Response of periodic systems to electromagnetic fields: multipole expansion, microscopic charge-current density, polarization and magnetization

by

Sylvia D. Swiecicki

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Physics
University of Toronto

© Copyright 2017 by Sylvia D. Swiecicki
Polarization and magnetization fields are central concepts in the theory of light-matter interaction, because they describe the distribution of charges and currents induced in the material medium due to an application of electromagnetic fields. Here I develop strategies to treat polarization and magnetization fields in periodic structures and I analyze the radiation spectrum emitted by them. The methods are based on the multipole expansion of charge-current distributions and on its extensions to account for the motion of charges between lattice sites.

In the first part of this thesis I treat the optical response of 2d nanoparticle arrays, with the focus on collective effects that arise due to higher order multipolar response. First, I evaluate the interaction constants that describe the collective radiative interactions between nanoparticles in magneto-electric quadrupolar arrays, in an approach that gives the radiative damping in an exact analytic form. Then I find the radiation spectrum from an array of gold spheres with large electric quadrupolar response. The spectrum is dominated by surface-lattice resonances (SLRs) which are of a mixed multipolar character, with significant contributions from multipoles that would be negligible if the sphere was isolated. I link the SLRs to an excitation of the normal modes of the array, and propose a simplified model of the normal mode dispersion relations to explain the SLR properties.

In the second part of this thesis I develop strategies to describe charges and currents and polarization and magnetization fields in regular solids. The formalism is based on gauge-invariant dynamical equations of electron Green functions. I derive and categorize the descriptions of dynamics introduced via Peierls transformation of the Green function. Then I apply the formalism to treat the linear response of an insulator to electromagnetic fields. I find the microscopic charge and current densities in a form such that their expansion around lattice sites follows in a natural way. Restricting myself to static and uniform fields I introduce microscopic polarization and magnetization fields, and derive all the response coefficients within one formalism that makes no reference to energy or thermodynamic potentials.
Acknowledgements

First and foremost, I would like to express my sincere gratitude to my advisor prof. John Sipe for his continuous support. John always had the time to discuss my research with me, and he gave a lot of valuable and insightful feedback on my work. John’s patience and his never-ending enthusiasm made him a great supervisor. His guidance and expertise helped me in all the time of research and in writing of this thesis.

Additionally, I would like to thank my PhD committee members, professor Daniel James, professor Luyi Yang, professor Joseph Thywissen, and professor Kenneth Burch for their insightful questions and comments. I would also like to thank all the group members that I have met over the past six years – Luke, Federico, Jin Luo, Temok, Rodrigo, Kiran, Zachary, Daniel, Zaheen, and Nicolas – I enjoyed getting to know all of them.

I am very grateful to my roommate and my best friend Jennifer for being a great company over the past six years. Jennifer made my time in grad school much more cheerful and enjoyable, and was always the person to turn to in good and bad times. I hope our friendship will continue to be as strong after we graduate.

Finally, I would like to thank my family for their continuous love and support. They always had faith in me, and encouraged me to continue with my research and my goals when things looked dire. Their support gave me the strength to overcome the difficult moments of this PhD.
## Contents

1 Introduction 1

2 Preliminary expressions 8
   2.1 Classical polarization and magnetization fields 8
   2.2 Power-Zienau-Wooley transformation 13

3 Optical theorem in magneto-electric quadrupolar arrays 18

4 Surface - lattice resonances in 2d arrays of spheres 19
   4.1 Introduction 19
   4.2 Multipolar model 21
      4.2.1 Isolated sphere 21
      4.2.2 2D triangular lattice 23
   4.3 Resonance spectrum and the multipolar couplings 27
   4.4 Surface-lattice modes 31
      4.4.1 Energy balance 32
      4.4.2 Dispersion relations 33
   4.5 Simplified model 36
   4.6 Surface-lattice modes and the sensitivity of SLRs to non-resonant multipoles 39
   4.7 Dependence of the spectrum on illumination conditions 41
   4.8 Direct coupling to surface-lattice modes of the array 44
   4.9 Conclusions 48

5 Gauge-invariant Green function dynamics 50
   5.1 Introduction 50
   5.2 General formalism 52
      5.2.1 Generalized LF 52
      5.2.2 GLF Green function dynamics 55
   5.3 The link to the LF approach 57
   5.4 The link to PZW 59
   5.5 Hamiltonians and Green functions 63
   5.6 Summary 66
6 Microscopic charge-current densities, polarizations and magnetizations 68
  6.1 Introduction ...................................................................................... 68
  6.2 Linear response to time and spatially dependent fields ....................... 71
    6.2.1 Construction of the multipole Green function ............................... 73
  6.3 Polarization and magnetization fields associated with lattice sites .......... 78
  6.4 Multipole expansion of polarization and magnetization fields ............... 83
  6.5 Discussion and future outlook ............................................................. 86

7 Conclusions ....................................................................................... 88

A Free-space Green functions .................................................................. 92

B Sum rules .......................................................................................... 94

Bibliography ......................................................................................... 97
List of Figures

1.1 Two choices of a unit cell in a crystal ................................................................. 6

4.1 Extinction cross section of an isolated sphere and contributions from multipoles: ED - electric dipole, MD - magnetic dipole, EQ - electric quadrupole, OC - electric octupole. Dotted line shows the total absorption cross section ...................................................... 21

4.2 Illumination conditions and reciprocal lattice. The incident wave vector is \( \mathbf{v}_+ \), the angle of incidence is \( \theta_0 \), and the lattice vectors are \( \mathbf{a}_1, \mathbf{a}_2 \). The reciprocal lattice, indicated by squares, is spanned by the basis vectors (10) and (01); we also indicate the reciprocal vector (11) explicitly. The black continuous and dashed lines in the diagram of the reciprocal lattice are the boundaries of the first and the second Brillouin zone respectively .......................................................... 23

4.3 The light line (red) and the Rayleigh line (black) for light incident with wave vector component in the plane \( \kappa_0 \) in the direction \( \hat{\kappa}_a \). The shading indicates the incident plane waves, specified by \( (\kappa_0, \tilde{\omega}) \), for which diffraction is possible. I consider resonance spectra in frequency range restricted by the two dotted lines ................................................. 28

4.4 Resonant structure of the specular reflection for s-polarized light and \( \kappa_0 \propto \hat{\kappa}_a \). Arrows identify the angles at which the Rayleigh line is crossed and diffraction appears. Plots are shifted by 1 to improve clarity ................................................................. 29

4.5 Total specular reflection at \( \lambda = 750 \text{nm} \) (continuous line) and the contributions from the two sets of multipole moments (dashed lines); the dotted lines show the SLMs supported by the array. The arrow indicates an angle at which the Rayleigh line is crossed .... 29

4.6 Comparison of the total reflectivity predicted by the multipolar model, which includes multipole contributions from the electric and magnetic dipoles and the electric quadrupole, with the exact result obtained from the layered-multipole scattering method .................................................. 30

4.7 The contribution to specular reflection from the set (4.33), compared with the contribution from the electric dipole only. Arrows identify the angles at which the Rayleigh line is crossed and diffraction appears ................................................................. 31

4.8 The contribution to specular reflection from the set (4.34), compared with the contribution from the electric quadrupole only. Arrows indicate the angle at which the Rayleigh line is crossed and diffraction appears ................................................................. 32

4.9 For a fixed \( \omega \), the SLMs at \( \kappa_{(10)} = \kappa_0 + \mathbf{K}_{(10)} \) and \( \kappa_{(11)} = \kappa_0 + \mathbf{K}_{(11)} \). The solid blue line indicates the “light circle”, the set of \( \kappa \) for which at chosen frequency \( \kappa = n\omega/c \). The dotted blue lines indicate the “Rayleigh circles”, which are light circles centered at reciprocal lattice sites other than that at the origin; diffraction associated with a particular \( \mathbf{K} \) occurs when \( \kappa_0 \) crosses its Rayleigh circle ................................................................. 34
4.10 Ideal dispersion relations of the two SLMs. Blue lines identify the light line (solid) and the Rayleigh line (dashed) of the SLM. The terminations of the dispersion relations are indicated by circles.

4.11 Scenario in the reciprocal lattice at the cut-off frequency $\omega_{c\ell}$. The SLM wave vector $\kappa = \tilde{\omega}_{c\ell} n \hat{\kappa}_a + K$, for $K = K_{(10)}$ or $K = K_{(11)}$, resides on a Rayleigh circle.

4.12 Dispersion relations found within the multipolar model (continuous lines) and within the simplified model (dashed lines). Insets show the “dressed” polarizabilities of the dipolar (left) and multipolar (right) SLMs with and without the coupling to magnetic dipole. When the coupling is included, the polarizabilities vanish at $\tilde{\omega}_v$ and $\tilde{\omega}_v'$ respectively.

4.13 Resonant structure of the specular reflection for p-polarized light with $\kappa_0 \propto \hat{\kappa}_a$. Arrows indicate the angle at which the Rayleigh line is crossed and diffraction appears. Plots are shifted by 1 to improve clarity.

4.14 Specular reflection for s-polarized light at $\lambda = 750$ nm for different values of $\phi$. Inset on the left shows the first Brillouin zone with the two symmetry directions indicated by $\hat{\kappa}_a$ and $\hat{\kappa}_b$; compare with the inset in Fig. 4.2, where the second Brillouin zone is also indicated. Inset on the right shows the disappearance of multipolar SLR at small angle $\phi$.

4.15 The array at a distance $l$ above an interface. The interface is between two semi-infinite homogenous media, with the background having the index of refraction $n = 1.45$, and the substrate having the index of refraction $n_s = 1.75$. The array is illuminated by a plane wave incident from the substrate.

4.16 First and second Brillouin zone of the reciprocal lattice. For illumination with in-plane wave vector $\kappa_{inc}$ along $\hat{\kappa}_b$, the SLR is found at an in-plane wave vector $\kappa_{SLR}$ and the SLM at $\kappa_{SLM} = \kappa_{SLR} + K_{(01)}$.

4.17 Specular reflection from the array at a distance $l_0 = 337.2$ nm from an interface, at $\lambda = 900$ nm. The dashed and dotted lines indicate the angles at which incident light would couple directly and with the help of the grating to the SLM of the array if it were free-standing in a background medium. The inset shows deviations from these angles of the resonance features, measured as a distance in reciprocal space, as $l$ increases from its value at $l_0$.

5.1 The dashed line is the path in space-time used by Levanda and Fleurov; the solid line indicates a generalization.

5.2 A path in space-time going through a special point in space.

5.3 The path in space-time used to introduce $\gamma(x)$.

5.4 The geometry of the Aharonov-Bohm effect.

5.5 The path not making use of a special point.

5.6 The function $J$ as defined in (5.71) for $R_s/r_e = 0.75$.

6.1 The path going through a special point $R$ leading to the PZW transformation. In the simplest case the straight line paths between points are chosen.

6.2 The path defining the global Peierls phase $\Gamma_{GL}(x, y; T)$, represented as a sum of the path leading to the PZW transformation plus a correction.
Chapter 1

Introduction

Understanding and describing the interaction of light and matter is the fundamental problem of optical physics. From the early years of the twentieth century the response of material media to applied electromagnetic fields has been described in terms of polarization and magnetization fields of the medium [1]. Polarization and magnetization fields describe the distribution of charges and currents within material media averaged over atomic scale variations, and in simple models they can be approximated by the electric dipole and magnetic dipole moments per unit volume of the molecules [2]. When the material medium has a non-vanishing polarization or magnetization, either due to permanent molecular moments or induced because of an application of external fields, the medium becomes a source of electromagnetic fields itself, with polarization and magnetization fields acting as source terms. The macroscopic electromagnetic fields that exist in the presence of a material medium are described in classical electromagnetism by the usual macroscopic Maxwell equations. In order to solve these equations it is necessary to link the polarization and magnetization fields to the macroscopic electromagnetic fields within the material. Such a relation is identified by the material response coefficients – the so called susceptibility tensors – which are thus central concepts in the description of the optical response of materials. The values of the material susceptibilities that are used in optical response calculations are often extracted from experimental data. Nevertheless, to fully understand the optical response of a material medium a theory of polarization and magnetization is needed, which would link susceptibilities to the microscopic properties of the medium such as multipole moments of molecules or the band structure of a crystal.

The first microscopic treatments of material media date back to the beginnings of the 20th century [1]. In those early treatments a solid is considered as a collection of point charges that are assumed to be either free or bound to a molecule. Free charges – conduction electrons or ions – move throughout medium over macroscopic distances. Bound electrons, on the other hand, remain at all times close to a molecule with which they are identified. No transfer of bound electrons from one molecule to another is considered, so that each molecule together with all the electrons associated with it constitutes an electrically stable unit. A center of each stable unit is identified, which is typically chosen at the center of mass or the center of a charge distribution of a molecule, and the electric and magnetic dipole moments of each unit are calculated with respect to that center. The distribution of molecular multipole moments and the electromagnetic fields in the medium vary rapidly over microscopic length scales, and thus they are averaged over a volume that is microscopically large but macroscopically small. Macroscopic Maxwell equations follow, with polarization and magnetization fields given by average electric and magnetic dipole
moments of molecules per unit volume. This early treatment was subsequently extended to include the
response due to higher multipole moments up to any order required. Higher order multipole moments can
become important in a description of an optical response, when the dipole moments vanish for reasons
of symmetry, or when the size of molecules is a considerable fraction of a wavelength of light [3, 4].

Polarization and magnetization fields within the simple model described above are defined in terms
of molecular multipole moments, irrespective of how these moments are induced. But in order to find
susceptibility tensors, the magnitude of the molecular moments needs to be linked to electromagnetic
fields within the medium. I consider first constitutive relations in a typical situation of a non-magnetic
material responding to weak electromagnetic fields at frequency $\omega$. In this case magnetization field
can be neglected, and polarization can be taken as responding locally to electric field [2],
\[
P_i(x) = \chi_{ij}E_j(x),
\]
where $P(x)$ is the frequency component of polarization field, $\chi$ is the frequency component of electric
susceptibility tensor, and $E(x)$ is the frequency component of Maxwell electric field at $\omega$; here and
henceforth I suppress the frequency dependence of frequency components for a simpler notation. A
simple microscopic treatment of the susceptibility tensor $\chi$ in eq. (1.1) is provided by the Lorentz point-
electron model [5]. Within this model, the response of electrons is approximated by damped harmonic
oscillators with a certain resonant frequency $\omega_0$. The resulting dipole moment $p$ of an oscillating electron
has a Lorentzian profile,
\[
p = \frac{e^2}{m} \frac{1}{\omega_0^2 - \omega^2 - i\Gamma\omega}E_{\text{loc}},
\]
where $\omega$ is the frequency of incident fields, $\Gamma$ is a damping constant, and $E_{\text{loc}}$ is the local electric field at
the position of the electron. The local electric field is in general different from the macroscopic electric
field that enters Maxwell equations and the constitutive relations; the difference between those fields is
known as the local field correction. Hence to identify the susceptibility tensor a relation between the
Maxwell field and the local field needs to be established. On physical grounds the local field corrections
arise because the contributions to the electric field from molecules that are close to the dipole cannot
be treated in an averaged way. A division of the contributions into those from molecules that are near
and those that are far constitutes the basis of the Clausius-Mosotti treatment [2]. Within this approach
the contributions from molecules that are within a macroscopically small sphere from the dipole are
treated in an exact manner, by noting that in cubic lattices these contributions vanish for symmetry
reasons. The contributions from outside the sphere are treated in an averaged way. These contributions
give rise to a Maxwell field modified by a local field correction that is proportional to polarizability of
the medium. With the local field established, the response of electrons described by the Lorentz model
(1.2) is spatially averaged, and the susceptibility tensor $\chi$ is found. It should be noted, however, that
the simple constitutive relation identified in eq. (1.1) does not always give a sufficient description of the
material response. When the response of material medium is non-local or the magnetic effects cannot
be neglected, more general constitutive relations need to be considered, which involve a response of
polarization and magnetization fields to electromagnetic fields and to their derivatives. And of course
the very model of electrons bound to individual nuclei, with their response described classically, is
suspicious from the perspective of condensed matter physics, where Bloch states extended through a
crystal often form a more natural basis for a calculation.

Yet recent work on metamaterials connects with this earlier work on atoms and molecules. Metamate-
Chapter 1. Introduction

Materials are artificial structures constructed by arranging nanoparticles in clusters or arrays. In these systems the electrons can be treated to good approximation as if each is associated with a single nanoparticle, and the optical response of each nanoparticle can be treated by introducing polarizabilities associated with its different multipole moments. Although at a fundamental level these must be calculated quantum mechanically, once they are identified an approach based on classical electromagnetism can be used to characterize the response of the system of nanoparticles. Over the past few years metamaterials have gained a lot of attention in the optics community due to a large degree of control that can be exerted over their optical properties [6]. The optical response of individual nanoparticles can be tailored by changing nanoparticle size, shape and composition [7, 8, 9, 10]. An additional degree of control is achieved in collections of nanoparticles, since radiative interactions between nanoparticles affect the frequency and line-width of resonances [11, 12, 13, 14, 15, 16]. Collective effects are especially prominent in the optical response of arrays at frequencies close to the onset of a new diffraction order [6]. At the onset of a new diffraction order the constants that describe radiative interactions between nanoparticles can diverge, and in the vicinity of frequencies that correspond to these divergences the so-called “surface-lattice resonances” are often observed [17, 18, 19, 20]. Surface-lattice resonances can be tailored by changing the array spacing and geometry [21, 12], and resonances with very narrow line-widths have been observed that are of interest for potential sensing applications and for field enhancement [22, 20].

Radiation from collections of nanoparticles is often analyzed using full-numerical simulations. However, an analytic multipolar model is often needed in order to understand the resonance spectrum. Within a multipolar model the exact optical response of nanoparticles is approximated by the response of their first few multipole moments. Generally speaking there are two approaches to constructing a multipole expansion of electromagnetic fields [2]. The first approach is an expansion in powers of a size of a charge-current distribution relative to the wavelength of light. Multipole moments introduced within this approach are defined in terms of powers of the position vector integrated with a charge distribution or an angular momentum of a current distribution; I will refer to them as “simple” moments. The second approach is an expansion in terms of polarity of the emitted radiation, and is introduced by decomposing electromagnetic fields into vector spherical harmonics. Multipole moments introduced in this approach involve integrals of spherical Bessel functions taken with charge and current distributions, and I will refer to them as “general” moments. The simple and general moments become equivalent in an extreme wavelength limit. When treating a response of a nanoparticle that is a considerable fraction of a wavelength of light, however, more terms need to be kept if the expansion is carried out in terms of the simple multipole moments, and a description in terms of general moments is preferred. Thus when I adopt a multipolar treatment of nanoparticles, I always use an expansion in terms of the general moments.

A multipolar treatment of a cluster of a few nanoparticles follows in a straightforward way if a multipolar description of individual nanoparticles is known. When nanoparticles are arranged in an infinite array, the description is more complicated. In an array each multipole moment responds to electromagnetic fields radiated from all the other multipoles in an array [23, 24, 25, 13, 26, 27]. These interactions are described by periodic Green function dyadics, which take a form of slowly convergent summations over lattice sites, and so acceleration techniques need to be used to evaluate them [28, 29]. The Ewald acceleration method is the most widely used technique, and its explicit implementations have been introduced for electric and magneto-electric arrays [30, 31, 32, 27, 13]. Ewald summation, however, does not explicitly show energy balance on a lattice, because the total radiation reaction field acting on
a multipole moment in an array is given in terms of an infinite series. Identifying radiation reaction in an exact analytic form facilitates an analysis of an optical response of the array. Nevertheless, the exact expressions for the radiation reaction fields have been obtained only for electric and magneto-electric rectangular arrays, using less known summation methods \[25, 33\].

The level of approximation that is necessary to treat an array within multipolar model depends on the size and composition of nanoparticles. A leading contribution to radiation from plasmonic arrays is due to electric dipole moments of nanoparticles \[34, 24, 35, 36\]. However, when nanoparticles are dielectric or when their size is a considerable fraction of the wavelength of light, contributions due to higher order multipole moments become important \[37, 38, 39, 7\]. Dielectric arrays suffer less severe Ohmic losses than plasmonic arrays, and thus are promising candidates for energy wave-guiding \[38, 40\]. Arrays of large nanoparticles support quadrupolar resonances that couple only weakly to light and thus have large quality factors \[7, 40\]. But despite some promising features of higher order multipolar response, multipolar treatments of nanoparticle arrays beyond the usual dipole approximation are rather limited, especially at a quadrupole level. Only few such treatments have been reported, and the scope of the analysis is restricted: often only illumination at normal incidence and along high symmetry directions was considered, and couplings between multipole moments of different polarity were neglected \[7, 40\]. The lack of a thorough analysis of higher order response from nanoparticle arrays is due to a complexity of the problem. There are two main difficulties. First, explicit implementations of quadrupolar periodic Green functions have not been reported. Secondly, the analysis of the system is difficult due to a large number of coupled multipole components.

In the first part of this thesis I analyze radiation from 2d arrays of nanoparticles. Nanoparticles are treated at the level of including their electric dipole moment, magnetic dipole moment, and electric quadrupole moment. In chapter 3, I evaluate all the periodic Green functions that describe radiative interactions between nanoparticles in electro-magnetic quadrupolar arrays of arbitrary geometry. The approach that I use gives simple analytic expressions for all the radiation reaction terms that determine energy balance in a system. Non-radiative contributions are given in a form of quickly converging series that are suitable for numerical computations. The content of this chapter is published in Journal of Optics \[41\]. In chapter 4, I use the Green functions representation found in chapter 3 to analyse surface-lattice resonances in 2d arrays of spheres of moderate sizes. The focus of this analysis is on magnetic and quadrupolar response of the spheres, and on effects that arise due to couplings between multipoles of different polarities. I consider different illumination conditions, link resonant response to an excitation of normal modes of an array, and propose a simplified model of the normal mode dispersion. The model is well suited to treat complex systems of coupled multipoles, and it explains a sudden termination of dispersion relations at low frequencies. The content of this chapter is published in Physical Review B \[42\].

In the second part of this thesis I develop strategies to treat polarization and magnetization fields beyond the simple models described in the previous paragraphs. Classical Lorentz-type models can give at most a qualitative description of the material response. In order to gain a deeper understanding of optical properties of materials a rigorous theory of polarization and magnetization is necessary. Rigorous quantum mechanical treatments of polarization and magnetization were developed essentially independently within the atomic physics and condensed matter communities. The former approach treats atoms and molecules interacting with time- and spatially-varying electromagnetic fields. The latter approach treats crystals that are in the ground state or interact with uniform electromagnetic fields. Below I give
an overview of these two established theories and discuss their limitations.

