Therefore, t-O approximations present in the literature seem safe for wide gap semiconductors, but the retarded approach is in general preferable.

Chapter 4

Temperature dependence of the (222) X-ray forbidden reflection in silicon

Crystals with an *fcc* lattice like silicon have Bragg scattering at $\mathbf{K} = 2\pi(h, k, l)/a$ for integers hkl all even or all odd. The two-atom basis of the diamond structure causes destructive interference whenever $h + k + l$ is an odd multiple of 2; for example, the (222) reflection is nominally forbidden. However, there is not total interference because of tetrahedral rather than spherical scattering symmetry. Such asymmetry arises from anharmonic vibrations and from bonding. Therefore, the weakly allowed (222) X-ray reflection in silicon is useful for studying bond charge. Temperature variation of the (222) X-ray intensity has been measured [170] and studied [5, 170] beyond that expected from anharmonicity. Previous theories have been somewhat ad hoc, not dealing fully with electron-phonon induced valence charge density thermal shifts. Namely, the valence charge was assumed to vibrate rigidly around the mid-point of the ions, and Debye-Waller (DW) factors were used. Here we obtain a full second-order perturbation theory expression for the change of the charge density with temperature, and use it to study the forbidden reflection in silicon. An acoustic sum rule is used to express the DW-type term in terms of Fan-type terms to simplify numerical calculations. An Appendix also includes a single-particle derivation, which coincides with the many-body approach. In future work, we will apply it to study the contribution of the electron-phonon interaction to pyroelectricity.

4.1 Introduction

As is well-known, in X-ray diffraction there is a peak of intensity when the difference between the incident and reflected wavevector is a reciprocal lattice vector of the Bravais lattice of the crystal. Buf if there is more than one atom per unit cell, destructive interference can drastically reduce the intensity of the peak. When considering spherical symmetry for the charge distribution of each atom, the intensity of the (222) peak in silicon becomes 0. It is said to be “forbidden” (see [171], p. 106 in Chapter 6). However, there is a small intensity, two orders of magnitude smaller than in a normal Bragg peak [172], due to tetrahedral scattering symmetry. Figure 24 shows the temperature dependence of the X-ray diffraction intensity of the (222) peak [170].

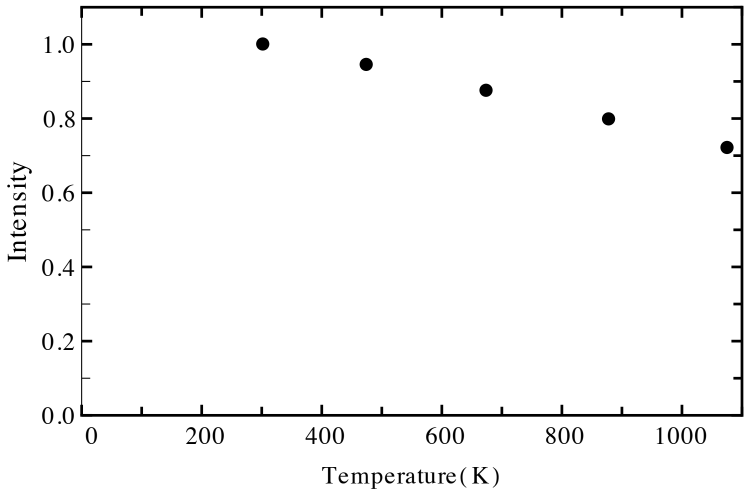


Figure 24: Experimental temperature dependence of the (222) intensity peak, normalized at room temperature. Extracted from [170].

The tetrahedral symmetry has two contributions: the distribution of the bond charge, and the anharmonicity of the movement of the ions (the core charge moves rigidly with the nuclei). The anharmonicity of the valence charge is of higher order. In Fig. 25 we can see in black the structure factor (square root of the intensity) of the experimental data shown in Fig. 24, normalized to 1 at 296K. The structure factor associated to anharmonicity is given by the green data points, and corresponds to a neutron diffraction experiment. Its absolute value increases with temperature, as opposed to the bond contribution. Although it is considerably smaller than the bond contribution (of the order of 5% at 1100K, and smaller at lower temperatures), it does contribute significantly to the *change* of the structure factor with temperature (close to a half of the bond charge contribution from about 300 to $1100K$). All the data of the figure was obtained from [170].

Here we focus on calculating the temperature dependence of the bond contribution to the (222) forbidden reflection. In non-forbidden peaks, the use of DW factors is justified in that the main contribution to the Bragg peaks comes from the core electrons, which move rigidly with the nuclei. We do not expect the valence charge to move rigidly, but this is what previous studies have assumed.

The temperature dependence of the (222) forbidden reflection in silicon was investigated [5,170,172–174] before the theory of temperature effects in crystals became well established. These studies considered the valence charge centered mid-way between nearest neighbors, and assumed it moved rigidly under different models (dominated by acoustic or optical modes) that gave different DW factors (equal or smaller than the ionic DW factor). However, the agreement with experiment was not satisfactory, and it was recognized that charge distribution might be changing as well, as opposed to just moving rigidly. Phillips [175] suggested that the amount of bonding charge decreases, and to this effect, Ref. [5] included a DW factor in the potential, trying to improve the previous models. In this way the authors

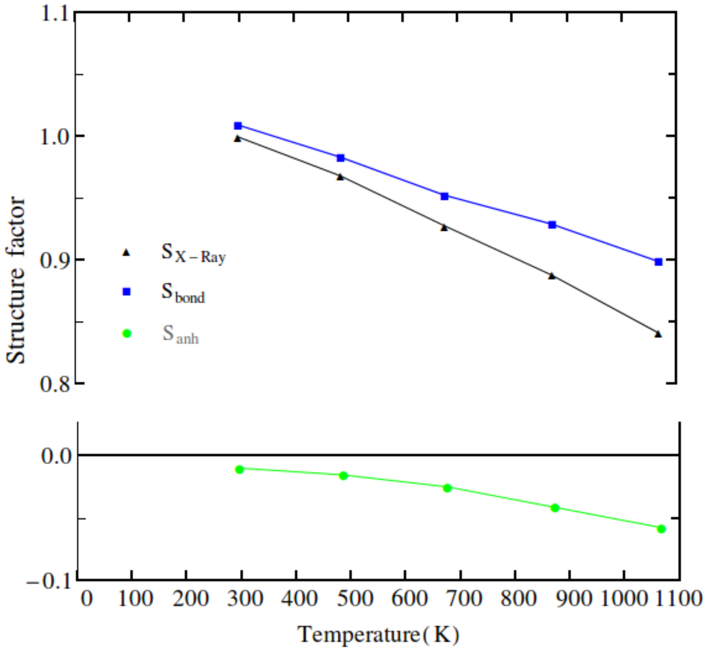


Figure 25: Temperature dependence of the (222) X-ray structure factor (square root of the intensity of Fig. 24), normalized to 1 at 296K. The structure factor due to anharmonicity, obtained from neutron diffraction, is given by the green curve. Subtracting it from the X-ray data, results in the blue curve, the contribution from the bond.

obtained a temperature dependent charge distribution, and then Fourier transformed it to obtain the scattering factor (see the following paragraphs). By considering an additional DW factor to take into account bond motion, a good agreement with experiment was obtained. However, this is not justified. The objective of this Chapter is to address the problem with a correct perturbative expression for the charge density.

First, we define the structure or scattering factor, an amplitude that considers the phase differences of X-rays scattered off different points of the charge distribution $\rho(\mathbf{r})$ [176]:

$$S_{\mathbf{K}} = \int \rho(\mathbf{r}) e^{i\mathbf{K}.\mathbf{r}}. \tag{4.1}$$

The explicit dependence of the charge density on the ionic positions is omitted. The treatment is in principle general, but we will be focusing on the valence charge. The intensity is proportional to the square of the absolute value of the scattering factor. Experimentally, the measurement takes place over a finite exposure time with a macroscopic sample, so the intensity is time (and therefore thermally) averaged:

$$I \;\propto \int d\mathbf{r}d\mathbf{r}' \langle \rho(\mathbf{r})\rho(\mathbf{r}')\rangle e^{i\mathbf{K}(\mathbf{r}-\mathbf{r}')} \tag{4.2}$$

As we show in the Appendix, for coherent scattering to second order, this thermal average factorizes, $\langle\rho(\mathbf{r})\rho(\mathbf{r}')\rangle = \langle\rho(\mathbf{r})\rangle\langle\rho(\mathbf{r}')\rangle$. Thus, $I \propto |\langle S_{\mathbf{K}}\rangle|^2$, and we calculate

$$S_{\mathbf{K}} = \int \langle\rho(\mathbf{r})\rangle e^{i\mathbf{K}.\mathbf{r}}, \tag{4.3}$$

which we denote with the same symbol as the structure factor in Eq. (4.1). In the next section, we obtain a temperature dependent expression $\langle\rho(\mathbf{r})\rangle_T$, which is the main contribution of this Chapter.

4.2 Thermal renormalization of charge density

The electronic charge density is the thermal expectation value of the charge density operator,

$$\rho(\mathbf{r},T) = \langle\hat{\psi}^\dagger(\mathbf{r})\hat{\psi}(\mathbf{r})\rangle_T, \tag{4.4}$$

$$\langle\hat{A}\rangle_T = \frac{\mathrm{Tr}e^{-(H-\mu N)/k_BT}\hat{A}}{\mathrm{Tr}e^{-(H-\mu N)/k_BT}}. \tag{4.5}$$

Here $\hat{\psi}^\dagger(\mathbf{r})$ is the operator that creates an electron at point $\mathbf{r}$. It can be expanded in any orthonormal complete basis set. The most natural such set is the Bloch-wave eigenstates of a self-consistent Kohn-Sham [53, 54] single particle Hamiltonian $H_0 = \sum \varepsilon_{\mathbf{k}n}c^\dagger_{\mathbf{k}n}c_{\mathbf{k}n}$,

$$\hat{\psi}^\dagger(\mathbf{r}) = \sum_{\mathbf{k}n} \psi^*_{\mathbf{k}n}(\mathbf{r})c^\dagger_{\mathbf{k}n} \tag{4.6}$$

where $c^\dagger_{\mathbf{k}n}$ creates an electron in the Bloch state $\psi_{\mathbf{k}n}(\mathbf{r})$. Then the charge density is

$$\rho(\mathbf{r},T) = 2\sum_{\mathbf{k}nn'} \psi^*_{\mathbf{k}n'}(\mathbf{r})\psi_{\mathbf{k}n}(\mathbf{r})\langle c^\dagger_{\mathbf{k}n'}c_{\mathbf{k}n}\rangle_T. \tag{4.7}$$

where the 2 comes from the implicit sum over spin indices in Eq. (4.6).

In a non-interacting theory, the factor $\langle c^\dagger_{\mathbf{k}n'}c_{\mathbf{k}n}\rangle$ is $f(\varepsilon_{\mathbf{k}n})\delta_{n,n'}$, where $f(\varepsilon_{\mathbf{k}n}) = (\exp((\varepsilon_{\mathbf{k}n} - \mu)/k_BT) + 1)^{-1}$ is the Fermi-Dirac occupation of the state $\mathbf{k}n$ with single particle energy $\varepsilon_{\mathbf{k}n}$. More generally it is a limit of a thermal Green's function [27],

$$\langle c^\dagger_{\mathbf{k}n'}c_{\mathbf{k}n}\rangle = -i\mathrm{Lim}_{t\to 0}G_<(\mathbf{k}nn',t). \tag{4.8}$$

The version of the Green's function convenient for perturbation theory is the Matsubara version, time-ordered on the imaginary time axis,

$$G(\mathbf{k}nn',\tau) = -\langle\hat{T}_\tau c_{\mathbf{k}n}(\tau)c^\dagger_{\mathbf{k}n'}(0)\rangle_T, \tag{4.9}$$

where $\hat{T}_\tau$ is the time-ordering operator (see Eq. (1.10)) that puts $c_{\mathbf{k}n}$ to the left of $c^\dagger_{\mathbf{k}n'}$ when $\tau > 0$ and *vice versa*, with a sign change, otherwise. For non-interacting electrons, the result, after a discrete Fourier transform (using antiperiodicity along the imaginary time τ axis, see Eq. (1.12)), has the familiar form

$$G_0(\mathbf{k}nn', i\omega_n) = (i\omega_n - \varepsilon_{\mathbf{k}n})^{-1}\delta_{n,n'}. \tag{4.10}$$

Here $\omega_n = (2n+1)\pi k_B T$ is a "Matsubara frequency" and $\hbar = 1$. The full Green's function $G(\mathbf{k}nn', i\omega_n)$ can be represented in a diagrammatic perturbation theory, and can be analytically continued to real frequencies ω, as discussed in Sec. 1.1.1. The answer has the form

$$G_R(\mathbf{k}nn', \omega)^{-1} = (\omega + i\eta - \varepsilon_{\mathbf{k}n})\delta_{n,n'} - \Sigma(\mathbf{k}nn', \omega) \tag{4.11}$$

This Green's function, when Fourier transformed back to real time t, is the retarded Green's function,

$$G_R(\mathbf{k}nn', t) = -i\theta(t)\langle\{c_{\mathbf{k}n}(t), c^\dagger_{\mathbf{k}n'}(0)\}\rangle_T. \tag{4.12}$$

This Green's function is very useful. Its self energy $\Sigma(\mathbf{k}nn', \omega)$ (especially the diagonal part $n = n'$), is the familiar version, and its imaginary part $-1/\pi \mathrm{Im} G_R(\mathbf{k}nn, \omega + i\eta)$ is the spectral function, experimentally accessible. In the time domain, G_R is not directly related to the function $G_<$ needed in Eq. (4.8). For the band diagonal ($n' = n$) part, the connection in the frequency domain was proved in Sec. 1.1.1 using Lehmann representations [27]:

$$G_<(\mathbf{k}nn, \omega) = -2if(\omega)\mathrm{Im}G_R(\mathbf{k}nn, \omega), \tag{4.13}$$

where f is again the Fermi-Dirac function. Extending this to off-diagonal parts raises the problem that phases of orbitals $\psi_{\mathbf{k}n}$ are arbitrary, but must cancel from the answer for $\rho(\mathbf{r})$. A formula that does this correctly is

$$G_<(\mathbf{k}nn', \omega) = -2if(\omega)\langle\mathbf{k}n'|\mathrm{Im}G^R(\omega + i\eta)|\mathbf{k}n\rangle. \tag{4.14}$$

This enables the charge density to be written by Fourier transforming back to the time domain, and taking the limit as t goes go 0,

$$\begin{aligned} \rho(\mathbf{r}, T) &= -4\sum_{\mathbf{k}nn'} \psi^*_{\mathbf{k}n'}(\mathbf{r})\psi_{\mathbf{k}n}(\mathbf{r}) \\ &\times \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega f(\omega)\langle\mathbf{k}n'|\mathrm{Im}G^R(\omega + i\eta)|\mathbf{k}n\rangle e^{i\omega 0^+}. \end{aligned} \tag{4.15}$$

This is in principle an exact formula for the charge density of a many body system. It is also a useful starting point for finding perturbative formulas. There is now some interest in T-dependent density functional theory [177], originating from Mermin's extension [178] of Kohn-Sham theory. Possibly this formula would have some application there.

