Electron-phonon interactions from first principles

Feliciano Giustino*

Department of Materials, University of Oxford, Parks Road, Oxford OX1 3PH, United Kingdom

This $\Psi_k$ Scientific Highlight is a reformatted preprint version of Ref. [1], and is reproduced with permission by the American Physical Society.

Abstract

This article reviews the theory of electron-phonon interactions in solids from the point of view of \textit{ab initio} calculations. While the electron-phonon interaction has been studied for almost a century, predictive non-empirical calculations have become feasible only during the past two decades. Today it is possible to calculate from first principles many materials properties related to the electron-phonon interaction, including the critical temperature of conventional superconductors, the carrier mobility in semiconductors, the temperature dependence of optical spectra in direct and indirect-gap semiconductors, the relaxation rates of photoexcited carriers, the electron mass renormalization in angle-resolved photoelectron spectra, and the non-adiabatic corrections to phonon dispersion relations. Here we review the theoretical and computational framework underlying modern electron-phonon calculations from first principles, as well as landmark investigations of the electron-phonon interaction in real materials. In the first part of the article we summarize the elementary theory of electron-phonon interactions and their calculations based on density-functional theory. In the second part we discuss a general field-theoretic formulation of the electron-phonon problem, and establish the connection with practical first-principles calculations. In the third part we review a number of recent investigations of electron-phonon interactions in the areas of vibrational spectroscopy, photoelectron spectroscopy, optical spectroscopy, transport, and superconductivity.

Contents

1 Introduction 4

2 Historical development 6

*feliciano.giustino@materials.ox.ac.uk
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2.2</td>
<td>Expressions for the electron self-energy used in <em>ab initio</em> calculations</td>
<td>49</td>
</tr>
<tr>
<td>5.2.3</td>
<td>Temperature-dependence of electronic band structures</td>
<td>52</td>
</tr>
<tr>
<td>5.2.4</td>
<td>Carrier lifetimes</td>
<td>54</td>
</tr>
<tr>
<td>5.2.5</td>
<td>Kinks and satellites</td>
<td>56</td>
</tr>
<tr>
<td>5.2.6</td>
<td>Model Hamiltonians, polarons, and the cumulant expansion</td>
<td>58</td>
</tr>
<tr>
<td>6</td>
<td>Efficient calculations of matrix elements and their integrals</td>
<td>59</td>
</tr>
<tr>
<td>6.1</td>
<td>Wannier interpolation</td>
<td>60</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Maximally-localized Wannier functions</td>
<td>60</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Interpolation of electron-phonon matrix elements</td>
<td>61</td>
</tr>
<tr>
<td>6.1.3</td>
<td>Electron-phonon matrix elements in polar materials</td>
<td>63</td>
</tr>
<tr>
<td>6.2</td>
<td>Fermi surface harmonics</td>
<td>64</td>
</tr>
<tr>
<td>7</td>
<td>Non-adiabatic vibrational frequencies and linewidths</td>
<td>66</td>
</tr>
<tr>
<td>8</td>
<td>Electron-phonon interactions in photoelectron spectroscopy</td>
<td>70</td>
</tr>
<tr>
<td>8.0.1</td>
<td>Electron mass enhancement in metals</td>
<td>75</td>
</tr>
<tr>
<td>9</td>
<td>Electron-phonon effects in the optical properties of semiconductors</td>
<td>76</td>
</tr>
<tr>
<td>9.1</td>
<td>Temperature dependence of band gaps and band structures</td>
<td>76</td>
</tr>
<tr>
<td>9.1.1</td>
<td>Perturbative calculations based on the Allen-Heine theory</td>
<td>76</td>
</tr>
<tr>
<td>9.1.2</td>
<td>Non-perturbative adiabatic calculations</td>
<td>80</td>
</tr>
<tr>
<td>9.2</td>
<td>Phonon-assisted optical absorption</td>
<td>81</td>
</tr>
<tr>
<td>10</td>
<td>Carrier dynamics and transport</td>
<td>83</td>
</tr>
<tr>
<td>10.1</td>
<td>Electron linewidths and lifetimes</td>
<td>83</td>
</tr>
<tr>
<td>10.2</td>
<td>Phonon-limited mobility</td>
<td>84</td>
</tr>
<tr>
<td>11</td>
<td>Phonon-mediated superconductors</td>
<td>86</td>
</tr>
<tr>
<td>11.1</td>
<td>McMillan/Allen-Dynes formula</td>
<td>86</td>
</tr>
<tr>
<td>11.2</td>
<td>Anisotropic Migdal-Eliashberg theory</td>
<td>88</td>
</tr>
<tr>
<td>11.3</td>
<td>Density functional theory for superconductors</td>
<td>90</td>
</tr>
<tr>
<td>12</td>
<td>Electron-phonon interactions beyond the local density approximation to DFT</td>
<td>91</td>
</tr>
<tr>
<td>13</td>
<td>Conclusions</td>
<td>94</td>
</tr>
</tbody>
</table>
1 Introduction

The interaction between fermions and bosons is one of the cornerstones of many-particle physics. It is therefore unsurprising that, despite being one of the most thoroughly studied chapters of solid state physics, the interaction between electrons and phonons in solids continues to attract unrelenting attention.

Electron-phonon interactions (EPIs) are ubiquitous in condensed matter and materials physics. For example, they underpin the temperature dependence of the electrical resistivity in metals and the carrier mobility in semiconductors, they give rise to conventional superconductivity, and contribute to optical absorption in indirect-gap semiconductors. In addition, EPIs enable the thermalization of hot carriers, determine the temperature dependence of electron energy bands in solids, and distort band structures and phonon dispersion relations of metals, leading to characteristic kinks and Kohn anomalies in photoemission and Raman/neutron spectra, respectively. EPIs also play a role in the areas of spintronics and quantum information, for example by coupling lattice and spin degrees of freedom in electromagnons, or by modulating the lifetimes of electron spins in color centers.

Given the fundamental and practical importance of electron-phonon interactions, it is perhaps surprising that the majority of theoretical studies in this area still rely on semi-empirical model Hamiltonians, especially in times when \textit{ab initio} calculations have become pervasive in every area of condensed matter and materials physics. The reason for this lag can be found in the complexity of electron-phonon calculations: while density functional theory (DFT) calculations of total energies and structural properties were already well established in the early 1980s \cite{2}, systematic \textit{ab initio} calculations of EPIs had to wait for the development of density functional perturbation theory (DFPT) for lattice dynamics between the late 1980s and the mid 1990s \cite{3,4,5}.

Despite this delayed start, the past two decades have witnessed tremendous progress in this area, and new exciting applications are becoming accessible as first-principles techniques for studying EPIs catch up with more established DFT methods. These advances are driving the evolution from \textit{qualitative} and \textit{descriptive} theories of electron-phonon effects in model solids to \textit{quantitative} and \textit{predictive} theories of real materials. As the methodology for calculating EPIs from first principles is rapidly reaching maturity, it appears that the time is ripe for reviewing this vast, complex and fascinating landscape.

One of the most authoritative reviews on the theory of EPIs is the classic book by Grimvall \cite{6}. This monumental work represents an unmissable reference for the specialist. However, as this book pre-dates the rise of \textit{ab initio} computational methods based on DFT, it inevitably misses the most recent developments in this area. The present article constitutes an attempt at filling this gap by reflecting on what DFT calculations can contribute to the study of electron-phonon
physics. In addition, this article is also an opportunity to establish a unified conceptual and mathematical framework in this incredibly diverse landscape, shed light on the key approximations, and identify some of the challenges and opportunities ahead.

As emphasized by the title ‘Electron-phonon interactions from first principles’, the aim of this article is to review the \textit{ab initio} theory of EPIs and to survey modern advances in \textit{ab initio} calculations of EPIs. The reader interested in the fundamentals of electron-phonon physics or in theoretical developments relating to model Hamiltonians is referred to the outstanding monographs by Ziman \cite{7}, Grimvall \cite{6}, Schrieffer \cite{8}, Mahan \cite{9}, and Alexandrov and Devreese \cite{10}.

Among significant recent advances that are covered in this review we mention the zero-point renormalization and the temperature dependence of electronic band structures; the calculation of phonon-assisted optical absorption spectra; the electron mass renormalization and the kinks in angle-resolved photoemission spectra; the thermalization of hot carriers in semiconductors; the calculation of phonon-limited mobility; the development of efficient computational techniques for calculating EPIs; and efforts to improve the predictive power of EPI calculations by going beyond standard density functional theory.

The review is organized as follows: Sec. \textcolor{blue}{2} provides an historical perspective on the development of theories of the EPI, from early semi-empirical approaches to modern first-principles calculations. In Sec. \textcolor{blue}{3} we examine the various components of DFT calculations of EPIs in solids, and set the formalism which will be used throughout this article. Section \textcolor{blue}{4} provides a synthesis of the most advanced field-theoretic approaches employed to study EPIs, and Sec. \textcolor{blue}{5} makes the link between the most general formalism and DFT calculations for real materials. In this section the reader will find a number of expressions which are useful for practical implementations. Section \textcolor{blue}{6} reviews advanced computational techniques for performing calculations of EPIs efficiently and accurately, such as Wannier interpolation and Fermi surface harmonics. Here we also discuss recent progress in the study of electron-phonon couplings in polar semiconductors. In Sec. \textcolor{blue}{7} we discuss recent calculations of phonons beyond the adiabatic Born-Oppenheimer approximation. Section \textcolor{blue}{8} reviews calculations of EPIs in the context of photoelectron spectroscopy. Section \textcolor{blue}{9} focuses on the optical properties of semiconductors and insulators, in particular the temperature dependence of the band structure and phonon-assisted optical processes. In Sec. \textcolor{blue}{10} we review calculations on the effects of EPIs on carrier dynamics and transport, including carrier thermalization rates and mobilities. Section \textcolor{blue}{11} discusses EPI calculations in the area of phonon-mediated superconductivity. Attempts at improving the accuracy and predictive power of \textit{ab initio} EPI calculations by using more sophisticated electronic structure methods are discussed in Sec. \textcolor{blue}{12}. Finally in Sec. \textcolor{blue}{13} we highlight the most pressing challenges in the study of EPIs from first principles, and we present our conclusions. We leave to the appendices some notational remarks and more technical discussions.
2 Historical development

The notion of ‘electron-phonon interactions’ is as old as the quantum theory of solids. In fact in the very same work where Bloch discussed the formal solutions of the Schrödinger equation in periodic potentials [11], Sec. V begins with the all-telling title: “The interaction of the electrons and the elastic waves of the lattice”. In this work the first quantum theory of the temperature-dependent electrical resistivity of metals was developed. It took only a few years for Bloch’s ‘elastic waves’ to be replaced by the brand-name ‘phonon’ by Frenkel [12], thus establishing a tradition that continues unaltered almost a century later [13].

In order to discuss the early approaches to the electron-phonon problem, it is useful to state right from the start the standard form of the Hamiltonian describing a coupled electron-phonon system:

\[
\hat{H} = \sum_{n\mathbf{k}} \varepsilon_{n\mathbf{k}} \hat{c}_{n\mathbf{k}}^\dagger \hat{c}_{n\mathbf{k}} + \sum_{\mathbf{q}\nu} \hbar \omega_{\mathbf{q}\nu} (\hat{a}_{\mathbf{q}\nu}^\dagger \hat{a}_{\mathbf{q}\nu} + 1/2)
\]

\[
+ \frac{N_p}{2} \sum_{\mathbf{k},\mathbf{q}} g_{mn\nu}(\mathbf{k}, \mathbf{q}) \hat{c}_{m\mathbf{k}+\mathbf{q}}^\dagger \hat{c}_{n\mathbf{k}} (\hat{a}_{\mathbf{q}\nu} + \hat{a}_{\mathbf{q}\nu}^\dagger)
\]

\[
+ \frac{N_p}{2} \sum_{\mathbf{k},\mathbf{q},\mathbf{q}'} g_{mn\nu\nu'}^\text{DW}(\mathbf{k}, \mathbf{q}, \mathbf{q'}) \hat{c}_{m\mathbf{k}+\mathbf{q}+\mathbf{q}'}^\dagger \hat{c}_{n\mathbf{k}} (\hat{a}_{\mathbf{q}\nu} + \hat{a}_{\mathbf{q}\nu}^\dagger) (\hat{a}_{\mathbf{q}'\nu'} + \hat{a}_{\mathbf{q}'\nu'}^\dagger). \quad (1)
\]

In this expression the first line describes the separate electron and phonon subsystems using the usual second-quantized formalism, while the second line specifies the mutual coupling between electrons and phonons to first order in the atomic displacements [9]. Here \(\varepsilon_{n\mathbf{k}}\) is the single-particle eigenvalue of an electron with crystal momentum \(\mathbf{k}\) in the band \(n\), \(\omega_{\mathbf{q}\nu}\) is the frequency of a lattice vibration with crystal momentum \(\mathbf{q}\) in the branch \(\nu\), and \(\hat{c}_{n\mathbf{k}}^\dagger/\hat{c}_{n\mathbf{k}}\) (\(\hat{a}_{\mathbf{q}\nu}^\dagger/\hat{a}_{\mathbf{q}\nu}\)) are the associated fermionic (bosonic) creation/destruction operators. \(N_p\) is the number of unit cells in the Born-von Kármán supercell (see Appendix A). The third line of Eq. (1) describes the electron-phonon coupling Hamiltonian to second order in the atomic displacements. This contribution is rarely found in the early literature (hence the square brackets), but it plays an important role in the theory of temperature-dependent band structures (Sec. 5.2.1). The matrix elements \(g_{mn\nu}(\mathbf{k}, \mathbf{q})\) and \(g_{mn\nu\nu'}^\text{DW}(\mathbf{k}, \mathbf{q}, \mathbf{q}')\) measure the strength of the coupling between the electron and the phonon subsystems, and have physical dimensions of an energy. Here the superscript ‘DW’ stands for Debye-Waller, and relates to the Debye-Waller self-energy to be discussed in Sec. 5.2.2. Complete details as well as a derivation of Eq. (1) will be provided in Sec. 3.

The formal simplicity of Eq. (1) conceals some important difficulties that one faces when attempting to use this equation for predictive calculations. For example, the electronic Hamiltonian relies on the assumption that the system under consideration can be described in terms of well-defined quasi-particle excitations. Similarly, the phonon term is meaningful only within the harmonic and the adiabatic approximations. More importantly, Eq. (1) does not provide us with any prescription for determining the numerical parameters \(\varepsilon_{n\mathbf{k}}, \omega_{\mathbf{q}\nu}, g_{mn\nu}(\mathbf{k}, \mathbf{q}),\) and \(g_{mn\nu\nu'}^\text{DW}(\mathbf{k}, \mathbf{q}, \mathbf{q}').\)

In a sense the history of the study of electron-phonon interactions is really the history of how to
calculate the parameters entering Eq. (1) using procedures that can be at once rigorous, reliable, and practical. As it will become clear in Sec. 4, despite enormous progress in this area, some conceptual difficulties still remain.

2.1 Early approaches to the electron-phonon interaction

2.1.1 Metals

A clear account of the theory of EPIs until the late 1950s is given by Ziman [7]. In the following we highlight only those aspects that are relevant to the subsequent discussion in this article.

Early studies of electron-phonon interactions in solids were motivated by the quest for a quantum theory of the electrical resistivity in metals [14]. The common denominator of most early approaches is that the electronic excitations in Eq. (1) were described using the free electron gas model,

$$\varepsilon_n^k = \frac{\hbar^2}{2m_e} |k|^2 - \varepsilon_F,$$

where $m_e$ is the electron mass and $\varepsilon_F$ is the Fermi energy; the lattice vibrations were described as acoustic waves using the Debye model,

$$\omega_{q\nu} = v_s |q|,$$

where $v_s$ is the speed of sound in the solid. Both approximations were reasonable given that the systems of interest included almost exclusively elemental metals, and primarily monovalent alkali and noble metals [15]. While these approximations were fairly straightforward, it was considerably more challenging to determine the EPI matrix elements $g_{mn\nu}(k, q)$ using realistic approximations.

The very first expression of the electron-phonon matrix element was derived by Bloch [11]; using contemporary notation it can be written as:

$$g_{mn\nu}(k, q) = -\frac{i}{2N_p M_\kappa \omega_{q\nu}} \frac{\hbar}{2N_p M_\kappa \omega_{q\nu}} q \cdot e_{\kappa\nu}(q) V_0.$$  \hspace{1cm} (2)

Here $M_\kappa$ is the mass of the $\kappa$-th nucleus, and $e_{\kappa\nu}(q)$ is the polarization of the acoustic wave corresponding to the wavevector $q$ and mode $\nu$. The term $V_0$ represents a unit-cell average of the ‘effective’ potential experienced by the electrons in the crystal. Equation (2) was meant to describe the scattering from an initial electronic state with wavevector $k$ to a final state with wavevector $k + q$, via an acoustic phonon of wavevector $q$ and frequency $\omega_{q\nu}$. The formula was developed for free electron metals, and neglects so-called ‘umklapp’ (folding) processes, i.e. scattering events whereby $k$ goes into $k + q + G$ with $G$ being a reciprocal lattice vector. A derivation of Eq. (2) is provided in Sec. 3.2.5. In order to determine $V_0$ Bloch [11] argued that the crystal may be described as a continuous deformable medium. Starting from this assumption he reached the conclusion that the average potential can be approximated as $V_0 = \frac{\hbar^2}{(16m_e a_0^2)}$ ($a_0$ is the Bohr radius). Even though Bloch’s matrix element is no longer in use, this model provides helpful insight into the nature of EPIs in monovalent metals. For example the so-called ‘polarization factor’ in Eq. (2), $q \cdot e_{\kappa\nu}(q)$, shows that (in the absence of umklapp processes) only longitudinal sound waves scatter electrons.

Nordheim proposed a refinement of Bloch’s model whereby the average potential $V_0$ in Eq. (2) is replaced by the Fourier component $V_\kappa(q)$ of the ionic Coulomb potential (see Sec. 3.2.5) [16]. The key assumption underlying this model is that the effective potential experienced by the electrons is simply the sum of the individual bare ionic potentials of each nucleus. When a nucleus is
displaced from its equilibrium position, the corresponding potential also shifts rigidly. This is
the so-called ‘rigid-ion’ approximation.

The main difficulty that arises with the rigid-ion model is that the Fourier transform of the
Coulomb potential diverges as \( q^{-2} \) for \( q = |q| \to 0 \); this leads to unrealistically strong EPs. In
order to circumvent this difficulty Mott and Jones proposed to truncate the ionic potential at
the boundary of the Wigner-Seitz unit cell of the crystal [15]. This choice represents the first
attempt at including the electronic screening of the nuclear potential in a rudimentary form. In
practice Mott and Jones calculated the Fourier transform of \( V_\kappa(r) \) by restricting the integration
over a Wigner-Seitz cell; the resulting potential is no longer singular at long wavelengths. A
detailed discussion of this model can be found in Ref. [7].

Despite some initial successes in the study of the electrical conductivity of metals, the descriptive
power of these early models was undermined by the complete neglect of the electronic response
to the ionic displacements. The first attempt at describing the effect of the electronic screening
was made by Bardeen [17]. In his model the average potential \( V_0 \) in Eq. (2) is replaced by:

\[
V_0 \to V_\kappa(q)/\epsilon(q),
\]

where \( \epsilon(q) \) is the Lindhard function [9]:

\[
\epsilon(q) = 1 + \left( k_{TF}/q \right)^2 F(q/2k_F).
\]

Here \( k_{TF} \) and \( k_F \) are the Thomas-Fermi screening wavevector and the Fermi wavevector, respec-
tively, and \( F(x) = 1/2 + (4x)^{-1}(1 - x^2) \log |1 + x|/|1 - x| \). A derivation of Bardeen’s model is
provided in Sec. 3.2.5. Since \( \epsilon(q) \to (k_{TF}/q)^2 \) for \( q \to 0 \), the singularity of the electron-nuclei
potential is removed in Bardeen’s matrix element. The work of Bardeen can be considered as a
precursor of modern ab initio approaches, insofar the calculation of the matrix element was car-
ried out using a self-consistent field method within the linearized Hartree theory. This strategy
is similar in spirit to modern DFPT calculations.

The key qualitative difference between the approach of Bardeen and modern techniques lies in
the neglect of exchange and correlation effects in the screening. A possible route to overcome
this limitation was proposed by Bardeen and Pines [18]. In this work the authors considered the
role of a screened exchange interaction in the electron-phonon problem (see Appendix B of their
work), however the mathematical complexity of the formalism prevented further progress along
this direction. Similar efforts were undertaken by Hone [19], and a more detailed account of the
early approximations to exchange and correlation can be found in Ref. [6].

The most interesting aspect of the work by Bardeen and Pines [18], as well as previous work
along the same lines by Nakajima [20], is that for the first time the electron-phonon problem was
addressed using a field-theoretic approach.

One interesting feature in the theory of Bardeen and Pines is that their field-theoretic formul-
ation naturally leads to a retarded electron-phonon vertex: the effective potential experienced
by electrons upon the displacement of nuclei depends on how fast this displacement takes place.
In this approach the effective potential \( V_0 \) in Eq. (2) is replaced by the dynamically screened
potential:
\[ V_0 \rightarrow V_\kappa(q)/\epsilon(q, \omega_{q\nu}). \] (5)

Here \( \epsilon(q, \omega) \) is the frequency-dependent Lindhard function \[9\], and the effect of electronic screening is evaluated at the phonon frequency, \( \omega = \omega_{q\nu} \). Somewhat surprisingly, this development was not followed up in the literature on \textit{ab initio} calculations of EPIs.

### 2.1.2 Semiconductors

While the investigation of electron-phonon effects was initially restricted to monovalent metals, the formal developments were soon extended to the case of more complex systems such as semiconductors. Carriers in semiconductors are typically confined within a narrow energy range near the band extrema; consequently it is expected that the dominant electron-phonon scattering mechanisms will involve long-wavelength phonons (\( q \rightarrow 0 \)). This concept was formalized by Bardeen and Shockley \[21, 22\], laying the foundations of the ‘deformation-potential’ method.

In the deformation potential approach it is assumed that the atomic displacements can be described by long-wavelength acoustic waves, and these can be related in turn to the elastic strain of the crystal. Using concepts from the effective mass theory, Bardeen and Shockley showed that in this approximation the potential \( V_0 \) in Eq. (2) can be replaced by:
\[ V_0 \rightarrow E_{1,nk} = \Omega \partial \epsilon_{nk}/\partial \Omega, \] (6)

where \( \Omega \) represents the volume of the unit cell, and the electron eigenvalues correspond to the valence or conduction band extrema. The derivation of this result can be found in Appendix B of Ref. \[22\]. The deformation potentials \( E_1 \) were obtained empirically; for example Bardeen and Shockley determined these values for the band extrema of silicon by fitting mobility data \[22\]. More complex scenarios such as anisotropic constant-energy surfaces in semiconductors were subsequently addressed by considering the effects of shear deformations \[23\]. While the concept of deformation potentials has become a classic in semiconductor physics, this method relies on a semi-empirical approach and lacks predictive power.

### 2.1.3 Ionic crystals

A class of materials that played an important role in the development of the theory of EPIs is that of ionic crystals. The qualitative difference between ionic solids and the systems discussed in Secs. 2.1.1 and 2.1.2 is that the atomic displacements can generate long-ranged electric fields; these fields provide a new scattering channel for electrons and holes.

The theory of polar electron-phonon coupling started with the investigation of the electron mean free path in ionic crystals, in search for a theoretical model of dielectric breakdown in insulators \[21, 25\]. The central idea of these models is that in insulators the density of free carriers is very low, therefore it is sensible to consider a single electron interacting with the polarization field of the ionic lattice.
The Fröhlich model is similar in spirit to the contemporary work of Bardeen for metals. The main difference is that Fröhlich considered the screening arising from the dielectric polarization of an insulating crystal, while Bardeen considered the screening arising from the response of the Fermi sea.

Fröhlich showed that in the case of isotropic ionic crystals the effective potential \( V_0 \) appearing in Eq. (2) must be replaced by [26]:

\[
V_0 \rightarrow -\frac{e^2 M_0 \omega_0^2 q^2}{\epsilon_0 \Omega} \left( \frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_0} \right)^{\frac{1}{2}} \frac{1}{|q|^2}.
\]  

(7)

In this expression \( e \) is the electron charge, \( \epsilon_0 \) is the dielectric permittivity of vacuum, \( \epsilon_0 \) and \( \epsilon_\infty \) are the static and the high-frequency relative permittivities, respectively. This result is derived in Sec. 6.1.3. Using Eqs. (7) and (2) we see that when \( \epsilon_0 > \epsilon_\infty \) the matrix element \( g_{mn\nu}(k, q) \) diverges as \( |q|^{-1} \) at long wavelengths. This singular behavior can lead to very strong EPIs, and provides the physical basis for the phenomenon of electron self-trapping in polarons [27, 28]. The initial studies in this area were rapidly followed by more refined approaches based on field-theoretic methods [29]. A comprehensive discussion of the various models can be found in the original review article by Fröhlich [30].

2.2 The pseudopotential method

The approximations underpinning the models discussed in Sec. 2.1 become inadequate when one tries to study EPIs for elements across the periodic table. This and other limitations stimulated the development of the pseudopotential method, starting in the late 1950s with the work of Phillips and Kleinman [31]. The theory of pseudopotentials is too vast to be summarized in a few lines, and the reader is referred to Chapter 11 of Ref. [2] for a thorough discussion. Here we only highlight the aspects that are relevant to the calculation of EPIs.

The genesis of the pseudopotential method is linked with the question on how the valence electrons of metals could be described using the electron gas model, even though the orthogonality to the core states imposes rapid fluctuations of the valence wavefunctions near the atomic cores. In order to address this question, it is useful to go through the key steps of the orthogonalized planewaves method [32]. In this method one considers planewaves \( |k + G\rangle \) for the wavevector \( k + G \), and projects out the component belonging the Hilbert subspace spanned by core electrons. This is done by defining \( |k + G\rangle_{\text{OPW}} = |k + G\rangle - \sum_c \langle \phi_c | k + G \rangle \langle \phi_c | \), where the \( |\phi_c\rangle \) represent the core states of all atoms in the system. The functions \( |k + G\rangle_{\text{OPW}} \) are by construction orthogonal to core states, therefore they can be used to expand the valence electron wavefunctions \( |\psi_{nk}\rangle \) using only a few basis elements: \( |\psi_{nk}\rangle = \sum_G c_k(G) |k + G\rangle_{\text{OPW}} \). In the language of pseudopotential theory \( |\psi_{nk}\rangle \) is referred to as the ‘all-electron’ wavefunction, while the function \( |\tilde{\psi}_{nk}\rangle = \sum_G \alpha_k(G) |k + G\rangle \) is referred to as the ‘pseudo’ wavefunction. The all-electron and the pseudo wavefunctions are simply related as follows:

\[
|\psi_{nk}\rangle = \hat{T} |\tilde{\psi}_{nk}\rangle, \quad \text{with} \quad \hat{T} = 1 - \sum_c |\phi_c\rangle \langle \phi_c |.
\]  

(8)
Here we used a modern notation borrowed from the projector-augmented wave (PAW) method of Blöchl [33]. By construction, the pseudo-wavefunction $|\tilde{\psi}_{nk}\rangle$ does not exhibit rapid fluctuations near the atomic cores. The projector operator $\hat{T}$ is now used to rewrite the single-particle Schrödinger equation for the all-electron wavefunction (e.g. the Kohn-Sham equations) in terms of the pseudo-wavefunctions. Using $\hat{H}|\psi_{nk}\rangle = \varepsilon_{nk}|\psi_{nk}\rangle$ and Eq. (8) we have:

$$\hat{T}^\dagger \hat{H} \hat{T} |\tilde{\psi}_{nk}\rangle = \varepsilon_{nk} |\tilde{\psi}_{nk}\rangle,$$

which is a generalized eigenvalue problem. By replacing the definition of $\hat{T}$ given above one finds $\hat{V}_{\text{rep}}$:

$$(\hat{H} + \hat{V}_{\text{rep}})|\tilde{\psi}_{nk}\rangle = \varepsilon_{nk} |\tilde{\psi}_{nk}\rangle,$$

with $\hat{V}_{\text{rep}} = \sum_c (\varepsilon_{nk} - \varepsilon_c)|\phi_c\rangle\langle\phi_c|$ and $\varepsilon_c$ being the eigenvalue of a core electron. Clearly the additional potential $\hat{V}_{\text{rep}}$ is strongly repulsive and is localized near the atomic cores. Cohen and Heine showed that this extra potential largely cancels the attractive potential of the nuclei [34]. This is the reason why valence electrons in metals behave almost like free electrons.

The practical consequence of these developments is that it is possible to define smooth effective ‘pseudo-potentials’ for systematic band structure calculations, whose form factors include only a few Fourier components [35–38].

The use of pseudopotentials in electron-phonon calculations started with the works of Sham and Ziman [39, 40]. Sham showed that, if the pseudopotential can be described by a local function, then the electron-phonon matrix element $g_{mn\nu}(\mathbf{k}, \mathbf{q})$ can be calculated by replacing the all-electron potentials and wavefunctions by the corresponding pseudo-potentials and pseudo-wavefunctions [39]. In this approach the pseudo-potentials move around rigidly with the ionic cores, therefore we are dealing effectively with an improved version of the rigid-ion approximation discussed in Sec. 2.1.

The pseudopotential method was employed by Shuey in order to calculate the electron-phonon matrix elements in germanium [41]. Shortly afterwards many calculations of electron-phonon interactions based on the pseudopotential method appeared in the literature, including work on the resistivity of metals [42–45], the electron mass-enhancement in metals [46–50], the superconducting transition temperatures within the McMillan formalism [51, 52], the mobility of semiconductors [53], and the temperature dependence of semiconductor band structures [54, 55]. These calculations were mostly based on phonon dispersion relations extracted from neutron scattering data, and the results were in reasonable agreement with experiment. It seems fair to say that the pseudopotential method enabled the evolution from qualitative to quantitative calculations of electron-phonon interactions.

Before proceeding we note that, although Eqs. (8) and (9) were introduced starting from the method of orthogonalized planewaves, there exists considerable freedom in the choice of the operator $\hat{T}$. In practice $\hat{T}$ can be chosen so as to make $\tilde{\psi}_{nk}$ as smooth as possible, while retaining information on the all-electron wavefunctions near the ionic cores. This was achieved by the PAW method [33]. Broadly speaking it is also possible to re-interpret the historical development of the pseudopotential method as the evolution of the projector $\hat{T}$. In fact Blöchl showed how the most
popular pseudopotential methods \cite{56-59} can be derived from the PAW method under specific approximations \cite{33}.

### 2.3 Ab initio self-consistent field calculations

Predictive calculations of EPIs became possible with the development of *ab initio* DFT techniques. The key advantage of DFT methods is the possibility of calculating electron band structures, phonon dispersion relations, and electron-phonon matrix elements entirely from first principles. Historically, DFT started with the works of Hohenberg and Kohn \cite{60} and Kohn and Sham \cite{61}. However, its widespread use had to wait for the development of accurate parametrizations of the exchange and correlation energy of the electron gas \cite{62-66}. An introduction to DFT techniques can be found in the books by Parr and Yang \cite{67} (advanced), Martin \cite{2} (intermediate), and Giustino \cite{68} (elementary).

The first calculation of electron-phonon interactions using DFT was carried out by Dacorogna *et al.* \cite{69} using a ‘frozen-phonon’ approach (see Sec. 3.2.3). In this work the authors computed electron bands, phonon dispersions, and electron-phonon matrix elements of Al entirely from first principles. Quoting from the original manuscript: “This calculation is *ab initio* since only information about the Al atom, i.e. the atomic number and atomic mass, is used as input”. Dacorogna *et al.* calculated the so-called electron-phonon coupling strength $\lambda_{q\nu}$ for several phonon branches $\nu$ and momenta $q$ throughout the Brillouin zone, as well as the phonon linewidths arising from the EPI (see Secs. 7 and 11.1). The average coupling strength was found to be in good agreement with that extracted from the superconducting transition temperature. In the approach of Refs. \cite{69-72} the electron-phonon matrix element was calculated using:

$$g_{mn\nu}(k, q) = \langle u_{mk+q} | \Delta_{q\nu} v^{\text{KS}} | u_{nk} \rangle_{uc},$$

with $u_{nk}$ and $u_{mk+q}$ being the Bloch-periodic components of the Kohn-Sham electron wavefunctions, $\Delta_{q\nu} v^{\text{KS}}$ being the phonon-induced variation of the self-consistent potential experienced by the electrons, and the integral extending over one unit cell. Equation (11) will be discussed in Sec. 3.2.2. The scattering potential $\Delta_{q\nu} v^{\text{KS}}$ was calculated by explicitly taking into account the re-arrangement of the electronic charge following a small displacement of the nuclei. The inclusion of the self-consistent response of the electrons constitutes a considerable step forward beyond the rigid-ion approximation of Sec. 2.2.

The next and most recent step in the evolution of electron-phonon calculations came with the development of DFPT for lattice dynamics \cite{3-5}. In contrast to frozen-phonon calculations, which may require large supercells, DFPT enables the calculations of vibrational frequencies and eigenmodes at arbitrary wavevectors in the Brillouin zone. This innovation was critical in the context of electron-phonon physics, since the calculation of many physical quantities requires the evaluation of nontrivial integrals over the Brillouin zone. The first calculations of EPIs using DPFT were reported by Savrasov *et al.* \cite{73}, Liu and Quong \cite{74}, Mauri *et al.* \cite{75}, and Bauer *et al.* \cite{76}. They calculated the electrical resistivity, thermal conductivity, mass enhancement and superconducting critical temperature of a number of elemental metals (e.g. Al, Au, Cu, Mo, Nb, Pb, Pd, Ta, V, and Te), and reported good agreement with experiment.
By the late 1990s most of the basic ingredients required for the \textit{ab initio} calculation of EPIs were available; subsequent studies focused on using these techniques for calculating a variety of materials properties, and on improving the efficiency and accuracy of the methodology. The most recent advances will be reviewed in Secs. 6-12.

3 Electron-phonon interaction in density-functional theory

In this section we review the basic formalism underlying the calculation of EPIs using DFT, and we establish the link with the Hamiltonian in Eq. (1). We start by introducing the standard formalism for lattice vibrations, and the electron-phonon coupling Hamiltonian. Then we briefly summarize established methods of DFPT for calculating electron-phonon matrix elements. For the time being we describe electrons and phonons as separate subsystems; a rigorous theoretical framework for addressing the coupled electron-phonon system will be discussed in Sec. 4.

3.1 Lattice vibrations in crystals

The formalism for studying lattice dynamics in crystals is covered in many excellent textbooks \cite{7, 77-80}. Here we introduce the notation and summarize those aspects which will be useful for subsequent discussions in this section and in Secs. 4 and 5.

We consider $M$ nuclei or ions in the unit cell. The position vector and Cartesian coordinates of the nucleus $\kappa$ in the primitive unit cell are denoted by $\tau_\kappa$ and $\tau_{\kappa \alpha}$, respectively. We describe the infinitely extended solid using Born-von Kármán (BvK) boundary conditions. In this approach, periodic boundary conditions are applied to a large supercell which contains $N_p$ unit cells, identified by the direct lattice vectors $R_p$, with $p = 1, \ldots, N_p$. The position of the nucleus $\kappa$ belonging to the unit cell $p$ is indicated by $\tau_{\kappa p} = R_p + \tau_\kappa$. The Bloch wavevectors $q$ are taken to define a uniform grid of $N_p$ points in one unit cell of the reciprocal lattice, and the vectors of the reciprocal lattice are indicated by $G$. In Appendix A we provide additional details on the notation, and we state the Fourier transforms between direct and reciprocal lattice.

Using standard DFT techniques it is possible to calculate the total potential energy of electrons and nuclei in the BvK supercell. This quantity is denoted as $U(\{\tau_{\kappa p}\})$, where the braces are a short-hand notation for the coordinates of all the ions. The total potential energy refers to electrons in their ground state, with the nuclei being represented as classical particles \textit{clamped} at the coordinates $\tau_{\kappa p}$. Every DFT software package available today provides the quantity $U$ as a standard output.

In order to study lattice vibrations, one begins by making the \textit{harmonic} approximation. Accordingly, the total potential energy is expanded to second order in the displacements $\Delta \tau_{\kappa \alpha p}$ of the ions in the BvK supercell away from their equilibrium positions $\tau_{\kappa 0}$:

$$U = U_0 + \frac{1}{2} \sum_{\kappa \alpha p} \frac{\partial^2 U}{\partial \tau_{\kappa \alpha p} \partial \tau_{\kappa' \alpha' p'}} \Delta \tau_{\kappa \alpha p} \Delta \tau_{\kappa' \alpha' p'},$$

(12)
where $U_0$ denotes the total energy calculated for the ions in their equilibrium positions, and the
derivatives are evaluated for the equilibrium structure. The second derivatives of the total energy
with respect to the nuclear coordinates define the matrix of \textit{‘interatomic force constants’}:

$$C_{\kappa\alpha p,\kappa'\alpha'p'} = \frac{\partial^2 U}{\partial \tau_{\kappa\alpha p} \partial \tau_{\kappa'\alpha'p'}}. \quad (13)$$

The Fourier transform of the \textit{interatomic force constants} yields the \textit{‘dynamical matrix’} \cite{81}:

$$D_{\kappa\alpha,\kappa'\alpha'}(q) = (M_\kappa M_{\kappa'})^{-\frac{1}{2}} \sum_p C_{\kappa\alpha 0,\kappa'\alpha'p} \exp(iq \cdot R_p), \quad (14)$$

where $M_\kappa$ is the mass of the $\kappa$-th ion. The superscript \textit{‘dm’} is there to distinguish this quantity from the many-body phonon propagators $D_{(12)}$ and $D_{\kappa\alpha p,\kappa'\alpha'p'}$ that will be introduced in Sec. 4.2. The dynamical matrix is Hermitian and therefore admits real eigenvalues, which we denote as $\omega_{qv}^2$:

$$\sum_{\kappa'\alpha'} D_{\kappa\alpha,\kappa'\alpha'}(q) e_{\kappa'\alpha',\nu}(q) = \omega_{qv}^2 e_{\kappa\alpha,\nu}(q). \quad (15)$$

In classical mechanics, each $\omega_{qv}$ corresponds to the vibrational frequency of an independent
harmonic oscillator. The hermiticity of the dynamical matrix allows us to choose the eigenvectors $e_{\kappa\alpha,\nu}(q)$ to be orthonormal for each $q$:

$$\sum_{\nu} e_{\kappa'\alpha',\nu}^* e_{\kappa\alpha,\nu}(q) = \delta_{\kappa\kappa'} \delta_{\alpha\alpha'}, \quad (16)$$

$$\sum_{\kappa\alpha} e_{\kappa\alpha,\nu}(q) e_{\kappa\alpha,\nu}(q) = \delta_{\nu\nu'}. \quad (17)$$

Here the index $\nu$ runs from 1 to $3M$. The column vectors $e_{\kappa\alpha,\nu}(q)$ for a given $\nu$ are called the \textit{‘normal modes of vibration’} or the \textit{‘polarization’} of the vibration wave. The following relations can be derived from Eq. (14):

$$\omega_{q'v}^2 = \omega_{qv}^2; \quad e_{\kappa\alpha,\nu}(-q) = e_{\kappa\alpha,\nu}^*(q). \quad (18)$$

These relations between normal modes carry a degree of arbitrariness in the choice of phases; here we have chosen to follow the same phase convention as in Ref. \cite{81}.