In atomic physics, polarization and magnetization are introduced using the Power-Zienau-Wooley (PZW) transformation [43, 44]. This is a highly successful approach that provides a complete quantum mechanical framework for a description of atoms and molecules in electromagnetic fields. It is introduced by transforming a wave function or a field operator of a molecule with the use of a Peierls phase [43, 45],

\[ \psi(x,t) = e^{-i\Phi_{\text{PZW}}(x,t)} \tilde{\psi}(x,t). \]  

(1.3)

Here \( \psi(x,t) \) is the usual field operator described by the minimal-coupling Hamiltonian, \( \tilde{\psi}(x,t) \) is the transformed PZW field operator, and \( \Phi_{\text{PZW}} \) is a Peierls phase. The Peierls phase is defined as an integral of a vector potential along a path in space from a special point of the transformation to the field variable. The special point of the transformation is usually chosen at the center of a molecule and it provides a reference point with respect to which all the expansions within the PZW formalism are carried out. After the Peierls transformation (1.3) is carried out, the molecule is described in terms of its microscopic polarization and magnetization fields. Microscopic polarization and magnetization fields contain the same information about the state of a molecule as microscopic charge and current distributions, but the use of polarization and magnetization fields is advantageous because they can be easily expanded around the special point of the transformation. The expansion leads to a description of a molecule in terms of multipole moment operators, which can be included up to any order required. Thus approximate descriptions of the molecule can be easily constructed, with the degree of approximation controlled by the order of multipole moments kept in the expansion. Charge and current densities are then found in terms of polarization and magnetization fields and their expansions using expressions of a form analogous to these known from classical electromagnetism. In addition to describing the charge-current densities, polarization and magnetization fields also describe the dynamics of the system. The PZW Hamiltonian involves polarization and magnetization fields coupled to electromagnetic fields. When the fields are expanded around the special point, the interaction Hamiltonian is given by molecular multipole moments coupled to electromagnetic fields and their derivatives evaluated at the center of a molecule. The ease with which a multipolar expansion can be constructed and the simple physical interpretation of terms that result is however not the only advantage of the PZW approach. As the dynamics is formulated directly in terms of electromagnetic fields rather than electromagnetic potentials, the sum rules that often need to be identified when working with the minimal coupling Hamiltonian are not needed. For the reasons described above PZW transformation is a highly successful theory describing the optical response of atoms and molecules.

A rigorous quantum-mechanical theory of polarization and magnetization of a solid was introduced only relatively recently [46, 47]. In a solid electrons are de-localized over a unit cell and can move from one lattice site to the next. For that reason a multipolar expansion cannot be carried out independently at each lattice site, and hence the PZW treatment is not applicable in a straightforward way. The approach adopted within condensed matter community was very different, as it was based primarily on macroscopic quantities and on thermodynamic considerations. A change in polarization under an adiabatic evolution of an insulating crystal was linked to the total current that flows through the bulk [48]. Magnetization was introduced starting from a macroscopic definition of a magnetization for a finite sample, and then the limit of an infinite crystal was taken [49]. Magnetization and a change in polarization between initial and final states of a crystal were shown to be bulk quantities, and they were linked to a crystal band structure. While the theory can be formulated in terms of Bloch states,
a particularly intuitive description of polarization and magnetization emerged when Wannier function basis was used instead. Wannier functions are Fourier transforms of the Bloch functions,

$$W_{nR}(x) = \sqrt{\frac{V_c}{(2\pi)^3}} \int_{BZ} d^3k e^{-ik \cdot R} \psi_{nk}(x).$$  

(1.4)

and they are exponentially localized around lattice sites provided that Bloch functions can be chosen that are smooth throughout the Brillouin zone [46]. Written in the Wannier basis, the expressions for polarization and magnetization bear some resemblance to the simple classical picture. Polarization per unit volume of an insulator in its ground state was linked to electric dipole moments of Wannier function centres [48],

$$P = \frac{e}{V_c} \sum_v \int d^3x W^*_v(x) x W_v(x),$$  

(1.5)

where $V_c$ is a volume of a unit cell and summation runs over all occupied valence bands in an insulator. A known feature of this theory is that polarization (1.5) is defined only up to a “quantum” of ambiguity, $\Delta P = eR/\hbar c$. This ambiguity follows from an ambiguity in how Wannier functions are associated with lattice sites, and can be understood by considering a simple toy model of a crystal [50]. The toy model consists of a 1d periodic arrangement of positive ions and negative electron clouds with a period of $R$ depicted in Fig. 1.1. The dipole moment of a unit cell of that crystal can be found in two ways: first by associating each electron cloud with an ion to the right, and then by associating it with an ion to the left. Those two calculations are equivalent to identifying two different unit cells, indicated with dotted lines in the top and bottom of Fig. 1.1 respectively. The two calculations lead to a dipole moment per unit cell equal to $eR/2$ and $-eR/2$ respectively, with the difference of $eR$ being a quantum of ambiguity. That quantum of ambiguity cancels out when a difference between polarization of a crystal in two states is calculated, thus the polarization differences are unambiguous. Magnetization per unit volume of a crystal in its ground state was found to be unambiguous and to involve two kinds of contributions [49]. The first contribution to magnetization is local as it was linked to circulation of Wannier function centres,

$$M_{loc} = \frac{e}{2\hbar V_c} \int d^3x W^*_v(x) x \times v W_v(x),$$  

(1.6)

where $v = i[H, x]/\hbar$ is the velocity operator. The second contribution is itinerant as it is due to a current that flows along the surface of the sample, but it was also shown to be associated with the bulk magnetization. These considerations were subsequently extended to a treatment of solids in uniform
electromagnetic fields [51, 52, 53], but in an approach that cannot be easily generalized to non-uniform perturbation. An alternative approach was reported to treat dynamics of charges in a crystal interacting with time- and spatially-dependent electromagnetic fields, that is based on Peierls transformation of electron correlation function [54, 55]. The Peierls transformation is carried out by employing a phase defined as an integral of a four-potential along a straight-line path in space-time, and the dynamical equations that govern the evolution of this transformed correlation function were found for arbitrary applied electromagnetic fields. However, this formulation of dynamics is so complex that it has not been widely adopted. The complexity is due to the choice of the path in space-time and because the approach does not make use of the averaging procedure that is identified by polarization and magnetization fields.

In the second part of this thesis I develop strategies to treat the response to electromagnetic fields of charges and currents and polarization and magnetization fields in a crystal at a microscopic level. The approach is based on a Peierls transformation of electron correlation function. There are two reasons that motivate this approach. First, the use of Peierls transformation leads to a gauge-invariant description of the system, which does not require an identification of sum rules. Second, when the system is described by a correlation function rather than an electron wave function a broader class of Peierls transformations can be considered. The use of a correlation function is thus advantageous even if calculations are done at a single particle level. In chapter 5, I develop and categorize descriptions of dynamics introduced with the use of Peierls-transformed electron correlation functions in crystals interacting with electromagnetic fields. I derive gauge-invariant dynamical equations satisfied by the correlation function for an arbitrary choice of an integration path in Peierls phase. Broadly speaking, all the paths can be seen to belong to one of the two categories. For paths in the first category the dynamics can be seen to follow from an effective Hamiltonian. For paths in the second category an effective global Hamiltonian cannot be introduced. By considering an Aharonov-Bohm scenario I argue that the latter paths are more useful in treating extended systems much larger than the wavelength of light. The content of this chapter is based on a publication in Annals of Physics [56]. In chapter 6, I use the formalism introduced in chapter 5 to treat a linear response of an insulator to electromagnetic fields. I find the charge and current densities in the crystal interacting with arbitrary electromagnetic fields. The expressions are in a form that allows for their easy expansion around lattice sites, but also accounts for motion of electrons from one lattice site to another. Taking then the limit of static and uniform fields I use those results to derive all the earlier expressions for polarization and magnetization within one unified framework, in an approach that makes no reference to macroscopic and thermodynamic quantities. The content of this chapter is published in Physical Review B [57]. I summarize and present the conclusions in chapter 7.

Chapters 3 and 4 of this thesis are written in the SI units. Chapters 5 and 6 are written in the Gaussian units. This choice of units reflects the usual notation used by the community that I address in the papers on which those chapters are based.
Chapter 2

Preliminary expressions

In this chapter I review two well-established treatments of material media. In the first section I give an overview the usual treatment of material media within classical electromagnetism. The macroscopic Maxwell equations are derived via a spatial averaging procedure from the microscopic equations, with the spatially averaged charge and current distributions expressed in terms of polarization and magnetization fields that are linked to molecular multipole moments. In the second section I give an overview of the Power-Zienau-Wooley (PZW) theory, which provides a rigorous quantum mechanical framework to treat the optical response of atoms and molecules at the microscopic level. The PZW theory is formulated in terms of microscopic polarization and magnetization fields, which play a two-fold role. First, they characterize the distribution of charges and currents within the molecule through expressions analogous to those introduced in classical electromagnetism. Second, they describe the dynamics of the system, which evolves according to a gauge-invariant Hamiltonian described by polarization and magnetization fields coupled to electromagnetic fields. The results of this chapter are given in Gaussian units, because the content of those sections is most relevant to work presented in chapters 5 and 6.

2.1 Classical polarization and magnetization fields

In this section I give an overview of the usual procedure used in classical electromagnetism to derive the macroscopic Maxwell equations [1, 2]. I start with the microscopic Maxwell equations,

\[ \nabla \cdot b(x, t) = 0, \]
\[ \nabla \cdot e(x, t) = 4\pi \rho(x, t), \]
\[ \nabla \times e(x, t) + \frac{1}{c} \frac{\partial b(x, t)}{\partial t} = 0, \]
\[ \nabla \times b(x, t) - \frac{1}{c} \frac{\partial e(x, t)}{\partial t} = \frac{4\pi}{c} j(x, t), \]

where \( e(x, t) \) and \( b(x, t) \) are the microscopic electric and magnetic fields, \( \rho(x, t) \) is the microscopic charge density and \( j(x, t) \) is the microscopic current density. Within the usual approach, the medium is assumed to consists of the free and bound charges [2, 5]. Free charges, such as conduction electrons, can move over macroscopic distances. Bound charges are assumed to remain at all times confined to the vicinity of the molecule with which they are identified. That is, no motion of bound charge from one
Chapter 2. Preliminary expressions

molecule to the next is considered. Within this picture, microscopic charge and current densities can thus be written as

\[
\rho(x, t) = \rho_{\text{free}}(x, t) + \rho_{\text{bound}}(x, t),
\]

(2.5)

\[
j(x, t) = j_{\text{free}}(x, t) + j_{\text{bound}}(x, t),
\]

(2.6)

where the free contributions \(\rho_{\text{free}}(x, t)\) and \(j_{\text{free}}(x, t)\) are due to density and motion of the free charges, while the bound contributions \(\rho_{\text{bound}}(x, t)\) and \(j_{\text{bound}}(x, t)\) are due to density and motion of the bound molecular charges. The bound charge and current density can be further decomposed into contributions from individual molecules,

\[
\rho_{\text{bound}}(x, t) = \sum_n \rho_n(x, t),
\]

(2.7)

\[
j_{\text{bound}}(x, t) = \sum_n j_n(x, t),
\]

(2.8)

where \(\rho_n(x, t)\) and \(j_n(x, t)\) are non-vanishing only in a close vicinity of the \(n\)-th molecule. As no transfer of charge between molecules is considered, contributions associated with each molecule individually satisfy the charge-current continuity equation,

\[
\partial_t \rho_n(x, t) + \nabla \cdot j_n(x, t) = 0.
\]

(2.9)

A spatial averaging of the microscopic equations (2.1-2.4) can be performed in a usual way [2, 1]. For an arbitrary microscopic field \(F(x, t)\), a corresponding macroscopic field is defined through spatial averaging

\[
F_{\text{mac}}(x, t) = \int d^3 x' f(x') F(x - x', t),
\]

(2.10)

where \(F_{\text{mac}}(x, t)\) is the macroscopic field and \(f(x)\) is an averaging function. The exact form of the averaging function is irrelevant, but \(f(x)\) needs to be real, non-negative, smooth and normalized to unity,

\[
\int d^3 x' f(x') = 1.
\]

(2.11)

Furthermore, \(f(x)\) peaks at \(x = 0\) and is non-vanishing only over distances \(|x| < \Delta\) such that

\[
a \ll \Delta \ll \lambda,
\]

(2.12)

where \(\lambda\) is the wavelength of light and \(a\) identifies the microscopic length scale in a system; in a treatment of a molecular crystal \(a\) would be a lattice constant. From the definition (2.10) it follows immediately that the operations of space- and time-differentiation commute with the spatial averaging,

\[
\frac{\partial}{\partial t} F_{\text{mac}}(x, t) = \int d^3 x' f(x') \frac{\partial}{\partial t} F(x - x', t),
\]

(2.13)

\[
\frac{\partial}{\partial x_i} F_{\text{mac}}(x, t) = \int d^3 x' f(x') \frac{\partial}{\partial x_i} F(x - x', t).
\]

(2.14)

Applying the spatial averaging procedure (2.10) to the microscopic equations (2.1-2.4), and using
(2.13,2.14) it follows that

\[ \nabla \cdot \mathbf{B}(\mathbf{x}, t) = 0, \quad (2.15) \]

\[ \nabla \cdot \mathbf{E}(\mathbf{x}, t) = 4\pi \rho_{\text{mac}}(\mathbf{x}, t), \quad (2.16) \]

\[ \nabla \times \mathbf{E}(\mathbf{x}, t) + \frac{1}{c} \partial_t \mathbf{B}(\mathbf{x}, t) = 0, \quad (2.17) \]

\[ \nabla \times \mathbf{B}(\mathbf{x}, t) - \frac{1}{c} \partial_t \mathbf{E}(\mathbf{x}, t) = \frac{4\pi}{c} \mathbf{j}_{\text{mac}}(\mathbf{x}, t), \quad (2.18) \]

where \( \mathbf{E}(\mathbf{x}, t) \) and \( \mathbf{B}(\mathbf{x}, t) \) are macroscopic electromagnetic fields,

\[ \mathbf{E}(\mathbf{x}, t) = \int d^3 x' f(x') \mathbf{e}(\mathbf{x} - \mathbf{x}', t), \quad (2.19) \]

\[ \mathbf{B}(\mathbf{x}, t) = \int d^3 x' f(x') \mathbf{b}(\mathbf{x} - \mathbf{x}', t), \quad (2.20) \]

and \( \rho_{\text{mac}}(\mathbf{x}, t) \) and \( \mathbf{j}_{\text{mac}}(\mathbf{x}, t) \) are macroscopic charge and current densities,

\[ \rho_{\text{mac}}(\mathbf{x}, t) = \int d^3 x' f(x') \rho(\mathbf{x} - \mathbf{x}', t), \quad (2.21) \]

\[ \mathbf{j}_{\text{mac}}(\mathbf{x}, t) = \int d^3 x' f(x') \mathbf{j}(\mathbf{x} - \mathbf{x}', t). \quad (2.22) \]

Using now the decomposition (2.5,2.6) in (2.21,2.22), the macroscopic charges and currents can be written as

\[ \rho_{\text{mac}}(\mathbf{x}, t) = \rho_{\text{free}}(\mathbf{x}, t) + \rho_{\text{bound}}(\mathbf{x}, t), \quad (2.23) \]

\[ \mathbf{j}_{\text{mac}}(\mathbf{x}, t) = \mathbf{j}_{\text{free}}(\mathbf{x}, t) + \mathbf{j}_{\text{bound}}(\mathbf{x}, t), \quad (2.24) \]

where the macroscopic free and bound contributions are the spatial averages of the corresponding microscopic quantities. The macroscopic free charges and currents in the medium are often known. In a clean cold semiconductor or in a molecular crystal the macroscopic free charges and currents typically vanish. In a conductor, the macroscopic free current is simply the conduction current while the free charge will typically vanish unless the flow of current is such that there is a net build up of conduction electrons in parts of the crystal. Therefore, the focus of the remainder of this section is on the description of the bound charges and currents.

The macroscopic bound charge and current densities are defined as the spatial averages of the corresponding microscopic densities. However, the exact microscopic densities are usually not known. Therefore, it is more useful to work instead in terms of the molecular multipole moments, to which the macroscopic densities can be linked. First, I outline how to link the bound charge density to molecular multipole moments in a usual way. Then I outline how to do a similar calculation for the bound current density. From (2.7) the total macroscopic bound charge is found to be a sum of contributions associated with each molecule,

\[ \rho_{\text{bound}}(\mathbf{x}, t) = \sum_n \rho_{\text{mac}}^n(\mathbf{x}, t), \quad (2.25) \]

where

\[ \rho_{\text{mac}}^n(\mathbf{x}, t) = \int d^3 x' f(x') \rho_n(\mathbf{x} - \mathbf{x}', t). \quad (2.26) \]
I identify the center of the \( n \)-th molecule as \( r_n \), which could for example be chosen at the center of mass of the molecule. Changing the integration variables so that the distances as expressed as relative distances with respect to the center of the molecule, the charge density takes the form

\[
\rho_{\text{mac}}^n(x, t) = \int d^3 \zeta f(x - r_n - \zeta) \rho_n(r_n + \zeta, t). \tag{2.27}
\]

The charge density \( \rho_n(r_n + \zeta, t) \) is non-vanishing only over distances \( \zeta \) comparable with the size of the molecule. Over those distances the averaging function does not vary appreciably as follows from (2.12), and so it can be expanded in Taylor series in parameter \( \zeta \),

\[
f(x - r_n - \zeta) = f(x - r_n) - \zeta \cdot \nabla f(x - r_n) + \frac{1}{2} \zeta^\alpha \zeta^\beta \frac{\partial^2}{\partial x^\alpha \partial x^\beta} f(x - r_n) + \ldots, \tag{2.28}
\]

where \( \alpha, \beta \) are Cartesian coordinates of the position vector and implicit summation over repeated indices is used. Using now the expansion (2.28) in (2.27), the charge density is found in the following form,

\[
\rho_{\text{mac}}^n(x, t) = Q_n f(x - r_n) - \mu_n(t) \cdot \nabla f(x - r_n) + q_n^{\alpha \beta}(t) \frac{\partial^2 f(x - r_n)}{\partial x^\alpha \partial x^\beta} + \ldots, \tag{2.29}
\]

where I explicitly indicated only the first three terms in the expansion, but in general more could be included. Here \( Q_n, \mu_n(t) \) and \( q_n^{\alpha \beta}(t) \) are the first three electric moments in the multipole expansion: the total charge, the electric dipole moment, and the electric quadrupole moment. These are defined in terms of the molecular density by the usual expressions,

\[
Q_n = \int d^3 \zeta \rho_n(\zeta, t), \tag{2.30}
\]

\[
\mu_n(t) = \int d^3 \zeta (\zeta - r_n) \rho_n(\zeta, t), \tag{2.31}
\]

\[
q_n^{\alpha \beta}(t) = \frac{1}{2} \int d^3 \zeta (\zeta^\alpha - r_n^\alpha)(\zeta^\beta - r_n^\beta) \rho_n(\zeta, t). \tag{2.32}
\]

The total charge of the molecule is time-independent and in a usual situation it vanishes; thus in the remainder of this section I put \( Q_n = 0 \). The multipole moments (2.30-2.32) are the simple moments that are generated when the expansion is done in powers of the size of a molecule. In the long wavelength limit, they correspond to the leading order correction of the generalized moments that describe the expansion in terms of the polarity of emitted radiation [2].

An analogous analysis can be performed to link the bound current to molecular multipole moments. The total bound current density is a sum of contributions associated with each molecule,

\[
\mathbf{j}_{\text{bound}}(x, t) = \sum_n \mathbf{j}_{\text{mac}}^n(x, t), \tag{2.33}
\]

where

\[
\mathbf{j}_{\text{mac}}^n(x, t) = \int d^3 x f(x') \mathbf{j}_n(x - x', t). \tag{2.34}
\]

Changing the integration variables the current can be written as

\[
\mathbf{j}_{\text{mac}}^n(x, t) = \int d^3 \zeta f(x - r_n - \zeta) \mathbf{j}_n(r_n + \zeta, t), \tag{2.35}
\]
and then after using the expansion (2.28) it takes the following form

\[ j_n(x, t) = f(x - r_n) \int d^3 \zeta j_n(\zeta, t) - \frac{\partial f(x - r_n)}{\partial x^i} \int d^3 \zeta (\zeta^i - r_n^i) j_n(\zeta, t) + \ldots \]  

(2.36)

The expression (2.36) can be rewritten in terms of molecular multipole moments, using the identities given below. After some algebra one can prove that

\[ \int d^3 \zeta j_n^i(\zeta, t) = \partial_t \mu_n^i(t), \]  

(2.37)

where in order to arrive at the right-hand side of (2.37) the charge-continuity equation (2.9) was used. Furthermore,

\[ \frac{1}{2} \int d^3 \zeta (\zeta^i - r_n^i) j_n^k(\zeta, t) - \frac{1}{2} \int d^3 \zeta (\zeta^k - r_n^k) j_n^i(\zeta, t) = \epsilon^{ik} c \nu_n^i(t), \]  

(2.38)

\[ \frac{1}{2} \int d^3 \zeta (\zeta^k - r_n^k) j_n^i(\zeta, t) + \frac{1}{2} \int d^3 \zeta (\zeta^i - r_n^i) j_n^k(\zeta, t) = \partial_t q_n^{ik}(t), \]  

(2.39)

where \( \nu_n \) is the magnetic dipole moment of the \( n \)-th molecule,

\[ \nu_n(t) = \frac{1}{2c} \int (\zeta - r_n) \times j_n(\zeta, t). \]  

(2.40)

Using expressions (2.37-2.39), the current density is found to be

\[ j_n(x, t) = f(x - r_n) \partial_t \mu_n(t) - \nabla f(x - r_n) \cdot \partial_t \vec{q}_n(t) + c \nabla f(x - r_n) \times \nu_n(t) + \ldots \]  

(2.41)

Again, only the first three terms in the expansion are explicitly given but in general more terms could be included.

The results of the previous two paragraphs can be used to find the macroscopic bound charge and current densities in terms of polarization and magnetization fields. From (2.29, 2.41) it follows that

\[ \rho_{\text{bound}}^{\text{mac}}(x, t) = -\nabla \cdot \mathcal{P}(x, t), \]  

(2.42)

\[ \jmath_{\text{bound}}^{\text{mac}}(x, t) = \frac{\partial \mathcal{P}(x, t)}{\partial t} + c \nabla \times \mathcal{M}(x, t), \]  

(2.43)

where the polarization and magnetization fields are

\[ \mathcal{P}(x, t) = P(x, t) - \nabla \cdot \vec{Q}(x, t) + \ldots, \]  

(2.44)

\[ \mathcal{M}(x, t) = M(x, t) + \ldots, \]  

(2.45)

with the polarization, magnetization and quadrupolarization given by

\[ P(x, t) = \sum_n f(x - r_n) \mu_n(t), \]  

(2.46)

\[ \vec{Q}(x, t) = \sum_n f(x - r_n) \vec{q}_n(t), \]  

(2.47)

\[ M(x, t) = \sum_n f(x - r_n) \nu_n(t). \]  

(2.48)
Expressions (2.42-2.48) identify the bound charges and currents in a medium in terms of polarization and magnetization fields, which are linked to molecular multipole moments.