4.3 Electron self energy from phonons

The retarded self energy $\Sigma(\mathbf{k}nn',\omega)$ for electrons interacting with phonons is well-known in lowest order. The interaction needed is the Born-Oppenheimer potential $V(\mathbf{r},\{\mathbf{R}_{li}\})$, expanded in displacements $\mathbf{u}_{li}=\mathbf{R}_{li}-\mathbf{R}^0_{li}$, to second order:

$$H_1' = \sum_{\mathbf{k}nm,\mathbf{q}s} V^{(1,\mathbf{q}s)}_{\mathbf{k}+\mathbf{q}m,\mathbf{k}n} c^\dagger_{\mathbf{k}+\mathbf{q}m} c_{\mathbf{k}n}(a_{\mathbf{q}s}+a^\dagger_{-\mathbf{q}s}) \tag{4.16}$$

$$H_2' = \sum_{\mathbf{k}nn',\mathbf{q}s} V^{(2,\mathbf{q}s)}_{\mathbf{k}n',\mathbf{k}n} c^\dagger_{\mathbf{k}n'} c_{\mathbf{k}n}(a_{\mathbf{q}s}+a^\dagger_{-\mathbf{q}s})(a_{-\mathbf{q}s}+a^\dagger_{\mathbf{q}s}) \tag{4.17}$$

The resulting retarded self-energy (the analytic continuation of the Matsubara $\Sigma(\mathbf{k}nn',i\omega_n)$ to the real axis from above) has two parts, as mentioned earlier, $\Sigma^{\mathrm{Fan}}+\Sigma^{\mathrm{DW}}$. The Fan term involves two factors of $V^{(1)}$ (second order perturbation theory), while the Debye-Waller term has only one factor of $V^{(2)}$ (first order perturbation theory). The DW process necessarily involves reabsorbing at the same vertex, the same phonon ($\mathbf{q}s$) that is emitted from the vertex. That is why only one wave vector $\mathbf{q}$ and mode s is kept in Eq. (4.17), and only the zero momentum transfer part of the $V^{(2)}$ interaction is shown. In fourth order perturbation theory, new pieces would occur where $V^{(2)}$ could make transitions from ($\mathbf{k}n$) to ($\mathbf{k}+\mathbf{q}_1+\mathbf{q}_2n'$), but these are not needed here since we will stop at second order. With the notation we are using in this Chapter, the formulas are [32]

$$\begin{aligned}\Sigma^{\mathrm{Fan}}(\mathbf{k}nn',\omega) &= \sum_{m,\mathbf{q}s} V^{(1,-\mathbf{q}s)}_{\mathbf{k}n,\mathbf{k}+\mathbf{q}m} V^{(1,\mathbf{q}s)}_{\mathbf{k}+\mathbf{q}m,\mathbf{k}n'} \\ \times &\left\{\frac{1-f(\varepsilon_{\mathbf{k}+\mathbf{q}m})+n_{\mathbf{q}s}}{\omega+i\eta-\varepsilon_{\mathbf{k}+\mathbf{q}m}-\omega_{\mathbf{q}s}} + \frac{f(\varepsilon_{\mathbf{k}+\mathbf{q}m})+n_{\mathbf{q}s}}{\omega+i\eta-\varepsilon_{\mathbf{k}+\mathbf{q}m}+\omega_{\mathbf{q}s}}\right\},\end{aligned} \tag{4.18}$$

$$\Sigma^{\mathrm{DW}}(\mathbf{k}nn',\omega) = \sum_{\mathbf{q}} V^{(2,\mathbf{q})}_{\mathbf{k}n,\mathbf{k}n'}(2n_{\mathbf{q}s}+1) \tag{4.19}$$

where $n_{\mathbf{q}s}=(\exp(\hbar\omega_{\mathbf{q}s}/k_BT)-1)^{-1}$ is the Bose-Einstein thermal occupation for phonon mode $\mathbf{q}s$.

4.4 $\Delta\rho$ in lowest order adiabatic approximation

To achieve the modest goal of a lowest order theory for the change $\Delta\rho$ of the charge density, consider the equation $G=G_0+G_0\Sigma G$ from which Eq. (4.11) was derived. Instead of solving this Dyson equation correctly for G, write it as the original expansion $G=G_0+G_0\Sigma G_0+\ldots$ from which it was derived, and truncate after first order. When doing so, the spectral function is not normalized to 1, and normalizing adds an extra term propotional to G_0.

Then Eq. (4.15) in lowest order gives

$$\Delta\rho(\mathbf{r},T) = 2\sum_{\mathbf{k}nn'} \psi^*_{\mathbf{k}n'}(\mathbf{r})\psi_{\mathbf{k}n}(\mathbf{r})\left[\Delta\rho^{\mathrm{Fan}}_{\mathbf{k}nn'}+\Delta\rho^{\mathrm{DW}}_{\mathbf{k}nn'}+\Delta\rho^{\mathrm{Norm}}_{\mathbf{k}nn'}\right], \tag{4.20}$$

$$\begin{aligned}\Delta\rho^{\rm Fan}_{\mathbf{k}nn'} &= -\frac{1}{\pi}\int_{-\infty}^{\infty} d\omega f(\omega){\rm Im}\left[\frac{1}{\omega+i\eta-\varepsilon_{\mathbf{k}n}}\right.\\ &\left.\times\Sigma^{\rm Fan}(\mathbf{k}nn',\omega)\frac{1}{\omega+i\eta-\varepsilon_{\mathbf{k}n'}}\right], \qquad (4.21)\end{aligned}$$

and similarly for the DW function $\Delta\rho^{\rm DW}_{\mathbf{k}nn'}$; the last term $\Delta\rho^{\rm Norm}_{\mathbf{k}nn'}$ comes from the term proportional to G_0. Remembering the discussion of Eq. (4.14), the $n \neq n'$ parts of $V^{(1,-\mathbf{q}s)}_{\mathbf{k}n,\mathbf{k}+\mathbf{q}m}V^{(1,\mathbf{q}s)}_{\mathbf{k}+\mathbf{q}m,\mathbf{k}n'}$ and $V^{(2,\mathbf{q}s)}_{\mathbf{k}n,\mathbf{k}n'}$ have random phases which should be ignored when taking the imaginary part. The DW term has ${\rm Im}(G_0\Sigma^{\rm DW}G_0)$. This can be evaluated as ${\rm Im}(G_0)\Sigma^{\rm DW}{\rm Re}(G_0) + {\rm Re}(G_0)\Sigma^{\rm DW}{\rm Im}(G_0)$. Then use ${\rm Im}G_0(k,\omega) = -\pi\delta(\omega-\varepsilon_k)$. From this it follows that

$$\Delta\rho^{\rm DW}_{\mathbf{k}nn'} = \sum_{\mathbf{q}} V^{(2,\mathbf{q}s)}_{\mathbf{k}n,\mathbf{k}n'}\frac{f(\varepsilon_{\mathbf{k}n})-f(\varepsilon_{\mathbf{k}n'})}{\varepsilon_{\mathbf{k}n}-\varepsilon_{\mathbf{k}n'}}(2n_{\mathbf{q}s}+1). \qquad (4.22)$$

and we see the term with $f(\varepsilon_{\mathbf{k}n'})$ is the complex conjugate of the term with $f(\varepsilon_{\mathbf{k}n})$.

The Fan contribution is more complicated because of the extra imaginary parts coming from the denominators within the $\{\}$ of Eq. (4.18). There are contributions to $\Delta\rho^{\rm Fan}_{\mathbf{k}nn'}$ analogous to Eq. (4.22) that involve the real part of the piece $\{\}$, and other contributions not analogous to Eq. (4.22) that involve the imaginary part of $\{\}$. The expressions simplify if we ignore the $\pm\omega_{\mathbf{q}s}$ in the denominators of $\{\}$. Then we get $\{\} = (2n_{\mathbf{q}s}+1)/(\omega+i\eta-\varepsilon_{\mathbf{k}+\mathbf{q}m})$. The terms analogous to Eq. (4.22) are

$$\begin{aligned}\Delta\rho^{\rm Fan,A}_{\mathbf{k}nn'} &= \sum_{\mathbf{q}m} V^{(1,-\mathbf{q}s)}_{\mathbf{k}n,\mathbf{k}+\mathbf{q}m}V^{(1,\mathbf{q}s)}_{\mathbf{k}+\mathbf{q}m,\mathbf{k}n'}\frac{1}{\varepsilon_{\mathbf{k}n}-\varepsilon_{\mathbf{k}n'}}\\ &\times\left[\frac{f(\varepsilon_{\mathbf{k}n})}{\varepsilon_{\mathbf{k}n}-\varepsilon_{\mathbf{k}+\mathbf{q}m}}-\frac{f(\varepsilon_{\mathbf{k}n'})}{\varepsilon_{\mathbf{k}n'}-\varepsilon_{\mathbf{k}+\mathbf{q}m}}\right](2n_{\mathbf{q}s}+1)\\ &. \qquad (4.23)\end{aligned}$$

and again the terms are the complex conjugate of each other. The terms not analogous to Eq. (4.22) are

$$\begin{aligned}\Delta\rho^{\rm Fan,B}_{\mathbf{k}nn'} &= \sum_{\mathbf{q}m} V^{(1,-\mathbf{q}s)}_{\mathbf{k}n,\mathbf{k}+\mathbf{q}m}V^{(1,\mathbf{q}s)}_{\mathbf{k}+\mathbf{q}m,\mathbf{k}n'}\\ &\times\left[\frac{f(\varepsilon_{\mathbf{k}+\mathbf{q}m})}{(\varepsilon_{\mathbf{k}n}-\varepsilon_{\mathbf{k}+\mathbf{q}m})(\varepsilon_{\mathbf{k}n'}-\varepsilon_{\mathbf{k}+\mathbf{q}m})}\right](2n_{\mathbf{q}s}+1). \qquad (4.24)\end{aligned}$$

Regarding the normalization of the spectral function for $G = G_0 + G_0\Sigma G_0$, the contribution from G_0 (with $n = n'$) is $-1/\pi\int d\omega G_0 = 1$, and the contribution from $G_0\Sigma G_0$ is

$$-\frac{1}{\pi}\mathrm{Im}\int_{-\infty}^{\infty} d\omega \frac{1}{\omega + i\eta - \varepsilon_{\mathbf{k}n}} \Sigma(\mathbf{k}nn, \omega) \frac{1}{\omega + i\eta - \varepsilon_{\mathbf{k}n}}. \tag{4.25}$$

This is very similar to (4.21), without $f(\omega)$, and we proceed in the same way to take the imaginary part. The part coming from the DW term is 0, because $\mathrm{Re}[1/(\varepsilon_{\mathbf{k}n} + i\eta - \varepsilon_{\mathbf{k}n})]$ is 0. For the Fan term, the result is the same as Eq. (4.24), with $n = n'$, and no Fermi-Dirac factor (the result in Eq. (4.23) is 0 when $n = n'$). Let

$$A_{\mathbf{k}n} = \sum_{m\mathbf{q}s} |V^{(1,-\mathbf{q}s)}_{\mathbf{k}n,\mathbf{k}+\mathbf{q}m}|^2 \frac{1}{(\varepsilon_{\mathbf{k}n} - \varepsilon_{\mathbf{k}+\mathbf{q}m})^2}(2n_{\mathbf{q}s} + 1). \tag{4.26}$$

Then $G = G_0(1 - A) + G_0\Sigma G_0$ is normalized to 1, and

$$\Delta\rho^{\mathrm{Norm}}_{\mathbf{k}nn'} = A_{\mathbf{k}n} f_n \delta_{nn'}. \tag{4.27}$$

The omission of $\pm\omega_{\mathbf{q}}$ in energy denominators is an adiabatic approximation, indicating that electrons are sensitive to the instantaneous static disorder of the vibrations, but not to the time-dependence of vibrational motion. It is known [20] that this approximation fails in the case of polar vibrations (LO modes) in insulators, where the (Fröhlich) coupling is particularly strong at small $|\mathbf{q}|$. The other case where an adiabatic approximation fails is in metals for $\varepsilon_{\mathbf{k}n}$ near the Fermi energy, and the temperature lower than $\omega_{\mathbf{q}s}$. The sharp Fermi edge singularity contributes an important resonance which goes away when the Fermi edge broadens. For a non-polar semiconductor like silicon, the topic of this Chapter, the adiabatic approximation should be fine. The full non-adiabatic theory can easily be worked out by keeping the $\pm\omega_{\mathbf{q}s}$ when integrating Eq. (4.21) to get $\rho(\mathbf{r}, T)$.

4.5 Results

The final expression that we want to calculate is included in the Appendix, Eq. (4.37). As in Chapters 2 and 3, the electronic energies, interatomic force constant and electron-phonon matrix elements are calculated using ABINIT [88, 89]. This expression also requires using the wavefunctions $\psi_{n\mathbf{k}}$, which are obtained in real space using the ABINIT postprocessing tool *cut3d*. The sum over the phonon modes s, the atoms i and j, phonon wave vectors $\mathbf{q}$ and bands is also present in Eq. (1.7). Equation (4.37) aditionally involves a sum over the electron wave vectors $\mathbf{k}$, and the sum over two additional band indices. Because of the Fermi-Dirac factor f, one of the indices is always restricted to the occupied bands (four of them in silicon), while the other two indices are unrestricted.

Hundreds of unoccupied bands are needed to obtain converged results when calculating the renormalization of electronic energies [137]. The sum over bands can be avoided by using a Sternheimer equation (see Sec. 1.5.1, Eq. (1.93)), making calculations considerably faster. In Eq. (2.1), the numerator is always positive, and the denominator does not change sign when n' takes unoccupied values. In Eq. (4.37), the situation is similar for the denominators. However, n and n' take in general different values, so the wavefunctions and electron-phonon matrix elements are not squared. We expect that the sum will require less unoccupied bands to obtain converged results because of the sign changes.