Using Eqs. (12) and (13) the Hamiltonian for nuclei considered as quantum particles can be
written as:

$$\hat{H}_p = \frac{1}{2} \sum_{\kappa'\alpha'p'} C_{\kappa\alpha p,\kappa'\alpha'p'} \Delta \tau_{\kappa\alpha p} \Delta \tau_{\kappa'\alpha'p'} - \sum_{\kappa\alpha p} \frac{\hbar^2}{2M_\kappa} \frac{\partial^2}{\partial \tau_{\kappa\alpha p}^2}, \quad (19)$$

where the ground-state energy $U_0$ has been omitted and the last term is the kinetic energy
operator. The Hamiltonian in the above expression corresponds to the energy of an entire BvK
supercell. Equation (19) relies on two approximations: (i) the harmonic approximation, which coincides with the truncation of Eq. (12) to second order in the displacements; and (ii) the Born-Oppenheimer \textit{adiabatic} approximation. This latter approximation is made when one calculates the interatomic force constants with the electrons in their ground state. The scope and validity of the adiabatic approximation will be discussed in detail in Sec. 5.1.1. We note incidentally that, strictly speaking, the Born-Oppenheimer approximation does not need to be invoked were one to use the generalization of DFT to multicomponent systems introduced in Refs. \cite{82,83}. 

14
For practical purposes it is convenient to rewrite Eq. (19) by introducing the quanta of lattice vibrations. This is accomplished by defining the standard creation (\(\hat{a}_q^{\dagger}\)) and destruction (\(\hat{a}_q\)) operators for each phonon of energy \(\hbar \omega_q\) and polarization \(e_{\kappa \alpha \nu} (q)\). This operation is not entirely trivial and is described in detail in Appendix[3]. The formal definition of the ladder operators is given in Eqs. (226)-(227). These operators obey the commutation relations

\[
[\hat{a}_q, \hat{a}_q^{\dagger}] = \delta_{\nu \nu}' \delta_{qq}'
\]

and

\[
[\hat{a}_q, \hat{a}_{q'}] = [\hat{a}_q^{\dagger}, \hat{a}_{q'}^{\dagger}] = 0,
\]

where \(\delta\) is the Kronecker symbol. From these relations we know that the quanta of the harmonic oscillations in crystals obey Bose-Einstein statistics. In Appendix[3] it is shown that the atomic displacements can be expressed in terms of the ladder operators as follows:

\[
\Delta \tau_{\kappa \alpha p} = \left( \frac{M_0}{N_p M_\kappa} \right)^{\frac{1}{2}} \sum_q e^{i q \cdot R_p} e_{\kappa \alpha \nu} (q) l_{q\nu} (\hat{a}_q + \hat{a}_q^{\dagger}),
\]

with \(l_{q\nu}\) being the ‘zero-point’ displacement amplitude:

\[
l_{q\nu} = [\hbar/(2M_0 \omega_{q\nu})]^{1/2}.
\]

Here \(M_0\) is an arbitrary reference mass which is introduced to ensure that \(l_{q\nu}\) has the dimensions of a length and is similar in magnitude to \(\Delta \tau_{\kappa \alpha p}\). Typically \(M_0\) is chosen to be the proton mass. Using Eqs. (13)-(21) the nuclear Hamiltonian can be written in terms of \(3Np\) independent harmonic oscillators as follows:

\[
\hat{H}_p = \sum_{q\nu} \hbar \omega_{q\nu} \left( \hat{a}_q^{\dagger} \hat{a}_q + 1/2 \right),
\]

where the sum is over all wavevectors. The ground-state wavefunction of this Hamiltonian is a product of Gaussians, and all other states can be generated by acting on the ground state with the operators \(\hat{a}_q^{\dagger}\). In the case of \(|q| = 0\) there are three normal modes for which \(\omega_{q\nu} = 0\). For these modes, which correspond to global translations of the crystal, the zero-point displacement \(l_{q\nu}\) is not defined. Throughout this article it is assumed that these modes are skipped in summations containing zero-point amplitudes. A detailed derivation of Eq. (22) and a discussion of the eigenstates of \(\hat{H}_p\) are provided in Appendix[3].

### 3.2 Electron-phonon coupling Hamiltonian

Having outlined the standard formalism for addressing lattice vibrations in crystals, we now proceed to make the connection between DFT calculations and the remaining terms of Eq. (1). The electronic band structure \(\epsilon_{nk}\) and electron-phonon matrix elements \(g_{mn\nu}(k, q)\) are almost invariably calculated by using the Kohn-Sham (KS) Hamiltonian [60][61]. A justification for these choices will be provided in Sec. 5; for now we limit ourselves to outline the key elements of practical calculations.

#### 3.2.1 Kohn-Sham Hamiltonian

Let us denote the Kohn-Sham eigenfunctions by \(\psi_{nk}(r)\), and use \(k\) to indicate both the wavevector and spin. We shall restrict ourselves to systems with collinear spins. The KS eigenfunctions
satisfy the equation \( \hat{H}^{KS} \psi_{nk}(r) = \varepsilon_{nk} \psi_{nk}(r) \), with the Hamiltonian given by:

\[
\hat{H}^{KS} = -\frac{\hbar^2}{2m_e} \nabla^2 + V^{KS}(r; \{\tau_{\kappa\alpha p}\}). (23)
\]

Here the potential \( V^{KS} \) is the sum of the nuclear (or ionic) contribution \( V^{en} \), the Hartree electronic screening \( V^{H} \), and the exchange and correlation potential \( V^{xc} \) [2]:

\[
V^{KS} = V^{en} + V^{H} + V^{xc}. (24)
\]

The potentials appearing in Eq. (24) are defined as follows. The electron-nuclei potential energy is given by:

\[
V^{en}(r; \{\tau_{\kappa\alpha p}\}) = \sum_{\kappa p, \mathbf{T}} V_{\kappa}(r - \tau_{\kappa p} - \mathbf{T}),
\]

where \( V_{\kappa}(r) \) is the interaction between an electron and the nucleus \( \kappa \) located at the center of the reference frame, and \( \mathbf{T} \) denotes a lattice vector of the BvK supercell. In the case of all-electron DFT calculations, \( V_{\kappa}(r) \) is the Coulomb interaction:

\[
V_{\kappa}(r) = -\frac{e^2}{4\pi\varepsilon_0 |r|},
\]

where \( Z_{\kappa} \) is the atomic number of the nucleus \( \kappa \). In the case of pseudopotential implementations \( V_{\kappa}(r) \) is a function that goes as in Eq. (26) at large \( |r| \), but remains finite at \( |r| = 0 \). Furthermore the nuclear charge is replaced by the ionic charge, that is the difference between the nuclear charge and the number of core electrons described by the pseudopotential. In all modern pseudopotential implementations \( V_{\kappa}(r) \) is nonlocal due to the separation of the angular momentum channels [2]. However, since this nonlocality is short-ranged and is inconsequential in the following discussion, it will be ignored here in order to maintain a light notation. The Hartree term is obtained from the electron density, \( n(r'; \{\tau_{\kappa\alpha p}\}) \):

\[
V^{H}(r; \{\tau_{\kappa\alpha p}\}) = \frac{e^2}{4\pi\varepsilon_0} \sum_{\mathbf{T}} \int_{sc} \frac{n(r'; \{\tau_{\kappa\alpha p}\})}{|r - r' - \mathbf{T}|} dr',
\]

where the integral extends over the supercell. The exchange and correlation potential is the functional derivative of the exchange and correlation energy with respect to the electron density [61]:

\[
V^{xc}(r; \{\tau_{\kappa\alpha p}\}) = \frac{\delta E^{xc}[n]}{\delta n}(n(r; \{\tau_{\kappa\alpha p}\})),
\]

The eigenfunctions \( \psi_{nk} \) of \( \hat{H}^{KS} \) can be expressed in the Bloch form:

\[
\psi_{nk}(r) = N_p^{-\frac{1}{2}} u_{nk}(r) e^{i\mathbf{k}\cdot\mathbf{r}},
\]

with \( u_{nk} \) a lattice-periodic function. The wavefunction \( \psi_{nk} \) is taken to be normalized to one in the supercell, while the periodic part \( u_{nk}(r) \) is normalized to one in the crystal unit cell. The electron density is \( n(r) = \sum_{nk} |\psi_{nk}(r)|^2 \), where \( v \) indicates occupied states. In order to determine \( \psi_{nk} \) and \( \varepsilon_{nk} \) the Kohn-Sham equations are solved self-consistently. This requires one to start from a reasonable guess for the electron density (for example a superposition of atomic electron densities), calculate the potentials in Eq. (24), and determine the solutions of the KS
Hamiltonian in Eq. (23). The electron density is re-calculated using these solutions, and the cycle is repeated until convergence.

In order to establish the link with Eq. (1), we can regard the KS Hamiltonian as an effective one-body operator, and make the transition to a second-quantized formalism by using the standard prescription [84]:

\[
\hat{H}_e = \sum_{nk,n'k'} \langle \psi_{nk} | H^{KS} | \psi_{n'k'} \rangle c_{nk}^\dagger c_{n'k'} = \sum_{nk} \varepsilon_{nk} c_{nk}^\dagger c_{nk}. \tag{30}
\]

This expression is useful for performing formal manipulations in the study of coupled electron-phonon systems. However, Eq. (30) implicitly introduces the drastic approximation that the electronic system can be described in terms of sharp quasiparticle excitations. A field-theoretic approach that does not rely on any such approximation is discussed in Sec. 4.

3.2.2 Electron-phonon coupling Hamiltonian to first- and second-order in the atomic displacements

Within the DFT Kohn-Sham formalism, the coupling Hamiltonian appearing in the second line of Eq. (1) is obtained by expanding the Kohn-Sham effective potential in terms of the nuclear displacements \( \Delta \tau_{kp} \) from their equilibrium positions \( \tau_{0}^{0} \). The potential to first order in the displacements is:

\[
V^{KS}(\{ \tau_{kp} \}) = V^{KS}(\{ \tau_{0}^{0} \}) + \sum_{\kappa \alpha p} \frac{\partial V^{KS}}{\partial \tau_{\kappa \alpha p}} \Delta \tau_{\kappa \alpha p}. \tag{31}
\]

This expression can be rewritten into normal mode coordinates using Eq. (20):

\[
V^{KS} = V^{KS}(\{ \tau_{0}^{0} \}) + N_p^{-1/2} \sum_{q \nu} \Delta_{q \nu} V^{KS}(\hat{a}_{q \nu} + \hat{a}_{-q \nu}^\dagger), \tag{32}
\]

having defined:

\[
\Delta_{q \nu} V^{KS} = e^{i \mathbf{q} \cdot \mathbf{r}} \Delta_{q \nu} V^{KS}, \tag{33}
\]

\[
\Delta_{q \nu} V^{KS} = l_{q \nu} \sum_{\kappa \alpha} (M_{\kappa} / M_\alpha)^{1/2} \varepsilon_{\kappa \alpha, \nu} V^{KS}(\hat{a}_{q \nu} + \hat{a}_{-q \nu}^\dagger), \tag{34}
\]

\[
\partial_{\kappa \alpha, q \nu} V^{KS} = \sum_{p} e^{-i \mathbf{q} \cdot (\mathbf{r} - \mathbf{R}_p)} \frac{\partial V^{KS}}{\partial \tau_{\kappa \alpha}} \bigg|_{\mathbf{r} - \mathbf{R}_p}. \tag{35}
\]

From the last expression we see that \( \partial_{\kappa \alpha, q \nu} V^{KS} \) and \( \Delta_{q \nu} V^{KS} \) are lattice-periodic functions. The transition to second quantization is performed as in Eq. (30) [84]:

\[
\hat{H}_{ep} = \sum_{nk,n'k'} \langle \psi_{nk} | V^{KS}(\{ \tau_{kp} \}) - V^{KS}(\{ \tau_{0}^{0} \}) | \psi_{n'k'} \rangle c_{nk}^\dagger c_{n'k'}, \tag{36}
\]

where the brakets indicate an integral over the supercell. After using Eqs. (29), (32)-(35), and (213) we have:

\[
\hat{H}_{ep} = N_p^{-1/2} \sum_{k q m n \nu} g_{mn \nu}(k, q) c_{nk}^\dagger c_{\nu k}^\dagger \hat{a}_{q \nu} + \hat{a}_{-q \nu}^\dagger, \tag{37}
\]
where the electron-phonon matrix element is given by:

\[ g_{mn\nu}(k, q) = \langle u_{m(k+q)} | \Delta_{q\nu} v^{KS} | u_{n(k)} \rangle_{uc}. \]  (38)

Here the subscript ‘uc’ indicates that the integral is carried out within one unit cell of the crystal.

The coupling Hamiltonian in Eq. (37) yields the energy of an entire supercell. In the case of the three translational modes at \(|q| = 0\) we set the matrix elements \(g_{mn\nu}(k, q)\) to zero, as a consequence of the acoustic sum rule (see discussion in Sec. 9.1.1).

Taken together, Eqs. (22), (30), and (37) constitute the starting point of most first-principles calculations of electron-phonon interactions. It remains to be seen how one calculates the electron-phonon matrix elements \(g_{mn\nu}(k, q)\); the most common procedures are described in Sec. 3.2.3.

Before proceeding, we discuss briefly the second-order coupling Hamiltonian which appears in the third and fourth lines of Eq. (1). The rationale for incorporating this extra term is that the expansion of the Kohn-Sham potential to first order in the atomic displacements, Eq. (31), is somewhat inconsistent with the choice of expanding the total potential energy in Eq. (12) to second order in the atomic displacements. This aspect was discussed by Allen and Heine [85] and Allen [86]. In order to obtain an electron-phonon coupling Hamiltonian including terms of second-order in the displacements, we must include the second derivatives of the Kohn-Sham potential in Eq. (31), and follow the same steps which led to Eq. (37). By calling the extra term \(\hat{H}_{ep}^{(2)}\) we have:

\[ \hat{H}_{ep}^{(2)} = N_p^{-1} \sum_{k, q, q'} g_{mn\nu'}^{DW}(k, q, q') \hat{c}^{\dagger}_{m(k+q+q')} \hat{c}_{n(k)} (\hat{a}_{q\nu} + \hat{a}_{-q\nu}^\dagger)(\hat{a}_{q'\nu'}^\dagger + \hat{a}_{-q'\nu'}), \]  (39)

where

\[ g_{mn\nu'}^{DW}(k, q, q') = \frac{1}{2} \langle u_{m(k+q+q')} | \Delta_{q\nu} \Delta_{q'\nu'} v^{KS} | u_{n(k)} \rangle_{uc}. \]  (40)

The variations \(\Delta_{q\nu}\) are the same as in Eqs. (33)-(35).

The second-order coupling Hamiltonian in Eq. (39) is considerably more involved than its first-order counterpart; the increased complexity partly explains why in the literature this term has largely been ignored. So far the Hamiltonian \(\hat{H}_{ep}^{(2)}\) has only been described using an approximation based on first-order perturbation theory [85]. In this special case, the only terms in Eq. (39) that can modify the electron excitation spectrum are those with \(q' = -q\). The corresponding energy shift is \(\Delta \epsilon_{nk} = N_p^{-1} \sum_{q\nu} g_{nn\nu'}^{DW}(k, -q)(2n_{q\nu} + 1)\), with \(n_{q\nu}\) being the number of phonons in each mode. We will come back to this point in Sec. 9.1.1.

### 3.2.3 Calculation of electron-phonon matrix elements using density-functional perturbation theory

In this section we review how the scattering potential \(\Delta_{q\nu} v^{KS}\) appearing in Eq. (38) is calculated in first-principles approaches. The most intuitive approach is to evaluate the derivatives appearing in Eq. (35) by using finite atomic displacements in a supercell:

\[ \frac{\partial V^{KS}}{\partial \tau_{k\alpha p}} \bigg|_{\tau_{k\alpha p}^0} \approx \left[ V^{KS}(r; \tau_{k\alpha p}^0 + b) - V^{KS}(r; \tau_{k\alpha p}^0) \right] / b. \]  (41)
In this expression $b$ is a small displacement of the order of the zero-point amplitude ($\sim 0.1 \, \text{Å}$), and the atom $\kappa$ in the unit cell $p$ is displaced along the direction $\alpha$. The first calculations of electron-phonon interactions within DFT employed a variant of this ‘supercell approach’ whereby all atoms are displaced according to a chosen vibrational eigenmode [69, 72], this strategy is usually referred to as the ‘frozen-phonon’ method.

One disadvantage of the frozen-phonon method is that the supercell may become impractically large when evaluating matrix elements corresponding to long-wavelength phonons. This difficulty can be circumvented by using DFPT [3–5]. The main strength of DFPT is that the scattering potential $\Delta_{q\nu}v^\text{KS}$ in Eq. (38) is obtained by performing calculations within a single unit cell. Since the computational workload of standard (non linear-scaling) DFT calculations scales as the cube of the number of electrons, the saving afforded by DFPT over the frozen-phonon method is proportional to $N_p^2$, and typically corresponds to a factor $>10^3$.

In the DFPT approach of Ref. [87] one calculates the lattice-periodic scattering potential $\partial_{\kappa \alpha, q}v^\text{KS}$ defined by Eq. (35). By differentiating Eq. (24) via Eq. (35) this potential is written as:

$$\partial_{\kappa \alpha, q}v^\text{KS} = \partial_{\kappa \alpha, q}v^\text{en} + \partial_{\kappa \alpha, q}v^\text{H} + \partial_{\kappa \alpha, q}v^\text{xc}. \quad (42)$$

The variation of the ionic potential is obtained from Eqs. (25) and (35). The result is conveniently expressed in reciprocal space:

$$\partial_{\kappa \alpha, q}v^\text{en}(G) = -i \langle q + G \rangle_\kappa V_\kappa(q + G) e^{-i(q+G)\cdot \tau_\kappa}, \quad (43)$$

where the convention for the Fourier transform is $f(G) = \Omega^{-1} \int_{uc} dr \, e^{-iG \cdot r} f(r)$, and $\Omega$ is the volume of the unit cell. In order to keep the presentation as general as possible we avoid indicating explicitly the non-locality of $V_\kappa$ which arises in pseudopotential implementations. The adaptation of this equation and the following ones to the case of nonlocal pseudopotentials, ultrasoft pseudopotentials, and the projector-augmented wave method can be found in Refs. [88], [89], and [90], respectively. The variation of the Hartree and exchange-correlation contributions to the Kohn-Sham potential is obtained from the self-consistent charge density response to the perturbation in Eq. (43). After a few manipulations using Eqs. (27) and (35) one obtains:

$$\partial_{\kappa \alpha, q}v^\text{H}(G) = \Omega v^\text{C}(q + G) \partial_{\kappa \alpha, q}n(G), \quad (44)$$

where $v^\text{C}(q) = \Omega^{-1} \int dr \, e^{-iq \cdot r} e^2/4\pi\epsilon_0 |r|$ is the Fourier transform of the Coulomb potential. For the exchange and correlation potential we use Eq. (28) and the Taylor expansion of a functional to find:

$$\partial_{\kappa \alpha, q}v^{\text{xc}}(G) = \Omega \sum_{G'} f^{\text{xc}}(q + G, q + G') \partial_{\kappa \alpha, q}n(G'), \quad (45)$$

where $f^{\text{xc}}$ indicates the standard exchange and correlation kernel, which is the second-order functional derivative of the exchange and correlation energy $E^{\text{xc}}$ with respect to the electron density [60]:

$$f^{\text{xc}}(r, r') = \frac{\delta^2 E^{\text{xc}}(n)}{\delta n(r) \delta n(r')} \bigg|_{n(r; \tau_p^0)}.$$

(46)

In the case of the local density approximation (LDA) to DFT the exchange and correlation kernel reduces to a local function [67], and Eq. (45) is more conveniently evaluated in real space. Today
DFPT calculations can be performed using one of several exchange and correlation kernels. The effect of the kernel on the calculation of lattice-dynamical properties of solids has been analyzed in several works [91–92]. The formal structure of the DFPT equations discussed in this section remains unchanged if we replace the DFT kernel in Eq. (46) by more sophisticated versions. For example both DFPT calculations based on Hubbard-corrected DFT [93] and DFPT coupled with dynamical mean-field theory [94] have been demonstrated.

It should be noted that in Eqs. (45)-(46) we are implicitly assuming a spin-unpolarized system. The adaptation of these equations as well as the other DFPT equations to the most general case of non-collinear spin systems can be found in Refs. [95–97].

From Eqs. (44) and (45) we see that the evaluation of \( g_{mnv}(k, q) \) goes through the calculation of the variation of the electron density induced by the perturbation \( \partial_{\alpha\alpha, q} v^{KS}(r) e^{iqr} \). Within DFPT such a variation is obtained by evaluating the change of the Kohn-Sham wavefunctions to first order in perturbation theory. After inspection of the perturbed Hamiltonian it becomes evident that the wavefunction change must be of the form \( \partial u_{nk,q} e^{iqr} \), with \( \partial u_{nk,q} \) a lattice-periodic function. Using this observation the first-order variation of the Kohn-Sham equations can be written as a Sternheimer equation [98]:

\[
\left( \hat{H}^{KS}_{k+q} - \varepsilon_{vk} \right) \partial u_{vk,q} = -\partial_{\alpha\alpha, q} v^{KS} u_{vk}, \tag{47}
\]

with \( \hat{H}^{KS}_{k+q} = e^{-i(k+q) \cdot r} \hat{H}^{KS} e^{i(k+q) \cdot r} \). In this equation the index \( v \) indicates an occupied state. For \( |q| = 0 \) one needs also to consider a shift of the energy eigenvalues which introduces an additional term \( \langle u_{vk}|\partial_{\alpha\alpha, 0} v|u_{vk}\rangle uc u_{vk} \) on the right-hand side of Eq. (47). In practice this term is canceled by the use of the projectors described in Eq. (48) below, unless one is dealing with metallic systems. This aspect is discussed in detail in Refs. [99] and [87]. The principal advantage of Eq. (47) over standard perturbation theory is that it does not involve unoccupied electronic states.

A practical problem arises when attempting to solve Eq. (47): the linear system on the left-hand side is ill-conditioned owing to small eigenvalues corresponding to \( \varepsilon_{vk} \approx \varepsilon_{v'k+q} \); furthermore in the case of accidental degeneracies, \( \varepsilon_{vk} = \varepsilon_{v'k+q} \), the system becomes singular. In order to make the system non-singular Giannozzi et al. noted that the variation of the electron density only involves the component of \( \partial u_{vk,q} \) belonging to the unoccupied manifold of Kohn-Sham states [88]. As a consequence, what is really needed is only \( \partial \tilde{u}_{vk,q} = (1 - \hat{P}^{occ}_{k+q}) \partial u_{vk,q} \), having denoted by \( \hat{P}^{occ}_{k+q} = \sum_v |u_{v'k+q}\rangle |u_{vk+q}\rangle \) the projector over the occupied states with wavevector \( k+q \). The equation for this ‘trimmed’ wavefunction variation is simply obtained by projecting both side of Eq. (47) onto \( (1 - \hat{P}^{occ}_{k+q}) \), and noting that \( \hat{P}^{occ}_{k+q} \) and \( \hat{H}^{KS}_{k+q} \) do commute:

\[
\left( \hat{H}^{KS}_{k+q} - \varepsilon_{vk} \right) \partial \tilde{u}_{vk,q} = -(1 - \hat{P}^{occ}_{k+q}) \partial_{\alpha\alpha, q} v^{KS} u_{vk}. \tag{48}
\]

At this point it is possible to remove all small or null eigenvalues of the operator on the left-hand side by adding a term \( \alpha \hat{P}^{occ}_{k+q} \) to the Hamiltonian. This term has no effect on the wavefunction variation, since \( \hat{P}^{occ}_{k+q} \partial \tilde{u}_{vk,q} = 0 \) by construction. The operator is made non-singular by choosing the parameter \( \alpha \) larger than the valence bandwidth [87]. From the wavefunction variation
obtained by solving Eq. (48), it is now possible to construct the density response associated with the wavevector \( q \):

\[
\partial n_{\alpha,q}(r) = 2 N_p^{-1} \sum_{\nu k} u^*_\nu k \partial \tilde{u}_{\nu k,q}.
\] (49)

For simplicity a spin-degenerate system has been assumed (a factor of 2 is implicitly included in the sum over \( k \)), and time-reversal symmetry has been used in order to make the expression more compact (yielding the factor of 2 on the right-hand side).

In practical DFPT calculations, Eq. (48) is solved using an iterative procedure which is similar to standard DFT total energy calculations. One sets the starting perturbation \( \partial n_{\alpha,q,\nu}^{KS} \) to be equal to the electron-nuclei potential in Eq. (43). By solving Eq. (48) for each occupied state \( \nu \) and each wavevector \( k \) using standard linear algebra techniques, one obtains the induced density in Eq. (49). The new density is now used to construct the variations of the Hartree and exchange-correlation potentials in Eqs. (44) and (45). These induced potentials are added to the electron-nuclei potential, yielding a ‘screened’ perturbation \( \partial n_{\alpha,q,\nu}^{KS} \) in Eq. (48). The cycle is repeated until the change of \( \partial n_{\alpha,q} \) between two successive cycles is smaller than a set tolerance.

It can be shown that the screened perturbation \( \partial n_{\alpha,q,\nu}^{KS} \) described in this section is also the key ingredient required for calculating the interatomic force constants in Eq. (13). As a practical consequence, every software implementation that supports DFPT calculations already contains all the information necessary for evaluating the electron-phonon matrix elements \( g_{nm\nu}(k, q) \).

All the quantities introduced in this section can equivalently be calculated using an alternative, variational formulation of density-functional perturbation theory. A thorough discussion of the connection between the Sternheimer approach and the variational approach to DFPT is given in Ref. [103]. The second-order matrix elements \( g_{mn,\nu,\nu'}^{DW}(k, q, q') \) given by Eq. (40) involve the second derivative of the Kohn-Sham potential with respect to the nuclear displacements. The evaluation of these quantities would require the solution of second-order Sternheimer equations for the second variations of the Kohn-Sham wavefunctions. The general structure of second-order Sternheimer equations can be found in Sec. IV.H of [103]. Since these calculations are rather involved, most practical implementations employ an approximation whereby the Debye-Waller matrix elements are expressed in terms of products of the standard matrix elements \( g_{mn\nu}(k, q) \). Such an alternative formulation was developed by Allen and Heine and Allen and Cardona, and will be discussed in Sec. 9.1.1.1. All recent ab initio calculations of electron-phonon interactions based on DFPT employed this latter approach.

### 3.2.4 The dielectric approach

Besides the DFPT method described in the previous section, it is also possible to calculate the screened perturbation \( \partial n_{\alpha,q}^{KS} \) using the so-called ‘dielectric approach.’ This latter approach did not find as widespread an application as those of Refs. [3][5], but it is useful to
establish a link between DFT calculations of electron-phonon matrix elements and the field-theoretic formulation to be discussed in Sec. 4.

For consistency with Sec. 3.2.3, we derive the key expressions of the dielectric approach starting from DFPT. To this aim we expand the variation of the wavefunction \( \partial \tilde{u}_{\nu k, q} \) using the complete set of states \( u_{nk+q} \) (with \( n \) referring to both occupied and empty Kohn-Sham states). Then we replace this expansion inside Eq. (48), project onto an arbitrary conduction state, and insert the result in Eq. (49). After taking into account time-reversal symmetry, these steps lead to the following result:

\[
\partial \kappa_{\alpha, q} n(r) = \int d \mathbf{r}' \chi^0_{q}(r, r') \partial \kappa_{\alpha, q} v_{KS}^{KS}(\mathbf{r}'),
\]

having defined:

\[
\chi^0_{q}(r, r') = N_p^{-1} \sum_{mnk} \left( f_{nk} - f_{mk+q} \right) \varepsilon_{nk} u^*_n(r) u_m(r) u^*_m(r+q) u_n(r+q). \tag{51}
\]

In this expression \( f_{nk} \) and \( f_{mk+q} \) are the occupations of each state, and the indices run over all bands. A factor of 2 for the spin degeneracy is implicitly included in the sum over \( k \). The quantity \( \chi^0_{q} \) in Eq. (51) is the lattice-periodic component for the wavevector \( q \) of the ‘independent-electron polarizability’ [113–116].

For ease of notation we can write Eq. (50) in symbolic form as \( \partial n = \chi^0 \partial v^{KS} \). Using the same symbolic notation it is also possible to formally rewrite Eqs. (42), (44), and (45) as follows:

\[
\partial v^{KS} = \partial v^{en} + (v^{C} + f^{xc}) \chi^0 \partial v^{KS},
\]

from which one obtains:

\[
\partial v^{KS} = (\epsilon^{Hxc})^{-1} \partial v^{en}, \tag{53}
\]

having defined the dielectric matrix:

\[
\epsilon^{Hxc} = 1 - (v^{C} + f^{xc}) \chi^0. \tag{54}
\]

The superscript ‘Hxc’ refers to the Hartree and exchange and correlation components of the screening. In the language of many-body perturbation theory \( \epsilon^{Hxc} \) is referred to as the ‘test electron’ dielectric matrix, hinting at the fact that the electron density redistribution in response to a perturbation arises both from classical electrostatics (the Hartree term \( v^{C} \chi^0 \)) and from quantum effects (the exchange and correlation term \( f^{xc} \chi^0 \)). If we neglect the kernel \( f^{xc} \) in this expression, then we obtain the ‘test charge’ dielectric matrix, which is most commonly known as the dielectric matrix in the random-phase approximation (RPA) [117]:

\[
\epsilon^{H} = 1 - v^{C} \chi^0. \tag{55}
\]

The symbolic expressions outlined here remain almost unchanged when using a reciprocal-space representation. As an example, Eq. (55) becomes simply:

\[
\epsilon^{H}_{GG'}(q) = \delta_{GG'} - \Omega^2 \sum_{G''} \chi^0_{G''G'}(q) v^{C}(q+G) \delta_{GG''}. \tag{56}
\]
Taken together Eqs. (38) and (53) show that the calculation of electron-phonon matrix elements using DFPT is equivalent to screening the bare electron-nucleus interaction using $\epsilon_{Hxc}$; in this case we say that the screening is described at the ‘RPA+xc’ level of approximation.

At this point it is worth to point out that so far we only considered the screening of static perturbations: in fact $\partial v^{en}$ was implicitly taken to be frequency-independent. Physically this choice corresponds to describing phonons as quasi-static perturbations, so that at each set of instantaneous atomic positions during a vibration cycle, the electrons have enough time to re-adjust and reach their ground state. This is a statement of the adiabatic approximation [118]. The importance of retardation effects in the electron-phonon problem was already recognized in the early work of Bardeen and Pines [18], but the first \textit{ab initio} calculations of these effects appeared much later [119, 119]. The formal framework required to incorporate retardation in the study of EPIs will be presented in Sec. 4.

3.2.5 Connection with early formulations

For completeness, we illustrate the link between electron-phonon matrix elements obtained within DFPT (Sec. 3.2.3) and the early approaches of Refs. [11, 17] (Sec. 2.1).

The Bloch matrix element can be derived as follows. We assume that the scattering potential is unscreened and corresponds to the bare pseudopotentials $V_\kappa$ in Eq. (43); that there is only one atom at the origin of the unit cell; and the Kohn-Sham wavefunctions can be approximated by free electrons, $u_{nk}(r) = \Omega^{-\frac{1}{2}} \exp(iG_n \cdot r)$. In the last expression, the subscript in $G_n$ is used in order to stress the one-to-one correspondence between the reciprocal lattice vectors and the energy bands of the free electron gas in the reduced zone scheme. Using these approximations in Eqs. (21), (34), (38), and (43), we find:

$$g_{mn\nu}(k, q) = -i \left[ \hbar / (2N_p M_\kappa \omega_{q\nu}) \right]^{\frac{1}{2}} V_\kappa(q + G_m - G_n)(q + G_m - G_n) \cdot e_{\kappa,\nu}(q).$$

(57)

By further neglecting umklapp processes ($G_m \neq G_n$) the previous result becomes (see Sec. 3.4 of Ref. [6]):

$$g_{mn\nu}(k, q) = -i \left[ \hbar / (2N_p M_\kappa \omega_{q\nu}) \right]^{\frac{1}{2}} q \cdot e_{\kappa,\nu}(q)V_\kappa(q).$$

(58)

The expression obtained by Bloch [11] and reproduced in Eq. (2) is simply obtained by replacing $V_\kappa(q)$ with the effective potential $V_0$. The Bardeen matrix element is more elaborate and can be derived as follows. We describe the screening of the bare ionic potential within the RPA approximation, and determine the dielectric matrix by replacing the Kohn-Sham wavefunctions by free electrons. Using $u_{nk}(r) = \Omega^{-\frac{1}{2}} \exp(iG_n \cdot r)$ and $\varepsilon_{nk} = \hbar^2(k + G_n)^2 / 2m_e - \varepsilon_F$ in Eq. (51), the polarizability reduces to:

$$\chi_{GG'}^\nu(q) = -m_e k_F \pi \hbar^2 \Omega \left[ \left| q + G \right| / 2k_F \right] \delta_{GG'},$$

(59)

where $F$ is the function defined below Eq. (4). The derivation of this result requires making the transition from the first Brillouin zone to the extended zone scheme. If we use Eq. (59)
inside Eq. (56), neglect the exchange and correlation kernel, and use the Fourier transform of the Coulomb potential, we find:

$$\epsilon_{GG'}(q) = \delta_{GG'} \left[ 1 + \left( \frac{k_{TF}^2}{|q+G|^2} \right) F(|q+G|/2k_F) \right],$$  \hspace{1cm} (60)

where the Thomas-Fermi screening length is given by $k_{TF} = \left[ \frac{4me^2k_F/(4\pi\varepsilon_0\pi\hbar^2)}{4} \right]^{1/2}$. Equation (60) is the well-known Lindhard dielectric matrix, and the diagonal matrix elements are the same as in Eq. (4) (see Refs. [9, 120] for in-depth discussions of the Lindhard function). By following the same steps that led to Eq. (58), replacing the bare ionic potential by its screened counterpart, and using Eq. (53) with $\epsilon$ instead of $\epsilon_{Hxc}$, we obtain:

$$g_{mn\nu}(k, q) = -i \frac{\hbar}{2NpM_\kappa \omega_{\kappa\nu}} \frac{1}{2} q \cdot e_{\kappa,\nu}(q) V_{\kappa}(q)/\epsilon(q).$$  \hspace{1cm} (61)

Here we considered only one atom at the center of the unit cell, and we neglected umklapp processes. This is essentially the result derived by Bardeen [17] and reproduced in Eq. (3).

4 Field-theoretic approach to the electron-phonon interaction

In Sec. 3 we discussed how the materials parameters entering the electron-phonon Hamiltonian in Eq. (1), namely $\epsilon_{n\kappa}$, $\omega_{\kappa\nu}$, and $g_{mn\nu}(k, q)$, can be calculated from first principles using DFT and DFPT. Today the formalism and techniques described in Sec. 3 constitute de facto the standard tool in quantitative studies of electron-phonon interactions in solids (see Secs. 7-12).

However, it should be noted that the DFT approach to EPIs does not rest on strong theoretical foundations. For one, the evaluation of the EPI matrix elements via Eq. (38) relies on the assumption that the interaction between electrons and nuclei is governed by the effective Kohn-Sham potential; therefore we can expect the matrix elements to be sensitive to the exchange and correlation functional (see Sec. 12). Furthermore, the very definition of phonons starting from Eq. (12) relies on the Born-Oppenheimer approximation, and one might ask whether this choice is accurate enough in metals and narrow-gap semiconductors (see Sec. 7). Finally, if one were to go beyond the Born-Oppenheimer approximation, then it would seem sensible to also incorporate retardation effects in the calculation of the EPI matrix elements.

On top of these practical points, and at a more fundamental level, we expect that the electron-phonon interaction will modify both the electronic structure and the lattice dynamics of a solid, and these modifications will in turn affect the coupling between electrons and phonons. It is therefore clear that a complete theory of interacting electrons and phonons must be self-consistent. In order to address these issues it is necessary to formulate the electron-phonon problem using a rigorous and general theory of interacting electrons and phonons in solids.

The most systematic and elegant approach is based on quantum field theory [121], and is tightly connected with the development of the GW method [122]. The first attempts in this direction were from Nakajima [20], Bardeen and Pines [18], Migdal [123], and Engelsberg and Schrieffer [124]. However, from the point of view of the present article, these works are of limited usefulness since they were mostly developed around the homogeneous electron gas.
A completely general formulation of the problem, which seamlessly applies to metals, semiconductors, and insulators, was first provided by Baym [125] and subsequently by Hedin and Lundqvist [126]. The formalism developed in these articles constitutes today the most complete theory of the electron-phonon problem. In fact, many aspects of this formalism are yet to be explored within the context of \textit{ab initio} calculations. After these seminal works several authors contributed to clarifying various aspects of the many-body theory of the coupled electron-phonon system, including Keating [127], Gillis [128], Sjölander and Johnson [129], Maksimov [130], Vogl [131], and more recently van Leeuwen [132] and Marini \textit{et al.} [133]. In particular, Ref. [132] focused on the issues of translational and rotational invariance of the resulting theory, while Ref. [133] analyzed the connection between many-body perturbation theory approaches and DFT calculations.

Since the mathematical notation of the original articles is obsolete and rather difficult to follow, in Secs. 4.1-4.4 we cover the theory in some detail using contemporary notation. The following derivations can be found across Refs. [125, 126, 130, 134]. Here we provide a synthesis of these contributions using a unified notation, and we fill the gaps wherever it is necessary. The presentation requires some familiarity with field operators (see for example Ref. [84] for a succinct introduction).

### 4.1 Operators and distinguishability

The starting point for studying EPIs using a field-theoretic approach is to define the Fock space and the field operators for electrons and nuclei. In the case of electrons the choice is unambiguous, since any many-body state can be represented as a linear combination of Slater determinants constructed using a basis of single-particle wavefunctions. In the case of nuclei the situation is slightly more ambiguous: in principle we might proceed in a very general way by choosing to focus on the nuclei as our quantum particles, as opposed to their displacements from equilibrium. In practice this choice leads to a dead end for two reasons. Firstly, the quantum statistics of nuclei would be dependent on their spin, therefore we would end up with an unwieldy mix of fermions and bosons depending on the solid. Secondly, the notion of ‘indistinguishable’ particles, which is central to second quantization, does not apply to nuclei in solids (at least in thermodynamic equilibrium and far from a solid-liquid phase transition). In fact, in many cases we can directly label the nuclei, for example by means of experimental probes such as scanning tunneling microscopy and electron diffraction. In order to avoid these issues, it is best to study the electron-phonon problem by considering (i) \textit{indistinguishable} electrons, for which it is convenient to use second-quantized operators; (ii) \textit{distinguishable} nuclei, for which it is best to use first quantization in the displacements; (iii) \textit{indistinguishable} phonons, resulting from the quantization of the nuclear displacements; in this latter case the distinction between first and second quantization is irrelevant. These aspects are briefly mentioned by Baym [125] and Maksimov [130].