2.2 Power-Zienau-Wooley transformation

In this section I give an overview of the Power-Zienau-Wooley (PZW) transformation. I focus here on a formulation of the transformation introduced in terms of field operators, because this formulation connects most directly with the work on gauge-invariant descriptions of dynamics presented in chapters 5 and 6.

I begin with the minimal coupling Hamiltonian [58],

\[ H(t) = \frac{1}{2m} \int d^3x \psi^\dagger(x,t) \left( \frac{\hbar}{i} \nabla - \frac{e}{c} A(x,t) \right)^2 \psi(x,t) \]

\[ + \int d^3x e \phi(x,t) \psi^\dagger(x,t) \psi(x,t) + H_{\text{rest}}(t), \]  \hspace{1cm} (2.49)

where \(\psi(x,t)\) is the electron field operator, \(A(x,t)\) is the vector potential and \(\phi(x,t)\) is the scalar potential. The Hamiltonian \(H_{\text{rest}}(t)\) is included to describe interactions that will not be considered explicitly; the only assumption on \(H_{\text{rest}}\) is that

\[ [\rho(x,t), H_{\text{rest}}(t)] = 0. \]  \hspace{1cm} (2.50)

The minimal-coupling Hamiltonian describes the evolution of the field operator in the usual way,

\[ i\hbar \frac{\partial \psi(x,t)}{\partial t} = [\psi(x,t), H(t)]. \]  \hspace{1cm} (2.51)

In this thesis I am interested in describing the charges and currents in the presence of electromagnetic fields. In terms of field operators, these are given by

\[ \rho(x,t) = e \psi^\dagger(x,t) \psi(x,t), \]  \hspace{1cm} (2.52)

\[ j_i(x,t) = \frac{e\hbar}{2mi} \left( \psi^\dagger(x,t) \frac{\partial \psi(x,t)}{\partial x_i} - \frac{\partial \psi^\dagger(x,t)}{\partial x_i} \psi(x,t) \right) - \frac{e^2}{mc} A_i(x,t) \psi^\dagger(x,t) \psi(x,t). \]  \hspace{1cm} (2.53)

The PZW transformation of field operators is introduced by means of a Peierls phase [43, 45],

\[ \tilde{\psi}(x,t) = e^{-i\Phi_{\text{PZW}}(x,t)} \psi(x,t), \]  \hspace{1cm} (2.54)

with the Peierls phase defined as

\[ \Phi_{\text{PZW}}(x,t) = \frac{e}{\hbar c} \int_{u_{\text{initial}}}^{u_{\text{final}}} A(p(w;x,R),t) \cdot \frac{\partial p(w;x,R)}{\partial u} du, \]  \hspace{1cm} (2.55)

where \(R\) is the special point of the transformation that is usually chosen at the center of a molecule, and the vector \(p(w;x,R)\) identifies a path in space. The path goes from \(R\) to the field variable at \(x\) as the parameter \(u\) varies from its initial value at \(u_{\text{initial}}\) to a final value at \(u_{\text{final}}\); that is \(p(u_{\text{initial}};x,R) = R\) and \(p(u_{\text{final}};x,R) = x\). In general, different paths from \(R\) to \(x\) can be chosen giving rise to different PZW transformations. There is thus a degree of freedom in the PZW transformation that is related to
the choice of the path [45].

To simplify the notation, it is convenient to introduce within the PZW theory two functions that involve line integrals of the path derivatives taken from the special point of the transformation to the field variable [43, 45]. These functions are defined as

$$s_i(w, x; R) = \int_{u_{\text{initial}}}^{u_{\text{final}}} du \frac{\partial p_i(u; x, R)}{\partial u} \delta(w - p(u; x, R)), \quad (2.56)$$

$$\alpha_{ji}(w, x; R) = \epsilon_{jrs} \int_{u_{\text{initial}}}^{u_{\text{final}}} du \frac{\partial p_r(u; x, R)}{\partial u} \frac{\partial p_s(u; x, R)}{\partial x^i} \delta(w - p(u; x, R)). \quad (2.57)$$

The functions $s_i(w, x; R)$ and $\alpha_{ji}(w, x; R)$ are not independent. They satisfy the following two identities

$$\frac{\partial s_i(w, x; R)}{\partial w_i} = -\delta(w - x) + \delta(w - R), \quad (2.58)$$

$$\epsilon_{ijr} \frac{\partial \alpha_{jk}(w, x; R)}{\partial w_p} = \delta_{jk} \delta(w - x) - \frac{\partial s_i(w, x; R)}{\partial x_k}, \quad (2.59)$$

as can be confirmed by a direct calculation. The first line integral can be used to define the Peierls phase in a simplified notation,

$$\Phi_{\text{PZW}}(x, t) = \frac{e}{\hbar c} \int d^3w (w, x; R) \cdot A(w, t), \quad (2.60)$$

as follows directly from eqs. (2.55, 2.56). But more importantly, the line integrals (2.56,2.57) are used to define the microscopic polarization and magnetization fields of the molecule.

Before introducing polarization and magnetization fields, the charge and current densities (2.52,2.53) need to be found in terms of the transformed field operators (2.54). The charge density follows immediately from (2.54, 2.52), and is given by

$$\rho(x, t) = e \bar{\psi}^\dagger(x, t) \bar{\psi}(x, t). \quad (2.61)$$

The current density involves the spatial derivative of the Peierls phase. Using the representation (2.60) of the Peierls phase and the identity (2.59), the derivative is found to be

$$\frac{\partial \Phi_{\text{PZW}}(x, t)}{\partial x^i} = \frac{e}{\hbar c} A_i(x, t) - \frac{e}{\hbar c} \int d^3w B_j(w, t) \alpha_{ji}(w, x; R), \quad (2.62)$$

where to arrive at the right hand side the expression $B(x, t) = \nabla \times A(x, t)$ was used. Using (2.62,2.53) the current density can be found in terms of the transformed field operators. It takes the form

$$j(x, t) = j^P(x, t) + j^D(x, t), \quad (2.63)$$

where $j^P(x, t)$ and $j^D(x, t)$ are the para- and dia-magnetic currents, given by

$$j^P_i(x, t) = \frac{e \hbar}{2 m_i} \left( \bar{\psi}^\dagger(x, t) \frac{\partial \bar{\psi}(x, t)}{\partial x^i} - \frac{\partial \bar{\psi}^\dagger(x, t)}{\partial x^i} \bar{\psi}(x, t) \right), \quad (2.64)$$

$$j^D_i(x, t) = -\frac{e^2}{mc} \bar{\psi}^\dagger(x, t) \bar{\psi}(x, t) \int d^3w \alpha_{ji}(w, x; R) B_j(w, t). \quad (2.65)$$

Having identified the expressions for charges and currents in terms of the new operators, polarization
and magnetization fields can be introduced. Polarization and magnetization fields are related to the charge and current densities by the expressions [45],

\[ p(x, t) = \int d^3y \ s(x, y; R)\rho(y, t), \]  
\[ m_{ij}(x, t) = \frac{1}{c} \int d^3y \ a_{jk}(x, y; R)j_k(y, t), \]  

With the use of identities (2.58,2.59), an inverse transformation that links the microscopic charge and current densities of a molecule to polarization and magnetization fields can be found,

\[ \rho(x, t) = \delta(x - R)Q - \nabla \cdot p(x, t), \]  
\[ j(x, t) = \frac{\partial p(x, t)}{\partial t} + c \nabla \times m(x, t), \]

where

\[ Q = \int d^3y \rho(y, t) \]

is a total electron charge of the molecule. Expressions (2.68,2.69) confirm that \( p(x, t) \) and \( m(x, t) \) introduced via (2.66,2.67) do in fact function as polarization and magnetization fields [43, 45].

In addition to characterizing the charge and current distributions, polarization and magnetization fields also characterize the effective Hamiltonian that governs the evolution of the transformed field operators [43, 45]. The evolution of the field operator (2.54) is given by

\[ i\hbar \frac{\partial \tilde{\psi}(x, t)}{\partial t} = \left[ \tilde{\psi}(x, t), H(t) \right] + \left[ i\hbar \frac{\partial \tilde{\psi}(x, t)}{\partial t} \right]_{\text{explicit}} \]  

where the second term on the right hand side is present because the transformed field operator explicitly depends on time. Effective Hamiltonian, \( H_{PZW} \), is identified by demanding that [43]

\[ i\hbar \frac{\partial \tilde{\psi}(x, t)}{\partial t} = \left[ \tilde{\psi}(x, t), H_{PZW}(t) \right]. \]  

Comparing (2.71) and (2.72), the effective Hamiltonian is identified as

\[ H_{PZW}(t) = H(t) + \int d^3x \tilde{\psi}^\dagger(x, t)\tilde{\psi}(x, t)\hbar \frac{\partial \Phi_{PZW}(x, t)}{\partial t} \]  

The effective Hamiltonian (2.73) can be written in terms of polarization and magnetization fields. To do that, we first rewrite the second term on the right-hand side of (2.73). Using (2.61) and (2.66) it follows that

\[ \int d^3x \tilde{\psi}^\dagger(x, t)\tilde{\psi}(x, t)\hbar \frac{\partial \Phi_{PZW}(x, t)}{\partial t} = \frac{1}{c} \int d^3x p(x, t) \cdot \nabla \phi(x, t). \]  

As a second step, the second term in (2.49) can be rewritten. Using (2.61,2.68) and integrating by parts it follows that

\[ \int d^3x e\phi(x, t)\psi^\dagger(x, t)\psi(x, t) = Q\phi(R, t) + \int d^3x p(x, t) \cdot \nabla \phi(x, t). \]  

Finally, the first term in (2.49) can be rewritten in terms of magnetization fields. Using (2.62) it follows
that
\[
\frac{1}{2m} \int d^3x \psi^\dagger(x, t) \left( \frac{i}{\hbar} \nabla - \frac{e}{c} A(x, t) \right)^2 \psi(x, t) = \frac{1}{2m} \int d^3x \bar{\psi}^\dagger(x, t) \sum_i \left( \frac{i}{\hbar} \frac{\partial}{\partial x_i} - \frac{e}{c} \int d^3y \alpha_{ij}(y, x; R) B_j(y, t) \right)^2 \bar{\psi}(x, t). \tag{2.76}
\]

Expanding the square term, using (2.67), and adding \(\mathcal{H}_{\text{rest}}(t)\),
\[
\frac{1}{2m} \int d^3x \psi^\dagger(x, t) \left( \frac{i}{\hbar} \nabla - \frac{e}{c} A(x, t) \right)^2 \psi(x, t) + \mathcal{H}_{\text{rest}}(t) = \mathcal{H}_0 - \int d^3y m^P(x, t) \cdot B(x, t) \tag{2.77}
\]
\[\quad - \frac{1}{2} \int d^3y m^D(x, t) \cdot B(x, t), \tag{2.78}\]
where \(\mathcal{H}_0\) is the Hamiltonian in the absence of electromagnetic fields, and the dia- and para-magnetization are given by (2.67) but with the total current replaced by the dia- and para-magnetic contributions (2.64, 2.65) respectively. Upon collecting terms together, the PZW Hamiltonian is found in its usual form [43, 45]
\[
\mathcal{H}_{\text{PZW}}(t) = \mathcal{H}_0 + \mathcal{H}_{\text{int}}(t), \tag{2.79}
\]
where the interaction term \(\mathcal{H}_{\text{int}}(t)\) is given by
\[
\mathcal{H}_{\text{int}}(t) = -\int d^3x \ p(x, t) \cdot E(x, t) - \int d^3x \ m^{(P)}(x, t) \cdot B(x, t)
\]
\[\quad - \frac{1}{2} \int d^3x \ m^{(D)}(x, t) \cdot B(x, t) + Q\phi(R, t). \tag{2.80}\]
The last term in (2.80) is a constant of time and thus does not affect the dynamics of the system.

To get more physical insight into the results given by the PZW theory, the polarization and magnetization fields can be expanded around the special point of the transformation [43, 45]. The expansion is generated by formally expanding the line integrals (2.56, 2.57). For simplicity now I assume that the integration path chosen in the Peierls phase (2.55) is a straight-line. The functions (2.56, 2.57) in this case take the form [45]
\[
s(w, x; R) = x_R \int_0^1 du \delta (w_R - u x_R), \tag{2.81}
\]
\[
\alpha_{ij}(w, x; R) = x_R^{\nu} \epsilon_{imj} \int_0^1 u du \delta (w_R - u x_R), \tag{2.82}
\]
where I put \(w_R = w - R, x_R = x - R\). Formally expanding the Dirac delta function,
\[
\delta(x_R - u y_R) = \delta(x_R) - u y_R \frac{\partial}{\partial x_R} \delta(x_R) + \ldots \tag{2.83}
\]
and using the expansion in (2.81, 2.82), the polarization and magnetization fields (2.66, 2.67) are found
to be

\[ p_i(x, t) = \mu_i(t)\delta(x - R) - q_{ij}(t)\frac{\partial \delta(x - R)}{\partial x^j} + \ldots \]  (2.84)
\[ m_i(x, t) = v_i(t)\delta(x - R) + \ldots \]  (2.85)

The multipole moments in (2.84, 2.85) are given by expressions (2.31, 2.32, 2.40) but with classical charge
and current densities replaced by the charge and current density operators (2.61, 2.63). When the micro-
scopic polarization and magnetization fields are spatially averaged, the resulting macroscopic polarization
and magnetization are of the form (2.44, 2.45), but with classical multipole moments replaced by mul-
tiple moment operators. Thus polarization and magnetization fields describe the charge and current
distributions in terms of molecular multipole moments. The expansions (2.84, 2.85) also lead to the
interaction Hamiltonian in a physically intuitive form. Using (2.84, 2.85) in (2.80) gives

\[ H_{\text{int}}(t) = -\mu_i(t) \cdot E(R, t) - q_{ij}(t)\partial_j E_i(R, t) - v^P(t) \cdot B(R, t) - \frac{1}{2} v^D(t) \cdot B(R, t) + \ldots \]  (2.86)

where the magnetic moments \( v^P(t) \) and \( v^D(t) \) are associated with the para- and dia-magnetic current
densities respectively. The interaction term in the PZW Hamiltonian describes the interaction of the
electric dipole moment with the electric field at the position of the molecule, the interaction of the
quadrupole moment with the gradient of the electric field at the position of the molecule, and so on for
the remaining multipole moments.
Chapter 3

Optical theorem in magneto-electric quadrupolar arrays


Summary of results:

I present all the periodic Green function dyadics that enter a description of a 2D array of emitters at the level that includes the electric dipole, magnetic dipole and electric quadrupole moment of each emitter. I find a concise analytic form for the radiative contributions to the periodic Green function dyadics that give rise to radiation reaction fields, and so the description of the scattered light explicitly satisfies the optical theorem; I give the non-radiative contributions that do not affect energy balance in the form of rapidly converging series. Finally, I present an approximation scheme for evaluating periodic Green function dyadics at long wavelengths that rigorously respects energy conservation. The scheme extends the range of validity of the usual static approximation by the inclusion of a simple dynamic correction.

Full text:

Please go to the journal’s website to read the contents of Chapter 3, at: https://doi.org/10.1088/2040-8986/aa7f57.
Chapter 4

Surface - lattice resonances in 2d arrays of spheres


4.1 Introduction

A collection of metallic nanoparticles under optical illumination can exhibit a resonance structure very different from the plasmon resonance of an isolated emitter [11]. The difference arises due to radiative interactions between the nanoparticles. When nanoparticles are arranged on a lattice, collective interactions give rise to rapid variations of the radiation intensity at wavelengths corresponding to the onset of a new diffraction order [6]. The dip in the reflection spectrum observed when a new diffraction order becomes evanescent – the so called Rayleigh anomaly – is often accompanied by a narrow surface-lattice resonance (SLR) [17, 18, 19, 20] exhibiting an asymmetric Fano profile [11]. Surface-lattice resonances have been studied in a wide range of systems, including 1d chains [23, 75, 10] and 2d arrays [34, 12, 21, 37, 7] of plasmonic and dielectric nanoparticles, nanoparticle arrays on a substrate [76, 77], and arrays involving few nanoantennas per unit cell [78, 79, 80].

The analytic description of the surface-lattice resonances usually relies on approximating each plasmonic nanoparticle as an electric dipole [23, 34, 75, 17, 12, 21, 76] or a few dipoles per unit cell in the case of a composite nanoparticle [78, 79]; in a description of dielectric nanoparticles the resonant magnetic dipole is also included [10, 37]. Although for many systems the dipole approximation is sufficient, higher multipole moments contribute to the radiation from nanoparticles of moderate sizes [81, 7, 9, 82, 40, 83, 84, 85]. In arrays of nanoparticles these higher order contributions can lead to SLRs with very high quality factors, as was observed in experiments on arrays of plasmonic nanorods [20] and predicted theoretically for an array of spheres with large quadrupolarizabilities [7, 83].

The electrodynamics responsible for the multipolar SLRs is still not well characterized, even in systems as simple as an array of spheres. Radiation from an array of spheres can be found essentially exactly using numerical methods, such as the layer-multipole scattering method [86, 87], but these do not give direct physical insight into the nature of the couplings between the nanoparticles. Such an insight can be gained when the optical response of an array is treated within analytic models that
identify the multipolar structure of resonance spectrum. However, previous analytical treatments \cite{7} of SLRs associated with higher order multipoles have been limited in their scope, considering only normally incident light, and only those multipole moments of each sphere that contribute significantly to radiation from an isolated sphere. Yet since the radiation from an array of spheres is different from that of an isolated sphere, with the different multipoles now coupled by radiation interactions that can depend strongly on illumination conditions, multipoles that have a negligible effect on the scattering of an isolated sphere can become important in the response of an array. Thus a more in-depth analysis of SLRs is needed.

In this chapter I use the multipolar model of chapter 3 to analyse surface-lattice resonances in 2d arrays of gold spheres. I take into account the electric dipole, magnetic dipole, and electric quadrupole moments, both resonant and non-resonant. I analyze the multipolar structure of the radiation spectrum over a broad range of illumination conditions. I show that strong interactions between multipoles in the vicinity of Rayleigh anomalies result in a significant contribution from moments that would be negligible if spheres were non-interacting. These non-resonant multipoles are especially important in a description of SLRs associated with higher multipole moments, for which the contribution from a non-resonant magnetic dipole – and, depending on the angle of incidence, other multipoles as well – strongly affects the position and profile of the SLR, as well as its very existence.

Surface-lattice resonances arise due to diffractive coupling of incident light to modes supported by the array \cite{88, 89, 90, 91, 92}; I refer to the modes for which this diffractive coupling occurs as surface-lattice modes (SLMs). However, the coupling mechanism and a link between properties of SLRs and those of SLMs has not been discussed in detail. This is because the analytic descriptions of SLRs have often only considered light at normal incidence \cite{75, 37, 7, 21}, and so even though the dispersion relations of modes supported by arrays of spheres has been studied extensively \cite{40, 38, 35, 39, 93, 65}, they have not been discussed in connection with SLRs. Here I discuss the connection between the modes and the resonance spectrum in detail. I identify SLMs of dipolar and multipolar character. I discuss the diffractive coupling to those modes that is signalled by SLRs, as well as a direct coupling that is also possible, with a careful consideration of energy balance in the system. SLMs are lossy due both to absorption in the spheres, and to the very radiative coupling that allows access to them by incident light. Using the representation of periodic Green function dyadics introduced in chapter 3, which allows me to identify exactly the part of the coupling that leads to the radiative loss, I study what might be called “ideal” SLMs by neglecting both the absorption and the radiative loss. The resulting ideal dispersion relations give direct insight into under what illumination conditions the SLRs appear, in the same way that the “ideal” dispersion of a surface plasmon at an air/metal interface, even though calculated with the neglect of absorption loss, can give insight into the response of the system under irradiation from a cladding prism separated from the interface by an air gap \cite{70, 71}. Having linked SLRs to ideal SLMs, I introduce a simplified model of the ideal SLM dispersion relations that explains their novel features, and I then use it to explain the sensitivity of SLRs to coupling and illumination conditions as well as the frequency cut-off below which SLRs are not found; these properties of SLRs are not easily understood without an appreciation of the SLM dispersion relations.

The chapter is organized as follows. In section 4.2 I give the multipolar model of chapter 3 for an array of spheres. In section 4.3 I present resonance spectrum for illumination along lattice symmetry direction. In section 4.4 I find the dispersion relations of the SLMs. In section 4.5 I propose a simplified model of the dispersion relations and I link it to the properties of SLRs in section 4.6. In section 4.7
I analyse the sensitivity of SLRs to the direction and polarization of light. In section 4.8 I discuss the direct coupling to SLMs. I conclude in section 4.9.

4.2 Multipolar model

4.2.1 Isolated sphere

As an example of typical nanoparticles I consider gold spheres of radius \( r = 100 \) nm. Adopting the optical constants of Johnson and Christy [94], and assuming the nanoparticles are embedded in a medium with index of refraction \( n = 1.45 \), I find from Mie theory [95] that the extinction of an isolated sphere illuminated by a plane wave is dominated by scattering at wavelengths above 600 nm (see Fig. 4.1). The main contribution to extinction comes from the electric dipole and the electric quadrupole moment, while the magnetic dipole moment gives only a small non-resonant contribution; see Fig. 4.1. Nevertheless, the magnetic dipole moment formally enters the multipole expansion at the same level as electric quadrupole [96], and it is important in describing the collective effects in radiation from a lattice of spheres, as I shall show. I thus develop a description using the full multipolar model of chapter 3, which includes the electric dipole, magnetic dipole, and electric quadrupole moments of the spheres.