In Fig. 26, we show a cross section of the charge density. The ions are placed in (0,0,0) and $a/4(1,1,1)$, where a is the lattice constant (5.40 $\mathring{A}$ in the calculations). The plane is along the diagonal, with $z = (x+y)/2$. The charge is denser where the color is more yellow (we only show the pseudo-valence charge, so all the charge has the same sign). Thus, the charge is concentrated in-between nearest neighbors. The two intense circular yellow regions are equivalent to the one in (0,0,0) and $a/4(1,1,1)$, but associated to ions outside of the plane.

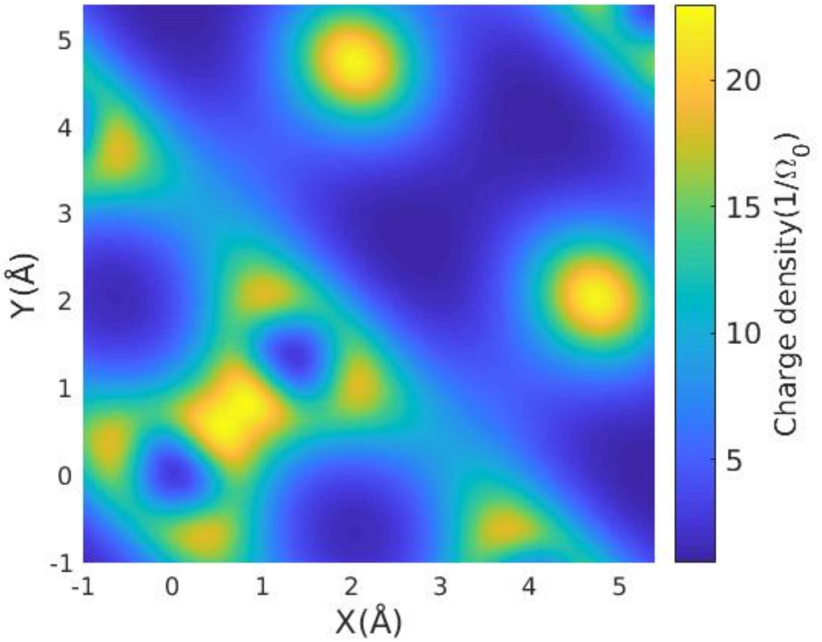


Figure 26: Electronic charge density across the plane $z = (x+y)/2$ (that is, that includes the origin and with normal (1,1,-2)), without EPI (lattice fixed at the equilibrium position). Ions are in positions (0,0,0) and $a/4(1,1,1)$. The x and y axis have units of $\mathring{A}$, and range from $-1\mathring{A}$ to $a = 5.40\mathring{A}$, the length of the unit cell in the calculations (points (0,0,0) and $a(1,1,1)$, the upper right corner of the figure, are equivalent). For the charge we still use atomic units ($e = 1$). Thus the intensity bar has units of $1/\Omega_0$, with $\Omega_0 = a^3/4$. We calculate the wavefunctions $\psi_{\mathbf{k}n}$, that determine the density $\rho = 2\sum_{\mathbf{k},n=1,4} |\psi_{\mathbf{k}n}|^2$, using ABINIT.

The electronic energies, IFC, wavefunctions and electron-phonon matrix elements appearing in Eq. (4.37) are well converged with an energy cutoff of 400 eV and a $6 \times 6 \times 6$ k-grid. Although the energy renormalization requires dense q-grids, here a $10 \times 10 \times 10$ q-grid provides good results. This is probably mitigated by the fact that the sum also involves the k-grid, helping with cancellations. The structure factor is fully converged with a $30 \times 30 \times 30$ q-grid, and we obtained the same results with a $10 \times 10 \times 10$ q-grid and $8 \times 8 \times 8$ k-grid. On the other hand, convergence with the number of bands proves hard.

In Fig. 27 we show the ZPR of the charge density for different number of bands. The interaction with the phonons smoothens out the charge density: the change in density places more electronic charge in the position of the ions (where the valence charge was less intense), and removes charge from the regions where the intensity was stronger in Fig. 26. As the number of bands increases, the highest value of the charge density for $n = 10, 20, 40, 48$ and 60 is 0.005, 0.0075, 0.032, 0.0379 and 0.0403 (units of $1/\Omega_0$), respectively. This corresponds to the density in the position of the ions increasing with the number of bands. The most negative values are -0.0189, -0.0244, -0.0264, -0.0252 and -0.0236, so there is not a clear trend. The variations in the "background" are smaller and seem better converged for a smaller number of bands.

Figure 28 shows the total relative change of the structure factor at $T = 1000K$, and also the contribution from the last term in Eq. (4.37) (there are seven terms in this equation, which we refer to as terms 1 to 7). Term 7 seems relatively well converged at about 20 bands, but the convergence of the total result is not clear. To identify if any particular term is harder to converge, we calculate each contribution separately.

In Fig. 29, the variation is of less than 0.002. In Fig. 30, the variation is less than 0.0003. So even at 10 bands, the result is well converged for our purposes. The largest variation comes from both terms 2 and 5, which vary by about 0.002 and 0.004, respectively. Terms 2 and 5 have matrix elements $\langle \mathbf{k}+\mathbf{q}n|\partial V/\partial u_{i\alpha}|\mathbf{k}n'\rangle$ where both bands are unrestricted. In Fig. 32, we show an average of the absolute value (seen as a vector with index α) squared of the matrix-elements. The average is over n', $\mathbf{k}$ and the $\mathbf{q}$ in the irreducible Brillouin zone, and we plot as a function of n. The matrix elements have an increasing trend as a function of n. This helps to understand why Terms 2 and 5 vary more when increasing the number of bands. Computational memory is a limitation at the moment to obtain converged results.

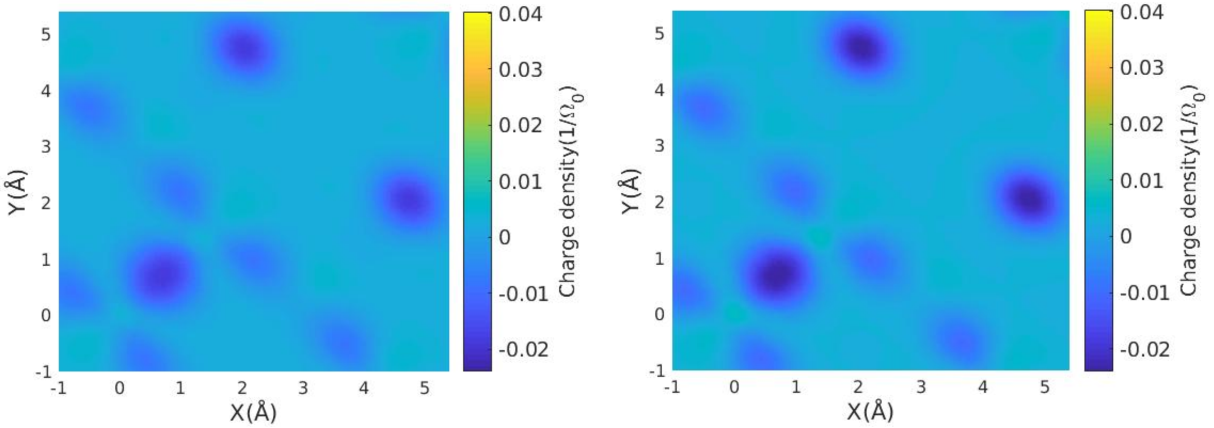
The temperature dependence for different number of bands can be seen in Fig. 33. As in previous calculations, the q-grid is $10 \times 10 \times 10$ and the k-grid is $6 \times 6 \times 6$. There is ZPR, since the normalization is with respect to the structure factor computed without phonon contributions (the Fourier transform of the charge density in Fig. 26). The calculated change is of about 2% at $1000K$. Some of these curves are also included in Fig. 34, where we compare with the experimental result, which shows a variation of the order of 10% [170].

4.6 Appendix

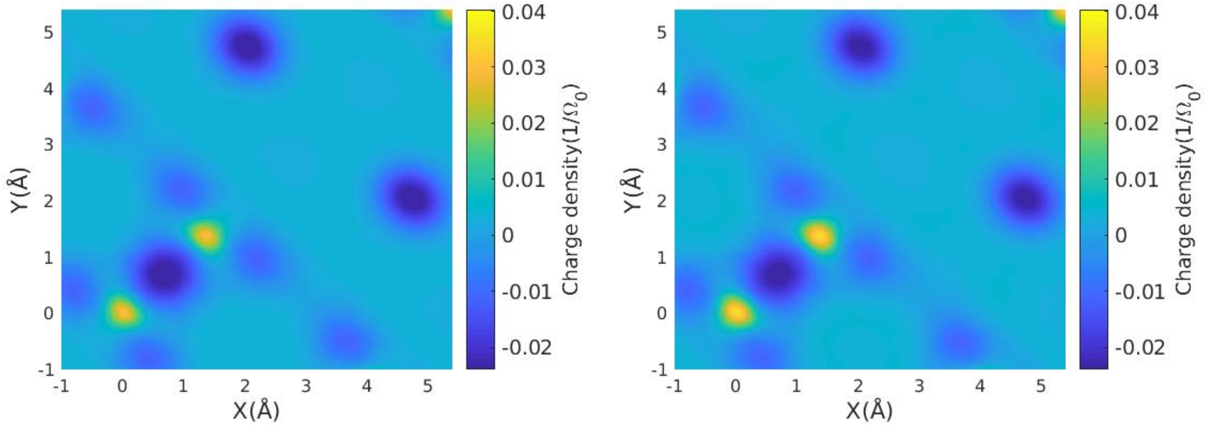
In the main section, we considered a many-body perspective and derived an expression for $\langle \rho(\mathbf{r})\rangle_T$. Here, using Rayleigh-Schrödinger perturbation theory, we consider a single particle picture and see it coincides with the many-body approach. In a single particle picture, the charge density is given by

$$\rho(\mathbf{r}) = 2\sum_{n=1}^{N_e/2}\sum_{\mathbf{k}} |\psi_{n\mathbf{k}}(\mathbf{r})|^2 \tag{4.28}$$

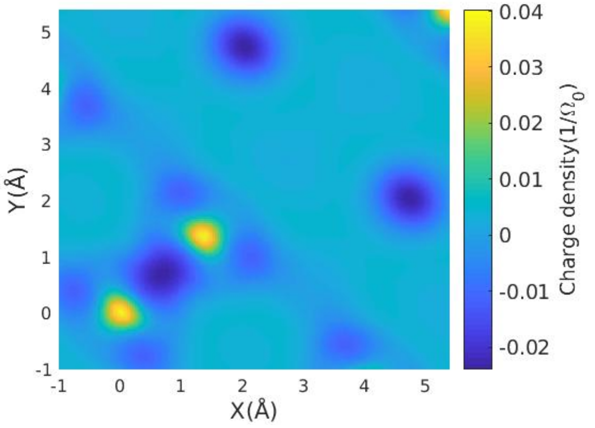
where N_e is the number of electrons. Thus (omitting from now the explicit range of n),



(a) $n = 10$ (b) $n = 20$



(c) $n = 40$ (d) $n = 48$



(e) $n = 60$

Figure 27: Density plots of the change of the charge density at $T = 0$, summing over a different number of bands. See caption of Fig. 26.

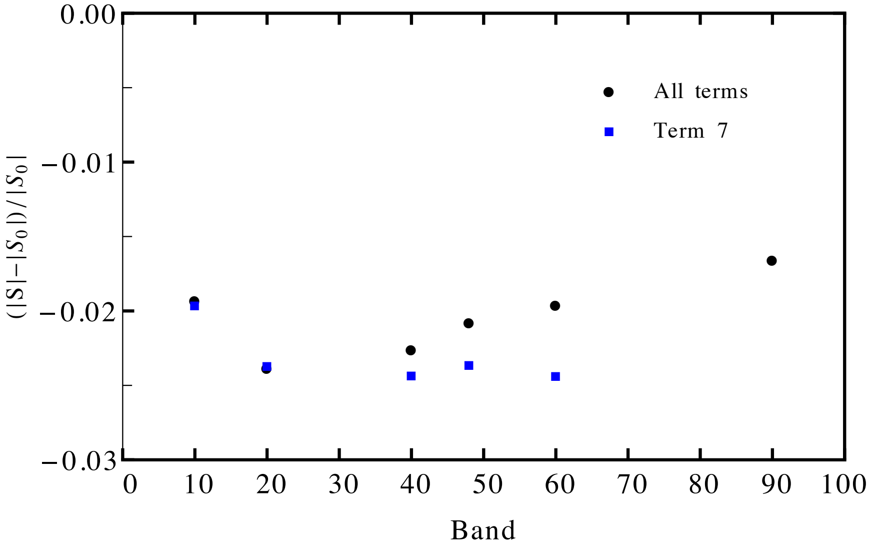


Figure 28: Relative change of the structure factor at $T = 1000K$ as a function of the total number of bands used in the sum Eq. (4.37). Up to about 40 bands, Terms 1 to 6 cancel each other, and most of the contribution comes from Term 7. Term 7 seems converged at about 20 bands, but convergence of the whole sum requires more bands.

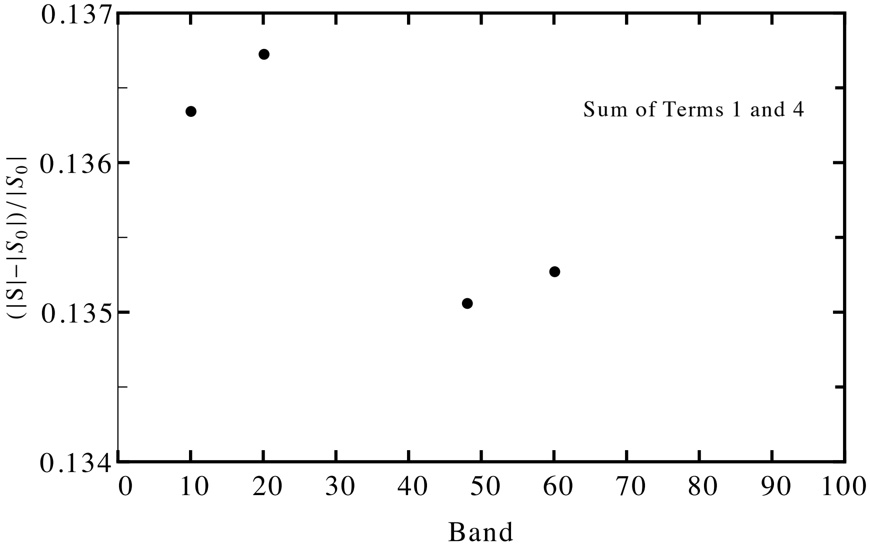


Figure 29: Relative change of the structure factor for the sum of Terms 1 and 4 at $T = 1000K$. The total variation is of less than 0.002 for the calculated bands.