With these choices, the dynamical variables of the problem are the electronic field operators $\hat{\psi}$ (discussed below) and the nuclear displacements from equilibrium $\Delta \hat{\tau}$ (discussed in Sec. 4.3). In this theory the equilibrium coordinates of the nuclei are regarded as \textit{external} parameters, and are
to be obtained for example from crystallography or DFT calculations. Throughout this section, we limit ourselves to consider equilibrium Green’s functions at zero temperature. As a result, all expectation values will be evaluated for the electron-nuclei ground state $|0\rangle$. The extension of the main results to finite temperature is presented in Sec. 5. We will not specify how to obtain the ground state, since the following discussion is independent on the precise shape of this state. In order to derive expressions that are useful for first-principles calculations, at the very end the ground state will be approximated using standard DFT wavefunctions and phonons (see Sec. 5).

The electronic field creation/destruction operators are denoted by $\hat{\psi}^{\dagger}(x)/\hat{\psi}(x)$, where the variable $x$ indicates both the position $r$ and the spin label $\sigma$. These operators obey the anti-commutation relations [43]: $\{\hat{\psi}(x), \hat{\psi}^{\dagger}(x')\} = \{\hat{\psi}^{\dagger}(x), \hat{\psi}(x')\} = 0$, $\{\hat{\psi}(x), \hat{\psi}^{\dagger}(x')\} = \delta(x-x')$. The most general non-relativistic Hamiltonian for a system of coupled electrons and nuclei can be written as:

$$\hat{H} = \hat{T}_e + \hat{T}_n + \hat{U}_{ee} + \hat{U}_{nn} + \hat{U}_{en},$$

where each term will be introduced hereafter. The electron kinetic energy is:

$$\hat{T}_e = -\frac{\hbar^2}{2m_e} \int dx \hat{\psi}^{\dagger}(x) \nabla^2 \hat{\psi}(x),$$

with $m_e$ being the electron mass, and the integrals $\int dx$ denoting the sum over spin and the integration over space. The electron-electron interaction is:

$$\hat{U}_{ee} = \frac{1}{2} \int dr \int dr' \hat{n}_e(r) [\hat{n}_e(r') - \delta(r-r')] v(r,r'),$$

where the electron particle density operator is given by $\hat{n}_e(r) = \sum_\sigma \hat{\psi}^{\dagger}(x)\hat{\psi}(x)$, and where $v(r,r') = e^2/(4\pi\epsilon_0|\mathbf{r} - \mathbf{r}'|)$ is the Coulomb interaction between two particles of charge $e$. In Eqs. (63) and (64) the integrals are over the entire crystal. This corresponds to considering a supercell of infinite size (therefore the lattice vectors $\mathbf{T}$ of the supercell drop out) and a dense sampling of wavevectors $\mathbf{q}$ in the Brillouin zone. This choice is useful in order to maintain the formalism as light as possible. Accordingly, all sums over $\mathbf{q}$ are replaced using $N_p^{-1} \sum_{\mathbf{q}} \rightarrow \Omega_{BZ}^{-1} \int d\mathbf{q}$, where the integral is over the Brillouin zone of volume $\Omega_{BZ}$. Similarly the closure relations in Eq. (213) are replaced by: $\int d\mathbf{q} \exp(i\mathbf{q} \cdot \mathbf{R}_p) = \Omega_{BZ} \delta_{p0}$ and $\sum_{\mathbf{p}} \exp(i\mathbf{q} \cdot \mathbf{R}_p) = \Omega_{BZ} \delta(\mathbf{q})$. The nuclear kinetic energy operator is the same as the last term in Eq. (19). Using the same notation as in Sec. 3 the nucleus-nucleus interaction energy is:

$$\hat{U}_{nn} = \frac{1}{2} \sum_{\kappa\kappa'\neq\kappa\kappa'} Z_\kappa Z_{\kappa'} v(\tau_{\kappa p}^0 + \Delta\hat{\tau}_{\kappa p}, \tau_{\kappa' p'}^0 + \Delta\hat{\tau}_{\kappa' p'}).$$

Here $\tau_{\kappa p}^0$ denotes the classical equilibrium position of each nucleus, and the displacement operators $\Delta\hat{\tau}_{\kappa p}$ will later be expressed in terms of the ladder operators from Appendix B. The electron-nucleus interaction energy is:

$$\hat{U}_{en} = \int dr \int dr' \hat{n}_e(r)\hat{n}_n(r')v(r,r'),$$

where the nuclear charge density operator is given by:

$$\hat{n}_n(r) = -\sum_{\kappa p} Z_\kappa \delta(r - \tau_{\kappa p}^0 - \Delta\hat{\tau}_{\kappa p}).$$
Here the density operators are expressed in units of the electron charge, so that the expectation value of the total charge density is $-e\langle 0|\hat{n}(r)|0\rangle$ with $\hat{n}(r) = \hat{n}_e(r) + \hat{n}_n(r)$.

We underline the asymmetry between Eqs. (64) and (65): in the case of electrons one considers the electrostatic energy of a continuous distribution of charge, and the unphysical self-interaction is removed by the Dirac delta; whereas in the case of nuclei, the particles are distinguishable therefore one has to take into account all pairwise interactions individually.

4.2 Electron Green’s function

4.2.1 Equation of motion and self-energy

In this section we focus on the electrons. By combining Eqs. (62)-(66) and using the anti-commutation relations for the field operators one finds the standard expression:

$$\hat{H} = \hat{T}_n + \hat{U}_{nn} + \int d\mathbf{x} \hat{\psi}^\dagger(\mathbf{x}) \left[ -\frac{\hbar^2}{2m_e} \nabla^2 + \hat{V}_n(\mathbf{r}) \right] \hat{\psi}(\mathbf{x}) + \frac{1}{2} \int d\mathbf{x} d\mathbf{x}' v(\mathbf{r}, \mathbf{r}') \hat{\psi}^\dagger(\mathbf{x}) \hat{\psi}^\dagger(\mathbf{x}') \hat{\psi}(\mathbf{x}) \hat{\psi}(\mathbf{x}')$$

where the nuclear potential $\hat{V}_n$ is given by:

$$\hat{V}_n(\mathbf{r}) = \int d\mathbf{r}' v(\mathbf{r}, \mathbf{r}') \hat{n}_n(\mathbf{r}')$$

In order to study the excitation spectrum of the many-body Hamiltonian $\hat{H}$ at equilibrium we need to determine the time-ordered one-electron Green’s function [134, 135]. At zero temperature this function is defined as:

$$G(\mathbf{x}t, \mathbf{x}'t') = -\frac{i}{\hbar} \langle 0| \hat{T} \psi(\mathbf{xt}) \psi^\dagger(\mathbf{x}'t') |0\rangle,$$

where $\hat{T}$ is Wick’s time-ordering operator for fermions, and ensures that the times of the subsequent operators increase towards the left. The formal definition of the Wick operator is:

$$\hat{T} \psi(\mathbf{xt}) \psi^\dagger(\mathbf{x}'t') = \theta(t-t') \psi(\mathbf{xt}) \psi^\dagger(\mathbf{x}'t') - \theta(t'-t) \psi^\dagger(\mathbf{x}'t') \psi(\mathbf{xt}),$$

where $\theta$ is the Heaviside function. Based on this definition we see that for $t > t'$ the Green’s function in Eq. (70) corresponds to the scalar product between the initial state $\psi^\dagger(\mathbf{x}'t')|0\rangle$ and the final state $\psi^\dagger(\mathbf{xt})|0\rangle$. This product is precisely the probability amplitude for finding an electron in the position $\mathbf{x}$ at the time $t$, after having introduced an electron in $\mathbf{x}'$ at an earlier time $t'$. In the case $t < t'$ the situation is reversed and the Green’s function describes the propagation of a hole created in the system at the time $t'$.

In order to determine $G(\mathbf{x}t, \mathbf{x}'t')$ we need to establish an equation of motion for the field operators. This can be done by describing the time-dependence of the operators within the Heisenberg picture:

$$\dot{\psi}(\mathbf{xt}) = e^{i\hat{H}/\hbar} \dot{\psi}(\mathbf{x}) e^{-i\hat{H}/\hbar},$$

where $\hat{H}$ was defined in Eq. (68). From this definition it follows immediately:

$$i\hbar \frac{\partial}{\partial t} \hat{\psi}(\mathbf{xt}) = [\hat{\psi}(\mathbf{xt}), \hat{H}].$$
By combining Eqs. (68) and (72) and using the anti-commutation relations for the field operators one obtains:

$$i\hbar \frac{\partial}{\partial t} \hat{\psi}(x,t) = \left[ -\frac{\hbar^2}{2m_e} \nabla^2 + \int dr' v(r,r') \hat{n}(r') \right] \hat{\psi}(x,t), \tag{73}$$

where the time-dependence in $\hat{n}(r' t)$ is to be understood in the Heisenberg sense, as in Eq. (71). This equation of motion allows us to write the corresponding equation for the electron Green’s function in Eq. (70):

$$\left[ i\hbar \frac{\partial}{\partial t} + \frac{\hbar^2}{2m_e} \nabla^2 - \varphi(r t) \right] G(x t, x' t') = \delta(x t, x' t') - \frac{i}{\hbar} \int dr'' dt'' v(r t, r'' t'') \langle \hat{T} \hat{n}(r'' t'') \hat{\psi}(x t) \hat{\psi}^+(x' t') \rangle. \tag{74}$$

Here $v(r t, r'' t'') = v(r, r'') \delta(t-t'')$, the brackets $\langle \cdots \rangle$ are a short-hand notation for $\langle 0 | \cdots | 0 \rangle$, and the additional term $\varphi$ is discussed below. In order to obtain Eq. (74) we used once again the anti-commutation relations, and we noted that the derivative of the Heaviside function is a Dirac delta.

The new term $\varphi(r t)$ which appeared in Eq. (74) is a scalar electric potential which couples to both electronic and nuclear charges. This potential has been introduced in order to perturb the system via the additional Hamiltonian $\tilde{H}_1(t) = \int dr \hat{n}(r t) \varphi( r t )$. The physical idea behind this choice is to use $\varphi( r t )$ in order to induce forced oscillations in the system. When the system resonates with the perturbation we know that the resonant frequency must correspond to a free oscillation, that is a many-body eigenmode. From a formal point of view, the potential $\varphi( r t )$ is introduced in order to exploit Schwinger’s functional derivative technique (Ref. 134, Appendix II) and is set to zero at the end of the derivation.

One complication arising from the introduction of $\varphi( r t )$ in Eq. (74) is that the time evolution in Eq. (71) is no longer valid, since the perturbed Hamiltonian now depends on the time variable. The way around this complication is to switch from the Heisenberg picture to the interaction picture. This change amounts to replacing the exponentials in Eq. (71) by the time-ordered Dyson series $\tilde{U}(t) = \hat{T} \exp \left( -i \hbar^{-1} \int_0^t \hat{H}(t') dt' \right)$ (135, p. 57). Since this would lead to an overlong derivation, we prefer to leave this aspect aside and refer the reader to Ref. 136 for a more comprehensive discussion.

In order to write Eq. (74) in a manageable form, we use the identity [134]:

$$\frac{\delta \langle \hat{T} \hat{a}(t_1) \hat{b}(t_2) \rangle}{\delta \varphi(r'' t'')} = -\frac{i}{\hbar} \langle \hat{T} \left[ \hat{n}(r'' t'') - \langle \hat{n}(r'' t'') \rangle \right] \hat{a}(t_1) \hat{b}(t_2) \rangle. \tag{75}$$

In this and the following expressions $\delta/\delta \varphi(r'' t'')$ denotes the functional derivative with respect to $\varphi(r'' t'')$, and should not be confused with the Dirac delta functions $\delta(x t, x' t')$. Equation (75) is proven in Refs. 134 (Appendix II) and 126 (Appendix B.a). After identifying $\hat{a}$ and $\hat{b}$ with $\hat{\psi}(x t)$ and $\hat{\psi}^+(x' t')$, respectively, Eq. (74) becomes:

$$\left[ i\hbar \frac{\partial}{\partial t} + \frac{\hbar^2}{2m_e} \nabla^2 - V_{\text{tot}}(r t) - i\hbar \int dr'' dt'' v(r t, r'' t'') \frac{\delta}{\delta \varphi(r'' t'')} \right] G(x t, x' t') = \delta(x t, x' t'), \tag{76}$$

where $\eta$ is a positive infinitesimal arising from the time-ordering, and $V_{\text{tot}}(r t)$ is the total potential acting on the electronic and nuclear charges, averaged over the many-body quantum state $| 0 \rangle$:

$$V_{\text{tot}}(r t) = \int dr' v( r, r' ) \langle \hat{n}(r' t) \rangle + \varphi( r t ). \tag{77}$$
Equation (76) was first derived by Kato et al. In order to avoid a proliferation of variables, it is common practice to replace the letters by integer numbers, using the convention \((xt) \rightarrow 1, (x't') \rightarrow 2, (rt + \eta) \rightarrow 1^+\), and so on. Using this convention the last two equations become:

\[
\left[ i\hbar \frac{\partial}{\partial t_1} + \frac{\hbar^2}{2m_e} \nabla^2(1) - V_{\text{tot}}(1) - i\hbar \int d3 \, v(1+3) \frac{\delta}{\delta \varphi(3)} \right] G(12) = \delta(12),
\]

and

\[
V_{\text{tot}}(1) = \int d2 \, v(12) (\hat{n}(2)) + \varphi(1).
\]

In these expressions the spin labels are implied for the Green’s function and for the Dirac delta.

At this point, a set of self-consistent equations for coupled electrons and phonons can be generated by eliminating the functional derivative in Eq. (78). For this purpose one first relates the total screened electrostatic potential \(V_{\text{tot}}\) to the external potential \(\varphi\) by introducing the inverse dielectric matrix \(\epsilon^{-1}\) as a functional derivative:

\[
\epsilon^{-1}(12) = \frac{\delta V_{\text{tot}}(1)}{\delta \varphi(2)}.
\]

The function \(\epsilon^{-1}(12)\) is the many-body counterpart of the dielectric matrix discussed in Sec. 3.2.4. The form given by Eq. (80) is the most general field-theoretic formulation for a system of interacting electrons and nuclei.

The next step is to rewrite \(\delta G/\delta \varphi\) inside Eq. (78) in terms of the inverse Green’s function, \(G^{-1}\). By using the fact that \(\delta \int d2 \, G(12) \, G^{-1}(23) = 0\) and the rule for the functional derivative of a product \([137]\) one obtains:

\[
\frac{\delta G(12)}{\delta \varphi(3)} = -\int d(45) G(14) \frac{\delta G^{-1}(45)}{\delta \varphi(3)} G(52).
\]

In order to eliminate any explicit reference to \(\varphi\) we can express the functional derivative on the right-hand side using the chain rule for functional differentiation \([137]\):

\[
\frac{\delta G^{-1}(45)}{\delta \varphi(3)} = \int d6 \frac{\delta G^{-1}(45)}{\delta V_{\text{tot}}(6)} \frac{\delta V_{\text{tot}}(6)}{\delta \varphi(3)}.
\]

It is customary to call ‘vertex’ the three-point quantity defined by:

\[
\Gamma(123) = -\delta G^{-1}(12) / \delta V_{\text{tot}}(3).
\]

By combining Eqs. (78) and (80)-(83) one finds:

\[
\left[ i\hbar \frac{\partial}{\partial t_1} + \frac{\hbar^2}{2m_e} \nabla^2(1) - V_{\text{tot}}(1) \right] G(12) - \int d3 \, \Sigma(13) \, G(32) = \delta(12),
\]

having introduced the so-called ‘electron self-energy’ \(\Sigma\):

\[
\Sigma(12) = i\hbar \int d(34) G(13) \Gamma(324) W(41^+),
\]

which in turn contains the ‘screened Coulomb interaction’ \(W\), defined as:

\[
W(12) = \int d3 \, \epsilon^{-1}(13) \, v(32) = \int d(3) \, v(13) \epsilon^{-1}(23).
\]
The last equality can be obtained by observing that $\delta \langle \hat{n}(1) \rangle / \delta \varphi(2) = \delta \langle \hat{n}(2) \rangle / \delta \varphi(1)$ after Eq. (75), therefore $W(12) = W(21)$.

Now, by inverting Eq. (84) and using Eq. (83), we can express the vertex $\Gamma$ in terms of $\Sigma$ and $G$:

$$\Gamma(123) = \delta(12)\delta(13) + \int d(4567) \frac{\delta \Sigma(12)}{\delta G(45)} G(46)G(75)\Gamma(673). \quad (87)$$

The derivation of this result is rather lengthy: it requires the use of the chain rule, in symbols $\delta/\delta V_{\text{tot}} = (\delta G/\delta V_{\text{tot}}) \delta/\delta G$, as well as Eq. (81) with $\Sigma$ and $V_{\text{tot}}$ instead of $G$ and $\varphi$, respectively. Equations (84)-(87) form a nonlinear system of equations for the electron Green’s function, $G$, the electron self-energy, $\Sigma$, the total screened Coulomb interaction, $W$, and the vertex, $\Gamma$. In order to close the loop it remains to specify the relation between $W$ and the other quantities. The next section is devoted to this aspect.

### 4.2.2 The screened Coulomb interaction

An equation for the screened Coulomb interaction can be found by combining Eqs. (79), (80), and (86):

$$W(12) = v(12) + \int d(34) v(13) \frac{\delta \langle \hat{n}(3) \rangle}{\delta V_{\text{tot}}(4)} W(42). \quad (88)$$

By defining the ‘polarization propagator’ as:

$$P(12) = \frac{\delta \langle \hat{n}(1) \rangle}{\delta V_{\text{tot}}(2)}, \quad (89)$$

the previous expression takes the usual form [122]:

$$W(12) = v(12) + \int d(34) v(13) P(34)W(42). \quad (90)$$

This result can be combined with Eq. (86) in order to express the dielectric matrix in terms of the polarization:

$$\epsilon(12) = \delta(12) - \int d(3) v(13) P(32). \quad (91)$$

We now consider the special case whereby the nuclei are regarded as classical point charges clamped to their equilibrium positions. In this situation, the variation of the charge density $\delta \langle \hat{n} \rangle$ in Eq. (89) corresponds to the re-distribution of the electronic charge in response to the perturbation $\delta V_{\text{tot}}$. In order to describe this special case it is convenient to introduce a new polarization propagator, $P_e$, associated with the electronic response only:

$$P_e(12) = \frac{\delta \langle \hat{n}_e(1) \rangle}{\delta V_{\text{tot}}(2)} = -i\hbar \sum_{\sigma_1} \int d(34) G(13)G(41^+)\Gamma(342). \quad (92)$$

The last equality in this expression is obtained by using Eq. (81) with $V_{\text{tot}}$ instead of $\varphi$, together with Eq. (83), and by considering that the electron density is related to the Green’s function via the relation:

$$\langle \hat{n}_e(1) \rangle = -i\hbar \sum_{\sigma_1} G(11^+). \quad (93)$$

30
In conjunction with $P_e$ it is natural to define the Coulomb interaction screened by the electronic polarization only:

$$W_e(12) = v(12) + \int d(34) v(13) P_e(34) W_e(42),$$

(94)
as well as the associated dielectric matrix, in analogy with Eq. (91):

$$\epsilon_e(12) = \delta(12) - \int d3 v(13) P_e(32).$$

(95)

Taken together, Eqs. (84)-(87) with $W$ replaced by $W_e$ constitute the well-known Hedin’s equations for a system of interacting electrons when the nuclei are clamped to their equilibrium positions [122].

In order to go back to the most general case whereby the nuclei are not clamped to their equilibrium positions, one has to describe the re-adjustment of both electronic and nuclear charge. To this aim we combine Eq. (80), (86), (88), (92), (94), and (95). The result is:

$$W(12) = W_e(12) + \int d(34) W_e(13) \frac{\delta \langle \hat{n}_n(3) \rangle}{\delta \varphi(4)} v(42).$$

(96)

An explicit expression for $\delta \langle \hat{n}_n \rangle / \delta \varphi$ can be obtained using the following reasoning. We go into the details since this is a delicate passage. Equation (75) provides a recipe for evaluating the variation of any operator with respect to a potential $\varphi(rt)$ which couples to the total charge density operator $\hat{n}(rt)$ via $\hat{H}_1(t) = \int dr \hat{n}(rt) \varphi(rt)$. Therefore we can replace $\hat{a} \hat{b}$ in Eq. (75) by $\hat{n}_n$ to obtain:

$$\frac{\delta \langle \hat{n}_n(1) \rangle}{\delta \varphi(2)} = - \frac{i}{\hbar} \langle \hat{T}[\hat{n}_n(2) - \langle \hat{n}_n(2) \rangle][\hat{n}_n(1) - \langle \hat{n}_n(1) \rangle] \rangle.$$  

(97)

In addition, if we introduce a second perturbation, $\hat{H}_2(t) = \int dr \hat{n}_n(rt) J(rt)$, which couples only to the nuclear charges, we can repeat the same reasoning as in Eq. (97) after replacing $\varphi$ by $J$ and $\hat{n}$ by $\hat{n}_n$:

$$\frac{\delta \langle \hat{n}_n(1) \rangle}{\delta J(2)} = - \frac{i}{\hbar} \langle \hat{T}[\hat{n}_n(2) - \langle \hat{n}_n(2) \rangle][\hat{n}_n(1) - \langle \hat{n}_n(1) \rangle] \rangle.$$  

(98)

The comparison between Eqs. (97) and (98) yields:

$$\frac{\delta \langle \hat{n}_n(1) \rangle}{\delta \varphi(2)} = \frac{\delta \langle \hat{n}_n(1) \rangle}{\delta J(1)}.$$  

(99)

This can be restated by using the chain rule, $\delta \langle \hat{n}_n \rangle / \delta J = \delta \langle \hat{n}_n \rangle / \delta V_{tot} \times \delta V_{tot} / \delta \langle n \rangle \times \delta \langle n \rangle / \delta J$:

$$\frac{\delta \langle \hat{n}_n(1) \rangle}{\delta J(2)} = \int d3 \epsilon^{-1}_e(34) \frac{\delta \langle \hat{n}_n(3) \rangle}{\delta J(2)}. $$  

(100)

The variation $\delta \langle \hat{n}_n \rangle / \delta J$ on the right-hand side can be expressed as in Eq. (98):

$$\frac{\delta \langle \hat{n}_n(1) \rangle}{\delta J(2)} = \frac{1}{\hbar} \langle \hat{T}[\hat{n}_n(2) - \langle \hat{n}_n(2) \rangle][\hat{n}_n(1)] \rangle,$$  

(101)

and since $\langle \hat{n}_n - \langle \hat{n}_n \rangle \rangle = 0$ this can also be rewritten as:

$$\delta \langle \hat{n}_n(1) \rangle / \delta J(2) = D(21), $$  

(102)
having defined:

\[ D(12) = -\frac{i}{\hbar} \langle \hat{T} \{ \hat{n}_n(1) - \langle \hat{n}_n(1) \rangle \} [\hat{n}_n(2) - \langle \hat{n}_n(2) \rangle] \rangle. \] (103)

This quantity is called the ‘density-density correlation function’ for the nuclei. Finally, we can combine Eqs. (96), (99), (100), and (103) to obtain:

\[ W(12) = W_e(12) + W_{ph}(12), \] (104)

where \( W_{ph} \) is the nuclear contribution to the screened Coulomb interaction, and is given by:

\[ W_{ph}(12) = \int d(34) W_e(13) D(34) W_e(24). \] (105)

This important result was derived first by Hedin and Lundqvist [126].

4.2.3 Nuclear contribution to the screened Coulomb interaction

In view of the forthcoming discussion, it is useful to derive a more explicit expression for the screened interaction \( W_{ph} \) in Eq. (105). Here we follow Refs. [125, 130]. The Taylor expansion of the Dirac delta to second order in the displacement \( \mathbf{u} \) reads:

\[ \delta(\mathbf{r} - \mathbf{u}) = \delta(\mathbf{r}) - \mathbf{u} \cdot \nabla \delta(\mathbf{r}) + \frac{1}{2} \mathbf{u} \cdot \nabla \mathbf{u} \cdot \nabla \delta(\mathbf{r}) \cdot \mathbf{u}, \] (106)

where \( \mathbf{u} \cdot \nabla \mathbf{u} \cdot \nabla \mathbf{u} \) is a short-hand notation for the second-order derivative \( \sum_{\alpha \alpha'} u_\alpha u_{\alpha'} \nabla_\alpha \nabla_{\alpha'} \). The above expression derives from the Fourier representation of the Dirac delta. Using Eq. (106) inside (67) we deduce:

\[ \hat{n}_n(\mathbf{r}) = n^0_n(\mathbf{r}) + \sum_{\kappa p} Z_\kappa \Delta \hat{\tau}_{\kappa p} \cdot \nabla \delta(\mathbf{r} - \tau^0_{\kappa p}) - \frac{1}{2} \sum_{\kappa p} Z_\kappa \Delta \hat{\tau}_{\kappa p} \cdot \nabla \nabla \delta(\mathbf{r} - \tau^0_{\kappa p}) \cdot \Delta \hat{\tau}_{\kappa p}, \] (107)

where \( n^0_n(\mathbf{r}) \) is the density of nuclear point charges at the classical equilibrium positions \( \tau^0_{\kappa p} \). After taking into account the time evolution in the Heisenberg picture as in Eq. (71), we can replace this expansion inside Eq. (103) to obtain:

\[ D(12) = \sum_{\kappa = \kappa'} \sum_{\alpha = \alpha'} Z_\kappa \nabla_{1,\alpha} \delta(\mathbf{r}_1 - \tau^0_{\kappa p}) D_{\kappa \alpha \kappa' \alpha'}(t_1 t_2) Z_{\kappa'} \nabla_{2,\alpha'} \delta(\mathbf{r}_2 - \tau^0_{\kappa' p}). \] (108)

On the right-hand side we introduced the ‘displacement-displacement correlation function’:

\[ D_{\kappa \alpha \kappa' \alpha'}(t t') = -\frac{i}{\hbar} \langle \hat{T} \Delta \hat{\tau}_{\kappa \alpha}(t) \Delta \hat{\tau}_{\kappa' \alpha'}(t') \rangle. \] (109)

If we insert the last two equations in Eq. (105) we find:

\[ W_{ph}(12) = \sum_{\kappa = \kappa'} \int d(34) \epsilon^{-1}_e(13) \nabla_{3,\alpha} V_\kappa(\mathbf{r}_3 - \tau^0_{\kappa p}) D_{\kappa \alpha \kappa' \alpha'}(t_3 t_4) \epsilon^{-1}_e(24) \nabla_{4,\alpha'} V_{\kappa'}(\mathbf{r}_4 - \tau^0_{\kappa' p}). \] (110)

In this expression \( V_\kappa \) is the bare Coulomb potential of a nucleus or its ionic pseudo-potential. At this point of the derivation, Hedin and Lundqvist introduce the approximation that the electronic dielectric matrix in Eq. (110) can be replaced by its static counterpart [126]. This choice
implies the Born-Oppenheimer adiabatic approximation. In view of maintaining the formalism as general as possible, we prefer to keep retardation effects, following the earlier works by Bardeen and Pines [18] and Baym [125]. We will come back to this aspect in Secs. 5.1.3 and 5.2.2.

We stress that the sole approximation used until this point is to truncate the density operator for the nuclei to the second order in the atomic displacements. This is the standard harmonic approximation. Apart from this approximation, which is useful to express $W_{\text{ph}}$ in a tractable form, no other assumptions are made. Gillis proposed a generalization of the results of Ref. [125] which does not use the harmonic approximation [128]. However, the resulting formalism is exceedingly complex, and has not been followed up.

4.3 Phonon Green’s function

In order to complete the set of self-consistent many-body equations for the coupled electron-phonon system, it remains to specify a prescription for calculating the displacement-displacement correlation function, $D_{\kappa \alpha p, \kappa' \alpha' p'}(t, t')$. This function is seldom referred to as the ‘phonon Green’s function’, although strictly speaking this name should be reserved for the quantity $-(i/\hbar)\langle \hat{T} \hat{a}_{q\nu}(t) \hat{a}_{q'\nu}'(t') \rangle$ which will be discussed in Sec. 5.1.1. In the following we describe the procedure originally devised by Baym [125], and subsequently analyzed by Keating [127], Hedin and Lundqvist [126], Gillis [128], and Maksimov [130].

The starting point is the equation of motion for the displacement operators $\Delta \hat{\tau}_{\kappa \alpha p}(t)$. In analogy with Eq. (72) we have: $i\hbar \partial \partial t \Delta \hat{\tau}_{\kappa \alpha p}(t) = [\Delta \hat{\tau}_{\kappa \alpha p}(t), \hat{H}]$. Since we are considering the harmonic approximation and we expect the nuclei to oscillate around their equilibrium positions, it is convenient to aim for an expression resembling Newton’s equation. This can be done by taking the time-derivative of the equation of motion:

$$M_{\kappa} \frac{\partial^2}{\partial t^2} \Delta \hat{\tau}_{\kappa \alpha p}(t) = -\frac{M_{\kappa}}{\hbar^2} [[\Delta \hat{\tau}_{\kappa \alpha p}(t), \hat{H}], \hat{H}].$$

(111)

After evaluating the commutators using Eqs. (62)-(66) and performing the derivatives with respect to the nuclear displacements by means of Eq. (106), we obtain (the steps are laborious but straightforward):

$$M_{\kappa} \frac{\partial^2}{\partial t^2} \Delta \hat{\tau}_{\kappa \alpha p}(t) = Z_{\kappa} \int d\mathbf{r} d\mathbf{r}' \hat{n}_{\kappa \alpha p}(r, \mathbf{r}) v(\mathbf{r}, \mathbf{r}') \left\{ \nabla' \delta(\mathbf{r}' - \mathbf{r}_{0,\kappa \alpha p}) + \nabla' \left[ \nabla' \delta(\mathbf{r}' - \mathbf{r}_{0,\kappa \alpha p}) \cdot \Delta \hat{\tau}_{\kappa \alpha p}(t) \right] \right\}.$$  

(112)

Here $\hat{n}_{\kappa \alpha p}(\mathbf{r})$ is the total charge density of electrons and nuclei, except for the contribution of the nucleus $\kappa$ in the unit cell $p$. In the second line $\nabla'$ indicates that the derivatives are taken with respect to the variable $\mathbf{r}'$. At this point, we can use the functional derivative technique as in Sec. 4.2 in order to determine an expression involving the displacement-displacement correlation function from Eq. (109). Here, instead of using $J(\mathbf{r})$ as in Sec. 4.2.2 for the nuclear density, it is convenient to work with the individual displacements, and introduce a third perturbation $\hat{H}_3(t) = \sum_{\kappa \alpha p} \mathbf{F}_{\kappa \alpha p}(t) \cdot \Delta \hat{\tau}_{\kappa \alpha p}(t)$. The extra terms $\mathbf{F}_{\kappa \alpha p}(t)$ have the meaning of external forces acting on the nuclei. Using this perturbation, it is possible to write the displacement-displacement
correlation function in a manner similar to Eq. (101):

$$\frac{\delta \langle \Delta \hat{r}_{\kappa\alpha p}(t) \rangle}{\delta F_{\kappa',\alpha'p'}(t')} = D_{\kappa\alpha p,\kappa',\alpha'p'}(tt').$$

(113)

This result was derived by Baym [125] using a finite-temperature formalism. As in the case of the electron Green’s function in Sec. 4.2, it can only be obtained by working in the interaction picture, and by taking into account the explicit time-dependence of the Hamiltonian $\hat{H} + \hat{H}_3(t)$. Also in the present case, we omit these details for the sake of conciseness.

If we take the expectation value of Eq. (112) in the ground state, after having added the new force term $-F_{\kappa p}(t)$, and carry out the functional derivative with respect to $F_{\kappa'p'}(t')$, we obtain:

$$M_{\kappa} \frac{\partial^2}{\partial t^2} D_{\kappa\alpha p,\kappa',\alpha'p'}(tt') = -\delta_{\kappa\alpha p,\kappa',\alpha'p'} \delta(tt') + Z_{\kappa} \int dt \int dr' \left[ -\frac{\delta \langle \hat{n}_{\kappa p}(t') \rangle}{\delta F_{\kappa',\alpha'p'}(t')} v(r, r') \nabla_{\alpha} \delta(r' - \tau_{\kappa p}^0) + \langle \hat{n}_{\kappa p}(t) \rangle v(r, r') \nabla_{\alpha} \nabla_{\gamma} \delta(r' - \tau_{\kappa p}^0) D_{\kappa\gamma p,\kappa',\alpha'p'}(tt') \right],$$

(114)

where the sum over the Cartesian directions $\gamma$ is implied. The derivation of this result is rather cumbersome and involves the following considerations: (i) the Dirac deltas in the first line come from the force terms $-F_{\kappa p}(t)$ added to the Hamiltonian; (ii) within the Harmonic approximation the expectation value $\langle \hat{n}_{\kappa p} \Delta \hat{r}_{\kappa p} \rangle$ can be replaced by $\langle \hat{n}_{\kappa p} \rangle \langle \Delta \hat{r}_{\kappa p} \rangle$ [125]; (iii) the expectation value $\langle \Delta \hat{r}_{\kappa p} \rangle$ can be set to zero, because at the end of the derivation one sets $|F_{\kappa p}| = 0$ hence the expectation values of the displacements vanish. We note that Hedin and Lundqvist omitted the last line of Eq. (114) in their derivation [126], but this term was correctly included by Baym [125] and Maksimov [130].

The remaining functional derivative in Eq. (114) can be expressed in terms of the nuclear charge density using the same strategy which led to Eq. (100). The result is:

$$\frac{\delta \langle \hat{n}_{\kappa p}(t) \rangle}{\delta F_{\kappa',\alpha'p'}(t')} = \int dr'' dt'' \epsilon_{1}(rt, r''t'') \frac{\delta \langle \hat{n}_{\kappa p}(r''t'') \rangle}{\delta F_{\kappa',\alpha'p'}(t')} - \sum_{\gamma} Z_{\kappa} D_{\kappa\gamma p,\kappa',\alpha'p'}(tt') \nabla_{\gamma} \delta(r - \tau_{\kappa p}^0).$$

(115)

By inserting this result inside Eq. (114) and using the expansion in Eq. (107), we finally obtain the equation of motion for the displacement-displacement correlation function:

$$M_{\kappa} \frac{\partial^2}{\partial t^2} D_{\kappa\alpha p,\kappa',\alpha'p'}(tt') = -\delta_{\kappa\alpha p,\kappa',\alpha'p'} \delta(tt') - \sum_{\kappa,\alpha'p'} \int dt'' \Pi_{\kappa\alpha p,\kappa',\alpha'p'}(tt'') D_{\kappa\alpha p,\kappa',\alpha'p'}(tt'').$$

(116)

The quantity $\Pi_{\kappa\alpha p,\kappa',\alpha'p'}(tt')$ in this expression is called the ‘phonon self-energy’ and is given by:

$$\Pi_{\kappa\alpha p,\kappa',\alpha'p'}(tt') = \int dr dr'' Z_{\kappa} \nabla_{\alpha} \delta(r - \tau_{\kappa p}^0) W_{\kappa}(rt, r''t') Z_{\kappa'} \nabla_{\alpha'} \delta(r' - \tau_{\kappa'p'}^0) + \delta_{\kappa p,\kappa'} \delta(tt') \nabla_{\alpha} \langle \hat{n}(r) \rangle v(r, r') Z_{\kappa'} \nabla_{\alpha'} \delta(r' - \tau_{\kappa'p'}^0).$$

(117)

The derivation of Eq. (117) is nontrivial and is not found consistently in the literature; it requires converting the derivatives with respect to the position variables $r, r'$ into derivatives with respect to the nuclear coordinates; integrating by parts in order to re-arrange the derivatives with respect to $r$ and $r'$; invoking the harmonic approximation; and considering that, after setting the forces
whereas the phonon self-energy in the frequency-domain at equilibrium reads:

\[
\Pi_{\kappa\alpha\kappa'\alpha'}(\omega) = \int d\mathbf{r}d\mathbf{r}' \left[ Z_\kappa \nabla_\alpha \delta(\mathbf{r} - \mathbf{r}_0^\kappa) W_e(\mathbf{r},\mathbf{r}',\omega) + \delta_{\kappa\kappa'} \nabla_\alpha \delta(\mathbf{r} - \mathbf{r}_0^\kappa) v(\mathbf{r},\mathbf{r}') \right] Z_{\kappa'} \nabla_{\alpha'} \delta(\mathbf{r}' - \mathbf{r}_0^\kappa).
\]

The second line in this expression is conveniently rewritten by making use of the acoustic sum rule. This sum rule is well-known in the theory of lattice dynamics of crystals, and can be generalized to the case of many-body Green’s function approaches as follows:

\[
\sum_{\kappa'\alpha'} \Pi_{\kappa\alpha\kappa'\alpha'}(\omega = 0) = 0 \quad \text{for any } \alpha, \alpha'.
\]

This relation was first derived by Baym by imposing the condition that the nuclei in the crystal must remain near their equilibrium positions due to fictitious restoring forces. Physically this condition corresponds to considering a crystal which is held fixed in the laboratory reference frame. In this approach, the crystal cannot translate or rotate as a whole. Similar relations were derived in Refs. [128, 129].

If we combine Eqs. (119) and (120), perform integrations by parts, and carry out the integrations in \( \mathbf{r} \) and \( \mathbf{r}' \) we obtain:

\[
\Pi_{\kappa\alpha\kappa'\alpha'}(\omega) = \sum_{\kappa'\alpha'} Z_\kappa Z_{\kappa'} \frac{\partial^2}{\partial r_\alpha \partial r_{\alpha'}} \left. \right|_{r = \mathbf{r}_0^\kappa, \mathbf{r}' = \mathbf{r}_0^\kappa'} \left[ \delta_{\kappa'\alpha',\kappa''\alpha''} W_e(\mathbf{r},\mathbf{r}',\omega) - \delta_{\kappa\kappa'} \delta_{\alpha\alpha'} W_e(\mathbf{r},\mathbf{r}',0) \right],
\]

which fulfills the sum rule in Eq. (120).