While Mie theory is usually derived in the context of a plane-wave illumination [97, 75, 40, 37, 7], even for an arbitrary incident field the Mie polarizabilities relate the multipole moments to electromagnetic fields and their gradients at the center of the sphere [98]. I consider a response of a sphere to electromagnetic fields of the form,

\[
E^{\text{inc}}(r, t) = E^{\text{inc}}(r)e^{-i\omega t} + c.c. \tag{4.1}
\]

\[
B^{\text{inc}}(r, t) = B^{\text{inc}}(r)e^{-i\omega t} + c.c. \tag{4.2}
\]

I identify the electric dipole moment \( \mathbf{p} \), magnetic dipole moment \( \mathbf{m} \), and electric quadrupole moment \( \mathbf{q} \) as Cartesian versions of the full multipole coefficients, defined in terms of integrals of the charge and current distribution taken with the spherical Bessel functions; these moments describe all the radiation

Figure 4.1: Extinction cross section of an isolated sphere and contributions from multipoles: ED - electric dipole, MD - magnetic dipole, EQ - electric quadrupole, OC - electric octupole. Dotted line shows the total absorption cross section.
pattern with the polarity up to the electric quadrupole level \[2\]. Writing \(E^{\text{inc}} = E^{\text{inc}(0)}\), etc., I have for an isolated sphere at the origin of coordinates,

\[
\begin{align*}
p &= \tilde{\alpha}^{pE} E^{\text{inc}}, \\
m &= \tilde{\alpha}^{mB} B^{\text{inc}}, \\
\vec{q} &= \tilde{\alpha}^{qF} \vec{F}^{\text{inc}},
\end{align*}
\]

where the dyadic \(F_{ij} = \frac{1}{2} \left( \frac{\partial E_j}{\partial x_i} + \frac{\partial E_i}{\partial x_j} \right)\) is the symmetrized gradient of the electric field, and the Mie polarizabilities are given by \[97, 40, 7\]

\[
\begin{align*}
\tilde{\alpha}^{pE} &= -i \frac{(\tilde{\omega} n)^3}{6\pi\epsilon_0 n^2} B^e_1, \\
\tilde{\alpha}^{mB} &= -i \frac{(\tilde{\omega} n)^3}{6\pi\epsilon_0 c^2} B^m_1, \\
\tilde{\alpha}^{qF} &= - \frac{1}{(\tilde{\omega} n)^5} 24\pi\epsilon_0 n^2 B^e_2,
\end{align*}
\]

where \(\tilde{\omega} = \omega/c\), and \(B^e_1, B^m_1\) are the amplitudes of the associated partial waves \[95\]. The Mie polarizabilities in (4.6-4.8) involve contributions that give rise to the radiative damping. I choose to work with the proper polarizabilities instead, the inverses of which are linked to the usual Mie polarizabilities by the relations,

\[
\begin{align*}
(\alpha^{pE})^{-1} &= (\tilde{\alpha}^{pE})^{-1} + i \frac{(\tilde{\omega} n)^3}{6\pi\epsilon_0 n^2}, \\
(\alpha^{mB})^{-1} &= (\tilde{\alpha}^{mB})^{-1} + i \frac{(\tilde{\omega} n)^3}{6\pi\epsilon_0 c^2}, \\
(\alpha^{qF})^{-1} &= (\tilde{\alpha}^{qF})^{-1} + i \frac{(\tilde{\omega} n)^5}{20\pi\epsilon_0 n^2}.
\end{align*}
\]

Identifying the sum of each incident field and the associated radiation reaction field

\[
\begin{align*}
E' &= E^{\text{inc}} + i \frac{(\tilde{\omega} n)^3}{6\pi\epsilon_0 n^2} p, \\
B' &= B^{\text{inc}} + i \frac{(\tilde{\omega} n)^3}{6\pi\epsilon_0 c^2} m, \\
\vec{F}' &= \vec{F}^{\text{inc}} + i \frac{(\tilde{\omega} n)^5}{20\pi\epsilon_0 n^2} \vec{q},
\end{align*}
\]

the response of an isolated sphere to fields at frequency \(\omega\) is given in terms of the proper polarizabilities by

\[
\begin{align*}
p &= \alpha^{pE} E', \\
m &= \alpha^{mB} B', \\
\vec{q} &= \alpha^{qF} \vec{F}'.
\end{align*}
\]

The expressions (4.15-4.17) are simpler than the corresponding expressions of chapter 3 given for an emitter of an arbitrary shape (see eq. (3)-(5) in \[41\]), because for spheres the polarizabilities of are now
scalars and some terms vanish for reasons of symmetry.

### 4.2.2 2D triangular lattice

I now consider a triangular array of these spheres as sketched in Fig. 4.2, with the basis lattice vectors of length $|a_1| = |a_2| = 475 \text{nm}$. The centers of the spheres are taken to lie in the $z = 0$ plane, and their positions are indicated by a lattice vector $R_n = n_1 a_1 + n_2 a_2$ where $n_1$ and $n_2$ are integers. The spheres are illuminated by a plane electromagnetic wave with wave vector $v_+ = \kappa_0 + w_0 \hat{z}$, $v_+ \cdot v_+ = (\tilde{\omega} n)^2$,

$$E^{\text{inc}}(r) = E^{\text{inc}} e^{i\kappa_0 \cdot R} e^{i w_0 z},$$

$$B^{\text{inc}}(r) = B^{\text{inc}} e^{i\kappa_0 \cdot R} e^{i w_0 z},$$

where $\kappa_0$ is the in-plane component of the incident wave vector, $|\kappa_0| = \tilde{\omega} \sin \theta_0$, $\theta_0$ is the angle of incidence, and $w_0 = \sqrt{(\tilde{\omega} n)^2 - \kappa_0^2} > 0$.

For this excitation, the induced multipole moments at the lattice site $R_n$, $(p_n, m_n, \vec{q}_n)$, are related to those at the origin, $(p, m, \vec{q})$, by

$$p_n = e^{i \kappa_0 \cdot R_n} p,$$

$$m_n = e^{i \kappa_0 \cdot R_n} m,$$

$$\vec{q}_n = e^{i \kappa_0 \cdot R_n} \vec{q},$$

and so the scattering problem can be formulated in terms of $p, m, \vec{q}$ only. The response of the array formulated in terms of the multipoles at the origin is described by the eqs. of chapter 3 (see eqs. (17)-(19) in [41]) with the proper polarizabilities being given by the eqs. (4.9-4.11). We have

$$p = \alpha^p E^{\text{tot}},$$

$$m = \alpha^m B^{\text{tot}},$$

$$\vec{q} = \alpha^{\vec{q}} \vec{F}^{\text{tot}},$$

where $E^{\text{tot}}, B^{\text{tot}}$ and $\vec{F}^{\text{tot}}$ are sums of the incident fields, the radiation reaction fields of the sphere at
the origin, and the fields radiated from all the other multipoles in an array,

\[
\begin{align*}
E^{\text{tot}} &= E^{\text{inc}} + G^{E_p} \cdot p + G^{E_m} \cdot m + G^{E_q} : \vec{q}, \\
B^{\text{tot}} &= B^{\text{inc}} + G^{B_p} \cdot p + G^{B_m} \cdot m + G^{B_q} : \vec{q}, \\
\dot{F}^{\text{tot}} &= \dot{F}^{\text{inc}} + G^{F_p} \cdot p + G^{F_m} \cdot m + G^{F_q} : \vec{q},
\end{align*}
\]

Periodic Green functions that enter the expressions for the total fields (4.26-4.28) are defined as in chapter 3 (see eqs. (23,24) in [41]). For example,

\[
G^{E_p} \cdot p = \lim_{z \to 0} \sum_{n \neq (0,0)} e^{i \kappa_0 R_n \vec{g}^{E_p}(-R_n + z\hat{z})} \cdot p
\]

\[
+ \frac{i(\tilde{\omega} n)^3}{6 \pi \varepsilon_0 n^2} p,
\]

specifies the electric field at the origin from the electric dipoles at all the other lattice sites together with the radiation reaction field from the dipole at the origin; the dyadic \(G^{E_p}(r)\) in the definition (4.29) is the usual free-space dyadic, where \(G^{E_p}(r) \cdot p\) gives the electric field at \(r\) due to an electric dipole \(p\) located at the origin. The other terms, \(G^{B_m} \cdot m\) and \(G^{F_q} : \vec{q}\) respectively, give the magnetic field at the origin from the magnetic dipoles at all the other lattice sites together with the radiation reaction field from magnetic dipole at the origin, and the symmetrized field gradient at the origin due to the electric quadrupoles at all the other lattice sites together with the radiation reaction from quadrupole at the origin. All the remaining periodic Green functions are defined as sums of fields at origin from corresponding multipoles at other lattice sites, but do not involve radiation reaction terms; all the periodic Green functions are implicit functions of \(\kappa_0\) and \(\omega\).

Periodic Green functions are evaluated using the representation derived in chapter 3. The periodic Green functions are implicit functions of \(\kappa_0\) and \(\omega\), and they involve Fourier contributions associated with each of the reciprocal lattice vectors \(K\), with non-analyticities as a function of \(\kappa_0\) associated with the appearance of each new diffracted order; these occur when \(\tilde{\omega} n = |\kappa_0 + K|\). Singular behavior in the dyadics as functions of \(\kappa_0\) arises because of their dependence, for all \(K\), on \(|\tilde{\omega}^2 n^2 - (\kappa_0 + K)^2|^{-1/2}\). Terms of this nature arise from the confinement of the radiating multipole moments to a plane, and they play a dominant role in the formation of SLRs and the modes associated with them [75, 78].

Equations (4.23-4.25) take a particularly simple form for an incident plane wave with wave vector component in the plane of the lattice, \(\kappa_0\), lying along one of the two types of symmetry directions of the lattice. These two excitation scenarios correspond to \(\kappa_0\) in the direction of \(\hat{r}_a\) or in the direction of \(\hat{r}_b\); see Fig. 4.2. For the first direction the wave vector \(\kappa_0\) is aligned with one of the lattice vectors, while in the second it is aligned with one of the reciprocal lattice vectors. For \(\kappa_0\) along either of these directions the self-consistent equations (4.23-4.25) decouple into two independent sets of equations. I consider the specific form that these equations take for s- and p-polarized incident light. I define the s- and p-polarization amplitudes in the usual way,

\[
E^{\text{inc}} = E^{\text{inc}}_s \hat{s} + E^{\text{inc}}_p \hat{p}_+,
\]

where for a plane wave with an in-plane wave vector component \(\kappa\) I define the polarization vectors in
general as

\[ \hat{s} = \hat{\kappa} \times \hat{z}, \]  
(4.31)

\[ \hat{p}_\pm = (\tilde{\omega} n)^{-1} [\kappa \hat{z} \mp w \hat{\kappa}], \]  
(4.32)

with \( w = \sqrt{\tilde{\omega} n^2 - \kappa^2} \), \( \text{Im}(w) > 0 \), but for use in (4.30) and the following equations in this paragraph I take \( \kappa = \kappa_0 \). For s-polarized incident light and \( \kappa_0 \) along either the direction \( \hat{\kappa}_a \) or \( \hat{\kappa}_b \), the equations (4.23-4.25) reduce to the form

\[
\begin{pmatrix}
  p_s \\
  m_z \\
  q_{\kappa s}
\end{pmatrix}
= 
\begin{pmatrix}
  E_{\text{inc}}^{s} \\
  B_{\text{inc}}^{s} \\
  F_{\text{inc}}^{s}
\end{pmatrix},
\]  
(4.33)

and

\[
\begin{pmatrix}
  m_{\kappa} \\
  q_{sz}
\end{pmatrix}
= 
\begin{pmatrix}
  B_{\kappa}^{s} \\
  F_{\text{inc}}^{s}
\end{pmatrix},
\]  
(4.34)

with the matrices \( S \) and \( S' \) given by

\[
S = 
\begin{pmatrix}
  (\alpha^{BE})^{-1} & 0 & 0 \\
  0 & (\alpha^{mB})^{-1} & 0 \\
  0 & 0 & (\alpha^{qF})^{-1}
\end{pmatrix}
- 
\begin{pmatrix}
  G_{q_{sz}}^{Ep} & G_{q_{sz}}^{Em} & 2G_{q_{sz}}^{Eq} \\
  G_{q_{szz}}^{BP} & G_{q_{szz}}^{Bm} & 2G_{q_{szz}}^{Bq} \\
  G_{q_{szz}}^{FP} & G_{q_{szz}}^{Fm} & 2G_{q_{szz}}^{Fq}
\end{pmatrix},
\]  
(4.35)

\[
S' = 
\begin{pmatrix}
  (\alpha^{mB})^{-1} & 0 & 0 \\
  0 & (\alpha^{qF})^{-1}
\end{pmatrix}
- 
\begin{pmatrix}
  G_{q_{sz}}^{BM} & 2G_{q_{sz}}^{Bq} \\
  G_{q_{szz}}^{FM} & 2G_{q_{szz}}^{Fq}
\end{pmatrix},
\]  
(4.36)

and where the matrix elements take on different values depending on \( \kappa, \omega \) and on whether \( \hat{\kappa}_0 = \hat{\kappa}_a \) or \( \hat{\kappa}_0 = \hat{\kappa}_b \). Similarly for p-polarized light the equations reduce to

\[
F = 
\begin{pmatrix}
  p_{\kappa} \\
  m_z \\
  q_{\kappa z}
\end{pmatrix}
= 
\begin{pmatrix}
  E_{\text{inc}}^{\kappa} \\
  B_{\text{inc}}^{\kappa} \\
  F_{\text{inc}}^{\kappa}
\end{pmatrix},
\]  
(4.37)

and

\[
F' = 
\begin{pmatrix}
  p_z \\
  m_s \\
  q_{sz}
\end{pmatrix}
= 
\begin{pmatrix}
  E_{\text{inc}}^{\kappa} \\
  B_{\text{inc}}^{\kappa} \\
  F_{\text{inc}}^{\kappa}
\end{pmatrix},
\]  
(4.38)
with the matrices $\mathbb{P}$ and $\mathbb{P}'$ given by

$$
\mathbb{P} = \begin{pmatrix} (\alpha^p)^{-1} & 0 & 0 \\ 0 & (\alpha^q)^{-1} & 0 \\ 0 & 0 & (\alpha^q)^{-1} \end{pmatrix} - \begin{pmatrix} \mathcal{G}^{Ep}_{\kappa\kappa} & \mathcal{G}^{Eq}_{\kappa\kappa} - \mathcal{G}^{Eq}_{\kappa z} & \mathcal{G}^{Eq}_{\kappa z} - \mathcal{G}^{Eq}_{\kappa z} \\ \mathcal{G}^{Fp}_{\kappa\kappa} & \mathcal{G}^{Fq}_{\kappa\kappa} - \mathcal{G}^{Fq}_{\kappa z} & \mathcal{G}^{Fq}_{\kappa z} - \mathcal{G}^{Fq}_{\kappa z} \\ \mathcal{G}^{Fp}_{\kappa\kappa} & \mathcal{G}^{Fq}_{\kappa\kappa} - \mathcal{G}^{Fq}_{\kappa z} & \mathcal{G}^{Fq}_{\kappa z} - \mathcal{G}^{Fq}_{\kappa z} \end{pmatrix},
$$

(4.39)

and $\mathbb{P}'$ by

$$
\mathbb{P}' = \begin{pmatrix} (\alpha^p)^{-1} & 0 & 0 \\ 0 & (\alpha^m)^{-1} & 0 \\ 0 & 0 & (\alpha^q)^{-1} \end{pmatrix} - \begin{pmatrix} \mathcal{G}^{Ep}_{\kappa z} & \mathcal{G}^{Em}_{\kappa z} & 2\mathcal{G}^{Eq}_{\kappa z} \\ \mathcal{G}^{Bp}_{\kappa z} & \mathcal{G}^{Bm}_{\kappa z} & 2\mathcal{G}^{Bq}_{\kappa z} \\ \mathcal{G}^{Fp}_{\kappa z} & \mathcal{G}^{Em}_{\kappa z} & 2\mathcal{G}^{Fq}_{\kappa z} \end{pmatrix},
$$

(4.40)

and where again their elements take on different values depending on whether $\mathbf{k}_0 = \mathbf{k}_a$ or $\mathbf{k}_0 = \mathbf{k}_b$.

Once the multipole moments $p, m, \mathbf{q}$ are found, I calculate the electromagnetic fields radiated by an array using the Green function formalism for s- and p-polarized light [72, 41]. The Fourier transform in the plane of the array (see eq. (108) in [41]) of the scattered field is identified by,

$$
E^\text{inc}(\kappa, z) = \frac{(2\pi)^2}{A_c} \sum_n \delta(\kappa - \kappa_n) f^E(\kappa, z),
$$

(4.41)

where $A_c$ is the area of a unit cell, and $n$ here and henceforth indicates a vector of integers indicating directions in reciprocal space; earlier, at the start of this section, I had used it to indicate directions in real space. I indicate the reciprocal lattice vector identified by $n$ as $\mathbf{K}_n$, and write $\kappa_n = \kappa_0 + \mathbf{K}_n$ for translation of the component of the incident wave vector in the lattice plane, $\kappa_0$, by the reciprocal lattice vector $\mathbf{K}_n$; in chapter 3 I used a convention for $\kappa_n$ that corresponds to choosing an opposite reciprocal vector, but here I find it more convenient to choose $\kappa_n$ as indicated. The vector $f^E(\kappa, z)$ is defined as

$$
f^E(\kappa, z) = \mathbf{g}^\text{Ep}(\kappa, z) \cdot \mathbf{p} + \mathbf{g}^\text{Em}(\kappa, z) \cdot \mathbf{m} + \mathbf{g}^\text{Eq}(\kappa, z) : \mathbf{q},
$$

(4.42)

where $\mathbf{g}^\text{Ep}(\kappa, z)$, $\mathbf{g}^\text{Em}(\kappa, z)$ and $\mathbf{g}^\text{Eq}(\kappa, z)$ are the Fourier transforms in the plane of the array of the usual free-space Green functions dyadics $\mathbf{g}(r)$, $\mathbf{g}(r)$, and $\mathbf{g}(r)$; see appendix A. The expressions for the reflection and transmission of an array immediately follow from the expressions for the incident electric field (4.18) and the field scattered by an array (4.41),

$$
R = \sum_{n \neq \kappa_0} R_n,
$$

(4.43)

$$
T = \sum_{n \neq \kappa_0} T_n,
$$

(4.44)

where the functions

$$
R_n = \left| \frac{1}{E^\text{inc}^2} \frac{w_n}{w_0} \frac{1}{A_c} f^E(\kappa_n, 0) \right|^2,
$$

(4.45)

$$
T_n = \left| \frac{1}{E^\text{inc}^2} \frac{w_n}{w_0} \frac{1}{A_c} f^E(\kappa_n, 0) \right|^2 + \left[ 1 + \frac{2}{A_c} \frac{\text{Re}(E^\text{inc})^\ast \cdot f^E(\kappa_0, 0)}{|E^\text{inc}|^2} \right] \delta_{n,0},
$$

(4.46)
describe the contribution to reflection and transmission from a beam diffracted at an angle \( \theta_n = \sin^{-1}\left(\frac{|\kappa_n|}{\bar{\omega}n}\right) \) with respect to the normal, the specular reflection and transmission are the components \( R_{\text{spec}} = R_0 \) and \( T_{\text{spec}} = T_0 \) respectively, and I identified \( w_n = \sqrt{(\bar{\omega}n)^2 - \kappa_n^2} \), as the z-component of a wave vector with the in-plane component \( \kappa_n \).

### 4.3 Resonance spectrum and the multipolar couplings

The radiation spectrum of an array of emitters depends on the properties of an isolated emitter as well as on the collective radiative interaction between emitters on a lattice. The collective interactions are especially pronounced at wavelengths close to the onset of diffraction, in the vicinity of which narrow and asymmetric surface-lattice resonances (SLRs) are observed [17, 18, 19, 20]. When the optical response of the system is dominated by one multipole moment [75, 23, 21], or the multipole moments can be treated as independent [37, 7, 10], the analysis of SLRs can be greatly simplified by introducing an effective polarizability of an array that incorporates both the single particle and the collective interactions. When the coupling between the multipole moments cannot be neglected, the analysis becomes more complicated.

Here I analyze the SLRs with the multipolar model (4.23-4.25), including electric dipole, magnetic dipole, and electric quadrupole moments. For different excitation geometries I consider angle scans at fixed frequencies rather than considering the more usual frequency dependence of the resonant spectrum [75, 23, 34, 37, 7]. This allows me to distinguish more clearly between the single emitter and the collective characteristics of the resonance spectrum; the polarizability of an isolated emitter for a chosen illumination scenario depends only on the wavelength of light, and thus the resonances observed when varying the incidence angle are driven by collective effects that are described by variations in the periodic Green functions.

For an arbitrary direction of \( \kappa_0 \) the behavior of the specularly reflected and transmitted light, and that of the diffracted light, can be very complicated. To focus on the underlying physics, in this section and the following three I consider s-polarized light incident along the first symmetry direction of the lattice, \( \kappa_0 \propto \hat{\kappa}_a \). For light incident along this direction multipole moments are given by the decoupled set of equations (4.33-4.34). This simplifies the description of the system, but the scenario is still rich enough to exhibit SLRs associated with both the dipole moments and the higher order moments of the spheres. While a decoupled set of equations also holds for light incident along the other symmetry direction of the lattice, \( \kappa_0 \propto \hat{\kappa}_b \), in that direction the response is more one-dimensional in nature, as I see in detail in section 4.8; in the direction \( \kappa_0 \propto \hat{\kappa}_a \) the two-dimensional nature of the lattice plays a central role in the behavior of the SLRs and the dispersion relations of modes associated with them, as I see in this section and the next. The extension to more general directions of the incident light, and to p-polarized light, is discussed in section 4.7.

First I discuss the range of wavelengths that I will consider. I identify a “light line” as the condition \( \bar{\omega} = \kappa_0/n \), characterizing light propagating parallel to the plane of the array. This is plotted in Fig. 4.3, where the part of the plane to the left of the light line is accessible to incident light. For \( \kappa_0 \) in the direction of \( \hat{\kappa}_a \), the first of the diffracted orders to appear as the angle of incidence is increased are (10) and (11) (see Fig. 4.2), occurring at angles \( \theta_{(10)} \) and \( \theta_{(11)} \) that satisfy the Rayleigh condition with
\[ \tilde{\omega} \approx 2\pi/\lambda \]

\[ K = K_{(10)} \quad \text{and} \quad K = K_{(11)} \]

I note that due to symmetry the angles satisfy \( \theta_{(10)} = \theta_{(11)} \equiv \theta_R \). The incident angle at which this occurs depends on the wavelength of incident light, and is identified in Fig. 4.3 by the “Rayleigh line”, defined as the set of points \((\kappa_0, \tilde{\omega})\) such that \( \kappa_0 = \tilde{\omega}\sin \theta_R \). For fixed wavelength, as the angle of incidence is increased and the Rayleigh line is crossed, the periodic Green functions exhibit singular behaviour. At the Rayleigh line the specular reflection typically, but not always, vanishes. In the standard theory of diffraction from gratings, that kind of behaviour at the onset of diffraction is referred to as a “Rayleigh anomaly” \([99, 100]\), and I adopt that notation here. For large enough wavelengths, where the Rayleigh line is to the right of the light line, no diffraction is possible; the cut-off for the system considered here is at the wavelength \( \lambda_{gz} \approx 1033 \text{nm} \) \((\tilde{\omega}_{gz} \equiv 2\pi/\lambda_{gz})\), where the diffraction would only arise at grazing incidence, with \( \theta_R = 90^0 \). At \( \lambda_{nr} \approx 596 \text{nm} \) \((\tilde{\omega}_{nr} \equiv 2\pi/\lambda_{nr})\) diffraction occurs at normal incidence, \( \kappa_0 = 0 \) and \( \theta_R = 0 \). I consider the range of wavelengths between \( \lambda_{nr} \) and \( \lambda_{gz} \); here and throughout this chapter the wavelengths I refer to are vacuum wavelengths.