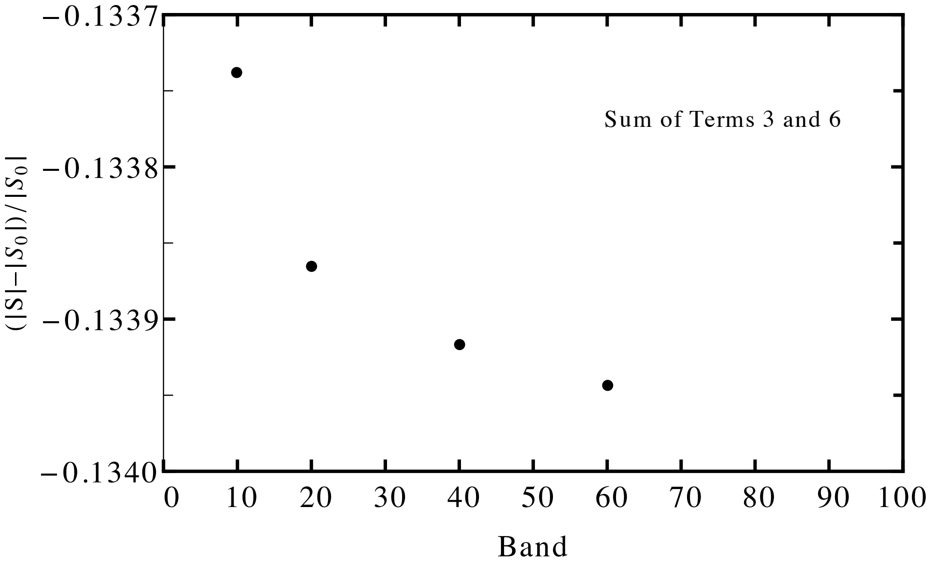


Figure 30: Relative change of the structure factor for the sum of Terms 3 and 6 at $T = 1000K$. The result is well converged for only 10 bands.

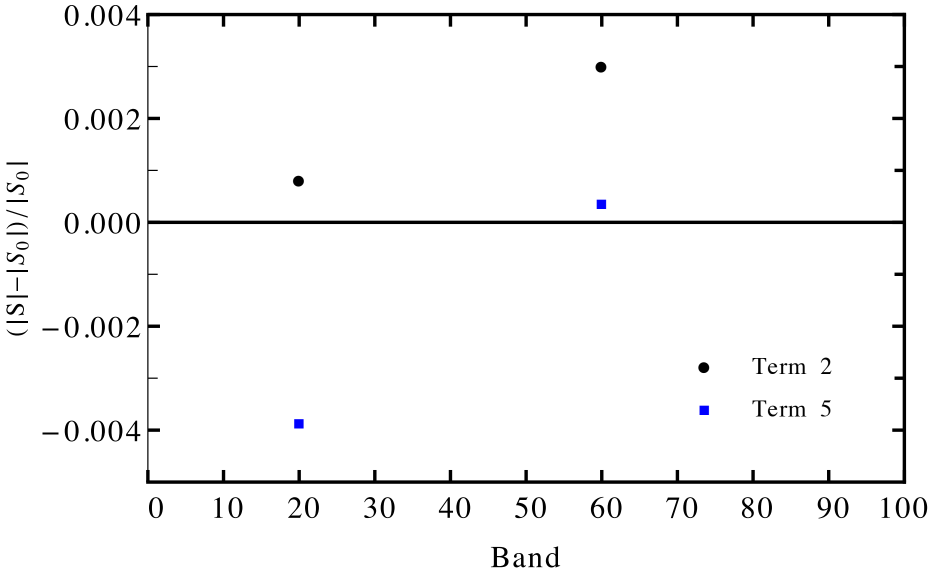


Figure 31: Relative change of the structure factor for Terms 2 and 5 at $T = 1000K$. More than 60 bands are needed to obtain a converged result. The total variation of the sum of these terms is 0.006, which is the main contribution to the lack of convergence of the total result.

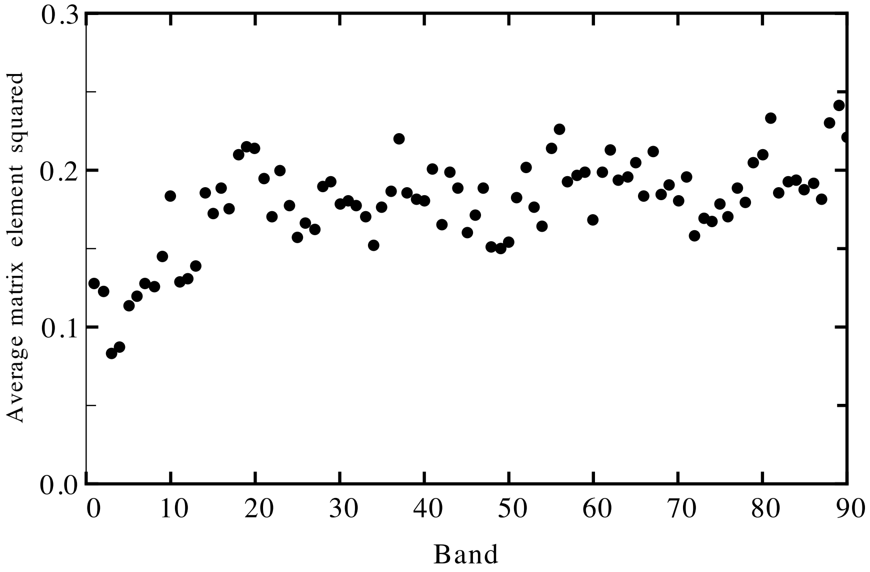


Figure 32: Average over **k**, **q** and n' of the absolute value (as a vector with index $\alpha = 1, 2, 3$) squared of the matrix-elements $\langle \mathbf{k}+\mathbf{q}n|\partial V/\partial u_{i\alpha}(\mathbf{q})|\mathbf{k}n'\rangle$, as a function of n. The increasing trend helps understand why Terms 2 and 5 have slower convergence, since they inolve matrix elements with both band indices unoccupied.

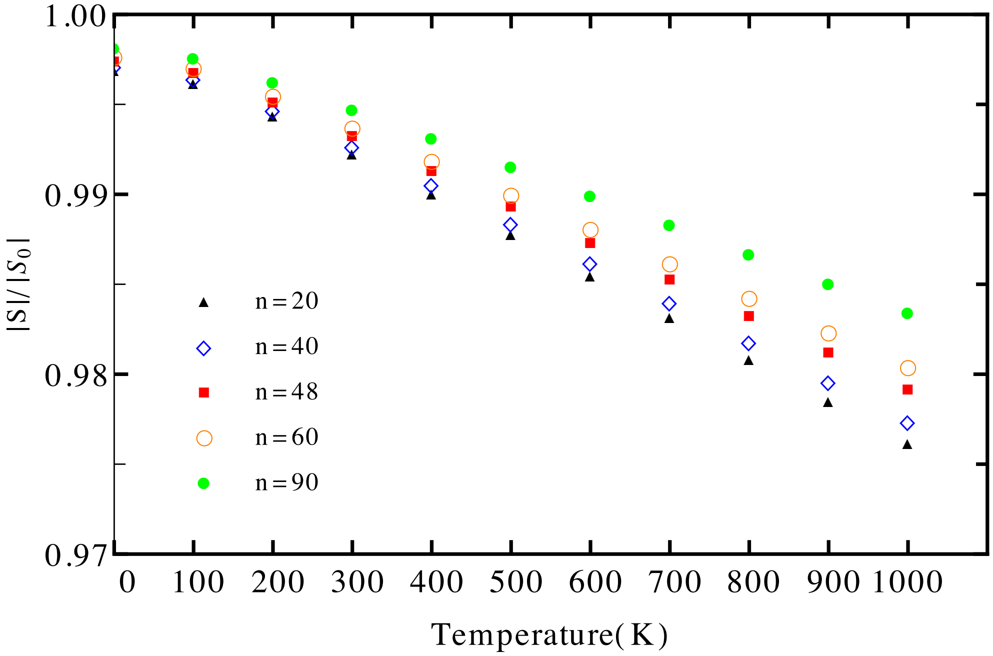


Figure 33: Temperature dependence of the (222) structure factor, normalized with respect to the structure factor with no phonon contributions.

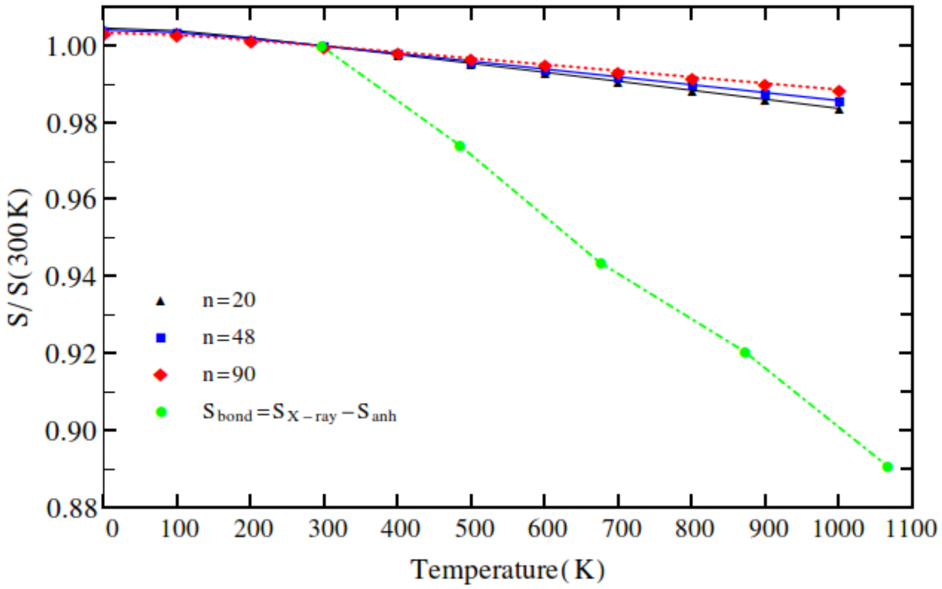


Figure 34: Temperature dependence of the calculated (222) structure factor, normalized here with respect to the structure factor at 300K, together with the experimental data [170] (blue curve in Fig. 25).

$$\Delta\rho(\mathbf{r}) = 2\sum_{n\mathbf{k}}[\Delta\psi_{n\mathbf{k}}(\mathbf{r})\psi_{n\mathbf{k}}(\mathbf{r})^* + \text{c.c}] + \Delta\psi_{n\mathbf{k}}\Delta\psi^*_{n\mathbf{k}}. \tag{4.29}$$

The eigenfunctions to second order are [57]

$$\begin{aligned}\Delta|n\mathbf{k}\rangle =& \sum_{n'\mathbf{k}'}\frac{\langle n'\mathbf{k}'|\Delta V|n\mathbf{k}\rangle}{\varepsilon_{n\mathbf{k}}-\varepsilon_{n'\mathbf{k}'}}|n'\mathbf{k}'\rangle + \sum_{n'\mathbf{k}',n''\mathbf{k}''}\frac{\langle n'\mathbf{k}'|\Delta V|n''\mathbf{k}''\rangle\langle n''\mathbf{k}''|\Delta V|n\mathbf{k}\rangle}{(\varepsilon_{n\mathbf{k}}-\varepsilon_{n'\mathbf{k}'})(\varepsilon_{n\mathbf{k}}-\varepsilon_{n''\mathbf{k}''})}|n'\mathbf{k}'\rangle \\ &-\sum_{n'\mathbf{k}'}\frac{\langle n\mathbf{k}|\Delta V|n\mathbf{k}\rangle\langle n'\mathbf{k}'|\Delta V|n\mathbf{k}\rangle}{(\varepsilon_{n\mathbf{k}}-\varepsilon_{n'\mathbf{k}'})^2}|n'\mathbf{k}'\rangle - \frac{1}{2}|n\mathbf{k}\rangle\sum_{n'\mathbf{k}'}\frac{|\langle n'\mathbf{k}'|\Delta V|n\mathbf{k}\rangle|^2}{(\varepsilon_{n\mathbf{k}}-\varepsilon_{n'\mathbf{k}'})^2}\end{aligned} \tag{4.30}$$

where $\Delta V = \Delta V^{(1)} + \Delta V^{(2)} = \frac{\partial H}{\partial u_{li\alpha}}u_{li\alpha} + \frac{1}{2}\frac{\partial^2 H}{\partial u_{li\alpha}\partial u_{mj\beta}}u_{li\alpha}u_{mj\beta}$. The last term, proportional to $|n\mathbf{k}\rangle$, is needed so that the wavefunction is normalized to 1 to second order, to compensate for the square of the first term. This is automatic to first order, because the square of $\Delta|n\mathbf{k}\rangle$ is of order 2 or higher. As for the third term, after multiplying by $\psi_{n\mathbf{k}}$ in Eq. (4.29) and taking the thermal average, it vanishes. This is because momentum conservation and the thermal average impose $\mathbf{q} = 0$, as opposed to the other terms that have a sum over $\mathbf{q}$. In the limit where the number of cells goes to infinity, this term is zero.

To second order in the ionic displacements, $\rho\rho' = (\rho_0 + \Delta\rho)(\rho'_0 + \Delta\rho') = \rho_0\rho'_0 + \rho_0\Delta\rho'^{(2)} + \Delta\rho^{(2)}\rho'_0 + \Delta\rho^{(1)}\Delta\rho'^{(1)}$. Let us see that the term $\Delta\rho^{(1)}\Delta\rho'^{(1)}$ does not contribute, as we mentioned in the main text after Eq. (4.2). That is, that the thermal average factorizes. The change of the charge density to first order is, from Eq. (4.29), $\sum_{n\mathbf{k}} \Delta\psi^{(1)}_{n\mathbf{k}}\psi^*_{n\mathbf{k}} + \text{c.c.}$ This is

$$\Delta\rho^{(1)} = 2\sum_{n\mathbf{k}n'\mathbf{q}} \frac{\langle n'\mathbf{k}+\mathbf{q}|\Delta V^{(1)}|n\mathbf{k}\rangle}{\varepsilon_{n\mathbf{k}} - \varepsilon_{n'\mathbf{k}+\mathbf{q}}}\psi_{n'\mathbf{k}+\mathbf{q}}\psi^*_{n\mathbf{k}} + \text{c.c} \tag{4.31}$$

When doing the integral in $\mathbf{r}$ in Eq. (4.2), the orthonormality of the wavefunctions imposes $\mathbf{q}+\mathbf{K} = 0$, fixing the value of $\mathbf{q}$. The same occurs when doing the integal in $\mathbf{r}'$. Thus, there is no sum over $\mathbf{q}$, and it vanishes just as the third term in Eq. (4.30).