Eqs. (118) and (121) completely define the nuclear dynamics in the harmonic approximation. After obtaining the displacement-displacement correlation function \( D_{\kappa\alpha\kappa'\alpha'}(t,t') \) by solving this set of equations, it is possible to construct the expectation value of the nuclear density using Eqs. (107) and (109):

\[
\langle \hat{n}_n(\mathbf{r}) \rangle = n_n^0(\mathbf{r}) - \frac{i\hbar}{2} \sum_{\kappa,\alpha'} Z_\kappa \frac{\partial^2}{\partial r_\alpha \partial r_{\alpha'}} D_{\kappa\alpha\kappa'\alpha'}(t,t).
\]

We should emphasize that, according to Eq. (121), the coupling of the nuclear displacements to the electrons is completely defined by the electronic dielectric matrix through \( W_e \). Similarly, the nuclei affect the electronic structure via the dielectric matrix which enters \( W_{\text{ph}} \) in Eq. (110) and via the nuclear density inside \( V_{\text{tot}} \) in Eq. (79). From these considerations it should be clear that the electronic dielectric matrix \( \epsilon_e(\mathbf{r},\mathbf{r}',\omega) \) plays an absolutely central role in the the field-theoretic approach to the electron-phonon problem.
adiabatic approximation, and therefore it encompasses insulators, intrinsic as well as doped
taggers. In particular we emphasize that this approach does

described non-perturbatively by means of a coupled set of nonlinear equations for the exact prop-
tions of a fully-interacting electron-phonon system, within the
Table 1 provides a closed set of self-consistent equations whose solution yields the Green’s func-
ions, (110), Screened Coulomb, nuclei $W_c(12) = v(12) + \int d(34) v(13) P_e(34) W_c(42)$

electronic polarization $P_e(12) = -i\hbar \sum_{\sigma \bar{\sigma}} G(13) G(41) \Gamma(342)$

electronic dielectric matrix $\epsilon_e(12) = \delta(12) - \int d(3) v(13) P_e(32)$

Vertex $\Gamma(123) = \delta(12) \delta(13) + \int d(4567) [\delta G(45) G(46) G(57) \Gamma(673)]$

Screened Coulomb, nuclei $W_{ph}(12) = \sum_{\kappa k p, \kappa' k' p', \alpha \alpha'} (34) \epsilon_e^{-1}(13) \nabla_{\kappa} V_{\alpha} (r_3 - \tau^0_{\alpha p})$

$\times D_{\kappa k p, \kappa' k' p', \alpha \alpha'} (t_3 t_4) \epsilon_e^{-1}(24) \nabla_{\kappa'} \nabla_{\alpha'} V_{\alpha'} (r_4 - \tau^0_{\alpha' p'})$

Phonon self-energy $\Pi_{\kappa k p, \kappa' k' p', \alpha \alpha'} (\omega) = \sum_{\kappa k p, \kappa' k' p', \alpha \alpha'} (\partial^2 / \partial \nu \partial \nu' )$

$\times [\delta_{\nu' \nu} W_{\alpha} (r, r', \nu) - \delta_{\kappa k p, \kappa' k' p'} W_{\alpha} (r, r', 0)]_{r = \tau^0_{\kappa p}, r' = \tau^0_{\kappa' p'}}$

Table 1: Self-consistent Hedin-Baym equations for the coupled electron-phonon system in the
harmonic approximation.

4.4 Hedin-Baym equations

Apart from making use of the harmonic approximation, the set of equations given by Eqs. (79),
(84), (85), (87), (92), (93), (94), (95), (110), (118), (121), and (122) describe the coupled electron-
phonon system entirely from first principles. This set of equations can be regarded as the most
sophisticated description of interacting electrons and phonons available today. Since the self-
consistent equations for the electrons were originally derived by Hedin [122], and those for the
nuclei were derived first by Baym [125], we will refer to the complete set as the Hedin-Baym
equations. Given the importance of these relations, we summarize them schematically in Table 1.

The standard Hedin’s equations for interacting electrons in the potential of clamped nuclei [122]
are immediately recovered from the Hedin-Baym equations by setting to zero the displacement-
displacement correlation function of the nuclei, $D_{\kappa p a, \kappa' p' a'} = 0$.

Table 1 provides a closed set of self-consistent equations whose solution yields the Green’s func-
tions of a fully-interacting electron-phonon system, within the harmonic approximation. We
stress that these relations are fundamentally different from diagrammatic approaches. In fact,
here the coupled electron-phonon system is not addressed using Feynman-Dyson perturbation
theory as it was done for example in Ref. [127]. Instead, in Table 1 electrons and phonons are
described non-perturbatively by means of a coupled set of nonlinear equations for the exact prop-
grators. In particular we emphasize that this approach does not require the Born-Oppenheimer
adiabatic approximation, and therefore it encompasses insulators, intrinsic as well as doped
Almost every property related to electron-phonon interactions in solids that can be calculated today from first principles can be derived from these equations. Examples to be discussed in Secs. 5-10 include the renormalization of the Fermi velocity, the band gap renormalization in semiconductors and insulators, the non-adiabatic corrections to vibrational frequencies, the Fröhlich interaction, and the lifetimes of electrons and phonons. The generalization of these results to the case of finite temperature should also be able to describe phonon-mediated superconductivity, although this phenomenon is best addressed by studying directly the propagation of Cooper pairs (see Sec. 11).

Baym’s theory can in principle be extended to go beyond the harmonic approximation [128]. However, the mathematical complexity of the resulting formalism is formidable, due to the appearance of many additional terms which are neglected in the harmonic approximation.

5 From a many-body formalism to practical calculations

The Hedin-Baym equations summarized in Table 1 define a rigorous formalism for studying interacting electrons and phonons in metals, semiconductors, and insulators entirely from first principles. However, a direct numerical solution of these equations for real materials is currently out of reach, and approximations are needed for practical calculations. The following sections establish the connection between the Hedin-Baym equations and standard expressions which are currently in use in ab initio calculations of electron-phonon interactions.

5.1 Effects of the electron-phonon interaction on phonons

5.1.1 Phonons in the Born-Oppenheimer adiabatic approximation

The vibrational eigenmodes of the nuclei can be identified with the resonances of the displacement-displacement correlation function $D_{\kappa p,\kappa' p'}(tt')$ in the frequency domain. If we denote by $M$ the diagonal matrix having the nuclear masses $M_\kappa$ along its diagonal, then the formal solution of Eq. (118) can be written as:

$$D(\omega) = \left[ M \omega^2 - \Pi(\omega) \right]^{-1},$$

(123)

where $D$ is the matrix with elements $D_{\kappa p,\kappa' p'}$. The resonant frequencies of the system correspond to the solutions of the nonlinear equations:

$$\Omega_\nu(\omega) - \omega = 0, \quad \text{with } \nu = 1, \ldots 3M,$$

(124)

where $\Omega_\nu^2(\omega)$ is an eigenvalue of $M^{-1/2} \Pi(\omega) M^{-1/2}$, parametric in the variable $\omega$.

As expected, the study of lattice vibrations within a field-theoretic framework resembles the standard eigenvalue problem reviewed in Sec. 3.1. In particular, the matrix $\Pi(\omega)$ represents the many-body counterpart of the matrix of interatomic force constants $C_{\kappa p,\kappa' p'}$ introduced in
Eq. (13). However, despite its formal simplicity, Eq. (124) conceals the full wealth of information associated with the many-body electronic screening $\epsilon_e(r, r', \omega)$ via Eq. (121). In fact, the phonon self-energy is generally complex and frequency-dependent. Therefore we can expect to find roots of Eq. (124) outside of the real frequency axis, as well as multiple roots for the same ‘eigenmode’.

The link between Eq. (124) and phonon calculations by means of DFT is established by noting that DFT relies on the Born-Oppenheimer adiabatic approximation. In the adiabatic approximation the nuclei are considered immobile during characteristic electronic timescales. Formally, this approximation is introduced by setting $\omega = 0$ in Eq. (121) \[127\]. In practice, this assumption corresponds to stating that $\epsilon_e(r, r', \omega)$ can be replaced by $\epsilon_e(r, r', 0)$ in the frequency range of the vibrational excitations. Obviously this is not always the case, and important exceptions will be discussed in Sec. 5.1.2.

In order to see more clearly the connection with the formalism discussed in Sec. 3.1, we partition the phonon self-energy into ‘adiabatic’ and ‘non-adiabatic’ contributions:

$$\Pi(\omega) = \Pi^A + \Pi^{NA}(\omega),$$

(125)

with $\Pi^A = \Pi(\omega = 0)$. As we will see below, the adiabatic term $\Pi^A$ will be taken to describe ‘non-interacting’ phonons, and the non-adiabatic self-energy $\Pi^{NA}$ will be used to describe the effects of electron-phonon interactions.

In the early literature it is common to find a different partitioning, whereby the non-interacting system is defined by the bare interatomic force constants, corresponding to nuclei in the absence of electrons \[6\]. This alternative choice is not useful in modern calculations, because the resulting non-interacting phonon dispersions are very different from the fully-interacting dispersions. The present choice of using instead adiabatic phonons as the non-interacting system, is more convenient in the context of modern \textit{ab initio} techniques, since calculations of adiabatic phonon spectra are routinely performed within DFPT.

In the remainder of this section we concentrate on the adiabatic term, and we defer the discussion of the non-adiabatic self-energy to Sec. 5.1.2. Using Eq. (121), we can rewrite the adiabatic self-energy as follows:

$$\Pi^A_{\kappa\alpha\kappa'\alpha'} = \sum_{\kappa''\alpha''} \left( \delta_{\kappa''\alpha''} \delta_{\kappa\alpha\kappa'} - \delta_{\kappa\alpha\kappa''} \right) \left[ \int dr \frac{\partial \langle \hat{n}_e(r) \rangle}{\partial \tau_{\kappa\alpha\kappa'}} \frac{\partial V^{en}(r)}{\partial \tau_{\kappa\alpha\kappa'}} + \frac{\partial^2 U_{nn}}{\partial \tau_{\kappa\alpha\kappa'} \partial \tau_{\kappa''\alpha''}} \right].$$

(126)

In this expression $U_{nn}$ is the nucleus-nucleus interaction energy from Eq. (65), $V^{en}$ is the electron-nuclei interaction from Eq. (25), and all the derivatives are taken at the equilibrium coordinates. The derivation of Eq. (126) requires the use of the identity:

$$\frac{\partial \langle \hat{n}_e(r) \rangle}{\partial \tau_{\kappa\alpha\kappa'}} = -Z_{\kappa} \int dr' [\epsilon_e^{-1}(r, r'; 0) - \delta(r, r')] \nabla' \delta(r' - \tau_{\kappa\alpha\kappa'}).$$

(127)

This identity follows from the same reasoning leading to Eq. (100), after considering an external potential which modifies the position of the nucleus $\kappa$ in the cell $p$. Equation (126) can be recast in a familiar form by exploiting the acoustic sum rule in Eq. (120). Indeed after a few tedious
but straightforward manipulations we obtain:

\[
\Pi_{\kappa p, \kappa' p'}^A = \int dr \frac{\partial \langle \hat{n}_e(r) \rangle}{\partial \tau_{\kappa' p'}} \frac{\partial V_{en}(r)}{\partial \tau_{\kappa p}} + \int dr \langle \hat{n}_e(r) \rangle \frac{\partial^2 V_{en}(r)}{\partial \tau_{\kappa p} \partial \tau_{\kappa' p'}} + \frac{\partial^2 U_{nn}}{\partial \tau_{\kappa p} \partial \tau_{\kappa' p'}}. \tag{128}
\]

In this form, one can see that the adiabatic self-energy gives precisely the interatomic force constants that we would obtain using the Born-Oppenheimer approximation and the Hellman-Feynman theorem, compare Eq. (128) for example with Ref. [87], p. 517.

The difference between the \(\Pi_{\kappa p, \kappa' p'}^A\) in Eq. (128) and the \(C_{\kappa p, \kappa' p'}\) in Eq. (13) is that, in the former case, the electron density response to atomic displacements is governed by the exact many-body dielectric matrix \(\epsilon_e(r, r', 0)\) and electron density \(\langle \hat{n}_e(r) \rangle\), as shown by Eqs. (127) and (128). As a result, \(\Pi_{\kappa p, \kappa' p'}^A\) corresponds to force constants and electron density dressed by all many-body interactions of the system (both electron-electron and electron-phonon interactions).

In contrast, when the force constants in Eq. (13) are calculated using DFT, the electron density response to an atomic displacement is evaluated using the RPA+xc screening, that is \(e^{H_{xc}}(r, r')\) from Sec. 3.2.4 and the ground-state electron density is calculated at clamped nuclei.

The use of the adiabatic approximation in the study of phonons carries the important advantage that the many-body force constants \(\Pi^A\) form a real and symmetric matrix. This can be seen by rewriting Eq. (121) for \(\omega = 0\), and using the relation \(W_e(r, r', \omega) = W_e(r', r, -\omega)\) which follows from the property \(W(12) = W(21)\) (see Sec. 4.2). Since \(\Pi^A\) is real and symmetric, all its eigenvalues are guaranteed to be real. In this approximation, the excitations of the lattice correspond to sharp resonances in the displacement-displacement correlation function \(D(\omega)\), and it is meaningful to talk about phonons as long-lived excitations of the system. In fact these excitations are infinitely long-lived in the harmonic approximation. In practical calculations, the many-body \(\Pi^A\) is invariably replaced by the DFT interatomic force constants, and in this case the agreement of the calculated phonon frequencies with experiment is excellent in most cases. Illustrative examples can be found among others in Refs. [87, 88, 91, 139, 146].

The most obvious criticism to the adiabatic approximation is that, in the case of metals, the assumption \(\epsilon_e(r, r', \omega) \approx \epsilon_e(r, r', 0)\) is inadequate. This can intuitively be understood by recalling that the dielectric function of the homogeneous electron gas diverges when \(\omega, q \to 0\) [9]. In practical calculations, this divergence is connected with vanishing denominators in Eq. (51) for \(q \to 0\). An approximate, yet very successful strategy for overcoming this problem, is to replace the occupation numbers in Eq. (51) by smoothing functions such as the Fermi-Dirac distribution, and to describe the singular terms analytically [99]. Most first-principles calculations of phonon dispersion relations in metals have been carried out using this strategy. Improvements upon this strategy will be discussed in Sec. 5.1.2.

The adiabatic approximation leads naturally to the definition of an ‘adiabatic’ propagator \(D^A(\omega)\), which can be obtained from Eq. (123) after replacing the phonon self-energy by its static limit:

\[
D^A(\omega) = \left[ M \omega^2 - \Pi^A \right]^{-1}. \tag{129}
\]

Now, if we identify \(\Pi_{\kappa p, \kappa' p'}^A\) with the interatomic force constant \(C_{\kappa p, \kappa' p'}\) in Eq. (13), we can obtain an explicit expression for the adiabatic phonon propagator in terms of the eigenmodes...
e_{\kappa\nu}(q) and eigenfrequencies $\omega_{q\nu}$ introduced in Sec. 3.1. To this end we invert Eq. (129) using Eqs. (14)-(17), and recall that the dynamical matrix is Hermitian and obeys the relation $D_{\alpha\beta}^{\text{dim}, \kappa\alpha'}(q) = D_{\alpha\beta}^{\text{dim}, \kappa\alpha'}(-q)$ \cite{81}. After tedious but straightforward steps we find:

$$D_{\kappa\alpha p, \kappa'\alpha' p'}^{A}(\omega) = \sum_{\nu} \int \frac{dq}{\Omega_{\text{BZ}}} S_{q\nu, \kappa\alpha p}^* S_{q\nu, \kappa'\alpha' p'} \frac{2\omega_{q\nu}}{\omega^2 - \omega_{q\nu}^2},$$

(130)

with the definition:

$$S_{q\nu, \kappa\alpha p} = e^{iq\cdot R_p} (2\Omega_{\kappa\nu} q_{\nu}^2)^{-1/2} e_{\kappa\nu}(q).$$

(131)

This result suggests that, as expected, the propagator should take a simple form in the eigen-modes representation. In fact, by using the inverse transform of Eq. (131) we have:

$$D_{q\nu, q'\nu'}^{A}(\omega) = \Omega_{\text{BZ}} \delta(q - q') D_{q\nu, q'\nu'}^{A}(\omega),$$

(132)

This result can alternatively be obtained starting from the ladder operators of Appendix B. In fact, after using Eqs. (20), (109), and (131) we find:

$$D_{q\nu, q'\nu'}^{A}(tt') = -i \langle \hat{T} [\hat{a}_{q\nu}(t) \hat{a}_{q'\nu'}(t') + \hat{a}_{-q\nu}(t) \hat{a}_{-q'\nu'}(t')]\rangle \delta_{\nu\nu'}.$$

(133)

An explicit evaluation of the right-hand side using the Heisenberg time evolution generated by the phonon Hamiltonian in Eq. (22) yields precisely Eq. (132), with the added advantage that it is easier to keep track of the time-ordering. The result is:

$$D_{q\nu, q'\nu'}^{A}(\omega) = \frac{1}{\omega - \omega_{q\nu} + i\eta} - \frac{1}{\omega + \omega_{q\nu} - i\eta},$$

(134)

with $\eta$ a positive real infinitesimal. This alternative approach is very common in textbooks, see for example Refs. \cite{8, 147}. However, it does not carry general validity in a field-theoretic framework since it rests on the adiabatic approximation.

### 5.1.2 Phonons beyond the adiabatic approximation

In order to go beyond the adiabatic approximation, it is necessary to determine the complete propagator $D(\omega)$ in Eq. (123). Formally this can be done by combining Eqs. (123) and (125) to obtain the following Dyson-like equation:

$$D(\omega) = D^{A}(\omega) + D^{A}(\omega) \Pi^{\text{NA}}(\omega) D(\omega).$$

(135)

In this form it is apparent that the non-adiabatic phonon self-energy $\Pi^{\text{NA}}(\omega)$ ‘dresses’ the non-interacting phonons obtained within the adiabatic approximation, as shown schematically in Fig. 1(a). It is convenient to rewrite the Dyson equation in such a way as to show more clearly the poles of the propagator. Using Eqs. (131)-(132) we find:

$$D_{q\nu, q'\nu'}^{-1}(\omega) = \frac{1}{2\omega_{q\nu}} \left[ \delta_{\nu\nu'}(\omega^2 - \omega_{q\nu}^2) - 2\omega_{q\nu} \Pi_{q\nu, q'\nu'}^{\text{NA}}(\omega) \right],$$

(136)

where $\Pi_{q\nu, q'\nu'}^{\text{NA}}$ and $D_{q\nu, q'\nu'}^{-1}$ are obtained using the transform of Eq. (131) and its inverse, respectively.
Figure 1: Diagrammatic representation of the phonon Green's function and self-energy. (a) Dyson equation for the phonon propagator, Eq. (135). The thick wavy line represents the fully-interacting, non-adiabatic propagator; the thin wavy line is the adiabatic propagator; the disc is the non-adiabatic self-energy. (b) Lowest-order diagrammatic expansion of the phonon self-energy in terms of the bare electron-phonon vertices and the RPA electronic polarization. The small dots are the bare electron-phonon coupling functions, and the thin lines are the non-interacting (for example Kohn-Sham) electron Green’s functions. This diagram is the simplest possible term which begins and ends with a phonon line. (c) Non-perturbative representation of the phonon self-energy in terms of the bare coupling, the dressed coupling (large gray disc), the fully-interacting electron’s Green’s functions (thick lines), and the vertex $\Gamma$ from Eq. (83). This diagram was proposed in Ref. [127] and describes the first line of Eq. (141). (d) Schematic representation of the relation between the dressed electron-phonon coupling $g$ and the bare coupling $g^b$, from Eq. (144). Ref. [131] reports a similar diagram, although with $g^b$ instead of $g$ in the second term of the r.h.s.; the difference stems from the present choice of using the irreducible polarization $P = GG\Gamma$ instead of the reducible polarization employed in Ref. [131]. We note that in Ref. [1] the placement of the large dark gray disc in (d) at one end of the bare Coulomb line is incorrect. The mistake has been corrected in the present figure.
From Eq. (136) we see that the non-adiabatic self-energy $\Pi^{NA}$ modifies the adiabatic phonon spectrum in four distinct ways: (i) the real part of the diagonal elements $\Pi^{NA}_{q\nu\nu}$ shifts the adiabatic frequencies; (ii) the imaginary part introduces spectral broadening; (iii) the off-diagonal elements of $\Pi^{NA}_{q\nu\nu'}$ introduce a coupling between the adiabatic vibrational eigenmodes; (iv) the frequency-dependence of $\Pi^{NA}_{q\nu\nu}(\omega)$ might lead to multiple poles for the same mode $\nu$, thereby introducing new structures in the phonon spectrum.

Today it is relatively common to calculate phonon linewidths arising from electron-phonon interactions \[76, 148\]. Recently it has also become possible to study the frequency renormalization due to non-adiabatic effects \[149, 150\].

The possibility of observing new features in vibrational spectra arising from the EPI has not been studied from first principles, but the underlying phenomenology should be similar to that of plasmon satellites in photoelectron spectra (see Sec. 5.2.6). Generally speaking we expect satellites whenever $\epsilon_e(r, r', \omega)$ exhibits dynamical structure close to vibrational frequencies. This can happen, for example, in the case of degenerate polar semiconductors, when phonon and plasmon energies are in resonance. In these cases, phonons and plasmons can combine into ‘coupled plasmon-phonon modes’ \[151\], which are the electronic analogue of photon polaritons (Ref. \[152\], p. 295). This phenomenon was predicted theoretically \[153\], and subsequently confirmed by Raman measurements on GaAs \[154\]. We speculate that it should be possible to obtain coupled plasmon-phonon modes from the frequency-dependence of the phonon self-energy in Eq. (136); it would be interesting to perform first-principles calculations in order to shed light on these aspects.

In practical calculations the non-adiabatic corrections to the adiabatic phonon spectrum are evaluated from Eq. (136) using first-order perturbation theory, by retaining only the diagonal elements of $\Pi^{NA}$. If we denote the complex zeros of $D_{q\nu\nu}^{-1}(\omega)$ by $\tilde{\Omega}_{q\nu} = \Omega_{q\nu} - i\gamma_{q\nu}$, in the case of non-degenerate eigenmodes Eq. (136) gives:

\[
\tilde{\Omega}_{q\nu}^2 = \omega_{q\nu}^2 + 2\omega_{q\nu}\Pi^{NA}_{q\nu\nu}(\tilde{\Omega}_{q\nu}),
\]

therefore:

\[
\gamma_{q\nu} = -\frac{\omega_{q\nu}}{\Omega_{q\nu}} \text{Im} \Pi^{NA}_{q\nu\nu}(\Omega_{q\nu} - i\gamma_{q\nu}), \tag{138}
\]

\[
\Omega_{q\nu}^2 = \omega_{q\nu}^2 + \gamma_{q\nu}^2 + 2\omega_{q\nu}\text{Re} \Pi^{NA}_{q\nu\nu}(\Omega_{q\nu} - i\gamma_{q\nu}). \tag{139}
\]

Apart from the small $\gamma_{q\nu}^2$ term in the second line, these expressions are identical to those provided in Refs. \[6, 148\]. Since non-adiabatic corrections are usually small as compared to the adiabatic phonon frequencies, the above expressions are often simplified further by using the additional approximations $|\Omega_{q\nu} - \omega_{q\nu}| \ll \omega_{q\nu}$ and $|\gamma_{q\nu}| \ll \omega_{q\nu}$, leading to $\gamma_{q\nu} \simeq -\text{Im} \Pi^{NA}_{q\nu\nu}(\omega_{q\nu})$ and $\Omega_{q\nu} \simeq \omega_{q\nu} + \text{Re} \Pi^{NA}_{q\nu\nu}(\omega_{q\nu})$. In these forms it becomes evident that the real part of the self-energy shifts the adiabatic phonon frequencies, and the imaginary part is responsible for the spectral broadening of the resonances. Using these expressions in Eq. (136) and going back to the time domain, it is seen that, as a result of the EPI, phonons acquire a finite lifetime given by $\tau^{\text{ph}}_{q\nu} = (2\gamma_{q\nu})^{-1}$. 
5.1.3 Expressions for the phonon self-energy used in \textit{ab initio} calculations

In the literature on electron-phonon interactions, the phonon self-energy $\Pi$ is almost invariably expressed in terms of an electron-phonon vertex $g$ and the electron Green’s function $G$ as $\Pi = |g|^2GG$ in symbolic notation, see for example Ref. [6], p. 195. While this has become common practice also in \textit{ab initio} calculations, the origin of this choice is not entirely transparent. One could derive the above expression directly from Eq. (1), using standard Green’s function techniques. However, this procedure does not answer the key question on \textit{how} to calculate the electron-phonon matrix elements $g$.

Closer inspection of the theory reveals that this is a rather nontrivial point. In fact, on the one hand, a straightforward expansion of the second-quantized Hamiltonian of Eq. (68) in terms of the nuclear coordinates leads to ‘bare’ electron-phonon matrix elements, $g^b$, which contain the bare Coulomb interaction between electrons and nuclei. On the other hand, if we go back to Sec. 3.2.2, we see that the electron-phonon matrix elements in DFT are ‘dressed’ by the self-consistent response of the electrons. The difference between bare and dressed vertex is not only quantitative, but also qualitative: for example in metals the bare vertex is long-ranged, while the screened vertex is short-ranged.

The relation between bare and dressed electron-phonon vertices and the derivation of explicit expressions for the phonon self-energy have been discussed by many authors, see for example Refs. [155, 156]. In short the argument is that the lowest-order Feynman diagram starting and ending with a phonon line must contain precisely two \textit{bare} electron-phonon vertices, as shown in Fig. 1(b). By construction this diagram corresponds to having $\Pi = |g^b|^2GG$. In order to make the transition from the bare vertex to the dressed vertex it is necessary to collect together all the proper electronic polarization diagrams around the vertex. However, these steps have been carried out only for the homogeneous electron gas [155, 156]. In the following we show how the dressed electron-phonon vertex emerges from a non-perturbative analysis based on the Hedin-Baym equations.

The non-adiabatic phonon self-energy $\Pi^{NA}$ introduced in Sec. 5.1.2 can be written explicitly by combining Eqs. (119) and (125):

$$\Pi^{NA}_{\kappa\alpha\nu,\kappa'\alpha'\nu'}(\omega) = \int d\mathbf{r} d\mathbf{r}' Z^{\kappa}Z_{\kappa'} \nabla_{\alpha} \delta (\mathbf{r} - \mathbf{r}'_\kappa) [W_e(\mathbf{r}, \mathbf{r}', \omega) - W_e(\mathbf{r}, \mathbf{r}', 0)] Z^{\kappa'}Z_{\alpha'} \nabla'_{\alpha'} \delta (\mathbf{r}' - \mathbf{r}'_{\kappa'}) \delta (\mathbf{r}_\kappa - \mathbf{r}'_{\kappa'}) \delta (\mathbf{r}_\kappa' - \mathbf{r}'_{\kappa'}) \frac{1}{\omega}.$$  \hspace{1cm} (140)

Using the Dyson equation for the screened Coulomb interaction, it can be seen that this expression does indeed contain electron-phonon matrix elements. In fact, by inserting Eq. (94) into Eq. (140) we find terms like $vP_eW_e$, and the electron-phonon matrix elements will arise from taking the gradients of $v$ and $W_e$, respectively. By working in the eigenmodes representation via Eq. (131), after lengthy manipulations this procedure yields:

$$\hbar \Pi^{NA}_{q\nu,q'\nu'}(\omega) = \int d\mathbf{r} d\mathbf{r}' g^b_{q\nu}(\mathbf{r}) P_e(\mathbf{r}, \mathbf{r}', \omega) g^c_{q'\nu'}(\mathbf{r}', \omega) - \int d\mathbf{r} d\mathbf{r}' g^b_{q\nu}(\mathbf{r}) P_e(\mathbf{r}, \mathbf{r}', 0) g^c_{q'\nu'}(\mathbf{r}', 0),$$ \hspace{1cm} (141)

where we introduced electron-phonon ‘coupling functions’ as follows. The \textit{bare} coupling $g^b$ is defined as:

$$g^b_{q\nu}(\mathbf{r}) = \Delta_{q\nu} V^{\text{en}}(\mathbf{r}),$$ \hspace{1cm} (142)
where $V_{\text{en}}$ is the potential of the nuclei from Eq. (25); in practical calculations this quantity is replaced by the usual ionic pseudo-potentials. The meaning of the variation $\Delta q_{\nu}$ is the same as in Eqs. (33)-(35). The dressed couplings $g$ and $g^{cc}$ are defined as:

\[
\begin{align*}
    g_{q\nu}(r, \omega) &= \int d'r' \epsilon_{e}^{-1}(r, r', \omega) g^{b}_{q\nu}(r'), \quad (143) \\
    g^{cc}_{q\nu}(r, \omega) &= \int d'r' \epsilon_{e}^{-1}(r, r', \omega) g^{b,\ast}_{q\nu}(r'). \quad (144)
\end{align*}
\]

Since the dielectric matrix is real at $\omega = 0$, we have the simple relation $g^{cc}_{q\nu}(r, 0) = g^{\ast}_{q\nu}(r, 0)$. In order to derive Eq. (141) it is best to carry out the algebra in the time domain. We emphasize that the result expressed by Eq. (141) is non-perturbative, and relies solely on the harmonic approximation.

Equation (141) is in agreement with the standard result for the homogeneous electron gas [155]. The same expression was also obtained by Keating [127] using a detailed diagrammatic analysis. Keating’s diagrammatic representation of the self-energy is shown in Fig. 1(c), and can be obtained from Eq. (141) by noting that, in symbolic notation, $P = GG\Gamma$ from Eq. (92), therefore $\Pi = g^{b}GG\Gamma g$. For completeness we also show in Fig. 1(d) a diagrammatic representation of the dressed electron-phonon coupling function $g$ as given by Eq. (144). This representation is obtained by observing that $g = \epsilon^{-1}g^{b}$, $\epsilon = 1 - vP$, and $P = GG\Gamma$, therefore $g = g^{b} + vGG\Gamma g$.

In view of practical first-principles calculations it is useful to have a simplified expression for the non-adiabatic phonon self-energy in Eq. (141). To this aim we make the following approximations:

(i) The vertex function in Eq. (92) is set to $\Gamma(123) = \delta(12)\delta(13)$. This is the same approximation at the heart of the GW method [122, 157, 158];

(ii) The fully-interacting electron Green’s function $G$ is replaced by its non-interacting counterpart, using the Kohn-Sham eigenstates/eigenvalues evaluated with the nuclei held in their equilibrium positions;

(iii) The fully-interacting dielectric matrix in Eq. (144) is replaced by the RPA+xc response obtained from a DFT calculation, as discussed in Sec. 3.2.2;

(iv) The frequency-dependence of the screened electron-phonon coupling defined in Eq. (144) is neglected: $g^{cc}_{q\nu}(r, \omega) \simeq g^{cc}_{q\nu}(r, 0) = g^{\ast}_{q\nu}(r, 0)$. This approximation is ubiquitous in the literature but it is never mentioned explicitly;

(v) For notational simplicity, we consider a spin-degenerate system with time-reversal symmetry; this simplification is easily removed.

Using these assumptions we can rewrite the component of Eq. (141) for $q = q'$ as:

\[
\hbar \Pi_{q\nu q'}^{NA}(\omega) = 2 \sum_{mn} \int \frac{d\mathbf{k}}{\Omega_{\text{BZ}}} g^{b}_{mn\nu}(\mathbf{k}, q) g^{\ast}_{mn\nu}(\mathbf{k}, q) \left[ \frac{f_{mk+q} - f_{nk}}{\epsilon_{mk+q} - \epsilon_{nk} - \hbar(\omega + i\eta)} - \frac{f_{mk+q} - f_{nk}}{\epsilon_{mk+q} - \epsilon_{nk}} \right]. \quad (145)
\]

We note that the components of the phonon self-energy for $q \neq q'$ vanish due to the periodicity of the lattice. In Eq. (145) the sums run over all Kohn-Sham states, with occupations $f_{nk}$, and $\eta$.
is a real positive infinitesimal. In this case we indicate explicitly the factor of 2 arising from the spin degeneracy. The matrix element \( g_{mn\nu}(k, q) \) is the same as in Eq. (38), and it is precisely the quantity calculated by most linear-response codes. The matrix element \( g^b_{mn\nu}(k, q) \) is obtained from \( g_{mn\nu}(k, q) \) by replacing the variation of the Kohn-Sham potential by the corresponding variation of the ionic (pseudo)potentials. The field-theoretic phonon self-energy given by Eq. (145) is in agreement with the expression derived in Ref. [150] starting from time-dependent density-functional perturbation theory.

The presence of both the bare electron-phonon matrix element and the screened matrix element in Eq. (145) has not been fully appreciated in the literature, and most \textit{ab initio} calculations employ an approximate self-energy whereby \( g^b \) is replaced by \( g \). The replacement of the bare matrix elements by their screened counterparts in the phonon self-energy goes a long way back, and can be found already in the seminal work by Allen [148]. As a result many investigators (including the author) calculated phonon lifetimes using the following expression [6]:

\[
\frac{1}{\tau_{\text{ph}q\nu}} = \frac{2\pi}{\hbar} \sum_{mn} \int \frac{dk}{\Omega_{\text{BZ}}} |g_{mn\nu}(k, q)|^2 (f_n k - f_m k + q) \delta(\varepsilon_{mnk+q} - \varepsilon_{mnk} - \hbar \omega_{q\nu}).
\]

(146)

This is obtained from Eq. (145) by taking the imaginary part and by making the replacement \( g^b \rightarrow g \). While Eq. (146) can be derived from the Fermi golden rule in an independent-particle approximation (see Ref. [159], Appendix B), the choice of the electron-phonon matrix elements is somewhat arbitrary. In future calculations of the phonon self-energy it will be important to assess the importance of using the correct vertex structure, that is replacing \( |g_{mn\nu}(k, q)|^2 \) by \( g^b_{mn\nu}(k, q) g^*_{mn\nu}(k, q) \) in Eq. (146).

In general, the effects of the non-adiabatic self-energy on the phonon spectrum are expected to be significant only in the case of metals and small-gap semiconductors. In fact, by combining Eqs. (138), (139), and (145) it is seen that \( \Pi^{\text{NA}} \) can be large only when occupied and empty single-particle states are separated by an energy of the order of the characteristic phonon energy. In such a case, we can expect a shift of the adiabatic phonon frequencies, and a concomitant broadening of the lines. A clear illustration of these effects was provided by Maksimov and Shulga, who analyzed a simplified model of a metal with linear bands near the Fermi level [160].

Calculations of phonon linewidths based on Eq. (146) have been reported by several authors [2] and have become commonplace in first-principles studies of electron-phonon physics. On the other hand, calculations of the non-adiabatic phonon frequencies using Eq. (145) have only been reported in Refs. [119, 149, 150, 167, 168], using the approximation that the bare vertex \( g^b \) can be replaced by the screened vertex \( g \). Examples of such calculations will be reviewed in Sec. 7.

Equation (145) suggests several avenues worth exploring in the future: firstly, the use of the bare vertex should not pose a challenge in practical calculations, since this quantity is already being calculated in linear-response DFT codes. Testing the impact of the bare vertex on phonon linewidths and frequency renormalizations will be important. Secondly, Eq. (145) contains off-diagonal couplings, which are usually ignored. It will be interesting to check the effect of using the complete matrix self-energy. Thirdly, the dynamical structure of the self-energy may con-
tain interesting information, such as for example spectral satellites and coupled phonon-plasmon modes. Lastly, the move from Eq. (141) to Eq. (145) involves the approximation that the frequency-dependence of the electron-phonon matrix elements can be neglected. The validity of this approximation is uncertain, and there are no reference ab initio calculations on this. However, we note that frequency-dependent electron-phonon matrix elements have been employed systematically in theoretical models of doped semiconductors (Ref. [9], Sec. 6.3).

Before closing this section we note that the formalism discussed here is based on zero-temperature Green’s functions. In order to extend the present results to finite temperature it is necessary to repeat all derivations using the Matsubara representation, and then perform the analytic continuation of the self-energy to the real frequency axis. Detailed derivations can be found in Refs. [125] and [128], and will not be repeated here. Fortunately it turns out that Eq. (145) can be extended to finite temperature by simply replacing the occupation factors $f_{nk}$ and $f_{mk+q}$ by the corresponding Fermi-Dirac distributions.

5.2 Effects of the electron-phonon interaction on electrons

5.2.1 Electron self-energy: Fan-Migdal and Debye-Waller terms

In Sec. 5.1 we discussed the link between the Hedin-Baym equations summarized in Table 1 and ab initio calculations of phonons. We first identified a Hermitian eigenvalue problem for the vibrational frequencies via the adiabatic approximation, and then we improved upon this description by means of a non-adiabatic self-energy. In this section, we adopt a similar strategy in order to discuss electronic excitation energies: first we identify an approximation to the Hedin-Baym equations which does not include any electron-phonon interactions, and then we introduce an electron self-energy to incorporate such interactions.

The single most common approximation in first-principles electronic structure calculations is to describe nuclei as classical particles clamped in their equilibrium positions. Within this approximation the expectation value of the nuclear charge density operator in Eq. (107), $\langle \hat{n}_n(r) \rangle$, is replaced by the first term in Eq. (107), $n_0^0(r)$. From Eq. (122) we see that this approximation formally corresponds to setting to zero the displacement-displacement correlation function of the nuclei. This observation suggests that, in order to unambiguously single out electron-phonon interactions in the Hedin-Baym equations, we need to define a non-interacting problem by setting $D_{\kappa\alpha p,\kappa'\alpha' p'} = 0$, and identify the electron-phonon interaction with the remainder. In the following, we write an equation of motion for the electrons analogous to Eq. (84), except with the nuclei clamped in their equilibrium positions; then we use a Dyson-like equation to recover the fully-interacting electron Green’s function.

The equation of motion for the electron Green’s function at clamped nuclei, which we denote

---

8Throughout this article, when $f_{nk}$ and $n_{qv}$ have the meaning of Fermi-Dirac and Bose-Einstein distributions, respectively, they are defined as follows: $f_{nk} = f[(\epsilon_{nk} - \epsilon_p)/k_B T]$ with $f(x) = 1/(e^x + 1)$ and $\epsilon_p$ being the Fermi energy; $n_{qv} = n(\hbar \omega_{qv}/k_B T)$ with $n(x) = 1/(e^x - 1)$. 

---
as $G^{cn}$, reads:

\[
\left[ i\hbar \frac{\partial}{\partial t_1} + \frac{\hbar^2}{2m_e} \nabla^2 (1) - V_{tot}^{cn}(1) \right] G^{cn}(12) - \int d3 \Sigma_{e}^{cn}(13) G^{cn}(32) = \delta(12). \tag{147}
\]

Here the potential $V_{tot}^{cn}$ differs from its counterpart $V_{tot}$ of Eq. (79) in that the total density of electrons and nuclei, $\langle \hat{n} \rangle$, is replaced by the density calculated at clamped nuclei, $\langle \hat{\tilde{n}}^{cn} \rangle$:

\[
V_{tot}^{cn}(1) = \int d2 v(1, 2) \langle \hat{\tilde{n}}^{cn}(2) \rangle, \tag{148}
\]

where we defined:

\[
\langle \hat{\tilde{n}}^{cn}(1) \rangle = -i\hbar \sum_{\sigma} G^{cn}(11^+) + n_0^{\tilde{n}}(r_1). \tag{149}
\]

The term $\Sigma_{e}^{cn}$ in Eq. (147) represents the electronic part of Hedin’s self-energy in Eq. (85), evaluated at clamped nuclei:

\[
\Sigma_{e}^{cn}(12) = i\hbar \int d(34) G^{cn}(13) \Gamma^{cn}(324) W_{e}^{cn}(41^+). \tag{150}
\]

In this expression, the vertex $\Gamma^{cn}$ and the screened Coulomb interaction $W_{e}^{cn}$ are both evaluated via the Hedin-Baym equations at clamped nuclei. Equations (147)-(150) lead directly to the well-known Hedin’s equations [122]. Hedin’s equations and the associated $GW$ method at clamped nuclei are addressed in a number of excellent reviews [126, 136, 157, 158] hence they will not be discussed here.