Although the intensity of the diffracted light shows no clear resonances in this range, the specularly reflected light exhibits one or two SLRs at angles of incidence less than the angle of the Rayleigh line, \( \theta < \theta_R \); see Fig. 4.4. To identify how the two independent sets of multipole moments \((4.33, 4.34)\) contribute to the SLRs, I compare the contributions they make to the specular reflectivity. These are shown in Fig. 4.5 for a narrow angular scan at \( \lambda = 750\text{nm} \), and I see that each of the resonances can be associated with one of two sets of moments: a broad peak is associated with the resonance for the multipole moments \((p_s, m_z, q_{sz})\) given by eq. \((4.33)\), while the very narrow peak is associated with the higher order multipoles \((m_{\kappa}, q_{sz})\) given by eq. \((4.34)\). I shall show that, as expected, the electric dipole makes the largest contribution to the first, and so it can be called the dipolar SLR; the second I call the multipolar SLR. The resonant peaks at other wavelengths can also be clearly identified with the sets of multipoles: the peak at 650 nm, for example, is a multipolar SLR while the broad peak at 900 nm is a
Before considering the physics of these resonances, I confirm that the multipolar model (4.23-4.25) describes the optical response of the array correctly. I have compared the multipolar model results with an essentially exact numerical treatment, obtained using the layer-multipole scattering method [87], for wavelengths in the range \( \lambda_{nr} \) to \( \lambda_{gz} \) and with scans over all incident angles. I find that our multipolar model captures all the resonances that appear in the exact numerical calculation. To show the degree of agreement in describing each of the resonances, in Fig. 4.6 I compare the multipolar model result at \( \lambda = 750 \text{ nm} \) with an essentially exact numerical treatment obtained using the layer-multipole scattering method [87]. The multipolar model gives an excellent description of the dipolar SLRs associated with the set of equations (4.33), and a good description of the multipolar SLRs associated with the set of equations (4.34). At large wavelengths \( \lambda \approx 805 \text{ nm} \) within the multipolar model and \( \lambda \approx 785 \text{ nm} \) in the exact calculation – the multipolar SLR merges with the Rayleigh anomaly and I observe a kink in reflectivity at \( \theta = \theta_R \) rather than a SLR. The model thus correctly predicts the disappearance of the dipolar SLR.
Chapter 4. Surface - lattice resonances in 2d arrays of spheres

Figure 4.6: Comparison of the total reflectivity predicted by the multipolar model, which includes multipole contributions from the electric and magnetic dipoles and the electric quadrupole, with the exact result obtained from the layered-multipole scattering method.

SLR, but there are some discrepancies between the exact calculation and the multipolar model in the description of set (4.34) close to the Rayleigh anomaly at long wavelengths. Nevertheless, even close to the anomaly the model correctly identifies the leading correction to each SLR beyond the simplest description with each SLR approximated by a response from one dominant multipole only. This point I discuss in detail below.

With the validity of truncating the multipole expansion at the level of the magnetic dipole and electric quadrupole confirmed for \( \lambda \lesssim 800 \text{ nm} \), I now investigate whether the description can be further simplified by approximating each set of multipole moments (4.33,4.34) by one dominant moment. Taking into consideration the dominant character of the electric dipole at long wavelengths, the electric dipole might be expected to make the dominant contribution to the resonances associated with the set of moments (4.33). Similarly, because of the negligible magnetic dipole response of an isolated sphere, it might seem that the inclusion of the magnetic dipole in the set (4.34) would lead to negligible effects. I shall show below that the first of these expectations is confirmed over a wide range of wavelengths, but the second is not.

I compare the total contribution to the specular reflection from the \((p_z, m_z, q_{qz})\) set with the reflection calculated within the electric dipole approximation; see Fig. 4.7. The dipole approximation is accurate for broad peaks observed at small wavelengths. At large wavelengths, \( \lambda \approx 900 \text{ nm} \), the inclusion of higher multipoles leads to some shift of the resonance towards the Rayleigh anomaly and to a non-zero reflectivity at \( \theta = \theta_R \); I will refer to the non-zero value of the reflectivity at the onset of diffraction as the suppression of the anomaly. Finally, I note that the range of wavelengths over which the resonance is observed is only weakly affected by the coupling of the dipole to the higher multipoles. Both within the full multipolar model and within the dipole approximation the SLR disappears at a cut-off wavelength \( \lambda_{cf} \) close to wavelength at which the Rayleigh line moves to the right of the light line in Fig. 4.3, \( \lambda_{cf} \approx \lambda_{gz} \); the difference in \( \lambda_{cf} \) within both approximations is of the order of few nanometers.

Next I compare the contribution to the specular reflection from the set \((m_z, q_{qz})\) with that of the electric quadrupole \( q_{qz} \) only, see Fig. 4.8. Again I observe that the inclusion of the magnetic response leads to a shift of the resonance and to the suppression of the Rayleigh anomaly, but this time the changes are much more pronounced and are significant over the whole range of wavelengths I consider.
Figure 4.7: The contribution to specular reflection from the set (4.33), compared with the contribution from the electric dipole only. Arrows identify the angles at which the Rayleigh line is crossed and diffraction appears.

I also note that the coupling to the magnetic dipole significantly affects the wavelengths over which the multipolar SLR is observed; if the magnetic response of the sphere were neglected, the merging of the SLR with the Rayleigh anomaly would not occur and the SLR would be found for all wavelengths \( \lambda \lesssim \lambda_{gx} \).

Thus even when there is a clearly dominant multipole moment, as for the set \((p_s, m_z, q_{\kappa s})\), the inclusion of other moments can lead to a change in the shape of the SLR and the suppression of the Rayleigh anomaly. And for the set \((m_{\kappa}, q_{sz})\), where one might expect the quadrupole moment to be dominant because of the small and non-resonant response of the magnetic dipole moment of an isolated sphere at these wavelengths, the inclusion of the magnetic dipole not only suppresses the Rayleigh anomaly, but significantly modifies the shape and the position of the SLR. To understand these features I first turn to a characterization of the SLRs as the excitation of effective modes of the array.

### 4.4 Surface-lattice modes

By way of background, consider first a simple planar system, such as a thin metal film surrounded by a dielectric with index of refraction \( n \). Over a range of frequencies the metal film will support long-range surface plasmons [101, 102]. For a real frequency \( \omega \) each possible long-range surface plasmon is characterized by a wave vector \( \kappa \) in the plane of the film. Because of absorption in the metal the plasmons decay as they propagate, and the corresponding wave vector \( \kappa \) is complex. The complex \( \kappa \) identifies a pole in the appropriate response coefficient of the structure, which very generally can be taken to identify the modes of a system. If absorption is neglected, \( \kappa \) moves onto the real axis. But since the fields are evanescent in the dielectric, \( |\kappa| > \tilde{\omega} n \), the excitation cannot be accessed by an incident plane wave, for which \( |\kappa_0| < \tilde{\omega} n \). In practice, one way to couple into such an excitation is with the aid of a grating structure [101, 102]. If the grating is characterized by a wave vector \( K \) such that \( \kappa \approx \kappa_0 + K \), coupling can be achieved.

Returning to the lattice of spheres, I analyze SLRs by considering the lattice as an effective planar structure plus a grating. With the neglect of loss, the SLRs observed in the specular reflection correspond
Figure 4.8: The contribution to specular reflection from the set (4.34), compared with the contribution from the electric quadrupole only. Arrows indicate the angle at which the Rayleigh line is crossed and diffraction appears.

to the coupling into modes of the structure, each characterized by a wave vector \( \kappa \) and using one of the \( \mathbf{K} \) of the lattice such that \(|\kappa_0 + \mathbf{K}| = \kappa\) \([88, 89, 90, 91]\). The strong coupling into the mode and back into the specularly reflected light leads to the large reflectivity. I refer to these modes as surface-lattice modes (SLMs). No evidence of SLM appears in the diffracted intensity, of course, since \( \kappa > \tilde{\omega}n \) and at \( \kappa \) there is no radiation into the dielectric.

The loss relevant here is due both to absorption in the metal and to coupling through \(-\mathbf{K}\) of the SLM back to the incident wave vector \( \kappa_0 \); the light at \( \kappa_0 + \mathbf{K} \) coupled back at \( \kappa_0 \) contributes to a loss of the SLM through the increased specular reflectivity that signals its very presence. So to identify a particular SLM in the lossless limit I need to neglect the elements of the response associated with the absorption in the metal and with the radiation at \( \kappa_0 \), identify the response to the incident field of the relevant combination of multipoles, and for each frequency \( \omega \) of interest identify the \( \kappa \) at which the response diverges. The resulting function \( \omega(\kappa) \) identifies the dispersion relation of the SLM in the lossless limit, what I call the “ideal” dispersion relation.

### 4.4.1 Energy balance

To construct these ideal dispersion relations the first step is then to isolate the loss mechanisms. The energy balance in a magneto-electric quadrupolar array has been discussed in detail in chapter 3, but I review here the main findings for more clarity. I first consider the energy balance in an array of dipoles. Within the dipole approximation, the response of an electric dipole to the incident field can be described in terms of an effective polarizability of the array \([6, 23, 34]\),

\[
\mathbf{p} = \tilde{\alpha}^{\text{eff}} \mathbf{E}^{\text{inc}},
\]

(4.49)

and where \( \tilde{\alpha} = \alpha^{\text{eff}} \tilde{U} \). The real part of the inverse effective polarizability (4.49) does not affect the energy balance in the system. The imaginary part of the inverse polarizability of the sphere, \( \tilde{\alpha}^{-1} \), determines the resonance broadening due to the absorption losses. The imaginary part of the periodic
Green function identifies radiation loss and is known in a analytic form,

\[ \text{Im} \mathbf{G}_n^\text{Ep} = \sum_{n \in |\kappa_n| < \omega_n} \text{Im} \mathbf{G}_n^\text{Ep}, \]

where \( \text{Im} \mathbf{G}_n^\text{Ep} \) is given by

\[ \text{Im} \mathbf{G}_n^\text{Ep} = \frac{\mu}{\epsilon_0 n^2} \frac{u_n^2 \hat{e}_n \hat{e}_n + (\omega_n)^2 \hat{s}_n \hat{s}_n + \kappa_n^2 \hat{z} \hat{z}}{2 w_n A_z}. \]

Each term in the sum (4.50) accounts for the radiation reaction associated with a diffracted beam radiating at an angle \( \theta_n = \sin^{-1}(|\kappa_n|/\omega_n) \); see eqs. (4.45,4.46).

These considerations generalize immediately to an array described by a full set of multipole moments (4.23-4.25). As in the dipole example, the imaginary part of the inverse proper polarizabilities determine the Ohmic losses. The radiative losses are determined by the imaginary part of the periodic Green functions \( \mathbf{g}_n^\text{Ep}, \mathbf{g}_n^\text{Bm}, \mathbf{g}_n^\text{Em}, \mathbf{g}_n^\text{Bp} \), and by the real part of the Green functions \( \mathbf{g}_n^\text{Eq}, \mathbf{g}_n^\text{Pp}, \mathbf{g}_n^\text{Eq}, \mathbf{g}_n^\text{Pm} \).

These contributions to the periodic Green functions are given in chapter 3 by simple analytic expressions, which in the pattern of (4.50) are given by a Fourier sum over wave vectors \( \kappa_n \), \(|\kappa_n| < \omega_n \), with the Fourier component at \( \kappa_n \) accounting for the radiation reaction associated with the beam radiating at an angle \( \theta_n \).

### 4.4.2 Dispersion relations

With the channels for energy loss due to both absorption and radiation identified, the SLMs of the array in the lossless limit can now be found. For the illumination conditions discussed in section 4.3, the matrices \( S \) and \( S' \), eqs. (4.33-4.34), describe the response as a function of \( \kappa_0 \), the in-plane component of the incident wave vector. The divergences in the response, and thus the modes, are associated with the vanishing of the determinants of the matrices in the limit of no loss. I now neglect absorption and radiation by dropping the imaginary parts of the inverse proper polarizabilities and the parts of the relevant Green functions associated with radiation reaction. Writing the determinants of \( S \) and \( S' \) in this limit as \( d(\kappa_0) \) and \( d'(\kappa_0) \) respectively, I find that these vanish for certain \( \kappa_0 \). In general \( d(\kappa_0) \) and \( d'(\kappa_0) \) are periodic functions over the Brillouin zone, so any vanishing of determinant for \( \kappa_0, |\kappa_0| < \omega_n \), is replicated with the periodicity of the Brillouin zone over all reciprocal space; thus each \( \kappa_0 \) I associate with a SLM at \( \kappa_n = \kappa_0 + K_n, |\kappa_n| > \omega_n \). Due to periodicity of the system there is an ambiguity in the choice of \( K_n \) and thus in the choice of \( \kappa_n \). Out of all the possible reciprocal vectors I choose the \( K_n \) that corresponds to \( \kappa_n \) closest to the light line, which for the parameter space that I consider corresponds to \( \kappa_n \) in the second Brillouin zone. I now claim that even when absorption and radiation are included, the SLRs can be understood as arising from a coupling of the incident field at \( \kappa_0 \) to the SLM at \( \kappa_n \) and then coupling back to \( \kappa_0 \) by the vector \( -K_n \). In Figure 4.5 I plot in dotted lines the location of the divergence in \( d(\kappa_0) \) that I find (associated with the set of moments \( (p_s, m_z, q_{ks}) \)), and the location of the divergence in \( d'(\kappa_0) \) that I find (associated with the set of moments \( (m_{ks}, q_{ks}) \)). There is excellent agreement between the location of the divergences in the lossless limit and the peaks that appear in the full calculation, both for the scenario depicted in Fig. 4.5 and at other wavelengths as well, justifying the physical picture that I presented.

Because of a high symmetry of the illumination conditions discussed in section 4.3, the value of \( \kappa_0 \)
within the light line where a divergence in \(d(\kappa_0)\) or \(d'(\kappa_0)\) occurs is in each case associated with two \(\kappa\) of equal magnitude \(\kappa\), \(\kappa_{(10)} = \kappa_0 + K_{(10)}\) and \(\kappa_{(11)} = \kappa_0 + K_{(11)}\), and thus two SLMs. The scenario is sketched in reciprocal space at a fixed \(\omega\) in Fig. 4.9. The light circle is the set of \(\kappa\) for which \(\kappa = \tilde{\omega}n\);

![Diagram of light and Rayleigh circles](image)

Figure 4.9: For a fixed \(\omega\), the SLMs at \(\kappa_{(10)} = \kappa_0 + K_{(10)}\) and \(\kappa_{(11)} = \kappa_0 + K_{(11)}\). The solid blue line indicates the “light circle”, the set of \(\kappa\) for which at chosen frequency \(\kappa = n\omega/c\). The dotted blue lines indicate the “Rayleigh circles”, which are light circles centered at reciprocal lattice sites other than that at the origin; diffraction associated with a particular \(K\) occurs when \(\kappa_0\) crosses its Rayleigh circle.

the point on the light circle crossed when moving in the direction of \(\hat{\kappa}_0\) identifies the point on the light line of Fig. 4.3 at the chosen frequency. The Rayleigh circles are the circles of the same radius \(\tilde{\omega}n\) centered at the other reciprocal lattice sites. It is easy to confirm that the light circle and the Rayleigh circles identify the curves in reciprocal space where the periodic Green functions undergo singular and, more generally, non-analytic behavior, as discussed in section 4.2.2. In Fig. 4.9, the first point on a Rayleigh circle that is crossed when moving through reciprocal space in the direction of \(\hat{\kappa}_0\) identifies the point on the Rayleigh line of Fig. 4.3 at the chosen frequency. This point is identified in Fig. 4.9 as the “crossing point” and is associated with (10) and (11) diffraction orders due to the symmetry. The SLRs found in the vicinity of the “crossing point” are associated with the SLMs that would result were loss ignored; those modes reside beyond the light circle as is indicated. For different frequencies \(\omega\) the value of \(\kappa = |\kappa_{(10)}| = |\kappa_{(11)}|\) associated with the SLMs is different, and I can plot out the dispersion relation \(\tilde{\omega}(\kappa)\). This is presented in Fig. 4.10, both for the SLM associated with the dipolar SLR, and that associated with the multipolar SLR.

The dispersion relations are what one would expect for those characterizing excitations bound to a surface or thin film, in that at lower wave numbers they are close to the light line, and they pull away at higher wave numbers. What is novel is their termination. The multipolar SLM terminates at a finite frequency on the light line, while the dipolar SLM terminates at a lower frequency away from the light line.

The termination of the dipolar SLM can be understood with the aid of the shaded region in Fig. 4.10. At each \(\omega\) it identifies the magnitudes \(\kappa = |\kappa|\) of the wave vectors \(\kappa\), which (a) are of the form \(\kappa = \kappa_0\hat{k}_a + K\) where \(K = K_{(10)}\) or \(K = K_{(11)}\) (b) are beyond the light line but within the second Brillouin zone, and (c) for which the periodic Green functions remain analytic, which is to say that the light circle and the Rayleigh circles are avoided. In general, the wave vectors on the Rayleigh circles can be excited...
with or without the help of the grating. The wave vectors on the Rayleigh circle that can be excited without the help of the grating are of the form \( \kappa = \kappa_0 \), for example the crossing point in Fig. 4.9, and they identify a point on the Rayleigh line of Fig. 4.3. The wave vectors on the Rayleigh circles that are excited with the help of the grating are of the form \( \kappa = \tilde{\omega} n \hat{\kappa}_a + K \), where \( K = K_{(10)} \) or \( K = K_{(11)} \); see for example Fig. 4.11, drawn for a lower frequency than Fig. 4.9. These wave vectors identify the onset of diffraction, and can thus be taken to define a new Rayleigh line that I plot in Fig. 4.10 as a dashed line; due to the symmetry at each frequency there are two such wave vectors of equal magnitude. In consideration of the SLMs it is the Rayleigh line of Fig. 4.10 that is important. Below \( \tilde{\omega}_{gz} \) this Rayleigh line restricts the shaded area in Fig. 4.10. When the dispersion relation of the SLM meets this Rayleigh line, at \( \tilde{\omega}_{cf} \), the singular change in the Green functions terminates the SLM. The situation in reciprocal space at this frequency is depicted in Fig. 4.11, where each SLM falls precisely on the Rayleigh circle of a reciprocal lattice vector. I note that were the magnetic dipole neglected there would be a small change of the dipolar dispersion relation, resulting in a small change in the cut-off frequency; however, the underlying description of the SLM termination would remain the same.

Figure 4.10: Ideal dispersion relations of the two SLMs. Blue lines identify the light line (solid) and the Rayleigh line (dashed) of the SLM. The terminations of the dispersion relations are indicated by circles.

Figure 4.11: Scenario in the reciprocal lattice at the cut-off frequency \( \omega_{cf} \). The SLM wave vector \( \kappa = \tilde{\omega}_{cf} n \hat{\kappa}_a + K \), for \( K = K_{(10)} \) or \( K = K_{(11)} \), resides on a Rayleigh circle.
The inclusion of magnetic response in the description of multipolar SLM, on the other hand, changes the cut-off frequency from a value just below \( \omega_{gz} \) to a value significantly above it; see Fig. 4.10. This dramatic change in cut-off frequency of the multipolar SLM cannot be identified with simple features of the reciprocal lattice geometry, as could the location of the cut-off frequency of the dipolar SLM. To understand the termination of the multipolar SLM, for which the coupling between the magnetic dipole and electric quadrupole terms plays an important role, and to understand in more detail how the singular change in the Green function terminates the dipolar SLM, it is useful to develop a simplified model for the SLMs based on expansions of the periodic Green functions.

4.5 Simplified model

With the neglect of loss due to absorption or diffraction, each SLM I have identified involves a set of wave vectors, each wave vector in the set differing from another in the set by a translation over a reciprocal lattice. For much of the parameter space considered in the last section, at least one wave vector in the set lies close to the light line, and I considered the dispersion relation of the SLM associated with it. In the vicinity of the light line periodic Green functions can diverge, and it is this divergent behaviour that plays a dominant role in how the dispersion relation of the SLM is established. In this section I establish a simplified model for the SLMs based on isolating the divergent contribution.

I begin by considering the dispersion relation of the dipolar SLM. For simplicity I neglect the contribution to the dipolar SLM from the electric quadrupole moment of the spheres, as the corrections beyond the electric dipole approximation come mainly from the magnetic response. In the absence of the quadrupole terms eq. (4.33) reduces to

\[
S_{pm} \begin{pmatrix} \psi_s \\ m_z \end{pmatrix} = \begin{pmatrix} E_{inc}^s \\ B_{inc}^z \end{pmatrix},
\]

(4.52)

where the matrix \( S_{pm} \) is given by,

\[
S_{pm} = \begin{pmatrix} (\alpha^p)^{-1} - G_{sz}^E & -G_{sz}^E \\ -G_{sz}^B & (\alpha^m)^{-1} - G_{zz}^E \end{pmatrix},
\]

(4.53)

and the components of the dyadics and vectors are written in the \((\hat{s}, \hat{\kappa}, \hat{z})\) basis associated with the incident light, that is here \( \hat{\kappa} = \hat{\kappa}_0 \), etc. To evaluate the periodic Green functions in an approximate way that respects the important physics, I note that each periodic Green function has a contribution from a Fourier sum that includes components at each \( \kappa_n = \kappa_0 + \hat{\kappa} \). Each of these components can diverge as \( w_{n}^{-1} = (\tilde{\omega}^2 - \kappa_n^2)^{-1/2} \) for \( \kappa_n \) close to the light line. Looking at the scenario shown in Fig. 4.9, for \( \kappa_0 \) close to the SLR I have \( \kappa_{(10)} \) and \( \kappa_{(11)} \) close to the light line; I treat these Fourier components separately. Due to symmetry \( w_{(10)} = w_{(11)} \equiv w_R \), where \( w_R \) is a function of two independent variables \( \kappa_0 \) and \( \tilde{\omega} \), and the Fourier components at \( \kappa_{(10)} \) and \( \kappa_{(11)} \) diverge as \( w_R^{-1} \); the next term in an expansion of these Fourier components in powers of \( w_R \) is of order \( w_R \), and I neglect it. The remaining Fourier components and the contribution from the multipoles at the origin are analytic in \( w_R \) as \( w_R \to 0 \), and the leading non-vanishing terms are of order \( w_R^0 \). I keep only this order and denote the sum of all such terms in the periodic Green functions as \( \tilde{G}_{sz}^{EP}, \tilde{G}_{sz}^{EM} \), etc. I then use these approximate expressions to identify the singular and the leading analytic terms in \( w_R \) as \( w_R \to 0 \) for each of the three independent
elements of the matrix $S_{pm}$, eq. (4.53), taken in the limit of no loss. The analytic contributions are
given by the functions independent of $w_R$,

$$
\beta^{-1}_p \equiv \text{Re} \left( \alpha^{\text{PE}} \right)^{-1} - \tilde{G}^{\text{EP}}_{ss},
$$
(4.54)

$$
\beta^{-1}_m \equiv \text{Re} \left( \alpha^{\text{MB}} \right)^{-1} - \tilde{G}^{\text{BM}}_{zz},
$$
(4.55)

$$
\beta^{-1}_{cp} \equiv \tilde{G}^{\text{EM}}_{sz} = -\tilde{G}^{\text{BP}}_{zs}.
$$
(4.56)

I denote the singular contribution to the electric dipole Green function as

$$
\mathcal{G}_{\text{sng}} = \text{Re} \frac{i}{\epsilon_0 n^2} \frac{(\tilde{\omega}n)^2 - K_s^2}{A_c w_R},
$$
(4.57)

where $K_s = |\mathbf{K}_{(10)} \cdot \hat{s}| = |\mathbf{K}_{(11)} \cdot \hat{s}|$. The singular contributions to the other Green function terms differ from $\mathcal{G}_{\text{sng}}$ by a factor $\eta$,

$$
\eta = \frac{n}{c} \frac{\tilde{\omega}n}{\sqrt{(\tilde{\omega}n)^2 - K_s^2}}.
$$

I can treat the $\beta$'s as “dressed” polarizabilities of a sphere on a lattice in the absence of absorption and radiation losses, since they do not diverge at the onset of the Rayleigh anomalies and they depend on frequency only. However, even if the $\alpha^{\text{PE}}$ and $\alpha^{\text{MB}}$ are to good approximation Lorentzian functions of frequency, in general $\beta_p$, $\beta_m$, and of course $\beta_{cp}$ will not be Lorentzian, due to the frequency dependence of the Green function terms. Nevertheless, the frequency dependence of the $\beta$ coefficients identifies the SLM dispersion relation, as I now show.