Now that we have seen that the cross terms do not contribute, we calculate the change in the charge density to second order $\Delta\rho^{(2)} = 2\sum_{n\mathbf{k}} \Delta\psi^{(1)}_{n\mathbf{k}}\Delta\psi^{*(1)}_{n\mathbf{k}} + [\Delta\psi^{(2)}_{n\mathbf{k}}\psi^*_{n\mathbf{k}} + \text{c.c}]$:

$$\begin{aligned}
\Delta\rho = 2 &\sideset{}{'}\sum_{n\mathbf{k}n'n''\mathbf{q}} \frac{\langle n\mathbf{k}|\Delta V^{(1)}|n''\mathbf{k}+\mathbf{q}\rangle\langle n'\mathbf{k}+\mathbf{q}|\Delta V^{(1)}|n\mathbf{k}\rangle}{(\varepsilon_{n\mathbf{k}} - \varepsilon_{n''\mathbf{k}+\mathbf{q}})(\varepsilon_{n\mathbf{k}} - \varepsilon_{n'\mathbf{k}+\mathbf{q}})}\psi_{n'\mathbf{k}+\mathbf{q}}\psi^*_{n''\mathbf{k}+\mathbf{q}} \\
&+ 2\left[\sum_{n\mathbf{k}n'} \frac{\langle n'\mathbf{k}|\Delta V^{(2)}|n\mathbf{k}\rangle}{\varepsilon_{n\mathbf{k}} - \varepsilon_{n'\mathbf{k}}}\psi_{n'\mathbf{k}}\psi^*_{n\mathbf{k}}\right. \\
&+ \sum_{n\mathbf{k}n'n''\mathbf{q}} \frac{\langle n'\mathbf{k}|\Delta V^{(1)}|n''\mathbf{k}+\mathbf{q}\rangle\langle n''\mathbf{k}+\mathbf{q}|\Delta V^{(1)}|n\mathbf{k}\rangle}{(\varepsilon_{n\mathbf{k}} - \varepsilon_{n'\mathbf{k}})(\varepsilon_{n\mathbf{k}} - \varepsilon_{n''\mathbf{k}+\mathbf{q}})}\psi_{n'\mathbf{k}}\psi^*_{n\mathbf{k}} \\
&\left.- \frac{1}{2}\sum_{n\neq n',\mathbf{k}\mathbf{q}} \frac{|\langle n'\mathbf{k}+\mathbf{q}|\Delta V|n\mathbf{k}\rangle|^2}{(\varepsilon_{n\mathbf{k}} - \varepsilon_{n'\mathbf{k}+\mathbf{q}})^2}\psi_{n\mathbf{k}}\psi^*_{n\mathbf{k}} + \text{c.c}\right]
\end{aligned} \tag{4.32}$$

where the prime indicates that in the first line, n, n' and n'' are not all the same. The terms with $n = n' = n''$ in line 1 get cancelled by the terms $n' = n$ in line 4, which are also excluded in the sum (this means $n' > 4$ in line 4 for silicon). Thus, there are no divergences associated with the positive (square) denominator as it approaches 0. The principal value is well defined for the other terms. Equation (4.32) assumes the thermal average has been done. It can be written as

$$
\begin{aligned}
\langle\Delta\rho\rangle = 2 \sideset{}{'}\sum_{\substack{nn'n''\mathbf{kq}\\ s\alpha\beta ij}} & \psi_{n\mathbf{k}}\psi^{*}_{n'\mathbf{k}}(2n_{\mathbf{q}s}+1)\times \\
& \times \Bigg\{ \frac{V^{(2,\mathbf{q})i\alpha,j\beta}_{n\mathbf{k},n'\mathbf{k}} f_{n'}}{\varepsilon_{n'\mathbf{k}}-\varepsilon_{n\mathbf{k}}} + \text{c.c} \\
& + V^{i\alpha,-\mathbf{q}}_{n\mathbf{k},n''\mathbf{k}+\mathbf{q}} V^{j\beta,\mathbf{q}}_{n''\mathbf{k}+\mathbf{q},n'\mathbf{k}} \Bigg[\frac{1}{\Delta\varepsilon^{(2)}_{\substack{n''\mathbf{k}+\mathbf{q},n\mathbf{k}\\ n''\mathbf{k}+\mathbf{q},n'\mathbf{k}}}} f_{n''} + \frac{1}{\Delta\varepsilon^{(2)}_{\substack{n\mathbf{k},n'\mathbf{k}\\ n\mathbf{k},n''\mathbf{k}+\mathbf{q}}}} f_{n} + \text{c.c} \\
& - \frac{1}{\Delta\varepsilon^{(2)}_{n\mathbf{k},n''\mathbf{k}+\mathbf{q}}} f_n \delta_{nn'} \Bigg] \Bigg\} \frac{e^{*}_{i\alpha}(\mathbf{q}s)e_{j\beta}(\mathbf{q}s)}{2\omega_{\mathbf{q}s}\sqrt{M_iM_j}},
\end{aligned}
\tag{4.33}
$$

$$
\begin{aligned}
\text{with}\ \ \Delta\varepsilon^{(2)}_{\substack{n\mathbf{k},n'\mathbf{k}'\\ m\mathbf{p},m'\mathbf{p}'}} &= (\varepsilon_{n\mathbf{k}}-\varepsilon_{n'\mathbf{k}'})(\varepsilon_{m\mathbf{p}}-\varepsilon_{m'\mathbf{p}'}), \\
V^{i\alpha,\mathbf{q}}_{n''\mathbf{k}+\mathbf{q},n'\mathbf{k}} &= \langle n''\mathbf{k}+\mathbf{q}|\tfrac{\partial H}{\partial u_{i\alpha}(\mathbf{q})}|n'\mathbf{k}\rangle, \\
V^{(2,\mathbf{q})i\alpha,j\beta}_{n\mathbf{k},n'\mathbf{k}} &= \langle n\mathbf{k}|\tfrac{\partial^2 H}{\partial u_{i\alpha}(-\mathbf{q})\partial u_{j\beta}(\mathbf{q})}|n'\mathbf{q}\rangle,
\end{aligned}
$$

the c.c affects the whole term (including the parts outside of the brackets), $e_{i\alpha}(\mathbf{q}s)$ are the polarization vectors defined in Sec. 1.3, and the derivatives that involve $\mathbf{q}$ are defined as in Eq. (1.8).

Let us now write an acoustic sum rule to replace the second derivative (the DW type term), which is computationally expensive, in terms of first derivatives. If both the electronic and ionic positions are displaced by the same amount, the charge density should not change. This can be written as $\rho(\mathbf{r},\{\mathbf{R}\}) = \rho(\mathbf{r}+\boldsymbol{\epsilon},\{\mathbf{R}+\boldsymbol{\epsilon}\})$. Schematically, to first order in ϵ,

$$
\begin{aligned}
\rho &= \rho + \frac{\partial\rho}{\partial r}\epsilon + \frac{\partial\rho}{\partial u}\epsilon \\
0 &= \frac{\partial\rho}{\partial r}\epsilon + \frac{\partial\rho}{\partial u}\epsilon
\end{aligned}
\tag{4.34}
$$

Writing indices explicitly and applying $\partial/\partial u_{li\alpha}$ on both sides,

$$
0 = \frac{\partial^2\rho}{\partial u_{li\alpha}\partial r_{\beta}} + \sum_{mj}\frac{\partial^2\rho}{\partial u_{li\alpha}\partial u_{mj\beta}}
\tag{4.35}
$$

This equation contains, except for the eigendisplacements, the derivative of $\Delta\rho^{(1)}$ with respect to $\mathbf{r}$, and $\Delta\rho^{(2)}$ (see Eqs. (4.31) and (4.32)). So the acoustic sum rule is:

$$
\begin{aligned}
0 =& \frac{1}{2}\sum_{n\mathbf{k}n'} \frac{\partial}{\partial r_\beta}\psi_{n\mathbf{k}}\psi^*_{n'\mathbf{k}}\frac{V^{li\alpha}_{n\mathbf{k},n'\mathbf{k}} f_{n'}}{\varepsilon_{n'\mathbf{k}}-\varepsilon_{n\mathbf{k}}} + \text{c.c} + (\alpha \leftrightarrow \beta) \\
&+ \sideset{}{'}\sum_{\substack{nn'n''\mathbf{k}\\ mj}} \psi_{n\mathbf{k}}\psi^*_{n'\mathbf{k}}\times \\
&\times \left\{ 2\frac{V^{(2)li\alpha,mj\beta}_{n\mathbf{k},n'\mathbf{k}} f_{n'}}{\varepsilon_{n'\mathbf{k}}-\varepsilon_{n\mathbf{k}}} + \text{c.c} \right. \\
&\quad V^{li\alpha}_{n\mathbf{k},n''\mathbf{k}} V^{mj\beta}_{n''\mathbf{k},n'\mathbf{k}} \left[\frac{1}{\Delta\varepsilon^{(2)}_{\substack{n''\mathbf{k},n\mathbf{k}\\ n''\mathbf{k},n'\mathbf{k}}}} f_{n''} + \frac{1}{\Delta\varepsilon^{(2)}_{\substack{n\mathbf{k},n'\mathbf{k}\\ n\mathbf{k},n''\mathbf{k}}}} f_n + \text{c.c} \right. \\
&\quad \left.\left. - \frac{1}{\Delta\varepsilon^{(2)}_{n\mathbf{k},n''\mathbf{k}}} f_n \delta_{nn'} + (li\alpha \leftrightarrow mj\beta) \right] \right\}
\end{aligned}
\tag{4.36}
$$

With the rigid ion approximation, the DW term in Eq. (4.33) can be written in terms of first derivative terms using Eq. (4.36), obtaining

$$
\begin{aligned}
\langle \Delta\rho \rangle = 2 & \sideset{}{'}\sum_{\substack{nn'n''\mathbf{kq}\\ s\alpha\beta ij}} \psi_{n\mathbf{k}}\psi^*_{n'\mathbf{k}}(2n_{\mathbf{q}s}+1)\times \\
& \times \left\{ V^{i\alpha,-\mathbf{q}}_{n\mathbf{k},n''\mathbf{k}+\mathbf{q}} V^{j\beta,\mathbf{q}}_{n''\mathbf{k}+\mathbf{q},n'\mathbf{k}} \left[\frac{1}{\Delta\varepsilon^{(2)}_{\substack{n''\mathbf{k}+\mathbf{q},n\mathbf{k}\\ n''\mathbf{k}+\mathbf{q},n'\mathbf{k}}}} f_{n''} + \frac{1}{\Delta\varepsilon^{(2)}_{\substack{n\mathbf{k},n'\mathbf{k}\\ n\mathbf{k},n''\mathbf{k}+\mathbf{q}}}} f_n + \text{c.c} \right. \right. \\
& \left. - \frac{1}{\Delta\varepsilon^{(2)}_{n\mathbf{k},n''\mathbf{k}+\mathbf{q}}} f_n \delta_{nn'} \right] \frac{u^*_{s,i\alpha}(\mathbf{q}) u_{s,j\beta}(\mathbf{q})}{2\omega_{\mathbf{q}s}} \\
& - \left[\frac{1}{2} V^{i\alpha\Gamma}_{n\mathbf{k},n''\mathbf{k}} V^{j\beta\Gamma}_{n''\mathbf{k},n'\mathbf{k}} \left(\frac{1}{\Delta\varepsilon^{(2)}_{\substack{n''\mathbf{k},n\mathbf{k}\\ n''\mathbf{k},n'\mathbf{k}}}} f_{n''} + \frac{1}{\Delta\varepsilon^{(2)}_{\substack{n\mathbf{k},n'\mathbf{k}\\ n\mathbf{k},n''\mathbf{k}}}} f_n + \text{c.c} - \frac{1}{\Delta\varepsilon^{(2)}_{n\mathbf{k},n''\mathbf{k}}} f_n \delta_{nn'} \right. \right. \\
& \left. \left. \left. + (i\alpha \leftrightarrow j\beta) \right) + \frac{1}{3}\frac{1}{4}\frac{\overleftarrow{\partial}}{\partial r_\beta} \frac{V^{i\alpha\Gamma}_{n\mathbf{k},n'\mathbf{k}}}{\varepsilon_{n'\mathbf{k}}-\varepsilon_{n\mathbf{k}}} f_{n'} + \text{c.c} + (\alpha \leftrightarrow \beta) \right] \frac{e^*_{s,i\alpha}(\mathbf{q}) e_{s,i\beta}(\mathbf{q})}{2\omega_{\mathbf{q}s}\sqrt{M_i M_j}} \right\}.
\end{aligned}
\tag{4.37}
$$

The c.c. in the first 3 terms affects all factors, while the c.c. and the swap of indices in the other terms (that come from the acoustic sum rule), affect everything except the eigenvectors.

To make the computation more efficient, we use the symmetries of the crystal in the phonon wavevector **q**, as implemented in ABINIT. Equation (4.37) includes a sum over **q**. We can write

$$\begin{aligned}
\langle \Delta\rho(\mathbf{r}) \rangle &= \sum_{\mathbf{q}} \Delta\tilde{\rho}(\mathbf{q}, \mathbf{r}) \\
&= \sum_{R} \sum_{\mathbf{q}\epsilon IBZ} \Delta\tilde{\rho}(R\mathbf{q}, \mathbf{r}) \\
&= \sum_{R} \sum_{\mathbf{q}\epsilon IBZ} \Delta\tilde{\rho}(\mathbf{q}, R^{-1}\mathbf{r})
\end{aligned} \tag{4.38}$$

Combining this equation with Eq. (4.37) we obtained the results included in the main text.

Conclusions

The electron-phonon interaction (EPI) is an active area of research that has regained popularity in the last few years. EPI calculations are quite demanding, and new techniques are being developed to improve accuracy and reduce computational costs. In Chapter 2, we described a simple method to calculate the energy renormalization at the bottom of the conduction band or top of the valence band in polar semiconductors. In these cases, the usual adiabatic approximation gives a divergent result for the Fröhlich term (intraband transition and longtidudinal phonon mode). The longitudinal optical phonon frequency ω_{LO} has to be kept in the energy denominators for small phonon wavevector $\mathbf{q}$. To avoid a very dense grid that captures the $1/q$ behavior of the Fröhlich matrix elements, we first used the adiabatic approximation with an $i\delta$ in the denominator. This ignores the central $\mathbf{q} = 0$ cell. A significant portion of the Fröhlich contribution is contained here, but the other terms can be safely ignored. By using an explicit analytical expression in effective mass approximation, valid in the central cell, the omitted Fröhlich contribution can be added. Outside of the central cell, the adiabatic approximation works well for all terms. A generalization to the anisotropic and degenerate case was described to treat the valence band. Adding both contributions, the temperature dependence of the band gap was determined for GaN-zincblende. In this way, a coarse grid can be used to calculate the energy renormalization of the band gap of polar semiconductors.