In order to recover the complete Hedin-Baym equation of motion, Eq. (84), starting from Eqs. (147)-(150), it is sufficient to introduce the Dyson equation:

\[
G(12) = G^{cn}(12) + \int d(34) G^{cn}(13) \Sigma^{ep}(34) G(42), \tag{151}
\]

together with the electron self-energy $\Sigma^{ep}$ arising from electron-phonon interactions:

\[
\Sigma^{ep} = \Sigma^{FM} + \Sigma^{DW} + \Sigma^{dGW}, \tag{152}
\]

where we have defined:

\[
\Sigma^{FM}(12) = i\hbar \int d(34) G(13) \Gamma(324) W_{ph}(41^+), \tag{153}
\]
\[
\Sigma^{DW}(12) = \int d3 v(13) \left[ \langle \hat{n}(3) \rangle - \langle \hat{\tilde{n}}^{cn}(3) \rangle \right] \delta(12), \tag{154}
\]
\[
\Sigma^{dGW}(12) = \Sigma_{e}(12) - \Sigma_{e}^{cn}(12). \tag{155}
\]

We emphasize that Eqs. (147)-(155) are just an alternative formulation of the Hedin-Baym equations in Table 1. The advantage of this formulation is that it better reflects standard practice, whereby the DFT equations and the $GW$ quasiparticle corrections are evaluated at clamped nuclei. Equations (147)-(155) are formally exact within the harmonic approximation.

A schematic representation of the Dyson equation for the electron Green’s function and the decomposition of the electron self-energy are given in Fig. 2. The self-energy contribution $\Sigma^{FM}$ in Eq. (153) is a dynamic correction to the electronic excitation energies, and is analogous to
Figure 2: Diagrammatic representation of the electron Green’s function and electron-phonon self-energy. (a) Dyson equation for the electron Green’s function, Eq. (151). The thick straight line represents the fully-dressed electron propagator, the thin straight line is the propagator calculated at clamped nuclei, and the disc is the electron-phonon self-energy. (b) Decomposition of the electron-phonon self-energy into Fan-Migdal self-energy, Eq. (153), Debye-Waller contribution, Eq. (154), and the remainder given by Eq. (155). (c) Fan-Migdal electron-phonon self-energy expressed in terms of the dressed electron-phonon coupling function (dark grey disc as in Fig. 1), the fully-interacting electron’s Green’s functions (thick straight line), the fully-interacting phonon propagator (thick wavy line), and the vertex $\Gamma$ from Eq. (83). (d) Debye-Waller contribution resulting from the fully interacting phonon propagator (thick wavy line) and the matrix element in Eq. (40) (hatched disc). (e) Correction to Hedin’s $GW$ self-energy arising from the modification of the electronic structure induced by the electron-phonon interaction. $W_e$ is the screened Coulomb interaction of Eq. (94) (bold dashed double line). $W_{e}^{cn}$ is the screened Coulomb interaction evaluated at clamped nuclei (thin dashed double line). $\Gamma_{cn}$ is the vertex of Eq. (83), but evaluated at clamped nuclei.
the $GW$ self-energy at clamped nuclei. Indeed, in the same way as the correlation part of the standard $GW$ self-energy describes the effect of the dynamic electronic polarization upon the addition of electrons or holes to the system, the term $GW_{\text{ph}}$ in Eq. (153) describes the effect of the dynamic polarization of the lattice.

In the semiconductors community, the self-energy obtained from Eq. (153) by setting $\Gamma(123) = \delta(13) \delta(23)$ is commonly referred to as the Fan self-energy [54, 85, 169, 170]. In the metals and superconductors communities, the same term is traditionally referred to as the self-energy in the Migdal approximation [8, 123, 124, 155]. By extension we refer to the self-energy $\Sigma_{\text{FM}}$ in Eq. (153) as the ‘Fan-Migdal’ (FM) self-energy.

The static term $\Sigma_{\text{DW}}$ in Eq. (154) corresponds to the difference between the self-consistent potential $V_{\text{tot}}$ calculated for the fully-interacting system, and the same potential evaluated with the nuclei clamped in their equilibrium positions, $V^{\text{cn}}_{\text{tot}}$. Intuitively this term corresponds to a time-independent correction to the ‘crystal potential’ that arises from the fuzziness of the nuclear charge density around the equilibrium nuclear positions. This term is similar to the one appearing in the study of the temperature dependence of X-ray diffraction and neutron diffraction spectra [79, 171], and is commonly referred to as the Debye-Waller (DW) term [172–174]. Hedin and Lundqvist did not include this term in their classic work [126], however this contribution was discussed in Refs. [85, 86].

The last term Eq. (152), $\Sigma_{\text{dGW}}$, is the correction to the standard Hedin self-energy arising from the fact that the fully-interacting electron Green’s function and density are slightly different from those evaluated at clamped nuclei, owing to the electron-phonon interaction. The magnitude of this term corresponds to a fraction of the $GWT$ quasiparticle corrections at clamped nuclei. Since $\Sigma_{\text{dGW}}$ has never been investigated so far, we will not discuss this term further.

5.2.2 Expressions for the electron self-energy used in ab initio calculations

A complete self-consistent solution of Eqs. (147)-(155) from first principles is not possible at present, and one has to replace the various entries of Eq. (152) by the best approximations available. In practice, one resorts to either DFT or to $GW$ calculations; recent progress will be reviewed in Secs. 8 and 9.

Using Eqs. (110), (130)-(131), and (144) we can rewrite the Fan-Migdal self-energy as follows:

$$\Sigma_{\text{FM}}(12) = i \sum_{\nu\nu'} \int \frac{d\omega}{2\pi} \frac{d\mathbf{q}}{\Omega_{\text{BZ}}} d(34) e^{-i\omega(t_4-t_4^+)} G(13) \Gamma(324) g_{\nu\nu'}^{cc}(\mathbf{r}_4,\omega) D_{\mathbf{q}\nu\nu'}(\omega) g_{\mathbf{q}\nu'}(\mathbf{r}_1,\omega).$$

This shows that the Fan-Migdal self-energy is, in symbolic notation, of the type $\Sigma = g^2 D G \Gamma$; a graphical representation of this term is given in Fig. 2(c). In order to make the last expression amenable to ab initio calculations, it is common to make the following approximations, which are similar to those introduced earlier for the phonon self-energy:

(i) The vertex $\Gamma(123)$ is set to $\delta(13) \delta(23)$;
(ii) The fully-interacting electron Green’s function is replaced by the Kohn-Sham Green’s function evaluated at clamped nuclei;

(iii) The fully-interacting phonon propagator $D_{q
u
u'}(\omega)$ is replaced by the adiabatic propagator $D^A_{q
u
u'}(\omega)$ from Eq. (134);

(iv) The screened electron-phonon vertex is evaluated using the RPA+$xc$ electronic screening from a DFT calculation;

(v) The frequency dependence of the electron-phonon coupling is neglected, $g_{q\nu}(r, \omega) \simeq g_{q\nu}(r, 0)$.

After using these approximations in Eq. (156), we obtain the following result for the $k = k'$ matrix elements of the Fan-Migdal self-energy in the basis of Kohn-Sham states:

$$\Sigma_{nmk}(\omega) = \frac{1}{\hbar} \sum_{m'\nu} \int \frac{d\mathbf{q}}{\Omega_{\text{BZ}}} g^*_{mn\nu}(\mathbf{k}, \mathbf{q}) g_{m'n\nu}(\mathbf{k}, \mathbf{q})$$

$$\times \left[ \frac{1 - f_{m\mathbf{k}+\mathbf{q}}}{\omega - \varepsilon_{m\mathbf{k}+\mathbf{q}}/\hbar - \omega_{q\nu} + i\eta} + \frac{f_{m\mathbf{k}+\mathbf{q}}}{\omega - \varepsilon_{m\mathbf{k}+\mathbf{q}}/\hbar + \omega_{q\nu} - i\eta} \right]. \quad (157)$$

Here $f_{m\mathbf{k}+\mathbf{q}} = 1$ for occupied Kohn-Sham states and 0 otherwise, and the matrix element $g_{mn\nu}(\mathbf{k}, \mathbf{q})$ is obtained from Eq. (38). As for the phonon self-energy in Eq. (145), also in this case the components of the electron self-energy for $k \neq k'$ vanish due to the periodicity of the lattice. The result in Eq. (157) is obtained by closing the contour of the frequency integration in the upper complex plane, owing to the $t_1^+$ in the exponential of Eq. (156). The infinitesimals inside the electron and phonon propagators, which reflect the time-ordering, are crucial to obtain the correct result [8]. The spin label is omitted in Eq. (157) since this contribution to the self-energy is diagonal in the spin indices.

The result above is only valid at zero temperature. The extension to finite temperature requires going through the Matsubara representation, and then continuing the self-energy from the imaginary axis to the real axis. The procedure is described in many textbooks, see for example Sec. 3.5 of [9]. The result is that at finite temperature the square brackets of Eq. (157) are to be modified as follows:

$$\left[ \frac{1 - f_{m\mathbf{k}+\mathbf{q}}}{\omega - \varepsilon_{m\mathbf{k}+\mathbf{q}}/\hbar - \omega_{q\nu} + i\eta} + \frac{f_{m\mathbf{k}+\mathbf{q}}}{\omega - \varepsilon_{m\mathbf{k}+\mathbf{q}}/\hbar + \omega_{q\nu} - i\eta} \right]$$

$$\rightarrow \left[ \frac{1 - f_{m\mathbf{k}+\mathbf{q}} + n_{q\nu}}{\omega - \varepsilon_{m\mathbf{k}+\mathbf{q}}/\hbar + \omega_{q\nu} - i\eta} + \frac{f_{m\mathbf{k}+\mathbf{q}} + n_{q\nu}}{\omega - \varepsilon_{m\mathbf{k}+\mathbf{q}}/\hbar - \omega_{q\nu} + i\eta} \right], \quad (158)$$

where $f_{m\mathbf{k}+\mathbf{q}}$ and $n_{q\nu}$ are now Fermi-Dirac and Bose-Einstein distribution functions, respectively (see footnote 3). The change of sign in the imaginary infinitesimal on the second fraction has to do with the fact that in the Matsubara formalism the analytic continuation from the imaginary frequency axis to the real axis through the upper complex plane leads to the so-called retarded self-energy, that is a self-energy with all poles below the real axis [9, 138].

The Debye-Waller term in Eq. (154) can also be written in a form which is convenient for practical calculations, by expanding the total density operator $\hat{n}(3)$ to second order in the atomic displacements. Using Eqs. (79) and (109) we find:

$$\Sigma^{\text{DW}}(12) = \delta(12) \frac{i\hbar}{2} \sum_{\kappa\alpha p} \frac{\partial^2 V_{\text{tot}}(1)}{\partial \tau_{\kappa\alpha}^0 \partial \tau_{\kappa'\alpha'}^0} D_{\kappa\alpha p,\kappa'\alpha' p'}(t_1^+, t_1). \quad (159)$$
In order to arrive at this result, it is necessary to make the additional approximation that the electronic field operators and the operators for the nuclear displacements are uncorrelated, that is $\langle \hat{n} \Delta \hat{\tau}_{\kappa \alpha \rho} \rangle = \langle \hat{n} \rangle \langle \Delta \hat{\tau}_{\kappa \alpha \rho} \rangle$, and similarly for the second power of the displacements. This requirement was noticed by Gillis [128], and is trivially satisfied if we describe phonons within the adiabatic approximation of Sec. 5.1.1. Equation (159) motivates the diagrammatic representation of the Debye-Waller self-energy shown in Fig. 2(d), whereby the phonon line begins and ends at the same time point. We note that Eq. (159) involves the variation of the screened potential $V_{\text{tot}}$; this result, which we here derived starting from Schwinger’s functional derivative technique, is also obtained when starting from a perturbative diagrammatic analysis [133]. The Debye-Waller self-energy can be simplified further if we make use of the following approximations, in the same spirit as for the Fan-Migdal self-energy:

(vi) The fully-interacting phonon propagator is replaced by the adiabatic propagator $D_{A q q'}^A(\omega)$ from Eq. (134);

(vii) The total many-body potential $V_{\text{tot}}$ of Eq. (79) is replaced by the Kohn-Sham potential $V_{\text{KS}}(r)$ evaluated at clamped nuclei. Strictly speaking, the Kohn-Sham effective potential includes also contributions from exchange and correlation, which in the Hedin-Baym equations are all contained in the electron self-energy. However, the present discussion holds unchanged if we add any local and frequency-independent potential to $V_{\text{tot}}$ in Eq. (84), while removing the same potential from the self-energy [127].

Using these simplifications together with Eqs. (130) and (131), we can write the $\Sigma_{\text{DW}}$ in the basis of Kohn-Sham eigenstates as follows:

$$\Sigma_{nn'k}^{\text{DW}} = \sum_{\nu} \int \frac{dq}{\Omega_{\text{BZ}}} g_{nn'\nu\nu}^{\text{DW}}(k, q, -q),$$

(160)

where the Debye-Waller matrix element $g_{nn'\nu\nu}^{\text{DW}}$ is obtained from Eq. (40), and the presence of only the diagonal terms $\nu = \nu'$ is a result of the Kronecker delta in Eq. (132). In going from Eq. (159) to Eq. (160) the frequency integration is performed by using Eq. (134), after closing the contour in the lower half plane. The resulting expression is diagonal in the spin indices.

The expression for the Debye-Waller term in Eq. (160) is only valid at zero temperature. In this case the extension to finite temperature is immediate since the self-energy does not involve the electron Green’s function, hence we only need to evaluate the canonical average of Eq. (133) at equal times. The result is that Eq. (160) is simply to be modified as follows:

$$g_{nn'\nu\nu}^{\text{DW}}(k, q, -q) \longrightarrow g_{nn'\nu\nu}^{\text{DW}}(k, q, -q)(2n_{q\nu} + 1),$$

(161)

with $n_{q\nu}$ being the Bose-Einstein occupations (see footnote 3).

The Debye-Waller contribution to the electron self-energy is almost invariably ignored in the literature on metals and superconductors, but it is well-known in the theory of temperature-dependent band structures of semiconductors [54, 85, 104, 105, 112]. Neglecting $\Sigma_{\text{DW}}$ in metals is partly justified by the fact that this term is frequency-independent, therefore it is expected to be a slowly-varying function over each Fermi surface sheet. A detailed first-principles analysis of this aspect is currently lacking.
5.2.3 Temperature-dependence of electronic band structures

Once determined the electron self-energy as in Sec. 5.2.2 it is possible to study the modification of the electronic structure induced by the EPI. To this aim it is convenient to rewrite Eq. (151) in the basis of Kohn-Sham eigenstates:

\[ G_{nn'}^{n} (\omega) = G_{n}^{cn} - \Sigma_{ep}^{nn'} (\omega). \]  

(162)

Assuming that the electronic structure problem at clamped nuclei has been solved using DFT or DFT+GW calculations, the Green’s function \( G_{n}^{cn} \) can be written in terms of simple poles at the Kohn-Sham or quasiparticle eigenvalues \( \tilde{\epsilon}_{nk} \). In this case Eq. (162) reduces to:

\[ G_{nn'}^{-1} (\omega) = (\hbar \omega - \tilde{\epsilon}_{nk}) \delta_{nn'} - \Sigma_{ep}^{nn'} (\omega), \]  

(163)

where \( \tilde{\epsilon}_{nk} = \epsilon_{nk} \pm i \hbar \eta \) with the upper/lower sign corresponding to occupied/unoccupied states. The spin indices are omitted since these self-energy contributions do not mix states with opposite spin.

In order to gain insight into the effects of the electron-phonon interaction, we start from the drastic approximation that \( \Sigma_{ep} \) only leads to a small shift of the quasiparticle poles, from the ‘non-interacting’ energies \( \epsilon_{nk} \) to the renormalized energies \( \tilde{E}_{nk} = E_{nk} + i \Gamma_{nk} \). In this approximation, the fully interacting Green’s function is expressed as a sum of simple poles, given by the zeros of Eq. (163):

\[ E_{nk} = \epsilon_{nk} + \text{Re} \Sigma_{ep}^{nn'} (\tilde{E}_{nk}/\hbar), \]  

(164)

\[ \Gamma_{nk} = \text{Im} \Sigma_{ep}^{nn'} (\tilde{E}_{nk}/\hbar). \]  

(165)

As in the case of vibrational frequencies in Eq. (137), we are considering for simplicity non-degenerate electronic states, and making the assumption that the off-diagonal elements of the self-energy \( \Sigma_{ep}^{nn'} \) with \( n \neq n' \) can be neglected. In more general situations the right-hand side of Eq. (163) needs to be diagonalized, or alternatively the off-diagonal terms \( \Sigma_{ep}^{nn'} \) need to be treated perturbatively. The energies \( E_{nk} \) obtained from Eq. (164) yield the band structure renormalized by the EPIs, to be discussed below. The imaginary part \( \Gamma_{nk} \) in Eq. (165) is connected with the quasiparticle lifetimes and will be discussed in Sec. 5.2.4.

Equation (164) is to be solved self-consistently for \( E_{nk} \) and \( \Gamma_{nk} \). When Eq. (164) is used in combination with the standard approximations to the Fan-Migdal and Debye-Waller self-energies given by Eqs. (157) and (160), the result that one obtains is equivalent to describing electron-phonon couplings to second order in Brillouin-Wigner perturbation theory [9]. Similarly one recovers the more basic Rayleigh-Schrödinger perturbation theory by making the replacements \( E_{nk} \rightarrow \epsilon_{nk} \) and \( \Gamma_{nk} \rightarrow 0 \) in Eq. (164).

By combining Eqs. (157)-(158), (160)-(161), and (164), we obtain the temperature-dependent
band structure renormalization’ arising from the EPI:

\[ E_{nk} = \varepsilon_{nk} + \frac{1}{\Omega_{BZ}} \sum_{\nu} \int d\mathbf{q} |g_{mn\nu}(\mathbf{k}, \mathbf{q})|^2 \]

\[ \times \text{Re} \left[ \frac{1 - f_{mk+\mathbf{q}} + n_{\mathbf{q}\nu}}{E_{nk} - \varepsilon_{mk+\mathbf{q}} - \hbar \omega_{\mathbf{q}\nu} + i\Gamma_{nk}} + \frac{f_{mk+\mathbf{q}} + n_{\mathbf{q}\nu}}{E_{nk} - \varepsilon_{mk+\mathbf{q}} + \hbar \omega_{\mathbf{q}\nu} + i\Gamma_{nk}} \right] \]

\[ + \sum_{\nu} \frac{d\mathbf{q}}{\Omega_{BZ}} g_{mn\nu}^{DW}(\mathbf{k}, \mathbf{q}, -\mathbf{q})(2n_{\mathbf{q}\nu} + 1). \]  \( (166) \)

For practical calculations it is important to bear in mind that this result rests on the approximations (i)-(vii) introduced at p. 49, as well as the harmonic approximation.

The theory of temperature-dependent band structures developed by Allen and Heine \[85\] makes two additional approximations on top of Eq. (166): Brillouin-Wigner perturbation theory is replaced by Rayleigh-Schrödinger perturbation theory; and the phonon energies in the denominators are neglected. Using these additional approximations Eq. (166) becomes:

\[ E_{nk} = \varepsilon_{nk} + \frac{1}{\Omega_{BZ}} \sum_{\nu} \int d\mathbf{q} |g_{mn\nu}(\mathbf{k}, \mathbf{q})|^2 \left( \sum_{m} \varepsilon_{nk} - \varepsilon_{mk+\mathbf{q}} + g_{mn\nu}^{DW}(\mathbf{k}, \mathbf{q}, -\mathbf{q})(2n_{\mathbf{q}\nu} + 1) \right), \]  \( (167) \)

which is referred to as the ‘adiabatic Allen-Heine formula’. By setting \( T = 0 \) the Bose-Einstein factors \( n_{\mathbf{q}\nu} \) vanish and we have the so-called ‘zero-point renormalization’ of the energy bands, \( \Delta E_{ZP}^{nk} = E_{nk}(T = 0) - \varepsilon_{nk} \). This is the modification of the electronic energies evaluated at clamped nuclei, which arises from the zero-point fluctuations of the atoms around their equilibrium sites.

An expression that is essentially identical to Eq. (167) can also be obtained directly from Eq. (1) using second-order Raleigh-Schrödinger perturbation theory in Fock space, following the same lines as in Ref. \( \text{[78]} \), p. 134. A detailed derivation of the formalism starting from Eq. (1) is given in Ref. \( \text{[175]} \).

Historically, the Allen-Heine theory \[85\] was developed by starting from a straightforward perturbation theory expansion of the electron energies in terms of the atomic displacements within the adiabatic approximation, followed by a canonical average of the displacements using Bose-Einstein statistics. It is reassuring that, after making a few well-defined approximations, a field-theoretic method leads to the same result.

Equation (167) was employed in many semi-empirical calculations\(^4\). More recently, this expression was used in the context of first-principles DFT calculations by Marini \[104\] and Giustino et al. \[105\]. DFT calculations of band structure renormalization based on Eqs. (166) or (167) are becoming increasingly popular, and the latest developments will be reviewed in Sec. 9.1.1.

The nature of the band structure renormalization by electron-phonon interactions can be understood at a qualitative level by considering a drastically simplified model, consisting of a semiconductor with parabolic and nondegenerate valence and conduction bands, with the band extrema coupled to all other states by a dispersionless phonon mode of frequency \( \omega_0 \). If the Debye-Waller

\(^4\) See for example Refs. \[54, 55, 176-179\]. Detailed reviews of early calculations and comparison to experiments can be found in Refs. \[170, 174, 180\].
matrix elements are much smaller than the Fan-Migdal matrix elements, then the dominant contributions to Eq. (167) arise from denominators such as 
\[ \epsilon_{nk} - \epsilon_{nk+q} \simeq \pm h^2 |q|^2 / 2 m^*_n, \]
where the upper/lower sign is for the valence/conduction band, and \( m^*_n \) are the effective masses. As a result the temperature dependence of the band gap will take the form:

\[ E_g(T) = E_{cn}^g - |\Delta E_{gZP}^Z| \left[ 1 + 2 n \left( \hbar \omega_0 / k_B T \right) \right], \]

(168)

where \( E_{cn}^g \) is the gap at clamped nuclei, and \( |\Delta E_{gZP}^Z| \) is the zero-point correction. The negative sign in the last expression arises from the opposite curvatures of the valence and conduction bands. In this example the band gap decreases with temperature: this is a well known effect in semiconductor physics, and is often referred to as the ‘Varshni effect’ [181]. The first measurements of such effects were performed by Becker and Fan [182], and stimulated the development of the first theory of temperature-dependent band gaps [169]. A schematic illustration of this qualitative model is provided in Fig. 3(a). The redshift of the gap as a function of temperature is seen in many albeit not all semiconductors. For example, copper halides [183] and lead halide perovskites [184] exhibit an ‘inverse Varshni’ effect, that is a blueshift of the gap with temperature; in addition some chalcopyrites exhibit a non-monotonic temperature dependence of the band gap [185]. We also point out that the qualitative model shown in Fig. 3(a) does not take into account the subtle temperature-dependence of the band gap renormalization at very low temperature. These effects were recently investigated by Allen and Nery [186].

### 5.2.4 Carrier lifetimes

While the real part of the poles in Eq. (164) describes the energy level renormalization induced by the electron-phonon coupling, the imaginary part \( \Gamma_{nk} \) in Eq. (165) carries information on the spectral broadening, which will be discussed in Sec. (5.2.5), and on quasiparticle lifetimes, which we discuss below.

After transforming \( G_{nn'k}(\omega) \) from Eq. (163) into the time domain it is seen that, for an electron or hole added to the system at time \( t \) in the state \( |nk\rangle \), the probability amplitude to persist in the same state decreases as \( \exp[-\Gamma_{nk}(t' - t)/\hbar] \). Using Eqs. (157) and (165) it can be seen that \( \Gamma_{nk} < 0 \) for an electron added to the system and \( \Gamma_{nk} > 0 \) for a hole. Therefore the average time spent by the particle in the state \( |nk\rangle \) is \( \tau_{nk} = \hbar / (2 |\Gamma_{nk}|) \).

A popular expression for the electron and hole lifetimes is obtained by making the replacement \( \tilde{E}_{nk} \to \epsilon_{nk} \) in Eq. (157), and by taking the absolute value of the imaginary part. We find:

\[ \frac{1}{\tau_{nk}} = \frac{2\pi}{\hbar} \sum_{m\nu} \int \frac{d\mathbf{q}}{\Omega_{BZ}} |g_{nm\nu}(\mathbf{k}, \mathbf{q})|^2 \times \left| (1 - f_{mk+q}) \delta(\epsilon_{nk} - \hbar \omega_{\nu q} - \epsilon_{mk+q}) - f_{mk+q} \delta(\epsilon_{nk} + \hbar \omega_{\nu q} - \epsilon_{mk+q}) \right|. \]

(169)

A more accurate expression is discussed after Eq. (174) in the next section. The extension of the above result to finite temperature is obtained by taking the absolute value of the imaginary
Figure 3: (Color online) Temperature-dependent band gap and lifetimes in an idealized semiconductor or insulator. (a) Temperature dependence of the band gap according to Eq. (168) (thick solid blue line). The straight thin black line is the asymptotic expansion at high temperature; this line intercepts the vertical axis at the band gap calculated with clamped nuclei, $E_{cn}^g$. The difference between the latter value and the band gap at $T=0$ including the EPI gives the zero-point renormalization, $\Delta E_{ZP}^g$. (b) Temperature dependence of the electron linewidth (solid blue line) and lifetimes (dashed red line) using the same model as in (a). The zero-point broadening is $\Gamma_{n,k}^{ZP}$. This simplified trend is only valid when the electron energy is at least one phonon energy away from a band extremum, so that both phonon emission and phonon absorption processes are allowed. The parameters of the model are: $E_{cn}^g = 1$ eV, $\Delta E_{ZP}^g = 100$ meV, $\hbar \omega_0 = 100$ meV, $\Gamma_{n,k}^{ZP} = 50$ meV; these values are representative of common semiconductors.

We emphasize the change of sign in the third line, resulting from the analytic continuation to the retarded self-energy. Equation (170) coincides with the expression that one would obtain by using the standard Fermi golden rule [6]. The intuitive interpretation of this result is that the quasiparticle lifetime is reduced by processes of phonon emission and absorption, corresponding to the second and third lines of Eq. (170), respectively. We note that in deriving Eq. (170) we did not consider the Debye-Waller self-energy; this is because the diagonal matrix elements of $\Sigma^{DW}$ are purely real, hence they do not contribute to the quasiparticle widths [187]. *Ab initio* calculations of carrier lifetimes using Eq. (170) were first reported by Eiguren et al. [188, 189]. These applications and more recent developments will be reviewed in Sec. 10.1.

If we evaluate Eq. (170) for the same simplified model introduced for the temperature renor-
malization, and we neglect the phonon energy in the Dirac delta functions, we obtain \( \Gamma_{nk}(T) = \Gamma_{nk}^{ZP} [1 + 2 n(\hbar \omega_0/\hbar g T)] \) where \( \Gamma_{nk}^{ZP} \) is the linewidth at \( T = 0 \). The dependence of the linewidth and the corresponding lifetime on temperature for this model are shown in Fig. 3(b). This trend is typical in semiconductors \[190, 191\].

5.2.5 Kinks and satellites

In many cases of interest, the use of Brillouin-Wigner perturbation theory as given by Eqs. (164)-(165) is not sufficient to provide an adequate description of EPIs, and it becomes necessary to go back to the complete Dyson equation, Eq. (163). Generally speaking a direct solution of the Dyson equation is important in all those cases where the electronic energy scales are comparable to phonon energies, namely in metals (including superconductors), narrow-gap semiconductors, and doped semiconductors. In order to study these systems, it is convenient to introduce an auxiliary function called the ‘spectral density function’, or simply spectral function.

In its simplest version the spectral function is defined as \[9, 138\]:

\[
A(k, \omega) = \frac{-1}{\pi} \sum_n \text{Im} G_{nnk}^{\text{ret}}(\omega),
\]

where the superscript ‘ret’ stands for ‘retarded’, and simply indicates that all poles of the Green’s function \( G_{nnk}(\omega) \) in the upper complex plane must be replaced by their complex conjugate.

The spectral function is positive definite and carries the meaning of a ‘many-body momentum-resolved density of states’ \[138\]. This is precisely the function that is probed by angle-resolved photoelectron spectroscopy experiments or ARPES \[192\]. Using Eq. (163) the spectral function can be rewritten as:

\[
A(k, \omega) = \sum_n \frac{-1}{\pi} \frac{\text{Im} \Sigma_{nnk}^{\text{ep}}(\omega)}{[\hbar \omega - \varepsilon_{nk} - \text{Re} \Sigma_{nnk}^{\text{ep}}(\omega)]^2 + [\text{Im} \Sigma_{nnk}^{\text{ep}}(\omega)]^2}.
\]

In order to obtain the correct spectral function, it is important to use the \textit{retarded} self-energy. This is done by using Eq. (158) for the Fan-Migdal term, while the static Debye-Waller term remains unchanged.

It is often convenient to approximate the spectral function as a sum of quasiparticle peaks. To this aim, one performs a linear expansion of Eq. (172) around each quasiparticle energy \( E_{nk} \), to obtain:

\[
A(k, \omega) = \sum_n Z_{nk} \frac{-1}{\pi} \frac{Z_{nk} \text{Im} \Sigma_{nnk}^{\text{ep}}(E_{nk}/\hbar)}{[\hbar \omega - E_{nk}]^2 + [Z_{nk} \text{Im} \Sigma_{nnk}^{\text{ep}}(E_{nk}/\hbar)]^2}.
\]

This is a sum of Lorentzians with strength \( Z_{nk} \) and width \( Z_{nk} \text{Im} \Sigma_{nnk}^{\text{ep}}(E_{nk}/\hbar) \). Here the ‘quasiparticle strength’ is defined as the homonymous quantity appearing in \textit{GW} calculations \[126\]:

\[
Z_{nk} = \left[ 1 - h^{-1} \partial \text{Re} \Sigma_{nnk}^{\text{ep}}(\omega)/\partial \omega \bigg|_{\omega=E_{nk}/\hbar} \right]^{-1}.
\]

The result expressed by Eq. (173) shows that, in a rigorous field-theoretic approach, the quasiparticle broadening and lifetime given by Eqs. (169) and (170) should be renormalized by \( Z_{nk} \) and \( Z_{nk}^{-1} \), respectively, and should be evaluated using the quasiparticle energy \( E_{nk} \) instead of \( \varepsilon_{nk} \).
Figure 4: (Color online) Two-dimensional maps of the electron spectral function $A(k, \omega)$ for electrons coupled to a dispersionless phonon of frequency $\omega_0$. The non-interacting bands are given by $\varepsilon(k) = -\varepsilon_F + \hbar^2 |k|^2 / 2m^*$, and the Fermi level coincides with the top of the energy window. The matrix element is $|g|^2 = \hbar \omega_0 / N_F$ when the electron energies differ by less than the cutoff $\varepsilon_{\text{max}}$, and zero otherwise ($N_F$ is the density of states at the Fermi level). (a) Spectral function for the case $\varepsilon_F = 10 \hbar \omega_0$ (white on blue/black), non-interacting band structure (solid line, yellow/light gray), and fully-interacting band structure within Brillouin-Wigner perturbation theory (solid line, red/dark gray). (b) Spectral function for the case $\varepsilon_F = 2 \hbar \omega_0$. The model parameters are: $m^* = 0.1 m_e$, $\hbar \omega_0 = 100$ meV, $\eta = 20$ meV, $\varepsilon_c = 5$ eV. For clarity the calculated spectral functions are cut off at the value $3$ eV$^{-1}$ and normalized. The self-energy is shifted by a constant so as to have $\Sigma_{\text{ep}}(0) = 0$; this correction guarantees the fulfillment of Luttinger’s theorem about the volume enclosed by the Fermi surface [193].

This result can also be derived from Eq. (165) by performing a Taylor expansion of the self-energy along the imaginary axis and using the Cauchy-Riemann conditions.

In order to illustrate the typical features of the spectral function, we consider a model system characterized by one parabolic conduction band. The occupied electronic states couple to all states within an energy cutoff via a dispersionless phonon mode and a constant electron-phonon matrix element. A simplified version of this model was discussed by Engelsberg and Schrieffer by considering a constant density of electronic states [124]. By evaluating the spectral function in Eq. (172) using the Fan-Migdal self-energy and neglecting the Debye-Waller term, we obtain the results shown in Fig. 4 for two sets of parameters.

In Fig. 4(a) the Fermi energy is much larger than the characteristic phonon energy. This case is representative of a metallic system with electron bands nearly linear around the Fermi level. Here the electron-phonon interaction leads to (i) a reduction of the band velocity in proximity of the Fermi level, and (ii) a broadening of the spectral function beyond the phonon energy $\hbar \omega_0$. A detailed analysis of these features for a slightly simpler model system, including a discussion...
of the analytic properties of the Green’s function, can be found in Ref. [124].

The solid line (red/dark gray) in Fig. 4(a) shows the renormalized band structure obtained from Brillouin-Wigner perturbation theory, Eq. (164). We see that these solutions track the maxima of the spectral function \(A(k, \omega)\). The renormalized bands exhibit a characteristic ‘S-shape’ near the Fermi level, corresponding to multiple solutions of Eq. (164) for the same wavevector \(k\). Starting from the late 1990s such S-shaped energy-momentum dispersion curves have been observed in a number of ARPES experiments, and have become known in the literature as the ‘photoemission kink’ [194]. First-principles calculations of kinks were first reported in Refs. [189, 195], and will be reviewed in Sec. 8.

In Fig. 4(b) the Fermi energy is comparable to the characteristic phonon energy. This case is representative of a degenerately doped semiconductor close to a conduction band minimum. Here the electron-phonon interaction leads to two distinct spectral features: (i) a parabolic band with a heavier mass, which is well described by the Brillouin-Wigner solutions (solid line, red/dark gray), and (ii) a polaron satellite that is visible further down. In this example, it is clear that Eq. (164) is unable to describe the satellite, and that the spectral function carries qualitatively new information about the system. Polaron satellites resembling Fig. 4(b) have been observed in ARPES experiments on doped oxides [196–199] and recently calculated from first principles [200].

5.2.6 Model Hamiltonians, polarons, and the cumulant expansion

At the end of this section it is worth mentioning complementary non first-principles approaches for studying the effects of EPIs on the electronic properties of solids. Model EPI Hamiltonians can be derived from Eq. (1) by choosing a priori explicit expressions for the electron band energies, the vibrational frequencies, and the coupling matrix elements. Examples of model Hamiltonians are those of Fröhlich [30], Holstein [201], Su, Schrieffer, and Heeger [202], the Hubbard-Holstein model [203], the Peierls-Hubbard model [204], the ‘\(t-J\)’ Holstein model [205], and the Su-Schrieffer-Heeger-Holstein model [206]. These models involve the tight-binding approximation, the Einstein phonon spectrum, and electron-phonon couplings to first order in the atomic displacements. Using these model Hamiltonians it is possible to go beyond the approximations introduced in Sec. 5.2.2 and obtain non-perturbative solutions by means of canonical Lang-Firsov transformations, path-integral methods, exact diagonalization, variational or quantum Monte Carlo techniques [10, 207]. These models have been used extensively to explore many aspects of polaron physics, for example the ground-state energy of polarons (weak or strong coupling), their spatial extent (large or small polarons), and transport properties (band-like or hopping-like).

Given the considerable body of literature on model EPI Hamiltonians, it is natural to ask whether one could bring ab initio calculations of EPIs to a similar level of sophistication. The main limitation of current first-principles approaches is that, given the complexity of the calculations, the electron self-energies are evaluated using the bare propagators, as in Eq. (157). As a consequence, higher-order interaction diagrams beyond the Migdal approximation [123] are omitted altogether.

A promising avenue for going beyond the Migdal approximation consists of introducing higher-
order diagrams via the ‘cumulant expansion’ approach \[126, 208, 209\]. In the cumulant expansion method, instead of calculating the electron Green’s function via a Dyson equation, one evaluates the time evolution of the Green’s function by formulating the problem in the interaction picture, in symbols: \( G_{nnk}(t) = \langle i/\hbar \rangle \exp\left[-i(\varepsilon_{nk}/\hbar)t + C_{nnk}(t)\right] \) \[209\]. The distinctive advantage of this approach is that the ‘cumulant’ \( C_{nnk}(t) \) can be obtained from a low-order self-energy, for example the Fan-Migdal self-energy in Eq. \[157\], and the exponential ‘resummation’ automatically generates higher-order diagrams (Ref. \[9\], pag. 523). Detailed discussions of the cumulant expansion formalism can be found in Refs. \[210, 211\].

The cumulant method provides an interesting point of contact between \textit{ab initio} and model Hamiltonian approaches. In fact, the cumulant expansion is closely related to the ‘momentum average approximation’ introduced by Berciu for studying the Green’s function of the Holstein polaron \[212\].

The cumulant expansion has proven successful in \textit{ab initio} calculations of electron-electron interactions, in particular valence band satellites in semiconductors \[213–220\]. In the context of EPIs, the \textit{ab initio} cumulant expansion method has been applied to elemental metals by Story \textit{et al.} \[221\], and to the ARPES spectra of \( n \)-doped TiO\(_2\) by Verdi \textit{et al.} \[200\]. In the latter work the cumulant method correctly reproduced the polaron satellites observed in the experiments of Moser \textit{et al.} \[196\].

The study of polarons using \textit{ab initio} many-body techniques is yet to begin, however a first calculation of the spectral function of Fröhlich polarons and an approximate polaron wavefunction have recently been reported \[200\].

6 Efficient calculations of matrix elements and their integrals

The study of EPIs from first principles requires evaluating Brillouin-zone integrals of functions that exhibit strong fluctuations. This requirement can be appreciated by inspecting Eqs. \(145\) and \(157\): there the denominators become large whenever the difference between two electronic eigenvalues approaches a phonon energy. As a result, while in DFT total energy calculations the Brillouin zone is typically discretized using meshes of the order of \(10 \times 10 \times 10\) points, the numerical convergence of EPI calculations requires much finer grids, sometimes with as many as \(10^6\) wavevectors \[112, 165\]. Determining vibrational frequencies \(\omega_{q\nu}\) and perturbations \(\Delta_{q\nu}v^{KS}(r)\) for such a large number of wavevectors is a prohibitive task, since every calculation is roughly as expensive as one total energy minimization.