SLMs are identified by wave vectors $\kappa$ at which, with absorption and diffraction neglected, the determinant of the matrix (4.53) vanishes. Within the dipole approximation the condition for SLM is given by

$$
\beta_p \mathcal{G}_{\text{sng}} = 1.
$$
(4.58)

When the coupling to magnetic dipole is included, the determinant of the matrix $S_{pm}$ (here taken in the absence of loss) is quadratic in periodic Green functions. Nevertheless, a cancellation of terms that are quadratic in $\mathcal{G}_{\text{sng}}$ results in a condition for SLM of the same form as before,

$$
\beta_{pm} \mathcal{G}_{\text{sng}} = 1,
$$
(4.59)

but with a dressed polarizability renormalized by the coupling to magnetic dipole. Here $\beta_{pm}^{-1} \equiv \beta_p^{-1} + \Delta_\beta$, where the correction due to the coupling is given by

$$
\Delta_\beta = \frac{(\eta \beta_p^{-1} - \beta_{cp}^{-1})^2}{2 \eta \beta_{cp}^{-1} - \eta^2 \beta_p^{-1} - \beta_m^{-1}}.
$$
(4.60)

Using eq. (4.57) in (4.58,4.59) I solve for the value of $i^{-1}w_R$ that satisfies the SLM condition respectively in the dipole approximation and when the magnetic dipole is taken into account; the solution characterizes the distance of the SLM at $\kappa = \kappa_{(10)}$ and $\kappa = \kappa_{(11)}$ in each of the approximations from the light line. I find that this quantity is proportional to the dressed polarizability $\beta$,

$$
\frac{1}{i} w_R = \frac{1}{\epsilon_0} \frac{(\tilde{\omega}n)^2 - K_s^2}{n^2 A_c \beta},
$$
(4.61)
with and without the coupling to magnetic dipole included, where respectively \( \beta = \beta_{pm} \) and \( \beta = \beta_{p} \). From eq. (4.61) it follows that the SLM merges with the light line when \( \beta \) vanishes.

I note if \( \alpha^{\text{PE}} \) is well approximated by a Lorentizan, \( (\alpha^{\text{PE}})^{-1} \approx A(\omega_{0}^{2} - \omega^{2}) - i\Gamma \), the inverse dressed polarizability in the dipole approximation with loss neglected is of the form

\[
\beta_{p}^{-1} = A(\omega_{0}^{2} - \omega^{2}) - \tilde{\beta}_{ss}^{\text{EP}}.
\]

(4.62)

The resulting polarizability \( \beta_{p} \) does not vanish, and so from (4.61) I see there is a finite distance of the SLM to the light line within the dipole approximation, and no termination of the SLM. However, when the coupling to the magnetic dipole is included, the polarizability \( \beta_{pm} \) vanishes when the correction (4.60) diverges. This is confirmed by the plots in Fig. 4.12 for the system considered here, where indeed

Figure 4.12: Dispersion relations found within the multipolar model (continuous lines) and within the simplified model (dashed lines). Insets show the “dressed” polarizabilities of the dipolar (left) and multipolar (right) SLMs with and without the coupling to magnetic dipole. When the coupling is included, the polarizabilities vanish at \( \tilde{\omega}_{v} \) and \( \tilde{\omega}'_{v} \) respectively.

\( \beta_{p} \) is finite while \( \beta_{pm} \) vanishes at a frequency labelled \( \omega_{v} \); see the inset. As the frequency is lowered, one might then expect that at \( \omega_{v} \) the normal SLM dispersion relation would reach the light line and the SLM would terminate.

Yet this scenario is pre-empted by the fact that at a frequency \( \omega_{cf} > \omega_{v} \) the position of the SLM reaches the Rayleigh circle, as discussed above (see Fig. 4.11), and as the Rayleigh circle is approached a new singular contribution to the periodic Green functions becomes important. When loss and the radiative contributions to the periodic Green function are ignored, this term vanishes until the Rayleigh circle is crossed, but it gives a divergent contribution at the crossing. This additional singular contribution to the Green function results in the vanishing of the SLM, which beyond the Rayleigh circle is no longer described by the simplified dispersion relations presented above, where only one singular term was taken into account.

The situation for the multipolar SLM is completely different, although a similar analysis can be applied. I introduce “dressed” polarizabilities that describe the leading analytic contributions to the
three independent elements of matrix $S'$ taken in the limit of no loss,

$$
\beta'^{-1}_q = \Re \left( \alpha'^q \right)^{-1} - 2 \tilde{G}^F_{szsz},
$$

$$
\beta'^{-1}_m = \Re \left( \alpha'^m \right)^{-1} - \tilde{G}^B_{kk},
$$

$$
\beta'^{-1}_{cp} = \tilde{G}^B_{sz} = -\tilde{G}^F_{sz},
$$

(4.63)

where $\tilde{G}^F_{szsz}$, $\tilde{G}^B_{kk}$, etc. denote the leading analytic contributions to periodic Green functions. The singular contribution to the quadrupole Green function I denote as

$$
\mathcal{G}'_{sng} = \Re \frac{i}{\epsilon_0 n^2} \frac{\omega_n^2 K_s^2}{4 A_c w_R},
$$

(4.64)

with the singular contributions to other Green functions differing by a parameter $\eta' = 1/(\tilde{\omega} c)$. In this notation the condition for the SLMs is given by

$$
\frac{1}{i} w_R = \frac{1}{\epsilon_0 n^2} \frac{\omega_n^2 K_s^2}{2 A_c} \tilde{\beta}',
$$

(4.65)

where $\beta' = \beta'_q$ when the coupling to magnetic dipole is neglected, and $\beta' = \beta'_{qm}$ when magnetic dipole is included in the description; again, $i^{-1} w_R$ characterizes the distance of the SLM dispersion relation from the light line. The quadrupole dressed polarizability renormalized by the coupling is given by

$$
\beta'^{-1}_{qm} = \beta'^{-1}_q + \Delta'_{qm},
$$

(4.66)

with the correction

$$
\Delta'_{qm} = \frac{(\eta' \beta'^{-1}_q - \beta'^{-1}_{cp})^2}{2 \eta' \beta'^{-1}_{cp} - \eta'^2 \beta'^{-1}_q - \frac{1}{2} \beta'^{-1}_m}.
$$

(4.67)

As in the discussion of the dipolar SLM, I find that when the SLM is approximated by one dominant multipole (here the electric quadrupole), the dressed polarizability does not vanish, $\beta'^{-1}_q \neq 0$; thus were only this multipole present there would be no termination of the SLM dispersion. However, when the coupling to the magnetic dipole is included the renormalized dressed polarizability $\beta'^{-1}_{qm}$ vanishes at a frequency $\omega'_v$; here $\omega'_v > \omega'_c$, and so the merging of the SLM dispersion relation with the light light is not pre-empted by the presence of a Rayleigh circle, and the dispersion relation terminates at $\omega'_v$, as shown by the green circle in Fig. 4.10.

I emphasize that the frequency at which the multipolar SLM vanishes, $\omega'_v$, strongly depends on all the details of the system, including both the lattice geometry and the polarizabilities of the spheres. This is in contrast to the situation for the dipolar SLM, which terminates before its dispersion relation reaches the light line. There the frequency at which the SLM vanishes, $\omega'_c$, depends mostly on the geometry of the reciprocal lattice which determines the Rayleigh anomaly line associated with the SLM of Fig. 4.10.

4.6 Surface-lattice modes and the sensitivity of SLRs to non-resonant multipoles

I have shown that the SLM dispersion relations identify the position of the SLRs and the range of frequencies over which they exist. Now I show that the sensitivity of the SLRs to the inclusion of non-
Chapter 4. Surface - lattice resonances in 2d arrays of spheres

resonant magnetic dipole in their description can be linked to the dispersion relations as well. The link I provide is only approximate, but nevertheless it gives an insight into a large sensitivity of the multipolar SLRs and of the dipolar SLRs at long wavelengths.

Besides their position, the SLRs are also characterized by their widths, which are mostly restricted by the rapid variations of the reflection close to the Rayleigh line; see Figs. 4.7, 4.8. For SLRs that are well described with the electric dipole approximation the reflection vanishes at the Rayleigh line, and the distance in reciprocal space from the peak to the Rayleigh line can be taken as an indication of the width [75]. When the contribution from more than one multipole becomes important the reflection at the Rayleigh line is in general finite, but it decreases rapidly as \( \kappa_0 \) moves beyond the line. Thus quite generally I can take the reciprocal space distance \( \Gamma = |\kappa_R - \kappa_{SLR}| \), where \( \kappa_{SLR} \) is the wave vector of an SLR and \( \kappa_R \) is the wave vector of the Rayleigh line, as an indicator of the width of the SLR. Even though the distance \( \Gamma \) gives only an approximate characterization of the SLR profile, it has an advantage over more sophisticated treatments in that it simultaneously identifies both the width and the position of an SLR. I refer to it as the “width parameter”, and a straight-forward calculation shows that it can be related to the wave vector of the SLM, \( \kappa_{SLM} = \kappa_{SLR} + K_n \) with \( n = (1, 0) \) or \( n = (1, 1) \), by

\[
\Gamma = d \left[ \frac{\kappa_{SLM} + \omega n}{\sqrt{\kappa_{SLM}^2 - K_s^2 + \sqrt{(\omega n)^2 - K_s^2}}} \right], \tag{4.68}
\]

where \( \kappa_{SLM} = |\kappa_{SLM}| \), \( d = \kappa_{SLM} - \omega n \) is the distance of the SLM to the light line, and \( K_s = |K_{(10)} \cdot \hat{s}| = |K_{(11)} \cdot \hat{s}| \) as in section 4.5.

I can then quantify the sensitivity of both the position and the width of the SLR to the inclusion of non-resonant multipoles by one simple parameter \( s \equiv (\Gamma^d - \Gamma^{cp})/\Gamma^d \). Here \( \Gamma^d \) is the width parameter predicted if only the dominant multipole is included in the analysis – in the problem considered here the electric dipole moment for the dipolar SLM and the electric quadrupole moment for the multipolar SLM – and is found by using \( \kappa_{SLM} \) in that limit in (4.68). Similarly, \( \Gamma^{cp} \) is the width parameter when the non-resonant multipole – in the problem considered here the magnetic dipole moment – is also taken into account, and is found by using the \( \kappa_{SLM} \) of that full calculation in (4.68). Using (4.68) to determine these quantities and noting that the term in brackets varies little when the magnetic dipole is included, I arrive at a simple relation

\[
s \approx 1 - \frac{d^{cp}}{d^d}, \tag{4.69}
\]

where \( d^d \) is the distance in reciprocal space of the position of the SLM to the light line when the magnetic dipole is neglected, and \( d^{cp} \) is the corresponding distance when the magnetic dipole is included in the calculation. The sensitivity of the SLR to the inclusion of the magnetic dipole moment of the spheres is thus most enhanced when, as a result of the inclusion of the magnetic dipole moment in the calculation, the position of the associated SLM reaches the light line, \( d^{cp} \to 0 \). Thus the merging of the dispersion relation of the multipolar SLM with the light line drives an extreme sensitivity of the multipolar SLR to the inclusion of the magnetic dipole moment. Such an extreme sensitivity is not observed for the dipolar SLR, since the merging of its dispersion relation with the light line is pre-empted by the appearance of the new diffracted order.

The SLM dispersion relations are also indicators of the sensitivity of SLRs to illumination conditions. This point I discuss in the next section.
4.7 Dependence of the spectrum on illumination conditions

I now turn to the dependence of the reflectance on the polarization and direction of the incident light. I first analyze an angular scan of the reflection with an in-plane wave vector $\mathbf{\kappa}_0 \propto \hat{\mathbf{\kappa}}_a$, as done in section 4.3-4.5, but now with the incident light p-polarized. While for s-polarization the two independent sets of moments were $(p_s, m_{\kappa}, q_{ss})$ and $(m_{\kappa}, q_{sz})$ (see eqs. (4.33,4.34)), for p-polarization they are $(p_{\kappa}, q_{\kappa\kappa}, q_{ss})$ and $(p_z, m_s, q_{nz})$. That is, both sets involve an electric dipole moment component, coupled to higher multipoles. The structure of the coupling between the multipoles within these sets is important in understanding the radiation from the array. For s-polarized light I could identify qualitatively different SLRs, a dipolar SLR associated with the set of moments $(p_s, m_{\kappa}, q_{ss})$ and a multipolar SLR associated with the set $(m_{\kappa}, q_{sz})$, but I shall see that the situation is more complicated here.

As in the discussion of illumination by s-polarized light, the Rayleigh line and the light line are identified in Fig. 4.3, and I consider the range of wavelengths between $\lambda_{nr}$ and $\lambda_{gz}$. The angular scan of the specular reflectance for wavelengths in this range is shown in Fig. 4.13. At angles before the onset of diffraction, $\theta_0 < \theta_R$, the reflection profile exhibits damped SLRs at long wavelengths, $\lambda \gtrsim 780$ nm, a clear broad SLR at intermediate wavelengths, $\lambda \approx (720 - 780)$ nm, and some resonant structure or kinks in reflectivity at short wavelengths, $\lambda \lesssim 720$ nm. To gain insight into the nature of these resonances I analyse the contributions to the reflection spectrum from the two sets of moments. From a calculation within the electric dipole approximation I find that at long wavelengths the response of both the in-plane $(p_{\kappa})$ and out-of-plane $(p_z)$ components of the electric dipole moment are resonant. The broad SLRs found within the full multipolar model at long wavelengths are of mostly dipole character and result from a mostly destructive interference of the resonances associated with the sets $(p_{\kappa}, q_{\kappa\kappa}, q_{ss})$ and $(p_z, m_s, q_{nz})$.

At intermediate wavelengths the response of the electric dipole component $p_{\kappa}$ is resonant while $p_z$ is not, and the reflection profile is dominated by an SLR associated with the set of moments $(p_{\kappa}, q_{\kappa\kappa}, q_{ss})$. At short wavelengths the response of neither $p_{\kappa}$ nor $p_z$ is resonant. Here the SLRs are of a mixed electric dipole-electric quadrupole character in a narrow range of wavelengths close to $\lambda \approx 650$ nm, and kinks in reflection arise at other short wavelengths. But generally there is no clear SLR structure that can be associated with higher order multipole moments, in contrast to the prominent structure of multipolar

![Figure 4.13: Resonant structure of the specular reflection for p-polarized light with $\mathbf{\kappa}_0 \propto \hat{\mathbf{\kappa}}_a$. Arrows indicate the angle at which the Rayleigh line is crossed and diffraction appears. Plots are shifted by 1 to improve clarity.](image-url)
SLRs found in the angular scans for s-polarized incident light, as shown in section 4.3.

To investigate this difference, I first compare the effects on the specular reflection of s- and p-polarized light due to the sets of multipoles \((p_s, m_s, q_{sz})\) and \((p_z, m_z, q_{sz})\), excited by s-polarized light and p-polarized light respectively (see Eqs. (4.33-4.38)); both of these involve an in-plane component of the dipole moment coupled to higher multipoles. The similar nature of the matrices \(S\) and \(P\) that govern these sets of multipoles results in qualitative similar angular scans of specular reflection, dominated by a broad SLR of dipolar nature over a wide range of wavelengths. In contrast, the nature of the matrices \(S'\) and \(P'\) governing of the response of the sets of multipoles \((m_n, q_{nz})\) and \((p_z, m_z, q_{nz})\), excited by s-polarized light and p-polarized light respectively, are quite different; the higher multipoles in the set \((p_z, m_s, q_{sz})\) are coupled to a dipole component, while no such component is present in the set \((m_n, q_{sz})\). Not surprisingly, the angular scans of the specular reflection due to the two sets of multipoles are markedly different at the long wavelengths at which the response of \(p_z\) is resonant. At these wavelengths the contribution to the specular reflection due to the multipoles \((p_z, m_s, q_{sz})\) leads to a broad response, much like the dipolar SLR that arises due to the moments \((p_z, m_z, q_{sz})\) that are excited by s-polarized light, while no dipole moment component is present in the set of multipoles \((m_n, q_{sz})\). But what is surprising is the markedly different angular scans of specular reflection due to the sets of multipoles \((p_z, m_z, q_{sz})\) and \((m_n, q_{sz})\), excited by p-polarized light and s-polarized light respectively, at intermediate and short wavelengths at which the response of the dipole moment component \(p_z\) is non-resonant.

The difference arises due to the coupling of the higher multipoles in the set \((p_z, m_s, q_{sz})\) to the non-resonant electric dipole component, as can be seen by comparing the radiation from this set with that which would be predicted were the dipole component \(p_z\) neglected in the calculation. In the latter calculation there is a clear structure with sharp features, similar to the multipolar SLR due to the set of moments \((m_n, q_{sz})\) observed in the response of the array to s-polarized light; this structure is absent in the full calculation of the radiation of the moments \((p_z, m_z, q_{sz})\). This extreme sensitivity of the multipolar response to changes in multipolar couplings can be understood by analyzing the SLMs of the array. I come back to this at the end of this section.

I now turn to a consideration of how the direction of the in-plane wave vector \(\kappa_0\) affects the angular scans of specular reflectance. The most significant change of reflectance with the direction of \(\kappa_0\) is on the sharp features associated with the multipolar SLR for s-polarized illumination, and since for p-polarized illumination no such sharp features arise I restrict my attention for the rest of this section on s-polarized illumination. I identify the direction of \(\kappa_0\) by the angle that it makes with that symmetry direction, \(\phi = \cos^{-1}(\hat{\kappa}_0 \cdot \hat{\kappa}_a)\), and I refer to \(\phi\) as a misalignment angle; see Fig. 4.14. As the lattice has a 6-fold rotational symmetry and an inversion symmetry, I need only consider angles in the interval \(\phi \in (0^\circ, 30^\circ)\); here \(\phi = 0^\circ\) corresponds to wave vector along the first symmetry direction, \(\kappa_0 \propto \hat{\kappa}_a\), and \(\phi = 30^\circ\) is equivalent to wave vector along the second symmetry direction, \(\kappa_0 \propto \hat{\kappa}_b\). I calculate angular scans of the specular reflectance by keeping the misalignment angle fixed and varying the angle of incidence.

These angular scans reveal that, while the broad dipolar SLR is only weakly dependent on the direction of \(\kappa_0\), the multipolar SLR is strongly anisotropic and observed only when \(\kappa_0\) is closely aligned with the symmetry direction of the lattice; see Fig. 4.14 for a comparison at \(\lambda = 750\) nm. To compare the anisotropy of multipolar response at different wavelengths I quantify it by the value of the misalignment angle at which the multipolar SLR merges with the Rayleigh anomaly and vanishes; I denote this angle by \(\phi_{an}\). At long wavelengths the multipolar response is extremely anisotropic, \(\phi_{an} \approx 0.1^\circ\) at \(\lambda \approx 750\) nm. At shorter wavelengths the anisotropy is slightly less pronounced, \(\phi_{an} \approx 5^\circ\) at \(\lambda \approx 650\) nm. Nevertheless,
the multipolar response of the array is strongly anisotropic over the whole range of wavelengths that I consider. This strong anisotropy is a result of sensitivity of the multipolar response to the changes in the structure of couplings between multipole moments. That structure changes qualitatively when the misalignment angle is varied from $\phi = 0$ to a small finite value, with the multipole moments decoupled into two independent sets at $\phi = 0$, but all the excited multipoles being coupled together for incidence direction misaligned with the lattice symmetry direction. Yet the degree of anisotropy that I observe is surprising, as I might expect the additional couplings that arise at small values of $\phi$ to be weak.

Finally I note that at $\phi = 30^0$ and equivalent angles (including along $\hat{\kappa}_b$) the multipolar SLR is absent; see the reflection spectrum at $\phi = 30^0$ in Fig. 4.14. At $\phi = 0$ and $\phi = 30^0$ the multipole moments are described by the same set of equations (4.33-4.34) and hence the structure of the couplings is the same. However, for the the wave vector oriented at an angle $\phi = 30^0$, the periodic Green functions that are associated with the higher multipole moments ($m_\kappa, q_\kappa z$) do not diverge at the onset of a new diffraction order, and the multipolar SLRs do not appear. This is associated with the transverse nature of the radiated electromagnetic field [78], and is easiest seen for the radiation from $m_\kappa$. I shall see in the next section that the diffracted field that arises first for incidence along $\hat{\kappa}_b$ (or an equivalent direction) is characterized by an in-plane wave vector opposite that of the incident field, and at diffraction onset the diffracted amplitude must vanish then since a dipole cannot radiate in the direction in which it points.

I have shown that multipolar response is very sensitive to even small changes in the structure of the couplings between multipoles that are driven by changes in either the polarization or in the direction of incident light. This sensitivity can be understood by analyzing SLMs of the array. Consider first the SLMs excited with s-polarized light incident along $\hat{\kappa}_a$; see Fig. 4.10 in section 4.4. The multipolar SLM is in a close proximity to the light line. Thus even a small change in the dispersion relation due to a misalignment of $\kappa_0$ with the symmetry direction $\hat{\kappa}_a$ results in the merging of the SLM with the light line and a disappearance of the SLR associated with it. The dipolar dispersion relation is further away from the light line, and a change in the dipolar dispersion relation results in a shift of the SLM rather than its termination, with the associated SLR of dipolar character existing over a wide range of illumination conditions. A similar reasoning explains the suppression of the multipolar response for p-polarized light when the coupling of the higher multipoles to non-resonant electric dipole is taken into account. Were the coupling to the electric dipole component $p_z$ neglected, a SLM associated with the set ($p_z, m_\kappa, q_\kappa z$) would be found over a wide range of wavelengths, at wave vectors in a close proximity to the light line;
when the coupling to \( p_z \) is taken into account the dispersion relation merges with the light line and the SLM is not found.