In Chapter 3, we used the cumulant approach to calculate the electronic spectral function of the polar semiconductors MgO and LiF, and of the Fröhlich model. ARPES experiments in recent years have been able to detect a sideband separated from the QP peak by around ω_{LO}, which agrees with cumulant spectral function results. According to diagrammatic quantum Monte Carlo calculations for the Fröhlich model, the position of the QP peak is also adequately given by the cumulant approach (which coincides with Rayleigh-Schrödinger perturbation theory). On the other hand, the Dyson approach considerably overestimates the distance from the QP peak to the satellite, and the renormalization of the QP peak is much too low. The weight of the QP peak is also better predicted by the cumulant approach. For the conduction bands of MgO and LiF, we compared the cumulant full *ab-initio* calculations with those of the Fröhlich polaron (with the parameters obtained from the corresponding *ab-initio* calculation). Despite the Fröhlich polaron corresponding to only one of the terms of the total result, the spectral functions are quite similar. The Fröhlich ZPR $-\alpha\omega_{LO}$ is also a significant portion of the total ZPR. The MgO valence spectral function was observed to be quite similar to the conduction spectral functions, so degeneracy does not seem to change the qualitative behavior. The Fröhlich spectral function at high values of α also helped to understand the broad peak observed in the LiF valence band. Since we used a retarded Green's function in our calculations, they can be easily extended to finite temperatures. We also showed that time-ordered versions of the cumulant give the same results for wide-gap semiconductors like MgO and LiF.

Finally, in Chapter 4 we studied the temperature dependence of the (222) X-ray forbidden reflection in silicon. Previous models had assumed rigidity of the movement of the valence charge, and included only Debye-Waller factors. We used a many-body approach to obtain a perturbative expression for the charge density, which coincides with a single particle picture in the adiabatic approximation. Since all current numerical DFPT implementations calculate only first order electron-phonon matrix elements, we used an acoustic sum rule to express the DW-type contribution in terms of Fan-type contributions. We then studied convergence of the structure factor, and obtained that the q-grid does not need to be very dense, as opposed to calculations like the ones in Chapter 3. Convergence in the number of bands is more challenging, but we still expect the sum to require fewer unoccupied bands that in energy renormalization calculations.

In the future, we plan to use the perturbative expression of the charge density to study pyroelectricity, that is, how polarization changes with temperature. We only wrote an adiabatic expression, but it is easy to keep the phonon frequencies in the many-body approach, giving similar but more cumbersome expressions. The non-adiabatic version is necessary in polar materials to avoid divergences and obtain accurate results. At a given temperature, the charge density can be obtained, and the corresponding polarization can then be calculated. Convergence is likely more difficult in polar semiconductors because of the non-analytic terms. It would also be interesting to determine if there is a general rule that relates the Fröhlich ZPR to the total ZPR. Further investigation is needed into the relation of the cumulant approach to the Monte Carlo calculations, and into how the latter are modified when considering all EPI effects in a crystal. We are also interested in studying other properties related to the Fröhlich polaron, like the ZPR in ice and isotope effects. The development of non-perturbative techniques that are also non-adiabatic would aid these calculations. As a final remark, it is exciting to see how the field has grown since I started my research three and a half years ago. While research continues to progress at a rapid pace, many interesting questions related to thermal effects and the Fröhlich interaction — and to the electron-phonon interaction more broadly — remain to be investigated.

Appendix

A.1 Cumulant calculations

In this section we are going to derive more explicitly some of the results of Chapter 3. Let us define $\phi_{\mathbf{q}s} = a_{\mathbf{q}s} + a^\dagger_{-\mathbf{q}s}$. The Hamiltonian for the electron-phonon problem is given by [1]

$$\begin{aligned} H &= H_e + H_n + H^{(1)}_{e-n} + H^{(2)}_{e-n} \\ &= \sum_{\mathbf{k}n} \varepsilon_{\mathbf{k}n} c^\dagger_{\mathbf{k}n} c_{\mathbf{k}n} + \sum_{\mathbf{q}s} \omega_{\mathbf{q}s} (a^\dagger_{\mathbf{q}s} a_{\mathbf{q}s} + 1/2) + \sum_{\substack{\mathbf{kq} \\ mns}} g_{mns}(\mathbf{k},\mathbf{q}) c^\dagger_{\mathbf{k}+\mathbf{q}m} c_{\mathbf{k}n} \phi_{\mathbf{q}s} \\ &+ \sum_{\substack{\mathbf{kq}' \\ mnss'}} g^{\mathrm{DW}}_{mnss'}(\mathbf{k},\mathbf{q},\mathbf{q}') c^\dagger_{\mathbf{k}+\mathbf{q}+\mathbf{q}'m} c_{\mathbf{k}n} \phi_{\mathbf{q}s} \phi_{\mathbf{q}'s'} \end{aligned} \tag{A.1}$$

The anticommutation relations for the electron and phonon creation and destruction operators are

$$\begin{aligned} \{c^\dagger_{\mathbf{k}n}, c_{\mathbf{k}'n'}\} &= \delta_{\mathbf{kk}'}\delta_{nn'} \\ [a^\dagger_{\mathbf{q}s}, a_{\mathbf{q}'s'}] &= \delta_{\mathbf{qq}'}\delta_{ss'} \end{aligned} \tag{A.2}$$

and zero for all other combinations.

Eq. (3.9):

Consider the following way to rewrite part of Eq. (3.8):

$$\begin{aligned} \omega - \varepsilon_{\mathbf{k}n} - \Re e\Sigma(\mathbf{k}n,\omega) &\sim \omega - \varepsilon_{\mathbf{k}n} - \Re e\Sigma(\mathbf{k}n, E^{\mathrm{DM}}_{\mathbf{k}n}) - \frac{\partial \Re e\Sigma(\mathbf{k}n, E^{\mathrm{DM}}_{\mathbf{k}n})}{\partial\omega}(\omega - E^{\mathrm{DM}}_{\mathbf{k}n}) \\ &= (Z^{\mathrm{DM}}_{\mathbf{k}n})^{-1}\omega - \varepsilon_{\mathbf{k}n} - \Re e\Sigma(\mathbf{k}n, E^{\mathrm{DM}}_{\mathbf{k}n}) + \frac{\partial \Re e\Sigma(\mathbf{k}n, E^{\mathrm{DM}}_{\mathbf{k}n})}{\partial\omega} E^{\mathrm{DM}}_{\mathbf{k}n} \\ &= (Z^{\mathrm{DM}}_{\mathbf{k}n})^{-1}\omega - E^{\mathrm{DM}}_{\mathbf{k}n} + \frac{\partial \Re e\Sigma(\mathbf{k}n, E^{\mathrm{DM}}_{\mathbf{k}n})}{\partial\omega} E^{\mathrm{DM}}_{\mathbf{k}n} \\ &= (Z^{\mathrm{DM}}_{\mathbf{k}n})^{-1}(\omega - E^{\mathrm{DM}}_{\mathbf{k}n}), \end{aligned} \tag{A.3}$$

where $Z^{\mathrm{DM}}_{\mathbf{k}n}$ is defined in Eq. (3.11) and we used Eq. (3.10) to get the third line. Evaluating also the imaginary part at the quasiparticle energy, leads to Eq. (3.9).

Eq. (3.16):

We use the spectral representation Eq. (3.15). First, we show it is normalized to 1:

$$\begin{aligned}
\int A(\mathbf{k}n,\omega)d\omega &= \sum_{mm'} \frac{e^{-\beta E_m}}{Z} \Big(\langle m'|c^\dagger_{\mathbf{k}n}|m\rangle\langle m|c_{\mathbf{k}n}|m'\rangle + \langle m'|c_{\mathbf{k}n}|m\rangle\langle m|c^\dagger_{\mathbf{k}n}|m'\rangle \Big) \\
&= \sum_{m} \frac{e^{-\beta E_m}}{Z} \Big(\langle m|c_{\mathbf{k}n}c^\dagger_{\mathbf{k}n}|m\rangle + \langle m|c^\dagger_{\mathbf{k}n}c_{\mathbf{k}n}|m\rangle \Big) \\
&= \langle \{c_{\mathbf{k}n}, c^\dagger_{\mathbf{k}n}\}\rangle \\
&= \langle 1\rangle = 1
\end{aligned} \tag{A.4}$$

Similarly,

$$\begin{aligned}
\int \omega A(\mathbf{k}n,\omega)d\omega &= \sum_{mm'} \frac{e^{-\beta E_m}}{Z} \Big(\langle m'|c^\dagger_{\mathbf{k}n}|m\rangle\langle m|c_{\mathbf{k}n}|m'\rangle(E_{m'}-E_m) \\
&\qquad + \langle m'|c_{\mathbf{k}n}|m\rangle\langle m|c^\dagger_{\mathbf{k}n}|m'\rangle(E_m-E_{m'}) \Big) \\
&= \sum_{mm'} \frac{e^{-\beta E_m}}{Z} \Big(\langle m'|c^\dagger_{\mathbf{k}n}|m\rangle\langle m|[c_{\mathbf{k}n},H]|m'\rangle + \langle m'|[c_{\mathbf{k}n},H]|m\rangle\langle m|c^\dagger_{\mathbf{k}n}|m'\rangle \Big) \\
&= \sum_{m} \frac{e^{-\beta E_m}}{Z} \Big(\langle m|[c_{\mathbf{k}n},H]c^\dagger_{\mathbf{k}n}|m\rangle + \langle m|c^\dagger_{\mathbf{k}n}[c_{\mathbf{k}n},H]|m\rangle \Big) \\
&= \langle \{[c_{\mathbf{k}n},H], c^\dagger_{\mathbf{k}n}\}\rangle
\end{aligned} \tag{A.5}$$

For the electronic part we have, using the commutation relations Eq. (A.2),

$$\begin{aligned}
\langle \{[c_{\mathbf{k}n},H_e], c^\dagger_{\mathbf{k}n}\}\rangle &= \langle \{[c_{\mathbf{k}n}, \sum_{\mathbf{k}'n'} \varepsilon_{\mathbf{k}'n'} c^\dagger_{\mathbf{k}'n'}c_{\mathbf{k}'n'}], c^\dagger_{\mathbf{k}n}\}\rangle \\
&= \sum_{\mathbf{k}'n'} \varepsilon_{\mathbf{k}'n'} \{\{c_{\mathbf{k}n}, c^\dagger_{\mathbf{k}'n'}\}c_{\mathbf{k}'n'} - c^\dagger_{\mathbf{k}'n'}\{c_{\mathbf{k}n}, c_{\mathbf{k}'n'}\}, c^\dagger_{\mathbf{k}n}\} \\
&= \sum_{\mathbf{k}'n'} \varepsilon_{\mathbf{k}'n'} \{\delta_{\mathbf{k}\mathbf{k}'}\delta_{nn'}c_{\mathbf{k}'n'}, c^\dagger_{\mathbf{k}n}\} \\
&= \varepsilon_{\mathbf{k}n}
\end{aligned} \tag{A.6}$$

where we used $[A,BC] = \{A,B\}C - B\{A,C\}$ to obtain the second line. The commutator with H_n is zero, since it only has phonon creation and destruction operators. As mentioned earlier, terms with odd powers of the lattice displacements have a zero thermal average. Finally, the DW contribution is

$$
\begin{aligned}
&\langle\{[c_{\mathbf{k}n}, \sum_{\substack{\mathbf{k}'\mathbf{q}\mathbf{q}'\\ mn'ss'}} g^{\mathrm{DW}}_{mnss'}(\mathbf{k},\mathbf{q},\mathbf{q}')c^{\dagger}_{\mathbf{k}'+\mathbf{q}+\mathbf{q}'m}c_{\mathbf{k}'n'}\phi_{\mathbf{q}s}\phi_{\mathbf{q}'s'}], c_{\mathbf{kn}}\}\rangle = \\
&\qquad = \sum_{\substack{\mathbf{k}'\mathbf{q}\mathbf{q}'\\ mn'ss'}} g^{\mathrm{DW}}_{mnss'}(\mathbf{k},\mathbf{q},\mathbf{q}')\{\{c_{\mathbf{k}n}, c^{\dagger}_{\mathbf{k}'+\mathbf{q}+\mathbf{q}'m}\}c_{\mathbf{k}'n'}, c_{\mathbf{kn}}\}\langle\phi_{\mathbf{q}s}\phi_{\mathbf{q}'s'}\rangle \\
&\qquad = \sum_{\substack{\mathbf{k}'\mathbf{q}\mathbf{q}'\\ mn'ss'}} g^{\mathrm{DW}}_{mnss'}(\mathbf{k},\mathbf{q},\mathbf{q}')\delta_{\mathbf{k},\mathbf{k}'+\mathbf{q}+\mathbf{q}'}\delta_{nm}\{c_{\mathbf{k}'n'}, c_{\mathbf{kn}}\}\langle\phi_{\mathbf{q}s}\phi_{\mathbf{q}'s'}\rangle \\
&\qquad = \sum_{\substack{\mathbf{k}'\mathbf{q}\mathbf{q}'\\ n'ss'}} g^{\mathrm{DW}}_{nnss'}(\mathbf{k},\mathbf{q},\mathbf{q}')\delta_{\mathbf{q},-\mathbf{q}'}\delta_{n'n}(2n_{\mathbf{q}s}+1)\delta_{ss'}\delta_{\mathbf{q},-\mathbf{q}'} \\
&\qquad = \sum_{\substack{\mathbf{k}\mathbf{q}\\ ns}} g^{\mathrm{DW}}_{nnss}(\mathbf{k},\mathbf{q},-\mathbf{q})(2n_{\mathbf{q}s}+1),
\end{aligned}
\tag{A.7}
$$

the usual result, as had been mentioned earlier.

Eq. (3.24):

While Eq. (3.21) is well defined because $e^{-i\omega t} + i\omega t - 1$ is proportional to ω^2, this is not the case when separating it into three pieces. This is why Eq. (3.24) has an additional $i\eta$ in the numerator. The Kramers-Kronig relation is

$$
\Sigma_1(\omega) = \frac{1}{\pi} P \int_{-\infty}^{\infty} d\omega' \frac{\Sigma_2(\omega')}{\omega' - \omega} \tag{A.8}
$$

where Σ_1 is the real part and Σ_2 the imaginary part. We cannot naively take the derivative because the principal value of $1/\omega'^2$ is ill defined (although this is basically what the final result Eq. (3.24) consists of). We first derive the Kramers-Kronig (K-K) relation, and then proceed similarly to derive an equation with a square in the denominator. Because of the residue theorem,

$$
\oint d\omega \frac{\Sigma(\omega')}{\omega' - \omega + i\eta} = 0 \tag{A.9}
$$

when closing the contour on the upper plane. Notice that this is possible for the retarded self-energy, because all the poles are in the lower-half plane. Using Eq. (1.9),

$$
P \int_{-\infty}^{\infty} d\omega \frac{\Sigma(\omega')}{\omega' - \omega} - i\pi\Sigma(\omega) = 0 \tag{A.10}
$$

Taking the real and imaginary part, one gets the usual K-K relations. The imaginary part of Eq. (A.10) at $\omega = 0$, using the definition of $\beta(\mathbf{k}n, \omega)$ Eq. (3.22), gives Eq. (3.23).