These difficulties stimulated the development of specialized numerical techniques for making calculations of EPIs affordable. In the following sections two such techniques are reviewed: electron-phonon Wannier interpolation and Fermi-surface harmonics.
6.1 Wannier interpolation

6.1.1 Maximally-localized Wannier functions

In addition to the standard description of electrons in solids in terms of Bloch waves, as in Eq. (29), it is possible to adopt an alternative point of view whereby electrons are described as linear combinations of localized orbitals called ‘Wannier functions’ [222]. The most general relation between Wannier functions and Bloch waves can be written as follows. One considers electron bands \( \varepsilon_{nk} \) with eigenfunctions \( \psi_{nk} \), where the index \( n \) is restricted to a set of bands that are separated from all other bands by finite energy gaps above and below. These bands are referred to as ‘composite energy bands’ [223]. Wannier functions are defined as:

\[
\begin{align*}
\w_{mp}(r) & = N_p^{-1} \sum_{nk} \mathcal{U}_{nmk} \psi_{nk}(r),
\end{align*}
\]

(175)

where \( U_{nmk} \) is a unitary matrix in the indices \( m \) and \( n \). From this definition and Eq. (213) it follows that Wannier functions are normalized in the supercell, \( \langle \w_{mp} | \w_{mp}' \rangle_{sc} = \delta_{mp,mp'} \). Furthermore, since \( u_{nk} \) is lattice-periodic, Wannier functions have the property \( \w_{mp}(r) = \w_{m0}(r - R_p) \).

The inverse transformation of Eq. (175) is obtained by using the unitary character of \( U_{nmk} \) together with Eq. (213):

\[
\begin{align*}
\psi_{nk}(r) & = \sum_{mp} \mathcal{U}_{nmk}^\dagger \mathcal{U}_{nmk} \w_{mp}(r).
\end{align*}
\]

(176)

The unitary matrix \( U_{nmk} \) is completely arbitrary, therefore there exists considerable freedom in the construction of Wannier functions. For example by requiring that \( U_{nm,-k} = U_{nmk}^\dagger \) one can make Wannier functions real-valued. Marzari and Vanderbilt exploited this degree of freedom to construct Wannier functions that are maximally localized [223].

A comprehensive and up-to-date review of the theory and applications of maximally-localized Wannier functions (MLWFs) can be found in Ref. [224]. Here we only recall that, in order to minimize the spatial extent of a function in a periodic solid, one needs to use a modified definition of the position operator, since the standard position operator is unbounded in an infinite crystal. This procedure is now well-established and it is linked to the development of the modern theory of dielectric polarization [225, 226]. Nowadays it is possible to determine MLWFs routinely [227]. The original algorithm of Marzari and Vanderbilt [223] was also extended to deal with situations where a composite set of bands cannot be identified. This happens notably in metals for electronic states near the Fermi energy. For these cases, Souza et al. developed a band ‘disentanglement’ procedure, which extracts a subset of composite bands out of a larger set of states [228].

For the purposes of the present article, the most important property of MLWFs is that they are exponentially localized in insulators, in the sense that \( |w_{m0}(r)| \sim |r|^{-\alpha} \exp(-h|r|) \) for large \( |r| \), with \( \alpha, h > 0 \) real parameters. This property was demonstrated in one spatial dimension in Refs. [229, 230], and in two and three dimensions in Ref. [231], under the condition that the system exhibits time-reversal symmetry. In the case of metallic systems, no exponential localization is expected. However, the Wannier functions obtained in metals using the disentanglement procedure of Ref. [228] are typically highly localized.
MLWFs are usually comparable in size to atomic orbitals, and this makes them ideally suited for Slater-Koster interpolation of band structures. This concept was successfully employed in a number of applications requiring accurate calculations of band velocities, effective masses, density of states, Brillouin-zone integrals, and transport coefficients.

6.1.2 Interpolation of electron-phonon matrix elements

Wannier functions were introduced in the study of EPIs by Giustino et al. The starting point is the definition of the electron-phonon matrix element in the Wannier representation:

\[ g_{mn\kappa\alpha}(\mathbf{R}_p, \mathbf{R}_p') = \langle m_0 | \frac{\partial V^{KS}}{\partial \tau_{\kappa\alpha}}(\mathbf{r} - \mathbf{R}_p') | n\rangle \langle w_{m0}(\mathbf{r}) | w_{n0}(\mathbf{r} - \mathbf{R}_p) \rangle_{sc}, \]  

(177)

where the subscript ‘sc’ indicates that the integral is over the BvK supercell. The relation between these quantities and the standard EPI matrix elements \( g_{mn\nu}(k, \mathbf{q}) \) is found by replacing Eq. (176) inside Eq. (38), and using Eqs. (34)-(35):

\[ g_{mn\nu}(k, \mathbf{q}) = \sum_{p'p} e^{i(kR_p + qR_{p'})} \sum_{m'n'\kappa\alpha} U_{mm'k+q} U_{n'n'k}^\dagger U_{\nu}^\dagger, \]  

(178)

where we defined \( u_{\kappa\alpha,\nu\nu'} = (h/2M_k\omega_{k\nu})^{1/2} \epsilon_{\kappa\alpha,\nu}(q) \) and \( \epsilon_{\kappa\alpha,\nu}(q) \) are the vibrational eigenmodes of Eq. (15). The inverse relation is:

\[ g_{mn\kappa\alpha}(\mathbf{R}_p, \mathbf{R}_p') = \frac{1}{N_pN_{p'}} \sum_{k,q} e^{-i(kR_p + qR_{p'})} \sum_{m'n'\nu} u_{\kappa\alpha,\nu\nu'}^{-1} U_{mm'k+q}^\dagger g_{m'n'\nu}(k, \mathbf{q}) U_{n'n'k}, \]  

(179)

with \( u_{\kappa\alpha,\nu\nu'}^{-1} = (h/2M_k\omega_{k\nu})^{-1/2} \epsilon_{\kappa\alpha,\nu}(q) \). The last two equations define a generalized Fourier transform of the electron-phonon matrix elements between reciprocal space and real space. In Eq. (179) we have \( N_p \) and \( N_{p'} \) to indicate that the BvK supercells for electronic band structures and phonon dispersions may not coincide.

If the quantity \( g_{mn\kappa\alpha}(\mathbf{R}_p, \mathbf{R}_p') \) decays rapidly as a function of \( |\mathbf{R}_p| \) and \( |\mathbf{R}_p'| \), then only a small number of matrix elements in the Wannier representation will be sufficient to generate \( g_{mn\nu}(k, \mathbf{q}) \) anywhere in the Brillouin zone by means of Eq. (178). The dependence of the matrix elements on \( \mathbf{R}_p \) and \( \mathbf{R}_p' \) can be analyzed by considering the following bound: \( |g_{mn\kappa\alpha}(\mathbf{R}_p, \mathbf{R}_p')| \leq \int_{sc} d\mathbf{r} |w_{m0}^*(\mathbf{r})w_{n0}(\mathbf{r} - \mathbf{R}_p)| \times \int_{sc} d\mathbf{r} |\partial V^{KS}/\partial \tau_{\kappa\alpha}(\mathbf{r} - \mathbf{R}_p)| \). The first term guarantees that the matrix element decays in the variable \( \mathbf{R}_p \) at least as fast as MLWFs. As a result the worst case scenario corresponds to the choice \( \mathbf{R}_p = 0 \). In this case, the matrix element \( |g_{mn\kappa\alpha}(0, \mathbf{R}_p')| \) decays with the variable \( \mathbf{R}_p' \) at the same rate as the "screened" electric dipole potential generated by the atomic displacement \( \Delta \tau_{\kappa\alpha} \). In non-polar semiconductors and insulators, owing to the analytical properties of the dielectric matrix, this potential decays at least as fast as a quadrupole, that is \( |\mathbf{R}_p'|^{-3} \). As a result, all matrix elements in reciprocal space are finite for \( \mathbf{q} \to 0 \) and hence amenable to interpolation. In the case of metals the asymptotic trend of \( \partial V^{KS}/\partial \tau_{\kappa\alpha} \) is dictated by Fermi-surface nesting, leading to Friedel oscillations that decay as

\footnote{We note that \( g_{mn\kappa\alpha}(\mathbf{R}_p, \mathbf{R}_p') \) has dimensions of energy by length, at variance with Eq. (38). For consistency here we use a definition that differs from that given in \cite{165} by a factor \( N_p \); this factor is inconsequential.}
Figure 5: (Color online) Spatial decay of the electron-phonon matrix elements of diamond in the Wannier representation: (a) $\max |g_{mn\kappa\alpha}(R_p,0)|$ vs. $|R_p|$, and (b) $\max |g_{mn\kappa\alpha}(0,R_{p'})| \text{ vs. } |R_{p'}|$. The maximum values are taken over all subscript indices, and the data are normalized to the largest value. The insets show the same quantities in logarithmic scale. The calculations were performed using the local-density approximation to DFT. Reproduced with permission from Ref. [165], copyright (2007) by the American Physical Society.

$|R_{p'}|^{-4}$ (Ref. [135], pp. 175–180). These oscillations are connected to the Kohn anomalies in the phonon dispersion relations [237]. In practical calculations, Friedel oscillations are usually not an issue since they are suppressed by the numerical smearing of the Fermi-Dirac occupations, and a Yukawa-type exponential decay is recovered. The case of polar materials is more subtle and will be discussed in Sec. 6.1.3. Figure 5 illustrates the spatial decay of $|g_{mn\kappa\alpha}(R_p,R_{p'})|$ as a function of $R_p$ and $R_{p'}$ for the prototypical case of diamond.

The interpolation strategy is entirely analogous to standard techniques for generating phonon dispersion relations using the interatomic force constants [102]: one first determines matrix elements in the Bloch representation using DFPT on a coarse grid in the Brillouin zone, as in Sec. 3.2.3. Then MLWFs are determined using the procedures of Refs. [223, 228]. This yields the rotation matrices $U_{mnk}$ to be used in Eq. (179). The Fourier transform to real space is performed via Eq. (179). At this point, one assumes that matrix elements outside of the Wigner-Seitz supercell defined by the coarse Brillouin-zone grid can be neglected, and uses Eq. (178) in order to obtain the matrix elements $g_{mn\nu}(k,q)$ on very fine grids. The last step requires the knowledge of the rotation matrices $U_{mnk}$ also on the fine grids; these matrices are obtained from the Wannier interpolation of the band structures, as described in Ref. [228]. The operation is computationally inexpensive and enables the calculation of millions of electron-phonon matrix elements. The procedure can now be applied routinely [238, 239]. Figure 6 shows the matrix elements obtained using this method, as compared to explicit DFPT calculations.

Wannier interpolation of electron-phonon matrix elements was successfully employed in a number of applications, ranging from metal and superconductors to semiconductors and nanoscale systems.

---

\footnote{See for example Refs. [105, 150, 195, 240, 251].}
Figure 6: (Color online) Comparison between Wannier-interpolated electron-phonon matrix elements and explicit DFPT calculations, for diamond. The interpolated matrix elements were calculated starting from a coarse $4^3$ Brillouin-zone grid (dotted line, black), a $6^3$ grid (dashed line, blue), and a $8^3$ grid (solid line, red). The dots indicate explicit DFPT calculations. In this example $|n\mathbf{k}\rangle$ is set to the valence band top at $\Gamma$; $|m\mathbf{k} + \mathbf{q}\rangle$ spans $\Lambda_3$, $\Delta_5$, and $\Sigma_2$ bands, and the phonon is set to the highest optical branch. Reproduced with permission from Ref. [165], copyright (2007) by the American Physical Society.

6.1.3 Electron-phonon matrix elements in polar materials

In the case of polar materials, that is systems exhibiting nonzero Born effective charges [113], the interpolation scheme discussed in Sec. 6.1.2 breaks down. In fact, in these systems the dominant contribution to the potential $\partial V^{KS}/\partial \tau_\kappa\alpha$ in Eq. (177) is a dipole, which decays as $|\mathbf{R}_\nu'|^{-2}$. As a consequence some of the matrix elements in reciprocal space diverge as $|\mathbf{q}|^{-1}$ for $\mathbf{q} \to 0$, and cannot be interpolated straightforwardly from a coarse grid to a fine grid. Physically this singularity corresponds to the ‘Fröhlich electron-phonon coupling’ [30].

The adaptation of the Wannier interpolation method to the case of polar materials was recently given by Sjakste et al. [250] and by Verdi and Giustino [249]. In both works the basic idea is to separate the matrix elements into a short-range contribution, $g_{mn\nu}(\mathbf{k}, \mathbf{q})$, which is amenable to standard Wannier interpolation, and a long-range contribution, $g_{mn\nu}^L(\mathbf{k}, \mathbf{q})$, which is singular and is dealt with analytically. The strategy is analogous to that in use for calculating LO-TO splittings in polar materials [102]. The starting point is to define the long-range component of the matrix elements by considering the potential generated by the Born charges of all the atoms, when displaced according to a given vibrational eigenmode. The derivation relies on standard
In this expression \( Z_k^* \) and \( \epsilon^\infty \) denote the Born effective charge tensors and the electronic permittivity tensor (that is, the permittivity evaluated at clamped nuclei). This expression is the generalization of Fröhlich’s model to the case of anisotropic crystalline lattices and multiple phonon modes [30]. The result can be derived alternatively using the analytical properties of the dielectric matrix [113] as discussed in Ref. [131].

In order to perform Wannier interpolation, one subtracts Eq. (180) from the matrix elements computed on a coarse grid, interpolates the remaining short-range part, and then adds back Eq. (180) on the fine grid. This process requires the interpolation of the brackets \( \langle \cdots \rangle_{sc} \) in the second line of Eq. (180). Verdi and Giustino showed that, for small \( q + G \), these brackets can be interpolated via the relation

\[
\langle \psi_{mk+q}|e^{i(q+G) \cdot r}|\psi_{nk}\rangle_{sc} = [U_{k+q}U^\dagger_{k}]_{mn},
\]

where the rotation matrices \( U_{mnk} \) are obtained as usual from the procedure of Refs. [223, 228]. Figure 7 shows an example of Wannier interpolation for the prototypical polar semiconductor TiO\(_2\): it is seen that the singularity is correctly captured by the modified interpolation method.

At the end of this section, we mention that other interpolation schemes are equally possible [252–255]. For example Eiguren and Draxl proposed to interpolate only the local component of \( \Delta V_{KS}^q \), while calculating explicitly the nonlocal part of the perturbation as well as the Kohn-Sham wavefunctions in the Bloch representation [252]. Furthermore, Eq. (178) remains unchanged if MLWFs are replaced by a basis of localized atomic orbitals, and all the concepts discussed in this section remain valid. An interpolation scheme using local orbitals was recently demonstrated in Ref. [255].

6.2 Fermi surface harmonics

In the study of metallic systems, one is often interested in describing EPIs only for electronic states in the vicinity of the Fermi surface. In these cases, besides the Wannier interpolation discussed in Secs. 6.1.2 and 6.1.3, it is possible to perform efficient calculations using ‘Fermi-surface harmonics’ (FSH). FSHs were introduced by Allen [256] and recently revisited by Eiguren and Gurtubay [257].

The basic idea underlying FSHs is to replace expensive three-dimensional Brillouin-zone integrals by inexpensive one-dimensional integrals in the energy variable. To this aim, Allen proposed to expand functions of the band index \( n \) and wavevector \( k \), say \( A_{nk} \), in products of pairs of functions,
Figure 7: (Color online) Wannier interpolation of electron-phonon matrix elements for anatase TiO$_2$. The initial state $|n\mathbf{k}\rangle$ is set to the bottom of the conduction band at $\Gamma$, the final state $|m\mathbf{k} + \mathbf{q}\rangle$ spans the bottom of the conduction band along high-symmetry lines, and the phonon is the highest LO mode. The dots correspond to explicit DFPT calculations. The red dashed line is the short-range component of the matrix elements, $g^S$. The solid curve in blue represents the matrix elements $g^S + g^L$, as obtained from the modified Wannier interpolation of Sec. 6.1.3. The interpolation was performed starting from a coarse $4 \times 4 \times 4$ unshifted grid. Reproduced with permission from Ref. [249], copyright (2015) by the American Physical Society.

one depending on the energy, $A_L(\varepsilon)$, and one depending on the wavevector, $\Phi_L(\mathbf{k})$:

$$A_{nk} = \sum_L A_L(\varepsilon_{nk}) \Phi_L(\mathbf{k}). \quad (181)$$

In this expression, the Fermi-surface harmonics $\Phi_L(\mathbf{k})$ (to be defined below) are constructed so as to obey the following orthogonality condition:

$$N_p^{-1} \sum_{nk} \delta(\varepsilon_{nk} - \varepsilon) \Phi_L(\mathbf{k}) \Phi_L'(\mathbf{k}) = N(\varepsilon) \delta_{LL'}, \quad (182)$$

where $N(\varepsilon) = N_p^{-1} \sum_{nk} \delta(\varepsilon_{nk} - \varepsilon)$ is the density of states. Using Eqs. (181)-(182) one finds:

$$A_L(\varepsilon) = N(\varepsilon)^{-1} N_p^{-1} \sum_{nk} \delta(\varepsilon_{nk} - \varepsilon) \Phi_L(\mathbf{k}) A_{nk}. \quad (183)$$

Allen showed that, in the FSH representation, a linear system such as $A_{nk} = N_p^{-1} \sum_{nk'} M_{nk,nn'} B_{n'}$ transforms into $A_L(\varepsilon) = \sum_{LL'} \int d\varepsilon' N(\varepsilon') M_{LL'}(\varepsilon, \varepsilon') B_L(\varepsilon')$ [256]. Linear systems of this kind are common in the solution of the Boltzmann transport equation (Sec. 10) and the Eliashberg equations for the superconducting gap (Sec. 11.2). If one could perform the expansion using only a few harmonics, then the transformation would be advantageous, since the integrals over the wavevectors would have been absorbed in the expansion coefficients.

In the original proposal of Ref. [256], the harmonics $\Phi_L(\mathbf{k})$ were defined as polynomials in the band velocities, however the completeness of the basis set was not established. In a recent
work, Eiguren and Gurtubay proposed to construct these functions as eigenstates of a modified Helmholtz equation \[257\]:

$$\mathbf{v}_k |\nabla^2_k \Phi_L(k) = \omega_L \Phi_L(k), \tag{184}$$

where $\mathbf{v}_k = \hbar^{-1} \nabla_k \varepsilon_{nk}$ is the band velocity for states at the Fermi surface, and $\omega_L$ is the eigenvalue for the harmonic $\Phi_L$. The new definition in Eq. (184) maintains the properties of the original FSHs, and carries the added advantage that the basis set is complete. In this case the subscript $L$ in $\Phi_L(k)$ labels the eigenstates of the Helmholtz equation. Eiguren and Gurtubay demonstrated the construction of ‘Helmholtz FSHs’ for prototypical metals such as Cu, Li, and MgB$_2$.

Recent examples of the application of Fermi surface harmonics to first-principles calculations of EPIs include work on the photoemission kink of YBa$_2$Cu$_3$O$_7$ \[258\], and on the Seebeck coefficient of Li \[259\].

7 Non-adiabatic vibrational frequencies and linewidths

As discussed in Secs. 5.1.2-5.1.3, the electron-phonon interaction can lead to a renormalization of the adiabatic vibrational frequencies and to a broadening of the spectral lines.

The first ab initio investigations of the effects of the non-adiabatic renormalization of phonon frequencies were reported by Lazzeri and Mauri \[119\] and Pisana et al. \[260\]. In these works the authors concentrated on the $E_{2g}$ phonon of graphene, which is found at the wavenumber $\omega/2\pi c = 1585$ cm$^{-1}$ at room temperature ($c$ is the speed of light). This phonon corresponds to an in-plane C–C stretching vibration with $q = 0$, and has been studied extensively via Raman spectroscopy. In the graphene literature this mode is referred to as the ‘Raman G band’. Figure 8 shows a comparison between calculated and measured $E_{2g}$ phonon frequencies, as a function of doping, from \[260\]. The calculations were performed (i) within the adiabatic approximation and (ii) by including the non-adiabatic frequency renormalization using Eq. (145).\[8\] From Fig. 8 we see that the adiabatic theory is unable to reproduce the experimental data. On the contrary, the calculations including non-adiabatic effects nicely follow the measured Raman shift. This is a clear example of the limits of the adiabatic Born-Oppenheimer approximation and a demonstration of the importance of the phonon self-energy in Eq. (145).

The fact that the adiabatic approximation is inadequate for the $E_{2g}$ phonon of doped graphene should have been expected from the discussion on p. 45. In fact, graphene is a zero-gap semiconductor, therefore electrons residing in the vicinity of the Dirac points can make ‘virtual’ transitions with $|q| = 0$ and energies comparable to that of the $E_{2g}$ mode. As a result, the condition underlying the adiabatic approximation, $|\varepsilon_{mk+q} - \varepsilon_{nk}| \gg \hbar \omega_{q\nu}$, does not hold in this case.

The importance of non-adiabatic effects was confirmed also in the case of metallic single-walled carbon nanotubes \[167, 168\]. In these works, the authors studied the phonon dispersion relations

---

\[8\] In the works reviewed in this section the authors used Eq. (145) with the bare matrix elements $g_{mn\nu}(k, q)$ replaced by the screened matrix elements $g_{mn\nu}(k, q)$. All calculations were performed within DFT, using either the LDA or gradient-corrected DFT functionals.
Figure 8: (Color online) Frequency of the Raman $G$ band of graphene vs. carrier concentration. The black filled disks are from Raman measurements of gated graphene on a silicon substrate at 295 K. The thick horizontal dashed line (red) shows the variation of the $E_{2g}$ mode frequency with doping, within the adiabatic approximation. The solid blue line shows the variation of the frequency calculated by including non-adiabatic frequency renormalization. Reproduced with permission from Ref. [260], copyright (2007) by Macmillan Publishers Ltd.

in the vicinity of $|\mathbf{q}| = 0$, and they found that the difference between adiabatic and non-adiabatic dispersions is concentrated around the zone center. This finding is consistent with earlier models of non-adiabatic effects. In fact Maksimov and Shulgash showed that, for metals with linear electron bands crossing the Fermi level, $\Pi^{\text{NA}}$ is only significant for wavevectors $|\mathbf{q}| \sim \omega/v_F$, where $\omega$ is the phonon energy and $v_F$ the Fermi velocity [160]. This result can be derived from Eq. (145).

In the previous examples, the non-adiabatic renormalization of the vibrational frequencies is measurable but very small, typically of the order of 1% of the corresponding adiabatic frequencies. Saitta et al. considered the question as to whether one could find materials exhibiting large non-adiabatic renormalizations, and considered several graphite intercalation compounds, namely LiC$_6$, LiC$_{12}$, KC$_8$, KC$_{24}$, RbC$_8$, CaC$_6$, SrC$_6$, BaC$_6$, as well as other metallic systems such as MgB$_2$, Mg, and Ti [149]. In order to calculate the non-adiabatic renormalization at a reduced computational cost, the real part of the phonon self-energy was approximated as follows: $\hbar \text{Re}\Pi^{\text{NA}}_{\mathbf{q}=\mathbf{0}, \nu} \simeq N_F \langle |g_{\nu \nu}(\mathbf{k}, \mathbf{q} = 0)|^2 \rangle_{\text{BZ}}$, where $\langle \cdot \cdot \rangle_{\text{BZ}}$ stands for the average taken over the wavevectors $\mathbf{k}$ in the Brillouin zone, and $N_F$ is the density of states at the Fermi level. This expression can be derived from Eq. (145) by replacing the bare matrix elements by their screened counterparts, and by neglecting the ‘interband’ contributions $m \neq n$ in the sum. Figure 9 shows
Figure 9: Comparison between measured ($\omega_{\text{exp}}$) and calculated ($\omega_{\text{th}}$) vibrational frequencies of the $E_{2g}$ mode of graphite intercalation compounds. Open symbols are adiabatic DFPT calculations, filled symbols are calculations including the non-adiabatic corrections. The line corresponds to $\omega_{\text{th}} = \omega_{\text{exp}}$. Reproduced with permission from Ref. [149], copyright (2008) by the American Physical Society.

a comparison between vibrational frequencies from experiment and those calculated with or without including the non-adiabatic self-energy. It is clear that the non-adiabatic frequencies are in much better agreement with experiment than the corresponding adiabatic calculations. Furthermore, in these compounds the renormalization can reach values as large as $\sim 300$ cm$^{-1}$. In contrast to this, the renormalization in Mg and Ti was found to be of only a few wavenumbers in cm$^{-1}$.

The case of MgB$_2$ proved more puzzling: here the calculated non-adiabatic frequency is 761 cm$^{-1}$, while experiments reported 600 cm$^{-1}$. In order to explain this discrepancy, Saitta et al. reasoned that a more accurate calculation would require taking into account the relaxation time of the electrons [149], as pointed out in Ref. [160]. This would act so as to partly wash out non-adiabatic effects. In the field-theoretic language of Sec. 5, this observation corresponds to stating that when one approximates the dressed electron propagator $G$ in Fig. 2(c) using the non-interacting Kohn-Sham Green’s function, one should include at the very least the effects of finite electron lifetimes (due to electron-electron, electron-impurity, and electron-phonon scattering), for example via the imaginary part of Eq. (157).

The calculations discussed so far in this section addressed the non-adiabatic renormalization of zone-center phonons. The generalization to calculations of complete phonon dispersions was made by Calandra et al. [150]. In their work, Calandra et al. employed Wannier interpolation (see Sec. 6) in order to calculate the non-adiabatic phonon self-energy of Eq. (145) throughout the Brillouin zone. Figure 10 shows a comparison between the standard DFPT phonon dispersion relations of CaC$_6$ and the dispersions obtained after incorporating the non-adiabatic self-energy. It is seen that also in this case non-adiabatic effects are most pronounced at small $q$, and can be
Figure 10: (Color online) Phonon dispersion relations of CaC$_6$ calculated using Wannier interpolation. The dashed lines (black) and the solid lines (red) represent the standard adiabatic calculation and the non-adiabatic phonon dispersions, respectively. Reproduced with permission from Ref. [150], copyright (2010) by the American Physical Society.

as large as 7% of the adiabatic frequency.

In their work Calandra et al. approximated the bare matrix element $g^b_{mn\nu}(k, q)$ appearing in Eq. (145) by the screened matrix element $g_{mn\nu}(k, q)$. This replacement was justified by reasoning that the error is of second-order in the induced electron density, hence it should be negligible.

The broadening of vibrational spectra arising from the electron-phonon interaction is almost invariably calculated from first principles using Eq. (146). Since the integration of the Dirac delta is computationally costly, it is common to rewrite that equation by neglecting the phonon energy in the delta function and by taking the low-temperature limit, as proposed in Ref. [148]:

$$
\frac{\gamma_{q\nu}}{\pi \omega_{q\nu}} \simeq 2 \sum_{mn} \int \frac{dk}{\Omega_{BZ}} |g_{mn\nu}(k, q)|^2 \delta(\varepsilon_{nk} - \varepsilon_F) \delta(\varepsilon_{mk+q} - \varepsilon_F),
$$

(185)

where $\varepsilon_F$ is the Fermi energy. Oddly enough, this is a sort of adiabatic approximation to the non-adiabatic theory. The main advantage is that this expression is positive definite, hence easier to converge numerically as compared to the complete expression in Eq. (146). The disadvantages are that the temperature dependence is lost, and that one cannot resolve fine features on the scale of a phonon energy. There exists a vast literature on first-principles calculations of phonon linewidths using Eq. (185), mostly in connection with electron-phonon superconductors.\footnote{See for example early frozen-phonon calculations [69, 72, 261] and more recent DFPT calculations [73, 76, 162, 166]. Earlier calculations not based on DFT are reviewed in Ref. [6].}
is now implemented in several large software projects [262, 263], and it is used routinely.

Phonon linewidths range from very small values such as $\sim 1$ meV in Nb [76] to large values such as $\sim 30$ meV in MgB$_2$ [162]. The agreement between calculations and neutron scattering or Raman measurements is usually reasonable.

Calculations of phonon linewidths using the more accurate expression in Eq. (146) are computationally more demanding and have been reported less extensively in the literature.

So far we considered the effect of EPIs on the frequencies and lifetimes of vibrational excitations in solids. Another important phenomenon which modifies frequencies and lifetimes is anharmonicity. Anharmonic effects result from third and higher-order terms in the Taylor expansion of the total potential energy $U$ in the atomic displacements (Sec. 3.1). These effects can be interpreted as additional interactions that couple the oscillators of the harmonic lattice; for example, third-order anharmonic effects reduce phonon lifetimes via energy up- or down-conversion processes involving three phonons.

Anharmonic effects can be described using a many-body field-theoretic formalism [265], in complete analogy with the theory of EPIs discussed in Sec. 4. The calculation of anharmonic effects from first principles goes through the evaluation of third- and fourth-order derivatives of the total potential energy $U$ in the adiabatic approximation. Third-order coefficients are routinely computed using DFPT [264, 266–269]. In these cases where the harmonic approximation fails completely, ‘self-consistent phonon’ techniques can be employed [270, 271]; recent implementations and calculations can be found in Refs. [272–275].

8 Electron-phonon interactions in photoelectron spectroscopy

In Sec. 5.2.5 we have seen how the electron-phonon interaction in metals can lead to band structure ‘kinks’ near the Fermi energy. This is illustrated by the model calculation in Fig. 4(a). The experimental investigation of these features started in the late 1990s, following the development of high-resolution angle-resolved photoelectron spectroscopy (ARPES). Since in ARPES only the component of the photoelectron momentum parallel to the sample surface is conserved [192], complete energy vs. wavevector dispersion relations can be measured directly only for 2D or quasi-2D materials. Accordingly, the first observations of kinks were reported for the surface states of elemental metals [11] and for the CuO$_2$ planes of copper oxide superconductors [12]. Ab initio calculations of ARPES kinks can be performed by using the diagonal part of the Fan-Migdal self-energy (the Debye-Waller self-energy will be discussed at the end of this section). To this aim, it is common to rewrite Eqs. (157)-(158) at finite temperature using a spectral representation:

$$
\Sigma_{\text{FM}}^{\text{nk}}(\omega, T) = \int_{-\infty}^{+\infty} d\varepsilon \int_{0}^{\infty} d\varepsilon' \alpha^2 F_{\text{nk}}(\varepsilon, \varepsilon') \left[ \frac{1-f(\varepsilon/k_BT)+n(\varepsilon'/k_BT)}{\hbar \omega - \varepsilon - \varepsilon' + i\hbar \eta} + \frac{f(\varepsilon/k_BT)+n(\varepsilon'/k_BT)}{\hbar \omega - \varepsilon - \varepsilon' + i\hbar \eta} \right].
$$

(186)
Here the function $\alpha^2 F$ is the so-called the ‘Eliashberg function’ and is defined as:

$$\alpha^2 F_{nk}(\epsilon, \epsilon') = \sum_{mn} \int d\mathbf{q} \frac{|g_{nmn}(\mathbf{k}, \mathbf{q})|^2 \delta(\epsilon - \epsilon_{nk+q})\delta(\epsilon' - \hbar\omega_{\mathbf{q}n})}{\Omega_{\mathbf{BZ}}}.$$  \hspace{1cm} (187)

This quantity is positive, temperature-independent, and contains all the materials-specific parameters. Physically it is proportional to the scattering rate of an electron in the state $|nk\rangle$ into any electronic state at the energy $\epsilon$, by emitting or absorbing any one phonon of energy $\epsilon'$. One complication related to the Eliashberg function is that in the literature many variants of Eq. (187) can be found, each stemming from specific approximations; some of these expressions are summarized in Ref. [6], pp. 107–109.

The first \textit{ab initio} calculations of the phonon-induced electron self-energies and photoemission kinks were reported by Eiguren \textit{et al.} for the surface state of the Be(0001) surface [189]. Since the evaluation of Eqs. (186) and (187) is computationally demanding, Eiguren \textit{et al.} employed simplified expressions which involve three approximations [189]: (i) the Eliashberg function is replaced by its isotropic average, $\alpha^2 F_{nk}(\epsilon, \epsilon') = \int d\mathbf{k} \delta(\epsilon_{nk} - \epsilon) \alpha^2 F_{nk}(\epsilon, \epsilon') / \int d\mathbf{k} \delta(\epsilon_{nk} - \epsilon)$; (ii) phonon energies are neglected next to electron energies (as in the adiabatic approximation), and (iii) particle-hole symmetry is assumed. Using these approximations the imaginary part of the Fan-Migdal self-energy becomes:

$$|\text{Im} \Sigma_{n}^{\text{FM}}(\omega, T)| = \pi \int_{0}^{\infty} d\epsilon' \alpha^2 F_{n}(\hbar\omega, \epsilon') \{1 + 2n(\epsilon'/k_B T) + f[(\hbar\omega + \epsilon')/k_B T] - f[(\hbar\omega - \epsilon')/k_B T]\},$$ \hspace{1cm} (188)

where the average of the self-energy is defined as for the Eliashberg function. The real part of the self-energy can be found starting from the same approximations, and is given in Ref. [6].

Figure 11 shows the self-energy of the surface state at the Be(0001) surface calculated by Eiguren \textit{et al.} using Eq. (188) [189]. The imaginary part resembles a step-function, with an onset around the energy threshold for phonon emission by a hole (40–80 meV in this case). At a qualitative level, this trend can be rationalized by replacing the Eliashberg function in Eq. (188) by a Dirac delta at a characteristic phonon energy $\hbar\omega_{ph}$. In this case, the hole self-energy becomes proportional to $f[(\hbar\omega + \hbar\omega_{ph})/k_B T]$. At $T = 0$ this is precisely a step function with onset at $-\hbar\omega_{ph}$. The real part of the self-energy vanishes for $|\omega| \gg \omega_{\text{max}}$, with $\omega_{\text{max}}$ being the largest phonon frequency. This can be seen in Fig. 11(b), and is a consequence of the approximation of particle-hole symmetry. Eiguren \textit{et al.} also determined the renormalization of the surface state band structure arising from electron-phonon interactions, using Eq. (164) [189]; this is shown in the inset of Fig. 11(a). Overall the calculations of Ref. [189] showed good agreement with photoelectron spectroscopy experiments, both in the shape and magnitude of the self-energy [280].

In addition to the above calculations, several studies of the electron-phonon self-energy at metal surfaces were reported, namely for the Cu(111) and Ag(111) surfaces [188], the Al(100), Ag(111), Cu(111), and Au(111) surfaces [281], and the W(110) surface [282]. Building on these studies, Eiguren \textit{et al.} performed a detailed analysis of the \textit{self-consistent} solutions of the complex Dyson equation for the quasiparticle energies, Eq. (164)-(165), and illustrated the key concepts in the cases of the W(110) surface and for the phonon-mediated superconductor MgB$_2$ [283].
Figure 11: (Color online) Calculated Fan-Migdal self-energy of the surface state at the Be(0001) surface. (a) Imaginary part of the self-energy, obtained from Eq. (188). The dashed (black) line is the self-energy evaluated using the DFT/LDA bands; the solid lines (color/grayscale) correspond to the self-energy calculated by taking into account the renormalization of the DFT band structure by the electron-phonon interaction. (b) Real part of the self-energy, using the same color/grayscale code as in (a). The inset in (a) compares the renormalized band structure (color) with the ‘bare’ DFT band (black dashed line). The inset of (b) shows the renormalization of the band velocity induced by the electron-phonon interaction. Reproduced with permission from Ref. [189], copyright (2003) by the American Physical Society.

Equation (188) or closely-related approximations were also employed in the study of electron and hole lifetimes in bulk Be [284], Pb [285], and Mg [286]; the photoemission kink in YBa$_2$Cu$_3$O$_7$ [258]; and the spectral function of Ca-intercalated graphite [287].

In the case of complex systems the validity of the approximations (i)–(iii) leading to Eq. (188) is not warranted, and a direct calculation of the Fan-Migdal self-energy using Eqs. (186) and (187) is necessary. Calculations of the complete self-energy were reported by Park et al. for graphene [240], by Giustino et al. for the high-temperature superconductor La$_{1-x}$Sr$_x$Cu$_2$O$_4$ [195], and by Margine et al. for Ca-intercalated bilayer graphene [288]. Figure 12 shows the calculated self-energy and spectral function of graphene calculated in Refs. [240, 289]. The structure of the self-energy is similar to that of Fig. 11 with one important exception: $\Sigma_{nnk}^{FM}(\omega)$ does not vanish a few phonon energies away from the Fermi level, but tends instead towards a
Figure 12: (Color online) (a), (b) Calculated real part of the Fan-Migdal self-energy in pristine and $n$-doped graphene, respectively (solid black lines). The doping level is $4 \times 10^{13}$ cm$^{-2}$. The dashed lines correspond to a simplified analytical model where particle-hole symmetry is assumed. (c), (d) Electron band velocity renormalization resulting from the self-energies in (a) and (b). All calculations in (a)-(d) were performed using DFT/LDA. Reproduced with permission from Ref. [240], copyright (2007) by the American Physical Society. (e) Calculated spectral function of $n$-doped graphene for one of the branches of the Dirac cone. $\omega_{\text{ph}}$ indicates the characteristic phonon energy leading to the photoemission kink; $E_D$ denotes the energy of the Dirac point. The calculations include $GW$ quasiparticle corrections. Reproduced with permission from Ref. [289], copyright (2009) by the American Chemical Society.

linear asymptote. Calandra and Mauri performed a combined *ab initio*/analytical study of the effects of the electron-phonon interaction on the electron bands of graphene and obtained very similar results [290]. A linear asymptote in the real-part of the self-energy is a general feature of systems which do not exhibit particle-hole symmetry. For another example see Ref. [195].

Using the Fan-Migdal self-energy, it is possible to calculate the renormalization of the band velocity induced by the electron-phonon interaction. Let us denote by $v_{nk} = h^{-1} \nabla_k \varepsilon_{nk}$ the DFT band velocity and $\mathbf{V}_{nk} = h^{-1} \nabla_k E_{nk}$ the band velocity after including electron-phonon interactions. Using Eq. (164) with $\Gamma_{nk} = 0$ we find that these two quantities are simply related by $\mathbf{V}_{nk} = Z_{nk} v_{nk} = v_{nk} / (1 + \lambda_{nk})$, where $Z_{nk}$ is the quasiparticle strength of Eq. (174), and $\lambda_{nk}$ is the ‘mass enhancement parameter’ or ‘electron-phonon coupling strength’ [6]:

$$\lambda_{nk} = Z_{nk}^{-1} - 1 = -h^{-1} \partial \text{Re}\Sigma_{nk}(\omega) / \partial \omega \bigg|_{\omega=E_{nk}/h}.$$  \hspace{1cm} (189)

In the study of EPIs in metals, the coupling strength $\lambda_{nk}$ is of significant interest since it is related to the superconducting transition temperature of phonon-mediated superconductors (see Secs. 8.0.1 and 11).

The velocity renormalization in graphene calculated using Eq. (189) is shown in Fig. 12(c) and
(d), while a calculation of the complete spectral function \( A(k, \omega) \) is shown in Fig. 12(e). Here the characteristic photoemission kink is visible between 150–200 meV but it is not very pronounced, since in this case \( \lambda_{nk} \sim 0.1 \) \[289\]. These results are in good agreement with measured photoelectron spectra \[291\] \[292\].