### 4.8 Direct coupling to surface-lattice modes of the array

In the previous sections I have discussed a diffractive coupling to SLMs of the array that is signalled by the appearance of SLRs in reflection. Yet the physical significance attributed to the SLMs leads immediately to the suggestion that it should be possible to construct a configuration where direct coupling to them would be possible. I do that here, and investigate the signatures of direct coupling on an angular scan of the specular reflectance.

Again the analogy with a thin metal film is a useful one. Suppose I have a metal film embedded in a background medium with index of refraction \( n \) above a substrate with a higher index of refraction, \( n_s > n \); see Fig. 4.15, but with the array of spheres imagined replaced by a metal film. When the light

![Figure 4.15: The array at a distance \( l \) above an interface. The interface is between two semi-infinite homogenous media, with the background having the index of refraction \( n = 1.45 \), and the substrate having the index of refraction \( n_s = 1.75 \). The array is illuminated by a plane wave incident from the substrate.](image)

is incident from the substrate above the critical angle, \( \tilde{\omega}n_s > \kappa_{\text{inc}} > \tilde{\omega}n \), where \( \kappa_{\text{inc}} \) is the in-plane component of the wave vector of light, a long-range surface plasmon [101] (LRSP) could be excited by an electromagnetic field that is evanescent in the background medium. Such an excitation would lead to absorption loss, and so a dip in the reflectivity would be seen at the angle where the coupling to the LRSP is optimal. Of course, the \( \kappa_{\text{inc}} \) where the dip would occur would not be determined exactly by the dispersion relation of the LRSP in the absence of the substrate – the presence of the substrate would modify the dispersion relation – but it would be close if the distance between the substrate and the metal film were large enough that the coupling were weak.

With this analogy in mind, I consider a 2D array of spheres a distance \( l \) above an interface at \( z = 0 \) separating two semi-infinite homogeneous media, where the index of refraction for the background medium at \( z > 0 \) is \( n = 1.45 \), and for the substrate at \( z < 0 \) it is \( n_s = 1.75 \). Suppose that when free-standing in the background medium the array exhibits an SLR at \( \kappa_{\text{SLR}} \), which signals coupling into a SLM at \( \kappa_{\text{SLM}} = \kappa_{\text{SLR}} + K \), where \( K \) is one of the reciprocal lattice vectors, and of course \( |\kappa_{\text{SLR}}| < \tilde{\omega}n \). In the presence of the substrate the SLM is shifted from its position \( \kappa_{\text{SLM}} \) for a free-standing array, but if that shift is small a field incident from the substrate with an in-plane wave vector \( \kappa_{\text{inc}} \) could certainly...
couple diffractively into the SLM when $\kappa_{\text{inc}} \approx \kappa_{\text{SLR}}$. But in addition to the diffractive coupling, a direct coupling to the SLM should also be possible at an in-plane wave vector $\kappa_{\text{inc}} \approx \kappa_{\text{SLM}}$ that corresponds to incidence at an angle above the critical angle, $|\kappa_{\text{inc}}| > \tilde{\omega} n$.

For simplicity and to illustrate the ideas more clearly I consider s-polarized light incident from the substrate with a wave vector component $\kappa_{\text{inc}}$ along the second symmetry direction $\hat{k}_b$ (see Figure 4.2). I consider a plane wave incident from the substrate,

$$E^{\text{inc}}(r) = \hat{s} E^{\text{inc}} e^{i \nu_{s+} r},$$

where $\kappa_{\text{inc}} = \tilde{\omega} n_s \hat{k}_b \sin \theta_s$, $\hat{s} = \hat{k}_{\text{inc}} \times \hat{z}$, and $\nu_{s \pm} = \kappa_{\text{inc}} \pm w_s(\kappa_{\text{inc}}) \hat{z}$ with $w_s(\kappa_{\text{inc}}) = \sqrt{\omega^2 n_s^2 - \kappa_{\text{inc}}^2}$. In the absence of the array of spheres, the fields that would be reflected and transmitted are given by

$$E^R(r) = r_0^s E^{\text{inc}} \hat{s} e^{i \nu_{s+} r},$$

$$E^T(r) = t_0^s E^{\text{inc}} \hat{s} e^{i \nu_{s-} r},$$

where $r_0^s$ and $t_0^s$ are the usual Fresnel coefficients for s-polarized light,

$$r_0^s = \frac{w_s(\kappa_{\text{inc}}) - w_0(\kappa_{\text{inc}})}{w_s(\kappa_{\text{inc}}) + w_0(\kappa_{\text{inc}})},$$

$$t_0^s = \frac{2w_s(\kappa_{\text{inc}})}{w_s(\kappa_{\text{inc}}) + w_0(\kappa_{\text{inc}})},$$

with $w_s(\kappa_{\text{inc}})$ as above, and $w_0(\kappa_{\text{inc}}) = \sqrt{\omega^2 n_s^2 - \kappa_{\text{inc}}^2}$.

I now include the presence of the array of spheres shown in Fig. 4.15. For this excitation scenario the optical response of the spheres is well described within the electric dipole approximation over a wide range of wavelengths, since as discussed in section 4.7 the narrow SLR associated with higher order multipole moments is not found. I thus model the spheres as electric dipoles and formulate the response of the array in terms of the dipole $p$ at the origin of the lattice, $R = 0, z = l$. The dipole is driven by an electric field that can be written as the sum of three terms: (a) the transmitted field (4.72) at the position of the dipole that would arise were the array not present, $E^T = E^T(R = 0, l)$, (b) a sum of the radiation reaction field and the field scattered from all the other spheres, $\hat{G}^{\text{Ep}} \cdot p$, where $\hat{G}^{\text{Ep}}$ is the periodic Green function introduced earlier, (c) the field that is radiated by the array and reflected back from the substrate, $\hat{G}_s^{\text{Ep}} \cdot p$, where $\hat{G}_s^{\text{Ep}}$ is the substrate correction to the dipole Green function [72],

$$\hat{G}_s^{\text{Ep}} = \frac{1}{A_c} \sum_n \frac{i \tilde{\omega}^2 e^{2 \nu_0 (|\kappa_{\text{inc}} + K_n|) l}}{2 \nu_0 w_0 (|\kappa_{\text{inc}} + K_n|)} A_n,$$

where I have defined a dyadic

$$A_n = r_n^s \hat{s}_n \hat{s}_n + t_n^p \hat{p}_n \hat{p}_n,$$

with

$$\hat{s}_n = |\kappa_{\text{inc}} + K_n|^{-1} (\kappa_{\text{inc}} + K_n) \times \hat{z},$$

$$\hat{p}_n = (\tilde{\omega} n)^{-1} |\kappa_{\text{inc}} + K_n| \hat{z} \mp (\tilde{\omega} n)^{-1} w_0 (|\kappa_{\text{inc}} + K_n|) \frac{(\kappa_{\text{inc}} + K_n) \times \hat{z}}{|\kappa_{\text{inc}} + K_n|^{-1}}.$$
Here \( r_s^n \) is the s-polarized reflection coefficient (see eq. (4.73)) with \( \kappa_{inc} \) replaced by \( |\kappa_{inc} + K_n| \), and \( r_p^n \) is the corresponding p-polarized Fresnel reflection coefficient,

\[
r_p^n = \frac{n^2 w_s (|\kappa_{inc} + K_n|) - n_s^2 w_0 (|\kappa_{inc} + K_n|)}{n^2 w_s (|\kappa_{inc} + K_n|) + n_s^2 w_0 (|\kappa_{inc} + K_n|)}.
\]

Adding all the contributions I arrive at the total driving field,

\[
E^{tot} = E^T + \left[ \hat{G}^E_p + \hat{G}_s^E \right] \cdot p,
\]

in terms of which the electric dipole is given by

\[
p = \alpha^PE^{tot}.
\]

I now calculate an angular scan of the specular reflectance of the structure at \( \lambda = 900 \) nm. For \( \kappa_{inc} \) in the direction of \( \hat{\kappa}_b \) the first reciprocal lattice vector of importance as \( |\kappa_{inc}| \) increases from zero is \( K_{(01)} \); see Fig. 4.16. I choose the array to be at a distance \( l = l_0 \) from the interface, where

\[
l_0 = (\kappa_{SLM}^2 - \omega^2 n_s^2)^{-1/2} = 337.2 \text{ nm} \text{ with } \kappa_{SLM} = \kappa_{SLR} + K_{(01)} \text{ being the wave vector of the SLM supported by the free-standing array. For this choice of the distance, the evanescent fields above the substrate at angles of incidence beyond the critical angle excite clear resonances in the array, yet the array is far enough from the interface so that its mode structure is not significantly modified.}

The specularly reflected light exhibits two resonant features of a distinct character as the angle of incidence in the substrate, \( \theta_s \), departs from zero. As \( \theta_s \) becomes increasingly positive, there is an SLR-like peak at an angle near that which corresponds to the SLR resonance of the free-standing array, \( \sin \theta_{SLR}^s = \kappa_{SLR}/\omega n_s \); I indicate this angle as a dotted line in Fig. 4.17, and it corresponds to diffractive coupling to the SLM, at an angle close to what is seen in the specular reflection of a free-standing array. From Fig. 4.16 I expect the SLM associated with this SLR to be at \( \kappa_{SLM} = \kappa_{SLR} + K_{(01)} \), which is accessible to light incident from the prism at a negative angle \( \theta_{SLM}^s \), \( \sin \theta_{SLM}^s = -\kappa_{SLM}/\omega n_s \). I indicate this angle as a dashed line in Fig. 4.17, and near this there is a dip in the specular reflectance analogous to the dip one sees in surface plasmon studies, for example, where the excitation of the surface mode leads...
to a large absorption and a decrease in the reflectance; this corresponds to direct coupling to the SLM. The third peak that appears in the angular range shown in Fig. 4.17 is near $-\theta_{SLR}^s$, and corresponds to the SLR that exists from symmetry at $\kappa_{SLR}$; associated with it is a dip at $\theta_{SLM}^s$, not in the angular range shown, which corresponds to direct coupling to the SLM at $-\kappa_{SLM}$.

The agreement between the resonant structures at angles near $\theta_{SLR}^s$ and $\theta_{SLM}^s$ and those angles themselves is not exact. This is not surprising, as the interaction between the array and the substrate modifies the effective polarizability of the array,

$$ \left( \leftrightarrow \alpha^{eff} \right)^{-1} = \left( \leftrightarrow \alpha^{pE} \right)^{-1} - \leftrightarrow G_{Ep} - \leftrightarrow G_{Ep}^s, $$

(4.82)

cf. eqs. (4.80,4.81), and thus it modifies the dispersion relation of the SLM as well. But as $l$ increases the contributions to $\leftrightarrow G_s$ due to in-plane wave vectors associated with evanescent fields in the medium becomes negligible, and I can expect better agreement of the positions of the resonant structures with $\theta_{SLR}^s$ and $\theta_{SLM}^s$. In the inset of Fig. 4.17 I show the deviation of the in-plane wave vectors associated with the resonant structures from the respective wave vectors $\kappa_{SLR}$ and $\kappa_{SLM}$, which I identify as $\Delta \kappa_{SLR}$ and $\Delta \kappa_{SLM}$ respectively. As $l$ increases I see that $\Delta \kappa_{SLR}$ indeed approaches zero, while $\Delta \kappa_{SLM}$ oscillates around a finite value. The oscillations are associated with the coupling of the evanescent fields exciting the array to fields propagating away from the array by diffraction; these fields can reflect off the interface, propagate back to the layer, and interact with it again, leading to interference effects. The finite value is a shift in the position of the dip arising because the angle at which the minimum dip occurs depends not only on the position of the SLM, but also on how it is driven. For a given field incident from the substrate, the evanescent field at array will be larger for smaller $|\theta_s|$, closer to the cut-off for total internal reflection. This results in a shift of the position of the dip to smaller $|\theta_s|$. Nonetheless, it is clear that the dip can be associated with the direct excitation of the SLM.
4.9 Conclusions

In conclusion, I have analyzed the multipolar structure of the reflection of radiation from an array of gold spheres of moderate size, and its dependence on frequency and illumination conditions. I included the electric dipole, magnetic dipole, and electric quadrupole moments of the spheres; full numerical calculations confirm that these are sufficient to capture the electrodynamics for the range of parameters I have considered. I have shown the importance of including the response of moments that would be non-resonant and negligible in magnitude were only an isolated sphere considered; such moments are sometimes neglected in treatments of arrays, and at first thought this might seem reasonable. However, the reflection of an array is dominated by the so-called surface lattice resonances (SLRs), which appear at angles close to those where new diffraction orders arise, and it is at these angles that the electromagnetic coupling between different moments in spheres in an array can diverge. Due to this coupling, moments that would be unimportant in the response of an isolated sphere – in the example considered here the magnetic dipole moment – can have important consequences on the angular and frequency dependence of the reflectivity of an array of spheres.

First, the coupling between the moments can suppress the SLRs associated primarily with moments higher than the electric dipole. In the lattice I considered, such an SLR only survives for s-polarized light incident in one symmetry direction, where it is due to a dominant quadrupole moment coupled only to the magnetic dipole moment. The introduction of additional couplings between the moments, due to a change in the polarization or direction of the incident field, results in a vanishing of this SLR. Second, the profile, position, and the cut-off frequency at which the SLRs vanish is affected by the coupling to non-resonant moments. This is most significant for the SLRs that do not involve the electric dipole moment in a significant way, for due to the weak coupling with the incident light these SLRs are close to the angles at which new diffraction orders arise. But even for the other SLRs the coupling effects are non-negligible.

When the interaction between the different multipoles is important the analysis of SLRs becomes complicated, but it can be simplified by identifying the “ideal” surface-lattice modes (SLMs) of the array. They should be contrasted with the waveguide-like excitations that exist with the neglect of absorption at wavelengths too large for diffraction to occur, and with the bound states in the continuum that do not radiate because of symmetry or interference between radiative channels [103]. Instead, the ideal surface-lattice modes I consider arise at wavelengths where diffraction occurs, and are identified by neglecting both absorption losses and diffractive losses. The dispersion relations of these ideal SLMs gives an extremely good prediction of where the SLRs should appear, even when absorption and diffraction are included; the physical picture is that diffraction allows coupling into them from the incident beam, and the coupling back leads to the SLR observed in reflection. I have shown that the dispersion relations can be understood within simplified models based on isolating the divergent and leading analytic contributions to the coupling between the moments, and that an understanding of the properties of the associated SLRs then follows. In particular, I identified differences in the effects that drive the termination of different ideal SLM dispersion relations, and have linked them to the difference in the sensitivity of the different SLRs to the couplings between the multipole moments and the illumination conditions.

Finally, I have shown that it should be possible to observe these ideal SLMs not only by coupling into them through diffraction, where they lead to the SLRs, but as well by directly coupling into them through a high-index prism. This is analogous to the coupling into surface plasmons or long-range surface
plasmons in Kretschmann or Otto configurations, and should be a new way to study these SLMs.

While I have considered only one lattice and one size of spheres, I have looked at the reflectivity as both a function of angle and wavelength over the range of parameters where the onset of diffraction is important, and where the isolated spheres would exhibit some resonant response. I have examined in detail the response for light incident along two high symmetry directions, one in which the electrodynamics is essentially one-dimensional in its important features, and one in which it is fully two-dimensional, and as well looked at how the situation changes when one moves away from these high symmetry directions. Thus the methods and concepts that I have applied here should be useful more generally to other systems as well.
Chapter 5

Gauge-invariant Green function dynamics


5.1 Introduction

In the previous chapters I treated the optical response of nanoparticle arrays within a multipolar model, with the multipole expansion of charges and currents carried out independently for each nanoparticle. Such an independent treatment of charges in nanoparticles was possible because the nanoparticles were spatially separated and there was no flow of electrons from one nanoparticle to the next. Therefore all the charges could be unambiguously identified with one of the lattice sites, and the multipole expansion for each charge was carried out with respect to that site. However, such an approach is not applicable in a straightforward way to a description of an array where nanoparticles are either electrically connected or they are close enough so that the motion of electrons between nanoparticles is possible. There, a more general description that takes into account the motion of electrons from one lattice sites to another would be necessary. A related problem can be seen to arise when treating the optical response of a crystal. Electrons in a crystal are delocalized and they cannot be unambiguously associated with one unit cell. Thus the methods developed for isolated molecules cannot be immediately applied to treat the optical response of a solid. Because of these issues, a rigorous theory of polarization and magnetization in a crystal was developed only relatively recently, and it was limited to uniform or nearly uniform electromagnetic fields [46, 47].

A natural requirement on any theory of the optical response of a crystal would be to demand that in an extreme limit of a molecular crystal the theory reduces to the multipolar treatment of individual molecules. In the previous chapters, the multipolar description was introduced using effective classical polarizabilities and susceptibility tensors. At the fundamental level, however, these should follow from a quantum-mechanical treatment of charges and currents in the material medium. A rigorous quantum-mechanical treatment of atoms and molecules was developed by Power, Zienau, and Wooley (PZW) [43], and elaborated and clarified by Healy [44]. While at a fundamental level the interaction of the electromagnetic field with charges is described by the minimal coupling Hamiltonian, in the PZW
transformation the interaction Hamiltonian is converted to one involving polarization and magnetization fields interacting directly with the electric and magnetic fields. The description is thus explicitly gauge invariant. Often one then expands the polarization and magnetization fields in terms of electric and magnetic multipole moments, yielding an approximate Hamiltonian that takes into account the variations of the electric and magnetic fields over the atom or molecule to the order required. Even before such an expansion is undertaken, a “special point” must be chosen to effect the PZW transformation. For an atom this special point is often taken to be the position of the nucleus, idealized as fixed; more generally it can be taken to be the position of the center of mass of an atom or molecule [104]. All the expansions within the PZW theory are done with respect to that special point.

Another strategy for constructing gauge-invariant quantum dynamical equations is the approach of Levanda and Fleurov (LF) [54, 55]. They showed that by concentrating on the equations of motion for the Green functions, rather than modifying the Hamiltonian formulation of those equations, a gauge-invariant description of the dynamics is possible. Their approach employs a generalized Peierls phase that involves an integration along a straight line in space-time between the two Green function variables; see also [105] and [106]. The LF approach is attractive because it does not rely on a special point, as does the PZW transformation. Thus this kind of approach, or a related one, could provide a basis for gauge-invariant descriptions of the interaction of radiation with charges in extended systems, such as solids, where the use of a special point does not seem to make physical sense. Gauge-invariant descriptions would be useful because the straightforward application of the minimal coupling Hamiltonian often requires the identification of sum rules to eliminate apparent divergences, even in problems as simple as the linear and nonlinear optical response of clean, cold semiconductors, with electrons treated in the independent particle approximation [107, 108, 109, 110].

These apparent divergences are associated with the difference between the mechanical momenta of the charges and their canonical momenta, described by a term involving the vector potential, and they seem to be avoided whenever a gauge-invariant description of the quantum dynamics can be constructed. There is a long tradition of work on special cases where this has been implemented. For situations where the response of the medium can be approximated as responding locally to the electric field, taken as uniform, an approach originally pioneered by Adams and Blount [111] can be applied. A quantity that plays the role of the matrix element of a dipole moment operator between Bloch states can be introduced, involving interband and intraband contributions [112]. The approach of the “modern theory of polarization” [113], where polarization and magnetization are associated with the appropriate moments of Wannier functions [48, 49], or extensions necessary in materials with nonzero Chern numbers [114], can be seen to be closely related to this, for these moments can be written in terms of the effective matrix elements of Adams and Blount and their generalizations. But to date these programs are restricted to the assumption of uniform or nearly uniform applied fields.

So an exploration of strategies based on the LF approach to treat the response of solids to more arbitrarily varying electromagnetic fields is promising. A natural first step is to link that approach with the PZW transformation, which is to date probably the most successful gauge-invariant description of the response of charges to the electromagnetic field. Yet when both methods are applied to an atom, the equivalence can be seen only after much manipulation, even when the electromagnetic field varies little over the atom. The first goal of this chapter is to clarify the link between the gauge-invariant descriptions of LF and PZW.

To do this I find it is necessary to extend the LF approach to treat a Peierls phase involving an
arbitrary path in space-time connecting the two Green function variables. This is done in section 5.2, where I construct a generalized LF approach (GLF). Then I recover the LF and PZW results by the choice of particular paths. This is done in sections 5.3 and 5.4.

Establishing those connections illustrates that there are two qualitatively different types of Peierls phase that can arise in a GLF approach. In the first type, which appears in recovering the PZW results, the generalized Peierls phase is given by the difference of the values that a single variable function takes at the two space-time points in the Green function. In the second type, which appears for example in the original LF approach, it cannot be written in that form. In section 5.4 I show that the Peierls phase of the first type leads to the gauge-invariant Green functions that can be understood as constructed from transformed field operators, which satisfy dynamical equations following from an effective Hamiltonian. In the special case of the path that leads to the PZW results, that transformation is precisely the usual PZW transformation, through which the polarization and magnetization fields arise.

The polarization and magnetization fields that arise in the PZW treatment of an atom or small molecule are examples of what can be called ideal polarization and magnetization fields, in that they are nonzero only over distances on the order of an atomic spacing from the charge and current densities that are responsible for them. A second goal of this chapter is to establish that, for systems where electrons cannot be considered as being bound to identifiable units such as atoms or molecules, it is in general impossible to construct effective Hamiltonians involving such ideal polarization and magnetization fields. This is done in section 5.5. I conclude that section by arguing that for such extended systems a more promising strategy is an approach based on a Peierls phase of the second type, at least in problems where the correlation lengths are short compared to the distance over which the electromagnetic field varies. I summarize the results in section 5.6.

5.2 General formalism

5.2.1 Generalized LF

In this section I extend the approach of Levanda and Fleurov [54] by generalizing the Peierls phase used in the transformation leading to gauge-invariant Green functions. I refer to the Green function formalism that results as “generalized Levanda-Fleurov” (GLF).