Again using the residue theorem on the upper-half plane, we can write

$$\int_{-\infty}^{\infty} d\omega \frac{\Sigma(\omega')}{(\omega' - \omega - i\eta)^2} = 2\pi i \Sigma'(\omega) \tag{A.11}$$

Taking the imaginary part on both sides,

$$\int_{-\infty}^{\infty} d\omega \Sigma_2(\omega) \Re e \frac{1}{(\omega' - \omega - i\delta)^2} + \Sigma_1(\omega') \frac{2\eta(\omega' - \omega)}{((\omega' - \omega)^2 + \eta^2)^2} = 2\pi \Sigma_1'(\omega) \tag{A.12}$$

As $\eta \to 0$, the function $f_\eta(x) = -2x\eta/(x^2 + \eta^2)^2\pi$ provides a representation of the derivative of the delta function. That is, $\Sigma_1(\omega)$ does not contribute because of asymmetry of $f_\eta(x)$, the first derivative gives

$$\Sigma_1'(\omega) \int_{\infty}^{\infty} d\omega' \frac{(\omega' - \omega)^2}{((\omega' - \omega)^2 + \eta^2)^2} = \pi \Sigma_1(\omega) \tag{A.13}$$

and higher order terms also give a zero contribution. Using this in Eq. (A.12) with $\omega = 0$, we get Eq. (3.24).

Eq. (3.30) and (3.31):

We consider two cases:
(i) $-(\omega_{LO} - \omega) < 0$: The steps are similar to those in Eq. (1.67). Here, the denominator is $\omega + i\eta - \epsilon_{\mathbf{k+q}n} - \omega_{LO}$ instead of $\varepsilon_{\mathbf{k}n} - \varepsilon_{\mathbf{k+q}n} - \omega_{LO}$. The result for the real part is the same as before, with the replacement $\omega_{LO} \to \omega_{LO} - \omega$ in the fourth line of Eq. (1.67). We can also take $\eta \to 0$, so there is no imaginary part.
(ii) $-(\omega_{LO} - \omega) > 0$: The integral for the real part is of the form

$$\int_0^{\infty} dx \frac{1}{x^2 - a^2} = \int_0^{\infty} dx \frac{1}{2a} \left(\frac{1}{x - a} - \frac{1}{x + a} \right) = \frac{1}{2a} \log \left| \frac{x - a}{x + a} \right| \Bigg|_0^{\infty} = 0 \tag{A.14}$$

It is convenient however to use the residue theorem (which could have been used in case (i) as well), to treat the imaginary part as well (compare with fourth line of Eq. (1.67)):

$$\begin{aligned} \Sigma(\mathbf{k}n, \omega) &= \int -|\tilde{M}|^2 \frac{\Omega_0}{(2\pi)^3} 4\pi \frac{1}{2} \int_{-\infty}^{\infty} dq 2m^* \frac{1}{q^2 - 2m^*(\omega - \omega_{LO}) - i\eta} \\ &= -|\tilde{M}|^2 \frac{\Omega_0}{(2\pi)^3} 4\pi \frac{1}{2} 2m^* 2\pi i \frac{1}{2\sqrt{2m^*(\omega - \omega_{LO})}} \end{aligned} \tag{A.15}$$

The real part is 0 as already shown in Eq. (A.14), and the imaginary part has $1/\sqrt{\omega - \omega_{LO}}$ instead of $1/\sqrt{\omega_{LO} - \omega}$.

Eqs. (3.18-3.22), Cumulant equations:

In the cumulant expansion, the Green's function is written as in Eq. (3.17), and to lowest order $G = G_0(1+C)$, where $G_0(\mathbf{k}n, t) = -i\theta(t)e^{-i\varepsilon_{\mathbf{k}n}t}$. To lowest order, the Green's function is also given schematically by $G = G_0 + G_0 \Sigma G_0$ (expression we also used in Chapter 4); it is easier to evalute the second term in momentum space using Feynman rules and then Fourier transform in frequency to the time domain. Equating both expansions, an expression for the cumulant is obtained:

$$C(\mathbf{k}n, t) = i\epsilon^{i\epsilon_{\mathbf{k}n}t} \int \frac{d\omega}{2\pi} e^{-i\omega t} [G_0(\mathbf{k}n, \omega)]^2 \Sigma(\mathbf{k}n, \omega) \tag{A.16}$$

Using Eq. (1.9), the self-energy can be written as

$$\Sigma(\mathbf{k}n, \omega) = \Sigma^{\mathrm{DW}}(\mathbf{k}n) - \int \frac{d\omega'}{\pi} \frac{\Im m \Sigma^{\mathrm{Fan}}(\mathbf{k}n, \omega')}{\omega - \omega' + i\eta} \tag{A.17}$$

For the DW piece, we have

$$\begin{aligned} i\epsilon^{i\varepsilon_{\mathbf{k}n}t} \Sigma^{\mathrm{DW}}(\mathbf{k}n) \int \frac{d\omega}{2\pi} \frac{e^{-i\omega t}}{(\omega - \varepsilon_{\mathbf{k}n} + i\eta)^2} &= ie^{i\varepsilon_{\mathbf{k}n}t} \Sigma^{\mathrm{DW}}(\mathbf{k}n) \frac{-2\pi i}{2\pi} (e^{-i\omega t})'|_{\omega=\epsilon_{\mathbf{k}n}} \\ &= e^{i\varepsilon_{\mathbf{k}n}t} \Sigma^{\mathrm{DW}}(\mathbf{k}n)(-it)e^{-i\varepsilon_{\mathbf{k}n}t} \\ &= -it\Sigma^{\mathrm{DW}}(\mathbf{k}n) \end{aligned} \tag{A.18}$$

For the Fan piece,

$$\begin{aligned} &-ie^{i\varepsilon_{\mathbf{k}n}t} \int \frac{d\omega'}{\pi} \Sigma^{\mathrm{Fan}}(\mathbf{k}n, \omega') \frac{d\omega}{2\pi} e^{-i\omega t} \frac{1}{(\omega - \varepsilon_{\mathbf{k}n} + i\eta)^2} \frac{1}{\omega - \omega' + i\eta} = \\ &\quad = -ie^{i\varepsilon_{\mathbf{k}n}t} \int \frac{d\omega'}{\pi} \Sigma^{\mathrm{Fan}}(\mathbf{k}n, \omega') \frac{2\pi i}{2\pi} \left[\frac{e^{-i\omega' t}}{(\omega' - \varepsilon_{\mathbf{k}n} + i\eta)^2} + \left(\frac{e^{-i\omega t}}{\omega - \omega' + i\eta} \right)' |_{\omega=\varepsilon_{\mathbf{k}n}} \right] \\ &\quad = -ie^{i\varepsilon_{\mathbf{k}n}t} \int \frac{d\omega'}{\pi} \Sigma^{\mathrm{Fan}}(\mathbf{k}n, \omega') \left[\frac{e^{-i\omega' t}}{(\omega' - \varepsilon_{\mathbf{k}n} + i\eta)^2} + \frac{-ite^{-i\varepsilon_{\mathbf{k}n}t}}{\varepsilon_{\mathbf{k}n} - \omega' + i\eta} - \frac{e^{-i\varepsilon_{\mathbf{k}n}t}}{(\varepsilon_{\mathbf{k}n} - \omega' + i\eta)^2} \right] \\ &\quad = -i \int \frac{d\omega'}{\pi} \Sigma^{\mathrm{Fan}}(\mathbf{k}n, \omega' + \varepsilon_{\mathbf{k}n}) \frac{e^{-i\omega' t} + it\omega' - 1}{\omega'^2} \end{aligned} \tag{A.19}$$

which corresponds to Eq. (3.21) and (3.22).

References

[1] F. Giustino, Rev. Mod. Phys. **89**, 015003 (2017).

[2] H. Fröhlich, Adv. Phys. **3**, 325 (1954).

[3] P. B. Allen and V. Heine, J. Phys. C **9**, 2305 (1976).

[4] P. B. Allen and M. Cardona, Phys. Rev. B **23**, 1495 (1981).

[5] J. R. Chelikowsky and M. L. Cohen, Phys. Rev. Lett. **33**, 1339 (1974).

[6] A. S. Alexandrov and J. T. Devreese, *Advances in polaron physics* (Springer, New York, 2010).

[7] H. Finkenrath, N. Uhle, and W. Waidelich, Solid State Commun. **7**, 11 (1969).

[8] V. L. Gurevich, I. G. Lang, and Y. A. Firsov, Fiz. Tverd. Tela **4**, 1252 (1962).

[9] T. Sohier, M. Calandra, and F. Mauri, Phys. Rev. B **94**, 085415 (2016).

[10] C. Verdi, F. Caruso, and F. Giustino, Nat. Comm. **8**, 15769 (2017).

[11] K. Miyata *et al.*, Sci. Adv. **3**, 033023 (2017).

[12] W. R. L. Lambrecht, C. Bhandari, and M. van Schilfgaarde, arXiv:1706.10252 (2017).

[13] J. Sjakste, N. Vast, M. Calandra, and F. Mauri, Phys. Rev. B **92**, 054307 (2015).

[14] C. Verdi and F. Giustino, Phys. Rev. Lett. **115**, 176401 (2015).

[15] J. P. Nery and P. B. Allen, Phys. Rev. B **94**, 115135 (2016).

[16] S. Moser *et al.*, Phys. Rev. Lett. **110**, 196403 (2013).

[17] C. Chen, J. Avila, E. Frantzeskakis, A. Levy, and M. C. Asensio, Nature Comm. **6**, 8585 EP (2015).

[18] R. Yukawa *et al.*, Phys. Rev. B **94**, 165313 (2016).

[19] C. Cancellieri *et al.*, Nature Comm. **7**, 10386 (2016).

[20] S. Poncé *et al.*, J. Chem. Phys. **143**, 102813 (2015).

[21] B. Monserrat, E. Engel, and R. Needs, Phys. Rev. B **92**, 140302 (2015).

[22] E. Cannuccia and A. Marini, Phys. Rev. Lett. **107**, 255501 (2011).

[23] E. Cannuccia and A. Marini, Eur. Phys. J. B **85**, 320 (2012).

[24] J. P. Nery *et al.*, Phys. Rev. B **97**, 115145 (2018).

[25] G. D. Mahan, *Many-Particle Physics* (Kluwer, New York, 2000).

[26] A. L. Fetter and J. D. Walecka, *Quantum theory of many-particle systems* (McGraw-Hill, New York, 1971).

[27] H. Bruus and K. Flensberg, *Many-body quantum theory in condensed matter physics* (Denmark, 2002).

[28] H. Fan, Phys. Rev. Lett. **78**, 808 (1950).

[29] E. Antončik, Czech. J. Phys. **5**, 449 (1955).

[30] S. Yu, The authors have paid to have an electronic image, which is now publicly available, 1965 (unpublished).

[31] M. Schlüter, G. Martinez, and C. M. L., Phys. Rev. B **12**, 650 (1975).

[32] P. B. Allen, Phys. Rev. B **18**, 5217 (1978).

[33] S. C. Yu, J. Phys. Chem. Solids **47**, 887 (1986).

[34] J. of Math. Phys. **2**, 232 (2018).

[35] N. A. Spaldin, J. Solid State Chem. **195**, 2 (2012).

[36] P. Vogl, Phys. Rev. B **13**, 694 (2018).

[37] X. Gonze and C. Lee, Phys. Rev. B **55**, 10355 (1997).

[38] P. B. Allen and J. P. Nery, Phys. Rev. B **95**, 035211 (2017).

[39] W. Cochran and R. A. Cowley, J. Phys. Chem. Solids **23**, 447 (1962).

[40] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford University Press, 1954).

[41] P. Y. Yu and M. Cardona, *Fundamenals of semiconductors* (Springer-Verlag, Berlin, 1996).

[42] R. H. Lydanne, R. G. Sachs, and E. Teller, Phys. Rev. **59**, 673 (1941).

[43] D. C. Elton and M. Fernández-Serra, Phys. Rev. **7**, 10193 (2016).

[44] D. C. Elton, Understanding the dielectric properties of water. Ph.D. thesis, Stony Brook University, 2016.

[45] R. P. Feynman, Phys. Rev. **97**, 660 (1955).

[46] I. G. Lang and Y. A. Firsov, Zh. Eksp. Teor. Fiz. **43**, 1843 (1962).

[47] A. Mishchenko, N. Prokof'ev, A. Sakamoto, and B. Svistunov, Phys. Rev. B **62**, 6317 (2000).

[48] H.-R. Trebin and U. Rössler, Phys. Stat. Sol. **70**, 717 (1975).

[49] K. Kim, W. R. L. Lambrecht, and B. Segall, Phys. Rev. B **56**, 7363 (1997).

[50] R. M. Pick, M. H. Cohen, and R. M. Martin, Phys. Rev. B **1**, 717 (1970).

[51] F. Giustino, M. L. Cohen, and S. G. Louie, Phys. Rev. B **76**, 165108 (2007).

[52] M. Calandra, G. Profeta, and F. Mauri, Phys. Rev. B **82**, 165111 (2010).

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

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

[55] R. M. Martin, *Electronic Structure: Basic Theory and Methods* (Cambrige University Press, Cambridge, 2004).

[56] S. Baroni, S. de Gironcoli, and A. Dal Corso, Rev. Mod. Phys. **73**, 515 (2002).

[57] J. J. Sakurai, *Modern Quantum Mechanics* (Addison-Wesley, 1994).

[58] X. Gonze, P. Boulanger, and M. Côté, Annalen der Physik (Leipzig) **523**, 168 (2011).

[59] E. I. Blount, Solid State Phys. **13**, 305 (1962).

[60] D. Olguín, A. Cantarero, and M. Cardona, Phys. Status Sol. **220**, 33 (2000).

[61] M. Cardona and M. L. W. Thewalt, Rev. Mod. Phys. **77**, 1173 (2005).

[62] J. Franck and E. G. Dymond, Trans. Faraday Soc. **21**, 536 (1926).

[63] E. U. Condon, Phys. Rev. **32**, 858 (1928).