Incidentally, we remark that in the analysis of ARPES data it is common to extract the coupling strength \( \lambda_{nk} \) directly from the ratio of the band velocities above and below the electron-phonon kink. However, this procedure is subject to a significant uncertainty, since the 'bare' velocity is not known and must be approximated by fitting the fully-interacting dispersions using \textit{ad hoc} models. For example, in the vicinity of Van Hove singularities this procedure leads to a significant overestimation of the electron-phonon coupling strength \( \lambda_{nk} \) \[241\] \[293\].

In addition to photoemission kinks, recent ARPES measurements revealed the existence of polaron satellites in doped oxides, namely TiO\(_2\) \[196\] and SrTiO\(_3\) \[197\] \[199\]. The phenomenology is similar to what was discussed in relation to Fig. 4(b). The first theoretical studies along this direction were reported by Story \textit{et al.}, who applied the cumulant expansion approach to the case of the electron-phonon self-energy \[221\]; by Antonius \textit{et al.} who identified satellites in the spectral functions of LiF and MgO \[294\]; and by Verdi \textit{et al.}, who calculated the ARPES spectra of \( n \)-doped TiO\(_2\) \[200\].

At the end of this section, it is worth coming back to the Debye-Waller self-energy. So far we only discussed the Fan-Migdal self-energy, starting from Eq. (186), and we ignored the Debye-Waller self-energy appearing in Eq. (152). This omission reflects the fact that, in the literature on electron-phonon interactions in metals, the DW term has always been disregarded. In order to rationalize this approximation, we rewrite the DW self-energy as follows, by combining Eqs. (160) and (40):

\[
\Sigma_{nnk}^{\text{DW}} = \langle u_{nk}|V_{\text{DW}}|u_{nk}\rangle_{\text{uc}},
\]  

with \( V_{\text{DW}}(r) = \Omega^{-1/2} \sum_{\nu} \int dq (n_{q\nu} + 1/2) \Delta_{q\nu} \Delta_{-q\nu} v^{KS}(r) \). The subscript ‘uc’ indicates that the integral is over one unit cell. From Eq. (190) we see that \( V_{\text{DW}} \) acts like a static local potential; indeed the first calculations including DW effects were performed by directly modifying the ionic pseudopotentials \[172\] \[173\]. From Eq. (190) we also see that the only variation in the DW self-energy comes from the Bloch amplitudes \( u_{nk} \). Let us consider the limiting situation of the homogeneous electron gas. In this case \( |u_{nk}(r)|^2 = 1/\Omega \) (Sec. 3.2.5), therefore \( \Sigma_{nnk}^{\text{DW}} \) is a constant, independent of \( k \). This behavior should be contrasted with the Fan-Migdal self-energy, which exhibits significant structure near the Fermi energy, as it can be seen in Fig. 4.

In more realistic situations, such as doped semiconductors, \( k \cdot p \) perturbation theory \[78\] can be used to show that \( \Sigma_{nnk}^{\text{DW}} \) varies smoothly as a function of \( k \) within the same band. In contrast with this scenario, \( \Sigma_{nnk}^{\text{DW}} \) exhibits large variations across different bands. This carries important consequences for the calculation of temperature-dependent band gaps (Sec. 9.1.1).
8.0.1 Electron mass enhancement in metals

We now come back to the mass enhancement parameter \( \lambda_{nk} \) introduced in Eq. (189), since this quantity played a central role in the development of the theory of EPIs in metals.

The notion of ‘mass enhancement’ becomes clear when we consider a parabolic band as in the model calculations of Fig. 4. Near the Fermi surface the non-interacting dispersions are given by \( \varepsilon_{nk} = \hbar \mathbf{k}_F \cdot \mathbf{v}(\mathbf{k} - \mathbf{k}_F)/m^* \), where \( \mathbf{k}_F \) is a wavevector on the Fermi surface, and the electron velocity is \( \mathbf{v}_{nk} = \hbar \mathbf{k}_F/m^* \). After taking into account the EPI, the electron velocity is renormalized to \( V_{nk} = v_{nk}/(1 + \lambda_{nk}) \). Since the magnitude of the Fermi momentum is unchanged (see caption of Fig. 4) this renormalization can be interpreted as an effective increase of the band mass from \( m^* \) to \( m_{ep} = m^*(1 + \lambda_{nk}) \). This reasoning holds for metals with parabolic bands and for doped semiconductors near band extrema, and does not take into account the Debye-Waller self-energy.

The electron mass enhancement is reflected into the increase of the heat capacity of metals at low temperature. In fact, below the Debye temperature the electronic contribution to the heat capacity dominates over the lattice contribution [80]. Since the heat capacity is proportional to the density of states at the Fermi level, and the density of states is inversely proportional to the band velocity, it follows that the heat capacity is directly proportional to the electron mass. This property can be used as a means to determine the strength of the electron-phonon coupling in simple metals from specific heat measurements [295].

The general theory of the effects of electron-phonon interactions on the heat capacity and other thermodynamic functions was developed by Eliashberg [296], Prange and Kadanoff [297], and Grimvall [47]. A field-theoretic analysis of the effect of EPIs on thermodynamic functions was developed by Eliashberg starting from the identities of Luttinger and Ward [298] in the zero-temperature limit. Eliashberg’s analysis was subsequently extended to all temperatures by Grimvall. Here we only quote Grimvall’s result relating the electronic entropy to the Fan-Migdal self-energy of Sec. 5.2.1:

\[
S_e = \frac{N_F k_B h^3}{(k_B T)^2} \int_0^\infty \frac{\omega \left[ \omega - h^{-1} \text{Re} \Sigma_{nk}^{\text{FM}}(\omega, T) \right]}{\cosh^2(\hbar \omega/2k_B T)} d\omega.
\]

(191)

In order to derive this relation, Grimvall started by expressing the thermodynamic potential of the coupled electron-phonon system in terms of the electron and phonon propagators and self-energies, and identified the electronic contribution by neglecting terms of order \((m_e/M_0)^{1/2}\) as well as electron-electron interactions (Ref. [47], Appendix).

Below the Debye temperature, an explicit expression for the entropy in Eq. (191) can be obtained by noting that the function \( \cosh^{-2}(\hbar \omega/2k_B T) \) is nonvanishing only for \( \hbar \omega \lesssim 2k_B T \); in this range Eq. (189) yields \( \Sigma_{nk}^{\text{FM}}(\omega, T) \simeq -\lambda_{nk} \omega \), therefore the integration in Eq. (191) can be carried out explicitly. As a result, the low-temperature heat capacity can be written as:

\[
C_e = T \frac{\partial S_e}{\partial T} = \frac{2}{3} \pi^2 k_B^2 N_F (1 + \lambda_{nk}) T.
\]

(192)

If we ignore the EPI by setting \( \lambda_{nk} = 0 \), this expression reduces to the standard textbook result for the free electron gas [80]. At high temperature Eq. (192) is no longer valid, and one has to
evaluate the integral in Eq. (191) using the complete frequency-dependent FM self-energy. The main result of this procedure is that at high temperature the electronic heat capacity is no longer renormalized by EPIs. A detailed discussion of this aspect is given in Ref. [47].

Early examples of DFT calculations of mass enhancement parameters and comparisons with specific heat measurements in simple metals can be found in Refs. [69, 73, 74].

9 Electron-phonon effects in the optical properties of semiconductors

9.1 Temperature dependence of band gaps and band structures

9.1.1 Perturbative calculations based on the Allen-Heine theory

In Sec. 5.2.3 we discussed how the electron-phonon interaction induces a ‘renormalization’ of the electronic energy levels, and thereby gives rise to ‘temperature-dependent band structures’.

These effects have been studied in detail using the Fan-Midgal and the Debye-Waller self-energies, either via the Raleigh-Schrödinger approximation to Eq. (166), or via its adiabatic counterpart given by Eq. (167).

Equation (167) was first employed in a number of calculations based on empirical pseudopotentials, following the seminal work of Allen and Heine [85] (see footnote 4). Allen and Cardona offer a clear introduction to the basic theory, a discussion of the computational methodology, as well as an historical perspective on earlier calculations [54].

The evaluation of the Debye-Waller contribution to the self-energy requires the calculation of the second-order variations of the Kohn-Sham potential with respect to the ionic displacements, Eq. (40). From a computational standpoint, this is challenging because one would have to use second-order DFPT, as discussed at the end of Sec. 3.2.3. In order to avoid this complication, it is common practice to recast all second-order derivatives as products of first-order derivatives. This strategy was introduced by Allen and Heine in the case of monoatomic crystals [85], and extended to polyatomic unit cells by Allen and Cardona [54]. The key observation behind this approach is that one can impose translational invariance of the theory to second order in the nuclear displacements. Specifically, the variation of the Kohn-Sham eigenvalues ensuing an arbitrary displacement of the nuclei should not change if all nuclei are further displaced by the same amount. Using time-independent perturbation theory, this condition yields the following two sum rules:

\[
\sum_{\kappa p} \langle\psi_{nk}|\partial_{\kappa\alpha p} V^{KS}|\psi_{nk}\rangle_{sc} = 0, \tag{193}
\]

\[
\sum_{\kappa' p'} \langle\psi_{nk}|\partial^2 V^{KS}_{\kappa\alpha p,\kappa'\alpha' p'}|\psi_{nk}\rangle_{sc} = -2Re \sum_{\kappa' p'} \sum_{m' q} \sum_{\kappa p} \langle\psi_{nk}|\partial_{\kappa\alpha p} V^{KS}|\psi_{mk+q}\rangle_{sc} \langle\psi_{mk+q}|\partial_{\kappa'\alpha' p'} V^{KS}|\psi_{nk}\rangle_{sc} \frac{\varepsilon_{mk+q} - \varepsilon_{nk}}{\varepsilon_{mk+q} - \varepsilon_{nk}}. \tag{194}
\]

Here \(\partial_{\kappa\alpha p} V^{KS}\) is a short-hand notation for \(\partial V^{KS}/\partial \tau_{\kappa\alpha p}\) and similarly for the second derivative; the primed summation indicates that eigenstates \(\psi_{mk+q}\) such that \(\varepsilon_{nk} = \varepsilon_{mk+q}\) are skipped.
The first sum rule is equivalent to stating that the electron-phonon matrix elements $g_{mn\nu}(k, q)$ associated with the three translational modes at $|q| = 0$ must vanish; this is an alternative formulation of the ‘acoustic sum rule’. The second sum rule, Eq. (194), suggests to express the matrix elements of the second derivatives of the potential in terms of first-order derivatives. However, Eq. (194) cannot be used as it stands, since it involves a sum of matrix elements on the left-hand side. In order to proceed, Allen and Heine employed the ‘rigid-ion’ approximation, whereby $V_{KS}$ is written as a sum of atom-centered contributions (see Sec. 2.1.1). Under this approximation all the terms $\kappa \rho \neq \kappa' \rho'$ on the left-hand side of Eq. (194) are neglected, and an explicit expression for $\langle \psi_{nk}| \partial^2 \kappa_{\alpha \rho} V_{KS} |\psi_{nk}\rangle_{sc}$ is obtained.

In view of practical DFT implementations, Giustino et al. used the sum rule in Eq. (194) in order to rewrite the Debye-Waller self-energy as follows [105]:

$$\Sigma_{mnk}^{DW} = -\sum_{\nu m}^\prime \int d\mathbf{q} \frac{2^{\text{DW}}(k, \mathbf{q})}{\Omega_{\text{BZ}}} \frac{\varepsilon_{nk} - \varepsilon_{mk}}{(2n_{q\nu} + 1)}.$$

(195)

Here $g_{mn\nu}^{\text{DW}}(k, q)$ is an ‘effective’ matrix element, and it is obtained from the standard DFPT matrix elements by means of inexpensive matrix multiplications:

$$g_{mn\nu}^{\text{DW}}(k, q) = \sum_{\kappa' \alpha' \rho'} \frac{\nu_{k'\alpha'\rho'}(q) h^*_{n\kappa\alpha}(k) h_{m\kappa'\alpha'}(k)}{2\omega_{q\nu} \epsilon_{k'\alpha'\rho'}(q)}.$$  

(196)

$$t_{k\alpha,k'\alpha'}^{\nu}(q) = \frac{\epsilon_{k\alpha\nu}(q) \epsilon_{k'\alpha'\nu}(q)}{M_k} + \frac{\epsilon_{k'\alpha'\nu}(q) \epsilon_{k\alpha\nu}(q)}{M_{k'}}.$$  

(197)

$$h_{mn,\kappa\alpha}(k) = \sum_{\nu} \left( M_k \omega_{0\nu} \right)^{1/2} \epsilon_{k\alpha\nu}(0) g_{mn\nu}(k, 0).$$  

(198)

In the case of the three translational modes at $|q| = 0$, these definitions are replaced by $g_{mn\nu}^{\text{DW}}(k, q) = 0$, see the discussion at the top of p. 98. The derivation of Eqs. (195)–(198) requires using Eqs. (20), (33)–(35), and (38), as well as taking the canonical average over the adiabatic nuclear quantum states.

Equation (195) involves a summation over unoccupied Kohn-Sham states, and so does the Fan-Migdal self-energy in Eq. (167). The numerical convergence of these sums is challenging, since one needs to evaluate a very large number of unoccupied electronic states. To address this issue, Gonze et al. developed a procedure whereby only a subset of unoccupied states is required, along the lines of the DFPT equations of Sec. 3.2.3 [106].

The first ab initio calculations using the formalism of Allen and Heine were reported by Marini, who investigated the temperature dependence of the optical absorption spectrum of silicon and boron nitride [104]. In this work Marini included excitonic effects by combining the Bethe-Salpeter formalism [158] with the Allen-Heine theory, and obtained good agreement with experiments by calculating the direct absorption peaks using DFT/LDA phonons and matrix elements (indirect optical absorption will be discussed in Sec. 9.2).

The second application of the Allen-Heine theory using DFT/LDA was reported by Giustino et al. for the case of diamond [105]. Here the authors investigated the temperature dependence of the direct band gap of diamond, and found that the Fan-Migdal and the Debye-Waller self-energies are of comparable magnitude. The calculations captured the characteristic Varshni effect
Figure 13: (Color online) Temperature dependence of the direct band gap of diamond calculated using the Allen-Heine theory. The upper curve shows the results obtained within DFPT at the LDA level. The lower curve was obtained via GW calculations in the frozen-phonon approach. Triangles are experimental data. The zero-point renormalization calculated by including GW quasiparticle corrections is 628 meV. Reproduced with permission from Ref. [110], copyright (2014) by the American Physical Society.

(Fig. 3), and were able to reproduce the measured redshift of the band gap in the temperature range 80-800 K. These calculations were based on the adiabatic version of the Allen-Heine theory, and employed Eqs. (196)-(198) for the Debye-Waller self-energy. The calculations of Ref. [105] confirmed the large (> 0.5 eV) zero-point renormalization of the direct gap of diamond predicted by Zollner et al. using the empirical pseudopotential method [178].

The unusually large zero-point correction to the electronic structure of diamond stimulated further work on this system: Cannuccia and Marini calculated the gap renormalization in diamond by employing both the adiabatic version of the Allen-Heine theory, as well as the non-adiabatic Green’s function approach, as described in Sec. 5.2.5 [299, 300]. Their calculations confirmed the large zero-point renormalization, and showed that the adiabatic theory underestimates the effect to some extent. Cannuccia and Marini also analyzed the quasiparticle renormalization and the spectral function [107, 300].

Antonius et al. revisited the electron-phonon interaction in diamond by assessing the reliability of the rigid-ion approximation and the importance of many-body GW quasiparticle corrections to the DFT/LDA band structure [110]. The main findings were that the rigid-ion approximation introduces a very small error in diamond, of the order of ~10 meV, while GW quasiparticle corrections can increase the electron-phonon renormalization of the band gap by as much as ~200 meV. The temperature dependence of the band gap of diamond calculated by Antonius et al. is shown in Fig. 13.

Further work on diamond was reported by Poncé et al., who compared ab initio calculations...
based on the Allen-Heine theory with explicit frozen-phonon calculations (Sec. 9.1.2) [109]. The corrections to the rigid-ion approximation were found to be smaller than 4 meV in all cases considered. Poncé et al. also reported a detailed assessment of the accuracy of the various levels of approximation in the calculation of the zero-point renormalization of energy levels, as well as a thorough comparison between the results of different first-principles implementations [108, 109].

The electron-phonon renormalization of band structures was also investigated in a number of other systems. For example Kawai et al. studied zinc-blend GaN by combining the Allen-Heine theory with the Bethe-Salpeter approach [111]. Poncé et al. investigated silicon, diamond, BN, α-AlN, and β-AlN using both the adiabatic version of the Allen-Heine theory and the non-adiabatic Green’s function method of Eqs. (157)-(158) [112]. Friedrich et al. investigated the zero-point renormalization in LiNbO$_3$ using the adiabatic Allen-Heine theory [301]. Villegas et al. studied the anomalous temperature dependence of the band gap of black phosphorous [302]. Antonius et al. investigated diamond, BN, LiF, and MgO, focusing on the dynamical aspects and the spectral function (see Sec. 8) [294]. Refs. [112, 294] were the first to report complete band structures at finite temperature.

In Ref. [112] the authors paid special attention to the numerical convergence of the self-energy integrals with respect to the limit $\eta \to 0$ of the broadening parameter; they noted that in the case of polar crystals the adiabatic correction to the electron energies of band extrema, as given by Eq. (167), diverges in the limit of dense Brillouin-zone sampling. This behavior stems from the polar singularity in the electron-phonon matrix elements, Eq. (180). In fact, near band extrema the bands are approximately parabolic, and the integrand in the adiabatic Fan-Migdal self-energy goes as $q^{-4}$ for $q \to 0$, while the volume element goes only as $dq = 4\pi q^2 dq$. This problem can be avoided by first performing the integration over $q$ in principal value, without neglecting phonon frequencies, and then taking the limit $\omega_{q\nu} \to 0$ so as to recover the adiabatic approximation (Sec. IV); in this way the adiabatic approximation can still be employed without incurring into a singularity in the calculations. A practical strategy to correctly perform the principal value integration in first-principles calculations was recently proposed by Nery and Allen [303]; here the authors treat the singularity via an explicit analytic integration near $q = 0$.

The complications arising in polar materials can also be avoided at once by using directly the more accurate expression in Eq. (166) based on Brillouin-Wigner perturbation theory, or even better by calculating the spectral functions as in Refs. [111, 294]. In particular, Eq. (166) shows that in more accurate approaches the infinitesimal $\eta$ should be replaced by the finite physical linewidth $\Gamma_{nk}$.

Although temperature-dependent band structures of polar materials were recently reported [111, 112, 294], the specific role of the Fröhlich coupling discussed in Sec. 6.1.3 received only little attention so far. The only ab initio investigations which specifically addressed the role of polar couplings in this context are those of Botti and Marques [304] and of Nery and Allen [303]. In order to understand the strategy of Botti and Marques, we refer to the Hedin-Baym equations in Sec. 4.2.1. They proposed that, instead of splitting the screened Coulomb interaction $W$ into electronic and nuclear contributions as in Eq. (104), one could try to directly calculate the screened Coulomb interaction $W$ including the lattice screening, as in Eqs. (85) and (86).
In order to make the calculations tractable, Botti and Marques evaluated the total dielectric matrices using a simplified model based on the Lyddane-Sachs-Teller relations. The resulting formalism combines GW calculations and experimentally measured LO-TO splittings. The zero-point renormalization of the band gaps calculated by Botti and Marques for LiF, LiCl, NaCl, and MgO were all $> 1$ eV. This is an interesting result and deserves further investigation. We note incidentally that the Allen-Heine theory and that of Botti and Marques can both be derived from the Hedin-Baym equations. Therefore the approach of Ref. [304] should effectively correspond to calculating the Fan-Migdal self-energy by retaining only the long-range part of the polar electron-phonon matrix elements. In the case of Ref. [303], Nery and Allen reported a Fröhlich contribution to the zero-point renormalization of the band gap of GaN of 45 meV, to be compared with the total renormalization of 150 meV arising from all modes.

9.1.2 Non-perturbative adiabatic calculations

An alternative approach to the calculation of temperature-dependent band structures consists of avoiding perturbation theory and electron-phonon matrix elements altogether, and replacing the entire methodology discussed in Sec. 9.1.1 by straightforward finite-differences calculations. To see how this alternative strategy works we perform a Taylor expansion of the Kohn-Sham eigenvalues $\varepsilon_{nk}$ to second order in the atomic displacements $\Delta r_{\kappa\alpha\nu}$, and then average the result on a nuclear wavefunction identified by the quantum numbers $\{n_{q\nu}\}$. After using Eq. (20) one obtains:

$$\langle \varepsilon_{nk} \{n_{q\nu} \} \rangle = \varepsilon_{nk} + \sum_{q\nu} \int dq \Omega_{\text{BZ}} \left( n_{q\nu} + 1/2 \right) \frac{\partial \varepsilon_{nk}}{\partial n_{q\nu}},$$

(199)

where we used the formal definition $\partial/\partial n_{q\nu} = \Delta r_{\kappa\alpha\nu}$, and the variations $\Delta r_{\kappa\alpha\nu}$ are the same as in Eqs. (33)-(35). The nuclear wavefunctions are obtained from the ground-state in Eq. (223) by applying the ladder operators, as discussed in Appendix B. The above expression can be generalized to finite temperature by considering a canonical average over all possible nuclear states. The result maintains the same form as in Eq. (199), except that we now have the Bose-Einstein occupations $n_{q\nu}(T)$ (see footnote B). Equation (199) is precisely the conceptual starting point of the Allen-Heine theory of temperature-dependent band structures, and appeared for the first time in [54]. If the variations $\Delta_{q\nu}\Delta_{q'\nu}\varepsilon_{nk}$ are calculated in second-order perturbation theory, one obtains precisely the formalism of Allen and Heine [55].

It has been proposed that the coefficients $\partial \varepsilon_{nk}/\partial n_{q\nu}$ could alternatively be obtained from the derivatives of the vibrational frequencies with respect to the electronic occupations, $\hbar \partial \omega_{q\nu}/\partial f_{nk}$ [109, 305, 306]. A formal derivation of the link between these alternative approaches can be found in Ref. [305], Appendix (the authors refer to this as Brooks’ theorem). Incidentally, the first ab initio calculation of temperature-dependent band gaps relied on this approach [306].

Most commonly, the coefficients $\partial \varepsilon_{nk}/\partial n_{q\nu}$ in Eq. (199) are evaluated using frozen-phonon supercell calculations, via the second derivative of the eigenvalue $\varepsilon_{nk}$ with respect to collective atomic displacements along the vibrational eigenmodes $e_{\kappa\alpha\nu}(q)$. This approach was employed by Capaz et al. to study the temperature dependence of the band gaps in carbon nanotubes (within a tight-binding model) [307], and by Han and Bester to obtain the zero-point renor-
malization and temperature dependence of the gaps of silicon and diamond quantum dots [308]. Recent examples include works on diamond, silicon, SiC [309], as well as CsSnI$_3$ [310].

Frozen-phonon supercell calculations based on Eq. (199) carry the advantage that the rigid-ion approximation, which is necessary to obtain Eqs. (195)-(198), is no longer required. Therefore this approach is more accurate in principle. In practice, however, the calculations are challenging as they require large supercells, and the derivatives must be evaluated for every vibrational mode of the supercell. Several computational strategies were developed to tackle this challenge. Patrick and Giustino proposed to perform the averages leading to Eq. (199) via importance-sampling Monte Carlo integration [311]. Monserrat described a constrained Monte Carlo scheme which improves the variance of the Monte Carlo estimator [312]. Recently Zacharias and Giustino showed that it is possible to perform these calculations more efficiently by replacing the stochastic sampling by a suitable choice of an ‘optimum’ configuration; the result becomes exact in the thermodynamic limit of large supercells [313]. In order to reduce the computational cost associated with the use of large supercells, Lloyd-Williams and Monserrat introduced ‘non-diagonal’ supercells, which allow one to access phonon wavevectors belonging to a uniform grid of $N_p$ points using supercells containing only $N_p^{1/3}$ unit cells [314].

The merit of these non-perturbative approaches is that they treat explicitly the nuclear wavefunctions, and enable exploring effects which go beyond the Allen-Heine theory. For example in Refs. [275, 315–317] the authors were able to investigate effects beyond the harmonic approximations in several systems, such as LiH, LiD, high-pressure He, molecular crystals of CH$_4$, NH$_3$, H$_2$O, HF, as well as Ice. In all these cases the authors found large zero-point effects on the band gaps.

Finally, we mention that the calculation of electronic properties at finite temperature via the Allen-Heine theory and its variants is closely related to what one would obtain using path-integral molecular dynamics simulations [318–320], or even classical molecular dynamics simulations at high enough temperatures [321].

### 9.2 Phonon-assisted optical absorption

In addition to modifying the electron energy levels in solids, the electron-phonon interaction plays an important role in the optical properties of semiconductors and insulators, as it is responsible for phonon-assisted optical transitions. Phonon-assisted processes could be analyzed by considering the many-body electronic screening function $\epsilon_{\alpha}(12)$ in Eq. (95), by using the electron Green’s function $G$ dressed by the electron-phonon self-energy $\Sigma^{\text{ep}}$ as in Eq. (151). Since this would require us a lengthy detour, here we simply reproduce the standard result of second-order time-
Figure 14: (Color online) Phonon-assisted optical absorption in silicon: comparison between first-principles calculations (solid lines, orange) and experiment (circles, blue). The calculations were performed using the theory of Hall et al., as given by Eq. (200) [324]. Spectra calculated at different temperatures were shifted horizontally so as to match the experimental onsets. Reproduced with permission from Ref. [244], copyright (2012) by the American Physical Society.

dependent perturbation theory [322, 323]:

\[
\alpha(\omega) = \frac{\pi e^2}{\varepsilon_0 c \Omega \omega n_r(\omega)} \int \frac{dk dq}{\Omega_B^2} \sum_{mn\nu s=\pm 1} \left( f_{nk} - f_{mk+q} \right) \times \left| \sum_p \left[ \frac{v_{np}(k)g_{pm\nu}(k, q) v_{pm}(k + q)}{\varepsilon_{pk} - \varepsilon_{nk} - \hbar \omega} + \frac{g_{np\nu}(k, q) v_{pm}(k + q)}{\varepsilon_{pk+q} - \varepsilon_{nk} + s\hbar \omega q \nu} \right] \right|^2 (n_{q \nu} + 1/2 + s/2) \delta(\varepsilon_{mk+q} - \varepsilon_{nk} - \hbar \omega + s\hbar \omega q \nu). \]

In this expression \(\alpha(\omega)\) is the absorption coefficient for visible light, \(e\) is the photon polarization, \(v_{mn}\) are the matrix elements of the electron velocity operator, and \(n_r(\omega)\) is the real part of the refractive index. The two denominators in the second line corresponds to indirect processes whereby a photon is absorbed and a phonon is absorbed or emitted (left), and processes whereby a phonon is absorbed or emitted, and subsequently a photon is absorbed (right). The above expression relies on the electric dipole approximation and is therefore valid for photon energies up to a few electronvolts. The theory leading to Eq. (200) was originally developed by Hall, Bardeen, and Blatt [324].

The first \textit{ab initio} calculation employing Eq. (200) was reported by Noffsinger \textit{et al.} for the prototypical case of silicon [244]. The authors employed DFT for computing phonons and electron-phonon matrix elements, and the \textit{GW} method for the quasiparticle band structures. The sampling of the Brillouin zone was achieved by means of the interpolation strategy de-
scribed in Sec. 6. Figure 14 shows that the calculations by Noffsinger et al. are in very good agreement with experiment throughout the energy range of indirect absorption.

Further work along similar lines was reported by Kioupakis et al., who calculated the indirect optical absorption by free carriers in GaN [325]; and Peelaers et al., who studied the indirect absorption by free carriers in the transparent conducting oxide SnO$_2$ [326]. Recently, the ab initio theory of phonon-assisted absorption was also extended to the case of indirect Auger recombination [327].

One limitation of the theory by Hall et al. is that the indirect absorption onset is independent of temperature [324]. This is seen by noting that the Dirac delta functions in Eq. (200) contain the band structure energies at clamped nuclei. The generalization to incorporate temperature-dependent band structures as discussed in Sec. 9.1.1 is nontrivial. Refs. [313, 328] showed that the electron-phonon renormalization of the band structure modifies the energies of real transitions but leaves unchanged the energies of virtual transitions; in other words the Allen-Heine renormalization should be incorporated only in the Dirac delta functions and in the first denominator in Eq. (200). In order to avoid these complications at once, Zacharias et al. [329] developed an alternative approach which relies on the 'semiclassical' approximation of Williams [330] and Lax [331]. In this approximation, the initial states in the optical transitions are described quantum-mechanically, and the final states are replaced by a quasiclassical continuum. In the formulation of Zacharias et al. the imaginary part of the temperature-dependent dielectric function takes the form [329]:

$$
\epsilon_2(\omega; T) = \frac{1}{Z} \sum_{\{n_{qv}\}} e^{-E_{\{n_{qv}\}}/k_B T} \langle \epsilon_2(\omega) \rangle_{\{n_{qv}\}},
$$

where $\epsilon_2(\omega)$ denotes the imaginary part of the dielectric function at clamped nuclei, and the expectation values have the same meaning as in Eq. (199). $E_{\{n_{qv}\}}$ is the energy of the quantum nuclear state specified by the quantum number $\{n_{qv}\}$ and $Z$ is the canonical partition function. Zacharias et al. demonstrated that this approach provides an adiabatic approximation to Eq. (200), and seamlessly includes the temperature dependence of the electronic structure within the Allen-Heine theory. Using techniques similar to those of Sec. 9.1.2 the authors calculated the indirect optical absorption lineshape of silicon at various temperatures and obtained very good agreement with experiment. These results were recently extended to the temperature-dependent optical spectra of diamond and gallium arsenide [313].

10 Carrier dynamics and transport

10.1 Electron linewidths and lifetimes

In Sec. 5.2.4 we have seen how the Fan-Migdal self-energy can be used in order to evaluate the quasiparticle lifetimes (or equivalently linewidths) resulting from the electron-phonon interaction. The first ab initio calculations of this kind were reported by Eiguen et al. in the study of the decay of metal surface states [188, 189], and by Sklyadneva et al. [284, 285] and Leonardo et al. [286] in the study of electron lifetimes of elemental metals. Some of these calculations and the
underlying approximations were reviewed in Sec. 8. Calculations of quasiparticle linewidths were also employed to study the temperature-dependent broadening of the optical spectra in semiconductors. For example Refs. [105, 112] investigated the broadening of the direct absorption edge of diamond and silicon, respectively. In both cases good agreement with experiment was obtained. More recently, the same approach was employed to study the broadening of photoluminescence peaks in lead-iodide perovskites [332]. In this case, it was found that the standard Fermi golden rule expression, Eq. (170), significantly overestimates the experimental data. The agreement with experiment is restored by taking into account the quasiparticle renormalization $Z_{n\mathbf{k}}$; see discussion following Eq. (174).

While these works were primarily concerned with the broadening of the spectral lines in photoemission or optical experiments, Eq. (170) can also be used to study carrier lifetimes in time-resolved experiments. The first ab initio study in this direction was reported from Sjakste et al. who investigated the thermalization of hot electrons in GaAs and the exciton lifetimes in GaP [333]. In the case of GaAs, Sjakste et al. found thermalization rates in quantitative agreement with time-resolved luminescence and transient optical absorption measurements. Work along similar lines was also reported for the intervalley scattering times in Ge [334]. Recently, the thermalization rates of hot electrons in GaAs were revisited by Bernardi et al. [251]. The authors employed Eq. (170) and the Wannier interpolation technique described in Sec. 6 in order to finely sample the electron-phonon scattering processes near the bottom of the conduction band, see Fig. 15. Based on these calculations they were able to interpret transient absorption measurements in terms of the carrier lifetimes within each valley. Another interesting application of Eq. (170) was reported by Bernardi et al., who investigated the rate of hot carrier thermalization in silicon within the context of photovoltaics applications [248].

Very recently Sangalli and Marini employed the lifetimes calculated using Eq. (170) in order to study carrier dynamics in silicon in real time [335, 336]. Strictly speaking, these developments lie outside of the scope of equilibrium Green’s functions discussed in Sec. 4 and require concepts based on non-equilibrium Green’s functions [137]. However, the basic ingredients of the electron-phonon calculations remain unchanged.

In all calculations discussed in this section, the electron-phonon matrix elements were obtained within DFT. However, in order to accurately describe electron-phonon scattering near band extrema in Refs. [248, 251, 332] the authors employed $GW$ quasiparticle band structures. This is important in order to obtain accurate band effective masses, which affect the carrier lifetimes via the density of states.

### 10.2 Phonon-limited mobility

The carrier lifetimes $\tau_{n\mathbf{k}}$ of Eq. (170) are also useful in the calculation of electrical mobility, conductivity, and resistivity, within the context of the semiclassical model of electron dynamics in solids [79]. In the semiclassical model, one describes the electronic response to an external perturbation by taking the fermionic occupations $f_{n\mathbf{k}}$ to represent the probability density function in the phase space defined by the unperturbed band structure. The probabilities $f_{n\mathbf{k}}$ are then
Figure 15: (Color online) Electron relaxation times in GaAs resulting from electron-phonon scattering. (a) Calculated relaxation times as a function of electron energy with respect to the conduction band bottom. The color code (gray shades) of the data points identifies the valley where each electronic state belongs. (b) Schematic representation of the conduction band valleys in GaAs. Reproduced with permission from Ref. [251], copyright (2015) by the National Academy of Sciences.

determined using a standard Boltzmann equation. A comprehensive discussion of these methods can be found in the classic book of Ziman [7].

Here we only touch upon the key result which is needed in \textit{ab initio} calculations of electrical conductivity. In the semiclassical model, the electrical current is calculated as

\[
J = -2e \Omega_{BZ}^{-1} \sum_n \int d\mathbf{k} v_{nk} f_{nk}. \tag{202}
\]

In the absence of external fields and thermal gradients, the occupations \( f_{nk} \) reduce to the standard Fermi-Dirac occupations at equilibrium, \( f^0_{nk} \), and the current vanishes identically. Upon introducing an electric field \( \mathbf{E} \), the electrons respond by adjusting their occupations. In this model it is assumed that the variation \( f_{nk} - f^0_{nk} \) is so small that the electronic density is essentially the same as in the unperturbed system. The modified occupations can be calculated using the linearized Boltzmann transport equation [7]:

\[
\frac{\partial f^0_{nk}}{\partial \varepsilon_{nk}} \mathbf{v}_{nk} \cdot (-e) \mathbf{E} = -\sum_\nu \int \frac{d\mathbf{q}}{\Omega_{BZ}} \Gamma_{mn\nu}(\mathbf{k}, \mathbf{q}) [(f_{nk} - f^0_{nk}) - (f_{mk+q} - f^0_{mk+q})], \tag{203}
\]

where the kernel \( \Gamma_{mn\nu}(\mathbf{k}, \mathbf{q}) \) is defined as:

\[
\Gamma_{mn\nu}(\mathbf{k}, \mathbf{q}) = \sum_{s=\pm 1} \frac{2\pi}{h} |g_{mn\nu}(\mathbf{k}, \mathbf{q})|^2 f^0_{nk}(1 - f^0_{mk+q})(n_{q\nu} + 1/2 - s/2) \delta(\varepsilon_{nk} + sh\omega_{q\nu} - \varepsilon_{mk+q}). \tag{203}
\]

The left-hand side of Eq. (202) represents the collisionless term of the Boltzmann equation, that is the change in occupations due to the particle drift under the electric field. The right-hand side represents the change of occupations resulting from electrons scattered in or out of the state \( |nk\rangle \) by phonon emission or absorption. The rates in Eq. (203) are simply derived from Fermi’s golden rule [6]. By solving Eq. (202) self-consistently for all \( f_{nk} \) it is possible to calculate the current, and from there the conductivity. The connection with the carrier lifetimes \( \tau_{nk} \) of Eq. (170) is
obtained within the so-called ‘energy relaxation time approximation’. In this approximation the incoming electrons are neglected in Eq. (202), that is the last term \((f_{mk+q} - f_{mk+q}^0)\) in the square brackets is ignored. As a result the entire right-hand side of the equation simplifies to \(-(f_{nk} - f_{nk}^0)/\tau_{nk}\).

The direct solution of Eq. (202) is computationally challenging, and fully ab initio calculations were reported only very recently by Li for Si, MoS\(_2\), and Al [337], and by Fiorentini and Bonini for n-doped Si [338]. Figure 16 shows that the mobility of n-type silicon calculated by Li is in good agreement with experiment. The theory overestimates the measured values to some extent, and this might have to do with the limitations of the DFT matrix elements (see Sec. 12). In addition to the carrier mobility, Fiorentini and Bonini employed the ab initio Boltzmann formalism to calculate thermoelectric properties, such as the Lorenz number and the Seebeck coefficient [338].

The first ab initio calculation of mobility was reported by Restrepo et al. for the case of silicon, within the energy relaxation time approximation [339]. Other recent calculations using various approximations to Eq. (202) focused on silicon [340, 341], graphene [247, 255, 342–345], MoS\(_2\) [255, 344, 346, 347], silicene [255, 347], SrTiO\(_3\) and KTiO\(_3\) [348, 349].

Ab initio calculations of the resistivity of metals are less challenging than for semiconductors, and started appearing already with the work of Bauer et al. [76]. Most calculations on metals are based on Ziman’s resistivity formula, see Ref. [6], p. 210. An interesting recent example can be found in the work by Xu and Verstraete on the transport coefficients of lithium [259]. We also highlight related work on phonon-limited transport in organic crystals [243, 243, 350–352].

11 Phonon-mediated superconductors

The last application of ab initio calculations of EPIs that we will consider is the study of phonon-mediated superconductivity [8]. This research field is so vast that any attempt at covering it in a few pages would not make justice to the subject. For this reason, it was decided to limit the discussion to those novel theoretical and methodological developments which are aiming at fully predictive calculations, namely the ‘anisotropic Migdal-Eliashberg theory’ (Sec. 11.2), and the ‘density functional theory for superconductors’ (Sec. 11.3). For completeness, in Sec. 11.1 we also summarize the most popular equations employed in the study of phonon-mediated superconductors. All calculations described in this section were performed at the DFT level.