[64] G. DeFilippis, V. Cataudella, A. S. Mishchenko, C. A. Perroni, and J. T. Devreese, Phys. Rev. Lett. **96**, 136405 (2006).

[65] R. S. Berry, S. A. Rice, and J. Ross, *Physical Chemistry* (Cambridge University Press, Oxford, 2000).

[66] P. B. Allen and J. C. K. Hui, Z. Phys. B: Cond. Mat. **37**, 33 (1980).

[67] H. Brooks, Advances in electronics and electron physics Vol. 7, pp. 121–124, Academic Press, New York, 1955.

[68] F. Giustino, S. Louie, and M. Cohen, Phys. Rev. Lett. **105**, 265501 (2010).

[69] Z. M. Gibbs *et al.*, Appl. Phys. Lett. **103**, 262109 (2013).

[70] S. Poncé *et al.*, Phys. Rev. B **90**, 214304 (2014).

[71] G. Antonius, S. Poncé, P. Boulanger, M. Côté, and X. Gonze, Phys. Rev. Lett. **112**, 215501 (2014).

[72] S. M. Story, J. J. Kas, F. D. Vila, M. J. Verstraete, and J. J. Rehr, Phys. Rev. B **90**, 195135 (2014).

[73] Z. M. Gibbs *et al.*, J. Phys.: Condens. Matter **125**, 136001 (2013).

[74] M. Verstraete, Phys. Rev. B **94**, 075120 (2016).

[75] M. A. Manjon, F. J. Hernández-Fenollosa *et al.*, Phys. J. B **40**, 453 (2004).

[76] M. Cardona and R. K. Kremer, Thin Solid Films **571**, 680 (2014).

[77] J. Bhosale *et al.*, Phys. Rev. B **86**, 195208 (2012).

[78] G. A. Slack, L. J. Schowalter, D. Morelli, and J. A. Freitas, J. Cryst. Growth **246**, 287 (2002).

[79] S. Nakamura, G. Fasol, and S. J. Pearton, *The Blue Laser Diode: The Complete Story* (Springer-Verlag, Berlin, 2000).

[80] M. A. Khan, M. S. Shur, Q. C. Chen, and J. N. Kuznia, Electron. Lett. **30**, 2175 (1994).

[81] S. N. Mohammad, A. A. Salvador, and H. Morkoc, Proc. IEEE **83**, 1306 (1995).

[82] K. Chung, C. H. Lee, and G. C. Yi, Science **330**, 665 (2010).

[83] S. Strite and H. Morkoc, J. Vac. Sci. and Technol. B **10**, 1237 (1992).

[84] O. Brandt, H. Yang, H. Kostial, and K. H. Ploog, Appl. Phys. Lett. **69**, 2707 (1996).

[85] M. E. Lin, G. Xue, G. L. Zhou, J. E. Greene, and H. Morkoc, Appl. Phys. Lett. **63**, 932 (1993).

[86] J. R. Müllhäuser, Properties of Zincblende GaN and (In,Ga,Al)N Heterostructures grown by Molecular Beam Epitaxy. Ph.D. thesis, Paul-Drude-Institut für Festkörperelektronik, Humboldt-Universität zu Berlin, 1998.

[87] J. Callaway, *Quantum Theory of the Solid State* (Academic Press, New York and London, 1974).

[88] X. Gonze *et al.*, Comput. Phys. Commun. **180**, 2582 (2009).

[89] X. Gonze *et al.*, Comput. Phys. Commun. **205**, 106 (2016).

[90] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).

[91] M. Fuchs and M. Scheffler, Phys. Commun. **119**, 67 (1999).

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

[93] M. Fenenberg *et al.*, Phys. Rev. B **85**, 155207 (2012).

[94] V. Bougrov, M. Levinshtein, S. Rumyantsve, and A. Zubrilov, *Gallium Nitride (GaN), in Properties of Advanced Semiconductor Materials: GaN, AlN, InN, BN, SiC, SiGe. Vol 7* (Wiley, New York, 2001), pp. 1–30.

[95] A. Baldereschi and N. O. Lipari, Phys. Rev. B **8**, 2967 (1973).

[96] A. R. Hutson, J. Appl. Phys. **32**, 2287 (1961).

[97] L. D. Landau, Sov. Phys. JETP **3**, 920 (1957).

[98] R. Martin, L. Reining, and D. Ceperley, *Interacting Electrons. Theory and Computational Approaches* (Cambrige University Press, Cambridge, 2016).

[99] M. L. Cohen and T. K. Bergstresser, Phys. Rev. **141**, 789 (1966).

[100] C. N. Berglund and W. E. Spicer, Phys. Rev. **136**, A1044 (1964).

[101] R. Y. Koyama and N. V. Smith, Phys. Rev. B **2**, 3049 (1970).

[102] A. Altland and B. Simons, *Condensed Matter Field Theory, Second Ed.* (Cambrige University Press, Cambridge, 2010).

[103] P. Steiner, H. Höchst, and S. Hüfner, Simple metals, in *Photoemission in Solids II*, edited by L. Ley and M. Cardona, chap. 7, pp. 349–372, Springer, Berlin, 1979.

[104] F. Aryasetiawan, L. Hedin, and K. Karlsson, Phys. Rev. Lett. **77**, 2268 (1996).

[105] M. Guzzo *et al.*, Phys. Rev. Lett. **107**, 166401 (2011).

[106] A. B. Migdal, Sov. Phys. JETP **7**, 996 (1958).

[107] H. Y. Fan, Phys. Rev. **82**, 900 (1951).

[108] D. Dunn, Can. J. Phys. **53**, 321 (1975).

[109] O. Gunnarsson, V. Meden, and K. Schönhammer, Phys. Rev. B **50**, 10462 (1994).

[110] We use here the original non-self-consistent formulation of Migdal, as also reported in Ref. [25], see e.g. Eq.(7.20) and Table 7.1, column E_{TD}/ω_0. This formulation might be improved by introducing self-consistency on the electronic eigenenergies (*i.e.* replacing G_0 by the self-consistent G), or even by introducing a simple self-consistent shift of the whole band following the one of the bottom of the band, as described e.g. by Hedin, see Eq.(21-22) in Ref. [112]. In the cumulant method as applied here, no self-consistency is needed to obtain polaron quasiparticle energies in reasonable agreement with reference data from Ref. [47].

[111] C.-O. Almbladh and L. Hedin, *Handbook of Synchrotron Radiation. Vol. 1* (North-Holland, Amsterdam, 1983), p. 607.

[112] L. Hedin, J. Phys.: Condens. Matter **11**, R489 (1999).

[113] B. Gumhalter, V. Kovač, F. Caruso, H. Lambert, and F. Giustino, Phys. Rev. B **94**, 035103 (2016).

[114] L. Hedin, Physica Scripta **21**, 477 (1980).

[115] G. D. Mahan, Phys. Rev. **145**, 602 (1966).

[116] A. S. Kheifets, V. A. Sashin, M. Vos, E. Weigold, and F. Aryasetiawan, Phys. Rev. B **68**, 233205 (2003).

[117] J. Lischner, D. Vigil-Fowler, and S. G. Louie, Phys. Rev. Lett. **110**, 146801 (2013).

[118] M. Guzzo *et al.*, Phys. Rev. B **89**, 085425 (2014).

[119] G. Mahan and W. Plummer, *Many-body effects in Photoemission. Handbook of Surface Science. Vol. 2. Electronic structure. Chap. 14* (Elsevier, Amsterdam, 2000), pp. 953–987.

[120] B. Holm and F. Aryasetiawan, Phys. Rev. B **56**, 12825 (1997).

[121] F. Caruso and F. Giustino, Eur. Phys. J. B **89** (2016).

[122] D. Vigil-Fowler, S. G. Louie, and J. Lischner, Phys. Rev. B **93**, 235446 (2016).

[123] B. A. McDougall, T. Balasubramanian, and E. Jensen, Phys. Rev. B **51**, 13891 (1995).

[124] A. Eiguren *et al.*, Phys. Rev. Lett. **88**, 066805 (2002).

[125] E. W. Plummer, J. Shi, S.-J. Tang, E. Rotenberg, and S. Kevan, Prog. Surf. Sci. **74**, 251 (2003).

[126] M. Hengsberger, D. Purdie, P. Segovia, M. Garnier, and Y. Baer, Phys. Rev. Lett. **83**, 592 (1999).

[127] S. LaShell, E. Jensen, and T. Balasubramanian, Phys. Rev. B **61**, 2371 (2000).

[128] T. Valla, A. V. Fedorov, P. D. Johnson, and S. L. Hulbert, Phys. Rev. Lett. **83**, 2085 (1999).

[129] F. Reinert *et al.*, Phys. Rev. Lett. **91**, 186406 (2003).

[130] J. J. Lee *et al.*, Nature **515**, 245 EP (2014).

[131] S. N. Rebec *et al.*, Phys. Rev. Lett. **118**, 067002 (2017).

[132] A. Marini, Phys. Rev. Lett. **101**, 106405 (2008).

[133] B. Monserrat, N. Drummond, and R. Needs, Phys. Rev. B **87**, 144302 (2013).

[134] H. Kawai, K. Yamashita, E. Cannuccia, and A. Marini, Phys. Rev. B **89**, 085202 (2014).

[135] B. Monserrat, N. Drummond, C. Pickard, and R. Needs, Phys. Rev. Lett. **112**, 055504 (2014).

[136] B. Monserrat and R. Needs, Phys. Rev. B **89**, 214304 (2014).

[137] S. Poncé *et al.*, Comp. Mat. Sci. **83**, 341 (2014).

[138] G. Antonius *et al.*, Phys. Rev. B **92**, 085137 (2015).

[139] M. Friedrich, A. Riefer, S. Sanna, W. Schmidt, and A. Schindlmayr, J. Phys.: Condens. Matter **27**, 385402 (2015).

[140] B. Monserrat and D. Vanderbilt, Phys. Rev. Lett. **117**, 226801 (2016).

[141] A. Molina-Sánchez, M. Palummo, A. Marini, and L. Wirtz, Phys. Rev. B **93**, 155435 (2016).

[142] C. Villegas, A. Rocha, and A. Marini, Nanoletters **16**, 5095 (2016).

[143] W. Saidi, S. Poncé, and B. Monserrat, J. Phys. Chem Lett. **7**, 5247 (2016).

[144] J. Devreese and A. Alexandrov, Rep. Prog. Phys. **72**, 066501 (2009).

[145] D. Emin, *Polarons* (Cambridge University Press, Cambridge, 2012).

[146] R. Kubo, J. Phys. Soc. Japan **17**, 1100 (1962).

[147] D. C. Langreth, Phys. Rev. B **1**, 471 (1970).

[148] L. Hedin, Nucl. Instrum. Methods Phys. Res. A **308**, 169 (1991).

[149] J. J. Kas, J. J. Rehr, and L. Reining, Phys. Rev. B **90**, 085112 (2014).

[150] J. S. Zhou *et al.*, J. Chem. Phys. **143**, 184109 (2015).

[151] F. Caruso, H. Lambert, and F. Giustino, Phys. Rev. B **114**, 146404 (2015).

[152] M. Z. Mayers, M. S. Hybertsen, and D. R. Reichman, Phys. Rev. B **94**, 081109 (2016).

[153] J. McClain *et al.*, Phys. Rev. B **93**, 235139 (2016).

[154] A. Marini, S. Poncé, and X. Gonze, Phys. Rev. B **91**, 224310 (2015).

[155] One can distinguish a physical η from a numerical δ, as in Eq.(16) of Ref. [20]. In the present paper, we consistently use δ to indicate the numerical broadening (Lorentzian). It replaces the η used in the different equations when the latter are numerically evaluated.

[156] S. N. Klimin, J. Tempere, and J. T. Devreese, Phys. Rev. B **94**, 125206 (2016).

[157] A. S. Mishchenko, N. Nagaosa, N. Prokof'ev, A. Sakamoto, and B. Svistunov, Phys. Rev. Lett. **91**, 236401 (2003).

[158] N. Mecholsky, L. Resca, I. Pegg, and M. Fornari, Phys. Rev. B **89**, 155131 (2014).

[159] Note however that the lack of structure at the broad maximum of the spectral function for large α is also a side effect of our numerical broadening of the self-energy, albeit quite small - of order $0.12\omega_{LO}$. Indeed, at zero broadening, the spectral function is made of a sum of functions, each of which has one inverse-square divergence appearing at a multi-phonon threshold, as described at the end of Sec. IV. These divergencies are spaced by ω_{LO}, corresponding to functions differing by one phonon. For such a function corresponding to n LO phonons, the numerical broadening is n times $0.12\omega_{LO}$, and thus exceeds the spacing of the multiphonon peaks beyond 8 phonons, which happens at the maximum of the spectral functions beyond $\alpha = 8$.

[160] T. O. Woodruff and W. Känzig, J. Phys. Chem. Solids **5**, 268 (1958).

[161] W. Känzig, Phys. Rev. **99**, 1890 (1955).

[162] T. G. Castner and W. Känzig, J. Phys. Chem. Solids **3**, 178 (1957).

[163] A. M. Stoneham, *Theory of Defects in Solids: Electronic Structure of Defects in Insulators and Semiconductors* (Oxford, London, 2001).

[164] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[165] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **78**, 1396(E) (1997).

[166] J. Laflamme Janssen *et al.*, Phys. Rev. B **93**, 205147 (2016).

[167] D. R. Hamann, Phys. Rev. B **88**, 085117 (2013).

[168] J. Zhou, Theory of electron spectroscopy beyond the state-of-the-art. An improved description of fermion-plasmon coupling in Green's function calculations. Ph.D. thesis, U. Paris-Saclay, 2015.

[169] This notation differs from the one in Ref. [168]. We use the "un" and "oc" labels in replacement of "e" and "h", in order to avoid using the latter with two different senses.

[170] J. B. Roberto, B. W. Batterman, and D. T. Keating, Phys. Rev. B **9**, 157 (2590).

[171] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Harcourt College Publishers, 1976).

[172] J. B. Roberto and B. W. Batterman, Phys. Rev. B **8**, 3220 (1970).

[173] R. Colella and A. Merlini, Phys. Stat. Sol. **18**, 157 (1966).

[174] I. Fujimoto, Phys. Rev. B **9**, 591 (1974).

[175] J. C. Phillips, Phys. Lett. **37A**, 434 (1971).

[176] P. Coppens, *X-Ray Charge Densities and Chemical Bonding* (Oxford University Press, 1997).

[177] K. Burke, J. C. Smith, and P. E. Grabowski, Phys. Rev. **137**, A1441 (1965).

[178] N. D. Mermin, Phys. Rev. **137**, A1441 (1965).