11.1 McMillan/Allen-Dynes formula

Most ab initio calculations on phonon-mediated superconductors rely on a semi-empirical expression for the critical temperature, first introduced by McMillan [353] and then refined by Allen and Dynes [354]:

\[
k_B T_c = \frac{\hbar \omega_{log}}{1.2} \exp \left[-\frac{1.04(1 + \lambda)}{\lambda - \mu^*(1 + 0.62 \lambda)}\right].
\]

Here \(T_c\) is the superconducting critical temperature, \(\omega_{log}\) is a ‘logarithmic’ average of the phonon frequencies [354], \(\lambda\) is the electron-phonon ‘coupling strength’, and \(\mu^*\) is a parameter describing
The Coulomb repulsion. The functional form of Eq. (204) was derived by McMillan by determining an approximate solution of the Eliashberg gap equations (see Sec. 11.2) \cite{353}. \( \lambda \) and \( \omega_{\text{log}} \) are calculated from the isotropic version of the Eliashberg function in Eq. (187) as follows \cite{6, 353–355}:

\[
\alpha^2 F(\omega) = \frac{1}{N_F} \int \frac{dk}{BZ} \int \frac{dq}{q} \sum_{mn\nu} |g_{mn\nu}(k, q)|^2 \delta(\varepsilon_{mk} - \varepsilon_F) \delta(\varepsilon_{mk+q} - \varepsilon_F) \delta(\hbar \omega - \hbar \omega_{\nu}),
\]

(205)

\[
\lambda = 2 \int_0^\infty \frac{\alpha^2 F(\omega)}{\omega} d\omega,
\]

(206)

\[
\omega_{\text{log}} = \exp \left[ \frac{2}{\lambda} \int_0^\infty d\omega \frac{\alpha^2 F(\omega)}{\omega} \log \omega \right],
\]

(207)

where \( N_F \) is the density of states at the Fermi level and the matrix elements \( g_{mn\nu}(k, q) \) are the same as in Eq. (38). The remaining parameter \( \mu^* \) \cite{356} is obtained as \( 1/\mu^* = 1/\mu + \log(\omega_p/\omega_{\text{ph}}) \), where \( \hbar \omega_p \) is the characteristic plasma energy of the system, \( \hbar \omega_{\text{ph}} \) the largest phonon energy, and \( \mu \) is the average electron-electron Coulomb repulsion across the Fermi surface. More specifically: \( \mu = N_F \langle \langle V_{nk,n'k'} \rangle \rangle_{\text{FS}} \), where \( \langle \cdot \rangle_{\text{FS}} \) denotes a double average over the Fermi surface, and \( V_{nk,n'k'} = \langle k'n', -k'n' | W | kn, -kn \rangle \), with \( W \) being the screened Coulomb interaction of Sec. 4.2.2 \cite{357, 358}.

The coupling strength \( \lambda \) is related to the mass enhancement parameter \( \lambda_{nk} \) discussed in Sec. 8.0.1. The main difference between \( \lambda \) and \( \lambda_{nk} \) is that the former represents an average over the Fermi
surface, while the latter refers to the Fermi velocity renormalization of a specific electron band. While these quantities are related, they do not coincide and hence cannot be used interchangeably.

Equations (204)-(207) involve a number of approximations. For example, it is assumed that the superconductor is isotropic and exhibits a single superconducting gap. Furthermore, almost invariably the effective Coulomb potential $\mu^*$ is treated as an adjustable parameter, on the grounds that it should be in the range $\mu^* = 0.1-0.2$. This procedure introduces a large uncertainty in the determination of $T_c$, especially at moderate coupling strengths.

### 11.2 Anisotropic Migdal-Eliashberg theory

A first-principles approach to the calculation of the superconducting critical temperature is provided by the anisotropic Migdal-Eliashberg theory [123, 359]. This is a field-theoretic approach to the superconducting pairing, formulated in the language of finite-temperature Green’s functions. A comprehensive presentation of the Migdal-Eliashberg theory is provided by Allen and Mitrovic [355]. Their article served as the starting point of current first-principles implementations of the theory.

In the Migdal-Eliashberg theory, one solves the two coupled equations:

$$ Z_{nk}(i\omega_j) = 1 + \frac{\pi k_B T}{N_F} \sum_{n'k'j'} \frac{\omega_{j'}/\omega_j}{\sqrt{h^2\omega_{j'}^2 + \Delta_{n'k'}^2(\omega_{j'})}} \lambda_{nk,n'k'}(i\omega_j - i\omega_{j'}) \delta(\varepsilon_{n'k'} - \varepsilon_F), \tag{208} $$

$$ Z_{nk}(i\omega_j) \Delta_{nk}(i\omega_j) = \frac{\pi k_B T}{N_F} \sum_{n'k'j'} \frac{\Delta_{n'k'}(i\omega_{j'})}{\sqrt{h^2\omega_{j'}^2 + \Delta_{n'k'}^2(\omega_{j'})}} \left[ \lambda_{nk,n'k'}(i\omega_j - i\omega_{j'}) - N_F V_{nk,n'k'} \right] \times \delta(\varepsilon_{n'k'} - \varepsilon_F), \tag{209} $$

where $\sum_{k'}$ stands for $\Omega_{BZ}^{-1} \int d{k'}$. In these equations, $T$ is the absolute temperature, $Z_{nk}(i\omega_j)$ is the quasiparticle renormalization function, and is analogous to $Z_{nk}$ in Eq. (189). $\Delta_{nk}(i\omega_j)$ is the superconducting gap function. The functions $Z_{nk}(i\omega_j)$ and $\Delta_{nk}(i\omega_j)$ are determined along the imaginary frequency axis, at the fermion Matsubara frequencies $i\omega_j = i(2j + 1)\pi k_B T/h$ with $j$ an integer. The anisotropic and frequency-dependent generalization of Eq. (206) to be used in the Migdal-Eliashberg equations is:

$$ \lambda_{nk,n'k'}(i\omega) = \frac{N_F}{h} \sum_{\nu} \frac{2\omega_{q\nu}}{\omega_{q\nu}^2 + \omega_{\nu}^2} |g_{mn'\nu}(k, q)|^2, \tag{210} $$

with $q = k' - k$. Equations (208)-(209) are to be solved self-consistently for each temperature $T$.

The superconducting critical temperature is then obtained as the highest temperature for which a nontrivial solution is obtained, that is a solution with $\Delta_{nk}(i\omega_j) \neq 0$. From the superconducting gap along the imaginary axis it is then possible to obtain the gap at real frequencies by analytic continuation [362], and from there one can compute various thermodynamic functions.

The first ab initio implementation of the anisotropic Migdal-Eliashberg theory was reported by Choi et al. in a study of the superconducting properties of MgB$_2$ [363–365]. The authors
succeeded to explain the anomalous heat capacity of MgB$_2$ in terms of two distinct superconducting gaps, and obtained a $T_c$ in good agreement with experiment. These calculations were later extended to MgB$_2$ under pressure [366] and other hypothetical borides [367]. Margine and Giustino demonstrated an implementation of the Migdal-Eliashberg theory based on the Wannier interpolation scheme of Sec. 6 and reported applications to Pb and MgB$_2$ [245]. The superconducting gap and superconducting density of states of MgB$_2$ calculated by Margine and Giustino are shown in Fig. 17. In all these calculations, the Coulomb repulsion was described empirically via $\mu^*$, and this partly accounts for the slight discrepancy between the calculated $T_c$ of 50 K and the experimental $T_c$ of 39 K [368]. Additional calculations based on the anisotropic Migdal-Eliashberg theory include a study of doped graphene [246], as well as investigations of Li-decorated monolayer graphene [369] and Ca-intercalated bilayer graphene [388]. In this latter work the authors incorporated Coulomb interactions from first principles, after calculating $\mu^*$ via the screened Coulomb interaction in the random-phase approximation. The calculated $T_c = 7$-8 K was in reasonable agreement with the experimental value of 4 K [370]. The Migdal-Eliashberg theory has also been extended to describe the superconducting state as a function of applied magnetic field; a complete ab initio implementation was successfully demonstrated with
an application to MgB$_2$ [371]. Very recently Sano et al. performed \textit{ab initio} Migdal-Eliashberg calculations including retardation effects on high-pressure sulfur hydrides, obtaining good agreement with experiment [372]. Interestingly in this work the authors also checked the effect of the zero-point renormalization of the electron bands within the Allen-Heine theory, and found that it accounts for a change in $T_c$ of up to 20 K.

11.3 Density functional theory for superconductors

Another promising \textit{ab initio} approach to the calculation of the superconducting critical temperature is the density functional theory for superconductors [373, 374]. The starting point of this approach is a generalization of the Hohenberg-Kohn theorem [60] to a system described by three densities: the electron density in the normal state, the density of superconducting pairs, and the nuclear density. Based on this premise, Lüders et al. mapped the fully-interacting system into an equivalent Kohn-Sham system [61] of non-interacting nuclei and non-interacting, yet superconducting, electrons [373]. The resulting Kohn-Sham equations for the electrons take the form of Bogoliubov-de Gennes equations [375], whereby electrons are paired by an effective gap function $\Delta(r, r')$.

In its simplest formulation, the density functional theory for superconductors determines the expectation value of the pairing field over Kohn-Sham eigenstates,

$$\Delta_{nk} = \langle u_{nk}(r) | \Delta(r, r') | u_{nk}(r') \rangle,$$

using the following gap equation:

$$\Delta_{nk} = -Z_{nk} \Delta_{nk} - \sum_{n'k'} K_{n,k,n'k'} \Delta_{n',k'} \tanh \left( \frac{E_{n',k'}}{2k_B T} \right),$$

where $E_{nk}^2 = \epsilon_{nk}^2 + |\Delta_{nk}|^2$. In this expression, the kernel $K$ contains information about the phonon-mediated pairing interaction and the Coulomb repulsion between electrons, $K = K^{\text{ep}} + K^{\text{ee}}$, and $Z$ contains information about the electron-phonon interaction. More specifically, $K^{\text{ep}}$ and $Z$ are evaluated starting from the electron-phonon matrix elements $g_{mn\nu}(k, q)$ and the DFT electron band structure and phonon dispersions, as in the Migdal-Eliashberg theory. $K^{\text{ee}}$ is approximated using the screened Coulomb interaction $V_{nk,n'k'}$ introduced below Eq. (207). Complete expressions for $Z$ and $K$ can be found in [374].

Equation (211) is reminiscent of the gap equation in the Bardeen-Cooper-Schrieffer (BCS) theory [8], with the difference that the \textit{ab initio} kernel $K$ replaces the model interaction of the BCS theory, and the function $Z$ introduces quasiparticle renormalization as in the Migdal-Eliashberg theory, see Eq. (208). At variance with the Migdal-Eliashberg theory, the gap function in the density functional theory for superconductors does not carry an explicitly frequency dependence. Nevertheless, retardation effects are fully included through the dependence of the kernels $Z$ and $K$ on the electron bands and the phonon dispersions. An important advantage of this theory is that the Coulomb potential $\mu^*$ is not required, since the electron-electron repulsion is seamlessly taken into account by means of the kernel $K^{\text{ee}}$.

The density functional theory for superconductors was successfully employed to study the superconducting properties of MgB$_2$ [376], Li, K, and Al under pressure [377, 378], Pb [379], Ca-intercalated graphite [380], high-pressure hydrogen [381, 382], CaBeSi [383], layered nitrides [384].
alkali-doped fullerides [385], compressed sulfur hydrides [386], and intercalated layered carbides, silicides, and germanides [387].

An interesting recent development of the theory was the determination of the superconducting order parameter in real space, \( \chi(r, r') \) = \( \langle \hat{\psi}^\uparrow(r) \hat{\psi}^\downarrow(r') \rangle \) [388]. In the density functional theory for superconductors, the order parameter is obtained from the superconducting gap using the relation \( \chi_{nk} = \Delta_{nk}/(2|E_{nk}| \tanh[E_{nk}/(2k_B T)] \). Figure 18 shows the order parameter calculated by Linscheid et al. [388] for both MgB\(_2\) and hole-doped graphane [389]. The plots show Friedel-like oscillations of the superconducting density as a function of the relative coordinates between two paired electrons.

Further developments of the superconducting density functional theory include the study of non-phononic pairing mechanisms, such as plasmon-assisted superconductivity [390], and the extension to magnetic systems [391, 392].

12 Electron-phonon interactions beyond the local density approximation to DFT

The calculations of electron-phonon interactions reviewed in Sec. 7-11 have in common the fact that most investigations used the local density approximation to DFT or a generalized gradient approximation (GGA) such as the PBE functional [393]. Although the LDA and the GGA do represent the workhorse of electron-phonon calculations from first principles, there is growing
evidence that these choices can lead to an underestimation of the electron-phonon coupling strength. At a conceptual level we can understand this point by rewriting the electron-phonon matrix element after combining Eqs. (38) and (142), (143):

$$ g_{mn\nu}(k, q) = \langle u_{m+k+q}|\int_0^1 \epsilon_{\nu}^{-1}(r, r', \omega)\Delta_{\nu\nu} v^{\nu}(r')|u_{n+k}\rangle_{uc}. $$

In DFT the many-body dielectric matrix $\epsilon$ appearing in this expression is replaced by the RPA+$xc$ screening $\epsilon^{Hxc}$ from Eq. (54). Given the DFT band gap problem, we expect $\epsilon^{Hxc}$ to overestimate the screening, thereby leading to matrix elements $g_{mn\nu}(k, q)$ which are underestimated to some extent.

Several groups investigated this point on quantitative grounds. Zhang et al. studied the electron-phonon coupling in a model copper oxide superconductor, CaCuO$_2$ [391]. By calculating the vibrational frequencies of the half-breathing Cu-O stretching mode, the authors established that the local spin-density approximation (LSDA) yields phonons which are too soft (65.3 meV) as compared to experiment (80.1 meV). In contrast, the introduction of Hubbard corrections in a LSDA+$U$ scheme restored agreement with experiment (80.9 meV). Since the electron-phonon matrix elements are connected to the phonon frequencies via the phonon self-energy, Eq. (145), a corresponding underestimation of the matrix elements can be expected. These results were supported by the work of Floris et al., who developed DFPT within LSDA+$U$, and applied their formalism to the phonon dispersions of antiferromagnetic MnO and NiO [93]. Here the authors found that the DFT underestimates measured LO energies by as much as 15 meV in MnO, while the use of LSDA+$U$ leads to good agreement with experiment. Related work was reported by Hong et al., who investigated the multiferroic perovskites CaMnO$_3$, SrMnO$_3$, BaMnO$_3$, LaCrO$_3$, LaFeO$_3$, and the double perovskite La$_2$CrFeO$_6$ [395]. Here the authors calculated the variation of the vibrational frequencies between the ferromagnetic and the antiferromagnetic phases of these compounds as a function of the Hubbard $U$ parameter, and compared DFT+$U$ calculations with hybrid-functional calculations.

Lazzeri et al. investigated the effect of quasiparticle GW corrections on the electron-phonon coupling of graphene and graphite, for the $A'$ phonon at $K$ and the $E_{2g}$ phonon at $\Gamma$ [396]. They evaluated the intraband electron-phonon matrix elements using a frozen-phonon approach, noting that $g_{mn\nu}(k, q = 0)$ represents precisely the shift of the Kohn-Sham energy $\epsilon_{nk}$ upon displacing the atoms according to the $\nu$-th phonon eigenmode at $q = 0$. Lazzeri et al. found that the matrix elements increase by almost 40% from DFT to GW [396]. The GW values led to slopes in the phonon dispersions near $K$ in very good agreement with inelastic X-ray scattering data [397]. Similar results, albeit less dramatic, were obtained by Grüneis et al. for the potassium-intercalated graphite KC$_8$ [398].

Laflamme Janssen et al. studied the electron-phonon coupling in the C$_{60}$ molecule as a model for superconducting alkali-doped fullerides [399]. They employed the PBE0 hybrid functional [400] with a fraction of exact exchange $\alpha = 30\%$, and obtained an enhancement of the total coupling strength $\lambda$ of 42% as compared to PBE. This work was followed up by Faber et al., who used the GW approximation and obtained a similar enhancement of 48% [401]. We also point out an earlier work by Saito based on the B3LYP functional, reporting similar results [402].

92
Yin et al. investigated the effects of using the GW approximation and the HSE hybrid functional \[403\] on the electron-phonon coupling in the superconducting bismuthates \(\text{Ba}_{1-x}\text{K}_x\text{BiO}_3\) and chloronitrides \(\beta\)-ZrNCl, as well as MgB\(_2\) \[404\]. In the case of \(\text{Ba}_{1-x}\text{K}_x\text{BiO}_3\) the authors obtained a three-fold increase in the coupling strength \(\lambda\) from PBE to HSE. This enhancement brought the critical temperature calculated using Eq. \((204)\) to 31 K, very close to the experimental value of 32 K. Similarly, in the case of \(\beta\)-ZrNCl, Yin et al. obtained a 50% increase of \(\lambda\), bringing the calculated critical temperature, 18 K, close to the experimental value of 16 K \[404\]. Instead, in the case of MgB\(_2\), they noticed only a slight increase of the electron-phonon coupling as compared to the standard LDA.

Another application of hybrid functionals to the study of EPIs was reported by Komelj and Krakauer \[405\]. Here the authors investigated the sensitivity of the superconducting critical temperature of the \(\text{H}_3\text{S}\) phase of sulfur hydride to the exchange and correlation functional. They found that the PBE0 functional enhances the critical temperature by up to 25% as compared to PBE, bringing \(T_c\) from 201-217 K to 253-270 K (the spread in values is related to the choice of the parameter \(\mu^*\)).

Mandal et al. reported work on the superconductor FeSe based on dynamical mean-field theory (DMFT) \[406\]. In this case DMFT yielded a three-fold enhancement of the coupling strength for selected modes.

As already mentioned in Sec. 9.1.1, Antonius et al. performed GW calculations of the electron-phonon coupling in diamond using a frozen-phonon approach \[110\]. They found that quasiparticle corrections lead to a uniform enhancement of the electron-phonon matrix elements. The net effect is an increase of the zero-point renormalization of the band gap by 40% as compared to standard LDA calculations. Monserrat confirmed this result and found a GW correction of comparable magnitude in the case of silicon \[407\]. However, Monserrat also found that the GW corrections to the zero-point band gap renormalization of LiF, MgO, and TiO\(_2\) are very small (~5% of the PBE value), therefore at present it is not possible to draw general conclusions.

Finally, we mention that Faber et al examined possible strategies for systematically incorporating GW corrections in electron-phonon calculations \[408\]. By using diamond, graphene, and C\(_{60}\) as test cases, the authors showed that a ‘constant screening’ approximation is able to reproduce complete GW results with an error below 10% at reduced computational cost. This approximation amounts to evaluating the variation of the Green’s function \(G\) in a frozen-phonon calculation, while retaining the screened Coulomb interaction \(W\) of the unperturbed ground state.

All these recent developments point to the need of moving beyond local exchange and correlation functionals in the study of electron-phonon interactions from first principles. In the future, it will be important to devise accurate computational methods for calculating not only the intraband electron-phonon matrix elements (as in the frozen-phonon method) but also matrix elements between all states and for scattering across the entire Brillouin zone.

For the sake of completeness we emphasize that the underestimation of the EPI matrix elements by semilocal DFT functionals does not propagate in the same way into different materials properties. This is readily understood by examining two fundamental quantities, the Allen-Heine
renormalization of electron bands, Eq. (167), and the adiabatic phonon frequencies, as obtained from Eqs. (126) and (127). In the former case the electronic screening enters as $\epsilon^{-2}$; in the latter case the screening contributes through a term which scales with $\epsilon^{-1/2}$. As a result, in the hypothetical case of a semiconductor for which DFT underestimated the electronic permittivity by 20%, we would have an error of $\sim 40\%$ in the energy renormalization, and of $\sim 10\%$ in the phonon frequencies. This example is an oversimplification of the problem, but it shows that different properties relating to the EPI could be affected to a very different degree by the inherent limitations of DFT functionals.

13 Conclusions

The study of electron-phonon interactions has a long and distinguished history, but it is only during the past two decades that quantitative and predictive calculations have become possible. First-principles calculations of electron-phonon couplings are finding an unprecedented variety of applications in many areas of condensed matter and materials physics, from spectroscopy to transport, from metals to semiconductors and superconductors. In this article we discussed the standard DFT formalism for performing calculations of electron-phonon interactions, we showed how most equations can be derived from a field-theoretic framework using a few well-defined approximations, and we reviewed recent applications of the theory to many materials of current interest.

As calculation methods improve relentlessly and quantitative comparisons between theory and experiment become increasingly refined, new and more complex questions arise. Much is still left to do, both in the fundamental theory of electron-phonon interactions, and in the development of more accurate and more efficient computational methods.

For one, we are still using theories where the coupling matrix elements are calculated using the adiabatic local density approximation to DFT. The need for moving beyond standard DFT and beyond the adiabatic approximation can hardly be overemphasized. Progress is being made on the incorporation of nonlocal corrections into electron-phonon matrix elements, for example using hybrid functionals or GW techniques, but very little is known about retardation effects. It is expected that such effects may be important in the study of heavily doped oxides and semiconductors, both in their normal and superconducting states (Ref. [9], Sec. 6.3.A), but ab initio investigations are currently missing. This is truly uncharted territory and a systematic effort in this direction is warranted.

In this article we emphasized that it is possible to formulate a compact, unified theory of electron-phonon interactions starting from a fully ab initio field-theoretic approach. The only assumption which is absolutely crucial to the theory is the harmonic approximation. Abandoning the harmonic approximation leads to the appearance of several new terms in the equations, and the resulting formalism becomes considerably more complex than in Table [1]. Despite these difficulties, given the importance of anharmonic effects in many systems of current interest, extending the theory to the case of anharmonic phonons and multi-phonon interactions constitutes a press-
ing challenge. *Ab initio* investigations of anharmonic effects on the temperature dependence of band gaps have recently been reported \[275, 294\]. Since these studies rely on non-perturbative adiabatic calculations in supercells, it would be highly desirable to establish a clear formal connection of these methods with the rigorous field-theoretic approach of Sec. [4]. Along the same line, it would be important to clarify the relation between many-body approaches, adiabatic supercells calculations, and more traditional classical or path-integral molecular dynamics simulations.

The study of electron-phonon interactions has long been dominated by Fröhlich-like Hamiltonians, whereby the electron-phonon coupling is retained only to linear order in the atomic displacements. This is the case for all the model Hamiltonians mentioned in Sec. [5.2.6]. It is now clear that quadratic couplings, leading to the Debye-Waller contributions in the optical spectra of semiconductors, are by no means negligible and should be investigated more systematically. For example, in the current literature it is invariably assumed that Debye-Waller contributions are negligible in metals near the Fermi surface; while this is probably the case for the simplest elemental metals, what happens in the case of multiple Fermi-surface sheets is far from clear, and should be tested by direct calculations.

The identification of the correct matrix elements to be calculated is not always a trivial task, as it was discussed for the case of the non-adiabatic phonon self-energy. In the future it will be important to pay attention to these aspects, especially in view of detailed comparison with experiment. For now, the issue on whether the phonon self-energy arising from EPIs should be calculated using bare or screened EPI matrix elements (Sec. [7]) is to be considered an open question, and calls for further investigation.

The theory and applications reviewed in this article focused on non-magnetic systems. The rationale for this choice is that a complete many-body theory of electron-phonon interactions for magnetic systems is not available yet. Recent investigations of spin-phonon couplings were conducted by assuming that the spin and the vibrational degrees of freedom can be decoupled, as in the Born-Oppenheimer approximation. Under this assumption it is possible to investigate how the spin configuration responds to a frozen phonon, or alternatively how the vibrational frequencies depend on the spin configuration (see for example Refs. [409–412]). In all these cases it would be desirable to employ a more rigorous many-body theory of spin-phonon interactions. The Hedin-Baym equations discussed in Sec. [4] maintain their validity in the case of spin-polarized systems, provided collinear spins are assumed. In more general situations, where it is important to consider noncollinear spins, external magnetic perturbations, or spin-dependent interactions such as spin-orbit and Rashba-Dresselhaus couplings, it becomes necessary to generalize the equations in Table [1]. Although such a generalization has not been reported yet, the work of Aryasetiawan and Biermann constitutes a promising starting point [113]. In that work the Schwinger functional derivative technique (see Sec. [4.2.1]) was used to extend Hedin’s equations at clamped nuclei to systems containing spin-dependent interactions. Generalizing Aryasetiawan and Biermann’s work to incorporate nuclear vibrations will be important for the study of electron-phonon interactions in many systems of current interest, from multifunctional materials to topological quantum matter.

At this time it is not possible to predict how this fast-moving field will evolve over the years to
come. However, the impressive progress made during the past decade gives us confidence that this interesting research area will continue to thrive, and will keep surprising us with fascinating challenges and exciting new opportunities.

Acknowledgments – I am indebted with Samuel Poncé, Carla Verdi, and Marios Zacharias for a critical reading of the manuscript, and with many colleagues who have kindly provided feedback on the arXiv preprint version of this article. This article is the result of many stimulating discussions with friends and colleagues over the last few years. I would like to thank in particular Philip Allen, Stefano Baroni, Lilia Boeri, Nicola Bonini, Fabio Caruso, Hyong Joon Choi, Marvin Cohen, Michel Côté, Andrea Dal Corso, Claudia Draxl, Asier Eiguren, Giulia Galli, Paolo Giannozzi, Stefano de Gironcoli, Xavier Gonze, Göran Grimvall, Eberhard Gross, Branko Gumhalter, Sohrab Ismail-Beigi, Emmanuil Kioupakis, Georg Kresse, Steven Louie, Roxana Margine, Andrea Marini, Miguel Marques, Nicola Marzari, Takashi Miyake, Bartomeu Monserrat, Jeffrey Neaton, Richard Needs, Jesse Noffsinger, Cheol-Hwan Park, Christopher Patrick, Warren Pickett, Samuel Poncé, Paolo Radaelli, Lucia Reining, John Rehr, Patrick Rinke, Angel Rubio, Matthias Scheffler, Young-Woo Son, Alexander Tkatchenko, Chris Van de Walle, Carla Verdi, Matthieu Verstraete, and Marios Zacharias. I also would like to thank Marco Bernardi, Asier Eiguren, Xavier Gonze, Eberhard Gross, Wu Li, Steven Louie, and Francesco Mauri for kindly allowing me to reproduce their figures. This work was supported by the Leverhulme Trust (Grant RL-2012-001), the Graphene Flagship (Horizon 2020 grant no. 696656-GrapheneCore1), and the UK Engineering and Physical Sciences Research Council (Grants No. EP/J009857/1 and No. EP/M020517/1).

\section*{A Born-von Kármán boundary conditions}

In this Appendix we provide more details on the notation related to the Born-von Kármán boundary conditions used throughout this article. The crystalline unit cell is defined by the primitive lattice vectors $a_i$ with $i = 1, 2, 3$, and the $p$-th unit cell is identified by the vector $R_p = \sum_i n_i a_i$ with $n_i$ integers between 0 and $N_i - 1$. The BvK supercell contains $N_p = N_1 \times N_2 \times N_3$ unit cells. The primitive vectors of the reciprocal lattice are denoted by $b_j$, and fulfill the duality condition $a_i \cdot b_j = 2\pi \delta_{ij}$. We consider Bloch wavevectors $q$ belonging to a uniform grid in one unit cell of the reciprocal lattice: $q = \sum_j (m_j/N_j)b_j$ with $m_j$ being integers between 0 and $N_j - 1$. This grid contains the same number of $q$-vectors as the number of unit cells in the BvK supercell. From these definitions the standard sum rules follow:

$$\sum_q \exp(iq \cdot R_p) = N_p \delta_{p0}, \quad \sum_p \exp(iq \cdot R_p) = N_q \delta_{q0}. \quad (213)$$

If $G$ is a reciprocal-lattice vector, the replacement of any of the $q$-vectors by $q + G$ in these expressions and in all expressions presented in this article is inconsequential, as $\exp(iG \cdot R_p) = 1$. Similarly any replacement of $R_p$ by $R_p + T$ where $T$ is a lattice vectors of the BvK supercell is inconsequential. Owing to these properties we are at liberty to replace the $q$-grid defined above with a Wigner-Seitz grid, i.e. the first Brillouin zone, and the supercell with a Wigner-Seitz supercell. These choices are useful for practical calculations in order to exploit the symmetry.
operations of the crystal, and to truncate the interatomic force constants, given by Eq. (13), outside a Wigner-Seitz supercell.

B Ladder operators in extended systems

In this Appendix we describe the construction of the phonon ladder operators \( \hat{a}_{\mathbf{q} \nu}/\hat{a}_{\mathbf{q} \nu}^\dagger \), and derive the phonon Hamiltonian given by Eq. (22). We show how the definition of the ladder operators depends on the behavior of the wavevector \( \mathbf{q} \) under inversion.

The normal modes introduced in Eq. (15) can be used to define a linear coordinate transformation of the ionic displacements as follows:

\[
z_{\mathbf{q} \nu} = N_p^{-\frac{1}{2}} \sum_{\kappa \alpha p} e^{-i \mathbf{q} \cdot \mathbf{R}_p} (M_\kappa/M_0)^{\frac{1}{2}} e_{\kappa \alpha \nu}^*(\mathbf{q}) \Delta \tau_{\kappa \alpha p}.
\]

Here \( z_{\mathbf{q} \nu} \) is referred to as ‘complex normal coordinate’ [111]. The exponential and the masses in Eq. (214) are chosen so as to obtain Eq. (22) starting from Eq. (12). Since there are \( 3MN_p \) degrees of freedom, and since the complex normal coordinates correspond to \( 2 \times 3MN_p \) real variables, this coordinate transformation carries some redundancy. Indeed by combining Eqs. (18) and (214) it is seen that:

\[
z_{-\mathbf{q} \nu} = z_{\mathbf{q} \nu}^*.
\]

The inverse relation of Eq. (214) is:

\[
\Delta \tau_{\kappa \alpha p} = N_p^{-\frac{1}{2}} (M_0/M_\kappa)^{\frac{1}{2}} \sum_{\mathbf{q} \nu} e^{i \mathbf{q} \cdot \mathbf{R}_p} e_{\kappa \alpha \nu}(\mathbf{q}) z_{\mathbf{q} \nu}.
\]

The right-hand side is real-valued after Eqs. (18) and (215). In preparation for the transition to a quantum description of lattice vibrations, it is useful to identify \( 3MN_p \) independent normal coordinates. This can be done by partitioning the grid of \( \mathbf{q} \)-vectors in three sets. We call \( \mathcal{A} \) the set of vectors which are invariant under inversion, that is \(-\mathbf{q} + \mathbf{G} = \mathbf{q}\) for some reciprocal lattice vector \( \mathbf{G} \) (including \(|\mathbf{G}| = 0\)). The center of the Brillouin zone and the centers of its faces belong to this set. The remaining vectors can be separated further in \( \mathcal{B} \) and \( \mathcal{C} \), in such a way that all the vectors in \( \mathcal{C} \) are obtained from those in \( \mathcal{B} \) by inversion (modulo a reciprocal lattice vector).

After defining \( z_{\mathbf{q} \nu} = x_{\mathbf{q} \nu} + iy_{\mathbf{q} \nu} \), Eq. (216) can be rewritten as:

\[
\Delta \tau_{\kappa \alpha p} = N_p^{-\frac{1}{2}} (M_0/M_\kappa)^{\frac{1}{2}} \left[ \sum_{\mathbf{q} \nu \in \mathcal{A},\nu} e_{\kappa \alpha \nu}(\mathbf{q}) x_{\mathbf{q} \nu} + 2 \text{Re} \sum_{\mathbf{q} \nu \in \mathcal{B},\nu} e^{i \mathbf{q} \cdot \mathbf{R}_p} e_{\kappa \alpha \nu}(\mathbf{q})(x_{\mathbf{q} \nu} + iy_{\mathbf{q} \nu}) \right].
\]

The \( \mathbf{q} \)-vectors of the set \( \mathcal{C} \) have been grouped together with those in \( \mathcal{B} \) by taking the real part in the second line. It can be verified that in this expression there are exactly \( 3MN_p \) real coordinates, therefore we can choose the \( x_{\mathbf{q} \nu} \) for \( \mathbf{q} \) in \( \mathcal{A} \) and the pairs \( x_{\mathbf{q} \nu}, y_{\mathbf{q} \nu} \) for \( \mathbf{q} \) in \( \mathcal{B} \) as the independent variables. These variables are referred to as ‘real normal coordinates’ [114].

Using Eqs. (12)-(18), (213), and (217) the nuclear Hamiltonian can be written in terms of \( 3MN_p \) independent harmonic oscillators in the real normal coordinates:

\[
\hat{H}_p = \frac{1}{2} \sum_{\mathbf{q} \nu \in \mathcal{B},\nu} \hbar \omega_{\mathbf{q} \nu} \left( -\frac{\partial^2}{\partial x_{\mathbf{q} \nu}^2} - \frac{\partial^2}{\partial y_{\mathbf{q} \nu}^2} + \bar{x}^2_{\mathbf{q} \nu} + \bar{y}^2_{\mathbf{q} \nu} \right) + \frac{1}{2} \sum_{\mathbf{q} \nu \in \mathcal{A},\nu} \hbar \omega_{\mathbf{q} \nu} \left( -\frac{\partial^2}{\partial \bar{x}^2_{\mathbf{q} \nu}} + \bar{x}^2_{\mathbf{q} \nu} \right),
\]
where for ease of notation we performed the scaling:

\[
\tilde{x}_{q\nu} = x_{q\nu}/2l_{q\nu} \quad \text{for } q \text{ in } A, \tag{219}
\]

\[
\tilde{x}_{q\nu} = x_{q\nu}/l_{q\nu}, \quad \tilde{y}_{q\nu} = y_{q\nu}/l_{q\nu} \quad \text{for } q \text{ in } B, \tag{220}
\]

with \(l_{q\nu}\) being the zero-point displacement amplitude of Eq. (21). In the case of \(|q| = 0\) there are three normal modes for which \(\omega_{q\nu} = 0\), and the corresponding potential terms \(\tilde{x}_{q\nu}^2\) must be removed from Eq. (218).

The eigenstates of Eq. (218) are found by introducing the real ladder operators for each normal coordinate [115]:

\[
\hat{a}_{q\nu,x} = \frac{1}{\sqrt{2}} \left( \tilde{x}_{q\nu} + \frac{\partial}{\partial \tilde{x}_{q\nu}} \right), \tag{221}
\]

and similarly for \(\hat{a}_{q\nu,y}\). With these definitions Eq. (218) becomes:

\[
\hat{H}_p = \sum_{q \in A, \nu} \hbar \omega_{q\nu} \left( \hat{a}_{q\nu,x}^\dagger \hat{a}_{q\nu,x} + 1/2 \right) + \sum_{q \in B, \nu} \hbar \omega_{q\nu} \left( \hat{a}_{q\nu,y}^\dagger \hat{a}_{q\nu,y} + 1 \right). \tag{222}
\]

The eigenstates of this Hamiltonian are products of simple harmonic oscillators [84], and the ground state is:

\[
\chi_0(\{\tau_{kp}\}) = A e^{-\frac{1}{2}(\sum_{q \in A, \nu} \tilde{x}_{q\nu}^2 + \sum_{q \in B, \nu} \tilde{y}_{q\nu}^2 + \tilde{y}_{q\nu}^2)}, \tag{223}
\]

with \(A\) a normalization constant. The relations between the positions \(\tau_{kp}\) and the normal coordinates \(\tilde{x}_{q\nu}, \tilde{y}_{q\nu}\) are given by Eqs. (214), (219), (220), and (21).

The eigenstates of \(\hat{H}_p\) can be generated by applying \(\hat{a}_{q\nu,x}^\dagger\) and \(\hat{a}_{q\nu,y}^\dagger\) to the ground state \(\chi_0\). However this approach is not entirely satisfactory, since we cannot assign separate quantum numbers to modes with wavevectors \(q\) or \(-q\). In order to avoid this inconvenience we observe that, for each normal mode, the first set of brackets in Eq. (218) defines an effective isotropic two-dimensional harmonic oscillator. The degenerate eigenstates of these oscillators can be combined to form eigenstates of the angular momentum; this leads to right and left circular quanta with the same energy and definite angular momentum [115]. This analogy motivates the consideration of the following linear combinations, for \(q\) in \(B\):

\[
\hat{a}_{q\nu}^+ = (\hat{a}_{q\nu,x} + i\hat{a}_{q\nu,y})/\sqrt{2}, \tag{224}
\]

\[
\hat{a}_{q\nu}^- = (\hat{a}_{q\nu,x} - i\hat{a}_{q\nu,y})/\sqrt{2}. \tag{225}
\]

Since both \(\hat{a}_{q\nu,x}\) and \(\hat{a}_{q\nu,y}\) lower the energy of an eigenstate by the same quantum of energy \(\hbar \omega_{q\nu}\), the resulting states are degenerate and their linear combinations are also eigenstates for the same eigenvalue. As a consequence we can generate all the eigenstates of the Hamiltonian \(\hat{H}_p\) by acting on the ground state \(\chi_0\) with the creation operators \(\hat{a}_{q\nu}^{+\dagger}\) and \(\hat{a}_{q\nu}^{-\dagger}\). In this reasoning the wavevectors \(q\) belong to \(B\); if we now consider Eqs. (215), (224), and (225) we see that formally we also have \(\hat{a}_{q\nu}^- = \hat{a}_{q\nu}^+\). Therefore it is natural to associate \(\hat{a}_{q\nu}^-\) to phonons propagating along the direction \(-q\).

These observations suggest replacing the real ladder operators of Eq. (221) by the complex ladder operators \(\hat{a}_{q\nu}^+\) and \(\hat{a}_{q\nu}^-\) for \(q\) in \(B\) and \(C\), respectively. In the case of \(q\) in \(A\) we keep the real
operators $\hat{a}_{q,\nu,x}$. These definitions can be turned into the compact expressions:

$$\hat{a}_q = \hat{a}_{q,\nu,x} \quad \text{for } q \text{ in } A, \quad (226)$$

$$\hat{a}_q = \frac{(\hat{a}_{q,\nu,x} + i\hat{a}_{q,\nu,y})}{\sqrt{2}} \quad \text{for } q \text{ in } B, C. \quad (227)$$

Using these operators the nuclear Hamiltonian of Eq. (218) takes the well-known form given by Eq. (22). Any eigenstate of $\hat{H}_p$ can now be generated as $\prod_{q,\nu} (n_{q,\nu})^{-\frac{1}{2}} (\hat{a}_q^\dagger)^{n_{q,\nu}} \chi_0$. In this form we see that it is possible to assign independently a number of phonons $n_{q,\nu}$ to each wavevector $q$ and each mode $\nu$. Using Eqs. (219)-(221) and (226)-(227) we also have the basic identity:

$$z_{q,\nu} = l_{q,\nu} (\hat{a}_q + \hat{a}_q^\dagger). \quad (228)$$

By combining this last expression with Eq. (216) we obtain Eq. (20).

References


[127] Keating, P. N. Dielectric Screening and the Phonon Spectra of Metallic and Nonmetallic


[130] Maksimov, E. G. A self-consistent description of the electron-phonon system in metals and


[133] Marini, A., Poncé, S. & Gonze, X. Many-body perturbation theory approach to the


[135] Fetter, A. L. & Walecka, J. D. *Quantum theory of many-particle systems* (Dover Publica-
tions, 2003).


[137] Kadanoff, L. P. & Baym, G. *Quantum statistical mechanics: Green’s function methods in
equilibrium and nonequilibrium problems.* Frontiers in physics (Benjamin, 1962).

in Statistical Physics* (Dover, 1975).

[139] Yin, M. T. & Cohen, M. L. *Ab initio* calculation of the phonon dispersion relation: Ap-

[140] Dal Corso, A., Baroni, S., Resta, R. & de Gironcoli, S. *Ab initio* calculation of phonon

[141] Lee, C., Ghosez, P. & Gonze, X. Lattice dynamics and dielectric properties of incipient