I begin by introducing an integral of the four-potential over a general path linking two space-time points \(y\) and \(x\),

\[
\Gamma(x, y) = \int_y^x \, dz^\nu A_\nu(z),
\]

(5.1)

I use the relativistic notation with the metric \(\eta = (-1, 1, 1, 1)\), contravariant components of four-potential \(A^\nu = (A^0, \mathbf{A})\), and an event vector \(x^\nu = (x^0, \mathbf{x})\). Equation (5.1) should be taken as a short-hand for a line integral in space-time over a path that is parametrized by a real variable \(u\): I write \(z^\nu = z^\nu(u; x, y)\) and as \(u\) varies from an initial value \(u_{\text{initial}}\) to a final value \(u_{\text{final}}\), for each arbitrary but fixed \(x\) and \(y\), the four-vector \(z\) varies from \(y\) to \(x\). That is, \(z^\nu(u_{\text{initial}}; x, y) = y^\nu\) and \(z^\nu(u_{\text{final}}; x, y) = x^\nu\) (see Figure 5.1). Thus (5.1) is taken to represent

\[
\Gamma(x, y) = \int_{u_{\text{initial}}}^{u_{\text{final}}} \, \frac{\partial z^\nu(u; x, y)}{\partial u} A_\nu(z(u; x, y)) du,
\]

(5.2)
Figure 5.1: The dashed line is the path in space-time used by Levanda and Fleurov; the solid line indicates a generalization

where I will often simply write $\partial z^\nu / \partial u$ for $\partial z^\nu (u; x, y) / \partial u$. Since a path $z^\nu (u; x, y)$ ($u_{\text{initial}} \leq u \leq u_{\text{final}}$) is assumed to be defined for each $y$ and $x$, for fixed $u$ between $u_{\text{initial}}$ and $u_{\text{final}}$ I can consider how the four-vector $z^\nu$ depends on $y$ and $x$. That is, in addition to $\partial z^\nu (u; x, y) / \partial u$, partial derivatives $\partial z^\nu (u; x, y) / \partial x^\mu$ and $\partial z^\nu (u; x, y) / \partial y^\mu$ can also be introduced; they are restricted only by the initial and final values of $z^\nu$, yielding the constraints $\partial z^\nu (u_{\text{initial}}; x, y) / \partial x^\mu = 0$, $\partial z^\nu (u_{\text{initial}}; x, y) / \partial y^\mu = \delta^\nu_\mu$, $\partial z^\nu (u_{\text{final}}; x, y) / \partial x^\mu = \delta^\nu_\mu$, and $\partial z^\nu (u_{\text{final}}; x, y) / \partial y^\mu = 0$. The functional dependence $z^\nu (u; x, y)$ on all its variables $(u, x^\nu, y^\nu)$ is assumed to be sufficiently smooth that mixed partial derivatives with values independent of the order of differentiation exist. On physical grounds I assume that the four-potential is differentiable, and in addition I assume that the path is chosen in such a way that the resulting $\Gamma (x, y)$ is differentiable with respect to both $x$ and $y$. Note that in general $\lim_{x \to y} \Gamma (x, y)$ need not vanish, since the specification of the path might involve a trajectory through space-time that starts at $y$ and finishes at $y$. Nonetheless, for such a closed path in space-time the resulting $\Gamma (y, y)$ can be written as a flux of the electromagnetic field tensor through the hypersurface associated with the path, so while $\Gamma (y, y)$ need not vanish it is guaranteed to be gauge-invariant [115].

The derivatives of $\Gamma (x, y)$ with respect to $x$ and $y$ will play a central role. From (5.2) I find

$$\frac{\partial \Gamma (x, y)}{\partial x^\mu} = I_\mu (x, y) + J_\mu (x, y), \quad (5.3)$$

where

$$I_\mu (x, y) = \left[ \int_{u_{\text{initial}}}^{u_{\text{final}}} \frac{\partial}{\partial u} \left( \frac{\partial z^\nu}{\partial x^\mu} \right) \right] A_\nu (z) du, \quad (5.4)$$

$$J_\mu (x, y) = \int_{u_{\text{initial}}}^{u_{\text{final}}} \left( \frac{\partial z^\nu}{\partial u} \right) \frac{\partial A_\nu (z)}{\partial x^\mu} du. \quad (5.5)$$

Integrating the first term by parts yields

$$I_\mu (x, y) = \left[ \int_{u_{\text{initial}}}^{u_{\text{final}}} \frac{\partial}{\partial u} \left( \frac{\partial z^\nu}{\partial x^\mu} \right) \right] A_\nu (z) du$$

$$= \left[ \frac{\partial z^\nu}{\partial x^\mu} A_\nu (z) \right]_{u_{\text{initial}}}^{u_{\text{final}}} - \int_{u_{\text{initial}}}^{u_{\text{final}}} \frac{\partial z^\nu}{\partial x^\mu} \frac{\partial A_\nu (z)}{\partial u} du$$

$$A_\mu (x) - \int_{u_{\text{initial}}}^{u_{\text{final}}} \frac{\partial z^\alpha}{\partial u} \frac{\partial z^\nu}{\partial x^\mu} \frac{\partial A_\nu (z)}{\partial z^\alpha} du, \quad (5.6)$$
which after interchanging the dummy indices can be written as

\[ I_\mu(x, y) = A_\mu(x) - \int_y^x dz_\nu \frac{\partial z^\alpha}{\partial x^\mu} \frac{\partial A_\alpha(z)}{\partial z^\nu}. \]  

(5.7)

In the term \( J_\mu(x, y) \), \( A_\nu(z) \) depends on \( x \) because the path depends on \( x \), so I can write

\[ J_\mu(x, y) = \int_{u_{\text{initial}}}^{u_{\text{final}}} \partial z^\mu \frac{\partial A_\nu(z)}{\partial z^\alpha} \frac{\partial z^\alpha}{\partial x^\mu} du = \int_y^x \frac{\partial z^\alpha}{\partial x^\mu} \frac{\partial A_\nu(z)}{\partial z^\alpha} dz^\nu. \]  

(5.8)

Combining (5.7) and (5.8) in (5.3) leads to

\[ \frac{\partial \Gamma(x, y)}{\partial x^\mu} = A_\mu(x) + X_\mu(x, y), \]  

(5.9)

where I have introduced the notation

\[ X_\mu(x, y) = \int_y^x \frac{\partial z^\alpha}{\partial x^\mu} F_{\alpha\nu}(z) dz^\nu. \]  

(5.10)

for the contribution to (5.9) arising because the differentiation of (5.1) with respect to \( x \) involves not only the vector potential at the endpoint \( x \), but also the variation of the path in space-time between \( x \) and \( y \). Here \( F_{\mu\nu} \) is electromagnetic field tensor, \( F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \). In terms of fields, the components are given by \( F_{0\nu}(x) = E_\nu(x) \) and \( F_{ij}(x) = \epsilon_{ijk} B_k(x) \), where \( \epsilon_{ijk} \) is the Levi-Civita symbol. In a similar way the derivative with respect to \( y \) is found to be

\[ \frac{\partial \Gamma(x, y)}{\partial y^\mu} = -A_\mu(y) + Y_\mu(x, y), \]  

(5.11)

where

\[ Y_\mu(x, y) = \int_y^x \frac{\partial z^\alpha}{\partial y^\mu} F_{\alpha\nu}(z) dz^\nu. \]  

(5.12)

I note that \( X(x, y) \) and \( Y(x, y) \) are not independent. From the equality of mixed partial derivatives, using (5.9) and (5.11) I find that

\[ \frac{\partial X_\mu(x, y)}{\partial y^\nu} = \frac{\partial Y_\nu(x, y)}{\partial x^\mu}. \]  

(5.13)

Equations (5.9, 5.11) show that the derivatives of \( \Gamma(x, y) \) contain two components, one of which involves the vector potential, and the other of which is gauge-invariant, depending only on the electromagnetic field tensor. To formally display the dependence on the electromagnetic field through an integral over all space-time rather than along a path, I introduce two tensors \( K^\alpha\nu_\mu(w; x, y) \) and \( L^\alpha\nu_\mu(w; x, y) \) that are antisymmetric in their upper indices,

\[ K^\alpha\nu_\mu(w; x, y) = \int_y^x \left( dz^\alpha \frac{\partial z^\alpha}{\partial x^\mu} - dz^\alpha \frac{\partial z^\nu}{\partial x^\mu} \right) \delta(w - z), \]  

(5.14)

\[ L^\alpha\nu_\mu(w; x, y) = \int_y^x \left( dz^\alpha \frac{\partial z^\alpha}{\partial y^\nu} - dz^\alpha \frac{\partial z^\nu}{\partial y^\nu} \right) \delta(w - z). \]  

(5.15)
With the help of (5.14) and (5.15), $X(x,y)$ and $Y(x,y)$ can be written in an equivalent form,

$$
X_{\mu}(x,y) = \frac{1}{2} \int d^4w K_{\mu}^{\alpha\nu}(w; x, y) F_{\alpha\nu}(w),
$$

(5.16)

$$
Y_{\mu}(x,y) = \frac{1}{2} \int d^4w L_{\mu}^{\alpha\nu}(w; x, y) F_{\alpha\nu}(w),
$$

(5.17)

where here $d^4w \equiv dw^0 dw^1 dw^2 dw^3$.

### 5.2.2 GLF Green function dynamics

I now use the generalized Peierls phase introduced in section 5.2.1 to derive gauge-invariant equations for the Green function dynamics. I consider a system of electrons interacting with time-dependent, classical electromagnetic fields and with an additional static potential energy $V(x)$ due to a fixed nucleus or lattice. Since my focus in this chapter is on describing the interaction of the electrons with the electromagnetic field, I adopt the independent-particle approximation for simplicity and neglect the spin degree of freedom. Then the Hamiltonian is

$$
H(t) = \frac{1}{2m} \int d^3x \psi^\dagger(x,t) \left( \frac{\hbar}{i} \nabla - \frac{e}{c} A(x,t) \right)^2 \psi(x,t) + \int d^3x (V(x) + e\phi(x,t)) \psi^\dagger(x,t)\psi(x,t),
$$

(5.18)

where $-A_0(x,t) = \phi(x,t)$ and $A(x,t)$ are a scalar and vector potential corresponding to the external electromagnetic fields. I write the potentials as depending on $(x,t)$ (sometimes $(x,t)$ when there is no ambiguity), or on $x$, depending on whether I use non-relativistic or relativistic notation. Although I first develop equations for the dynamics of time-ordered Green functions only, the derivation of equations for the others (i.e. anti-time ordered, greater and lesser functions) can be done in an analogous manner.

I recall the usual equations for the time-ordered Green function $G^t(x,y) = -i \langle T [\psi(x), \psi^\dagger(y)] \rangle$, which follow from the Heisenberg equations of motion for the field operators using (5.18). The time derivatives are given by

$$
-c \left( \frac{\hbar}{i} \frac{\partial}{\partial x^0} - \frac{e}{c} A_0(x) \right) G^t(x,y) = \left[ \frac{1}{2m} \left( \frac{\hbar}{i} \frac{\partial}{\partial x^j} - \frac{e}{c} A_j(x) \right)^2 + V(x) \right] G^t(x,y)
+ \hbar \delta(x-y),
$$

(5.19)

$$
+ c \left( \frac{\hbar}{i} \frac{\partial}{\partial y^0} + \frac{e}{c} A_0(y) \right) G^t(x,y) = \left[ \frac{1}{2m} \left( \frac{\hbar}{i} \frac{\partial}{\partial y^j} + \frac{e}{c} A_j(y) \right)^2 + V(y) \right] G^t(x,y)
+ \hbar \delta(x-y).
$$

(5.20)

Since the function $G^t(x,y)$ is not gauge-invariant, equations (5.19,5.20) explicitly depend on the electromagnetic potentials as expected. Among their other uses, the Green functions are central in determining the expectation values of single particle operators. Here I look at the electron charge and current den-
Green functions. I define obviously depend on the gauge chosen. The equalities follow immediately, where I am using the notation \( j^\mu(x) = (c\rho(x,t), j(x,t)) \). Here I denote \( x_+ = (ct_+, x_+) \) as a space-time point different than \( x = (ct, x) \), but after taking derivatives I put \( x_+ \rightarrow x \) and \( t_+ \) to be infinitesimally greater than \( t \). Although the results for \( j^\mu(x) \) are gauge-independent, the individual terms in (5.22) obviously depend on the gauge chosen.

Next I employ the generalized Peierls phase to formulate the dynamics in terms of gauge-invariant Green functions. I define

\[
\bar{G}^\mu(x, y) = e^{-i\Phi(x, y)} G^\mu(x, y),
\]

where

\[
\Phi(x, y) = \frac{e}{\hbar c} \left( \Gamma(x, y) - \frac{1}{2} \Gamma(x, x) - \frac{1}{2} \Gamma(y, y) \right).
\]

The equalities

\[
e^{-i\Phi(x, y)} \left( \frac{\hbar}{i} \frac{\partial}{\partial x^j} - \frac{e}{c} A_j(x) \right) e^{i\Phi(x, y)} = \frac{\hbar}{i} \frac{\partial}{\partial x^j} + \frac{e}{c} \bar{X}_j(x, y)
\]

and

\[
e^{-i\Phi(x, y)} \left( \frac{\hbar}{i} \frac{\partial}{\partial y^j} + \frac{e}{c} A_j(y) \right) e^{i\Phi(x, y)} = \frac{\hbar}{i} \frac{\partial}{\partial y^j} + \frac{e}{c} \bar{Y}_j(x, y)
\]

follow immediately, where

\[
\bar{X}_\mu(x, y) \equiv X_\mu(x, y) - \frac{1}{2} X_\mu(x, x) - \frac{1}{2} Y_\mu(x, x),
\]

\[
\bar{Y}_\mu(x, y) \equiv Y_\mu(x, y) - \frac{1}{2} X_\mu(y, y) - \frac{1}{2} Y_\mu(y, y),
\]

and I have used (5.9) and (5.11) to identify

\[
\frac{\partial \Gamma(z, z)}{\partial z^\mu} = X_\mu(z, z) + Y_\mu(z, z)
\]

for arbitrary \( z \). From (5.25, 5.26), and using the zeroth component of (5.9) and (5.11), I find the time derivatives of \( \bar{G}^\mu(x, y) \) to be

\[
-c \left( \frac{\hbar}{i} \frac{\partial}{\partial x^0} + \frac{e}{c} \bar{X}_0(x, y) \right) \bar{G}^\mu(x, y) = \left[ \frac{1}{2m} \left( \frac{\hbar}{i} \frac{\partial}{\partial x^0} + \frac{e}{c} \bar{X}_0(x, y) \right)^2 + V(x) \right] \bar{G}^\mu(x, y)
\]

\[
+ \hbar \delta(x - y),
\]

(5.29)

\[
c \left( \frac{\hbar}{i} \frac{\partial}{\partial y^0} + \frac{e}{c} \bar{Y}_0(x, y) \right) \bar{G}^\mu(x, y) = \left[ \frac{1}{2m} \left( \frac{\hbar}{i} \frac{\partial}{\partial y^0} + \frac{e}{c} \bar{Y}_0(x, y) \right)^2 + V(y) \right] \bar{G}^\mu(x, y)
\]

\[
+ \hbar \delta(x - y).
\]

(5.30)

This is the central result of this section. Equations (5.29, 5.30) depend on electromagnetic fields instead of potentials that entered equations (5.19,5.20) for the usual Green function \( G^\mu(x, y) \), and thus \( \bar{G}^\mu(x, y) \)
is a gauge-invariant quantity. However, it is clear that any (well-behaved) prescription for specifying the path \( z(u; x, y) \) can be employed. The gauge dependence of the usual Green function equations (5.19,5.20) is thus replaced by a “path dependence” in the construction of the Peierls phase, from which follows a path dependence of the functions \( X(x, y) \) and \( Y(x, y) \) and a path dependence of the new Green function equations (5.29,5.30), as well as of the Green function \( \hat{G}^t(x, y) \) itself. Expressions for the expectation values of the charge and current density in terms of the new Green function \( \hat{G}^t(x, y) \) are manifestly gauge-invariant. From (5.21) and (5.22) I have

\[
\langle j^0(x) \rangle = -iec\hat{G}^t(x, x_+),
\]

\[
\langle j^k(x) \rangle = -\frac{ie}{2m} \left( \hbar \frac{\partial}{\partial x^k} - \frac{\hbar}{i} \frac{\partial}{\partial x^k_+} \right) \hat{G}^t(x, x_+)
\]

\[
-\frac{ie}{2m} \left( \frac{\hbar}{c} X_k(x, x) - \frac{\hbar}{c} Y_k(x, x) \right) \hat{G}^t(x, x_+),
\]

where I have used \( \hat{X}_\mu(x, x) - \hat{Y}_\mu(x, x) = X_\mu(x, x) - Y_\mu(x, x) \) (recall (5.27,5.28)).

For completeness I give the kinetic equations satisfied by the remaining Green functions. I introduce a matrix Green function \( \hat{G}(x, y) \) defined as

\[
\begin{pmatrix}
G^t(x, y) & G^< (x, y) \\
G^>(x, y) & G^{at}(x, y)
\end{pmatrix} =
\begin{pmatrix}
-i \langle T[\psi(x), \psi^\dagger(y)] \rangle & i \langle \psi^\dagger(y) \psi(x) \rangle \\
-i \langle \psi(x) \psi^\dagger(y) \rangle & -i \langle AT[\psi(x), \psi^\dagger(y)] \rangle
\end{pmatrix},
\]

where \( AT \) is the anti-time ordering operator and I put

\[
\hat{G}(x, y) = e^{-i\Phi(x,y)} \hat{G}(x, y),
\]

for the gauge-invariant matrix Green function. Following the derivation of the kinetic equations for the time-ordered Green function I find

\[
-c \left( \frac{\hbar}{i} \frac{\partial}{\partial x^0} + \frac{e}{c} \hat{X}_0(x, y) \right) \tilde{G}(x, y) = \left[ \frac{1}{2m} \left( \frac{\hbar}{i} \frac{\partial}{\partial x^i} + \frac{e}{c} \hat{X}_j(x, y) \right)^2 + V(x) \right] \tilde{G}(x, y)
\]

\[
+ \hbar \delta(x-y) \hat{\sigma}_z,
\]

\[
c \left( \frac{\hbar}{i} \frac{\partial}{\partial y^0} + \frac{e}{c} \hat{Y}_0(x, y) \right) \tilde{G}(x, y) = \left[ \frac{1}{2m} \left( \frac{\hbar}{i} \frac{\partial}{\partial y^i} + \frac{e}{c} \hat{Y}_j(x, y) \right)^2 + V(y) \right] \tilde{G}(x, y)
\]

\[
+ \hbar \delta(x-y) \hat{\sigma}_z,
\]

where \( \hat{\sigma}_z \) is the Pauli matrix.

### 5.3 The link to the LF approach

In this section I outline the connection between the general formalism of section 5.2 and the approach of Levanda and Fleurov (LF). I choose the path in space-time \( z^\nu(u; x, y) \) to be a straight line, and employ the parametrization

\[
z^\nu(u; x, y) = \frac{x^\nu + y^\nu}{2} - u(y^\nu - x^\nu),
\]

(5.36)
with the parameter \( u \) ranging from \( u_{\text{initial}} = -1/2 \) to \( u_{\text{final}} = 1/2 \). The phase factor (5.2) takes the form
\[
\Gamma(x, y) = - \int_{-1/2}^{1/2} (y'' - x'') A_\nu \left( \frac{x + y}{2} + u' (y - x) \right) \, du',
\] (5.37)
where I have changed the integration variable \( u' = -u \). I introduce the center and difference coordinates defined as \( \mathcal{X}^\mu = (x^\mu + y^\mu)/2 \) and \( r^\mu = y^\mu - x^\mu \) respectively, in terms of which (5.37) yields
\[
\Gamma(\mathcal{X}, r) = - \int_{-1/2}^{1/2} r^\nu A_\nu (\mathcal{X} + u' r) \, du',
\] (5.38)
with the corresponding Green function transformation (see eq. (5.33)) given by
\[
\hat{G}(\mathcal{X}, r) = \exp \left[ \frac{ie}{\hbar c} \int_{-1/2}^{1/2} r^\nu A_\nu (\mathcal{X} + u' r) \, du' \right] \hat{G}(\mathcal{X}, r).
\] (5.39)
The Green function (5.39) is gauge-invariant, and for translationally invariant systems considered by Levanda and Fleurov it can be transformed to Wigner coordinates in the usual way, by Fourier transforming with respect to the difference coordinates; then it reduces to the Green function definition given by equation (7) of [54]. In the presence of a periodic potential, a procedure for a Wigner transformation of Green functions was discussed by Kita and Yamashita [116]. However, I emphasize that the Wigner transformation is not necessary to arrive at the gauge-invariant kinetic equations, as is clear from the discussion of section 5.2.

I am interested in the form that the kinetic equations (5.34,5.35) take for a straight-line path in space-time. Written in terms of the center and difference coordinates, (5.34,5.35) are
\[
c \left[ 2 \frac{\hbar}{i} \frac{\partial}{\partial \mathcal{X}^0} + \frac{e}{c} \vec{V}_0(\mathcal{X}, r) - \frac{e}{c} \vec{X}_0(\mathcal{X}, r) \right] \hat{G}(\mathcal{X}, r) = 2\hbar \delta(r) \hat{\sigma}_z \\
+ \left[ \frac{1}{2m} \left( \frac{\hbar}{i} \frac{\partial}{\partial \mathcal{X}^j} - \frac{\hbar}{i} \frac{\partial}{\partial r^j} + \frac{e}{c} \vec{Y}_j(\mathcal{X}, r) \right)^2 + V \left( \mathcal{X} - \frac{1}{2} r \right) \right] \hat{G}(\mathcal{X}, r) \\
+ \left[ \frac{1}{2m} \left( \frac{\hbar}{i} \frac{\partial}{\partial \mathcal{X}^j} + \frac{\hbar}{i} \frac{\partial}{\partial r^j} + \frac{e}{c} \vec{Y}_j(\mathcal{X}, r) \right)^2 + V \left( \mathcal{X} + \frac{1}{2} r \right) \right] \hat{G}(\mathcal{X}, r),
\] (5.40)
\[
- c \left[ \frac{\hbar}{i} \frac{\partial}{\partial \mathcal{X}^0} + \frac{e}{c} \vec{X}_0(\mathcal{X}, r) + \frac{e}{c} \vec{Y}_0(\mathcal{X}, r) \right] \hat{G}(\mathcal{X}, r) \\
= \left[ \frac{1}{2m} \left( \frac{\hbar}{i} \frac{\partial}{\partial \mathcal{X}^j} - \frac{\hbar}{i} \frac{\partial}{\partial r^j} + \frac{e}{c} \vec{X}_j(\mathcal{X}, r) \right)^2 + V \left( \mathcal{X} - \frac{1}{2} r \right) \right] \hat{G}(\mathcal{X}, r) \\
- \left[ \frac{1}{2m} \left( \frac{\hbar}{i} \frac{\partial}{\partial \mathcal{X}^j} + \frac{\hbar}{i} \frac{\partial}{\partial r^j} + \frac{e}{c} \vec{Y}_j(\mathcal{X}, r) \right)^2 + V \left( \mathcal{X} + \frac{1}{2} r \right) \right] \hat{G}(\mathcal{X}, r).
\] (5.41)
The functions \( \vec{X}_\mu(\mathcal{X}, r) \) and \( \vec{Y}_\mu(\mathcal{X}, r) \) can be found in a straightforward manner. Using the parametriza-
tion (5.36) in the definitions (5.27,5.28), I find

\[
\bar{X}_\mu(X,r) = X_\mu(X,r) = -r^{\nu} \int_{-1/2}^{1/2} \left( \frac{1}{2} - u \right) F_{\mu\nu}(X + ur) du, \tag{5.42}
\]

\[
\bar{Y}_\mu(X,r) = Y_\mu(X,r) = -r^{\nu} \int_{-1/2}^{1/2} \left( \frac{1}{2} + u \right) F_{\mu\nu}(X + ur) du. \tag{5.43}
\]

Here I note that when the form (5.42,5.43) of the functions \( \bar{X}_\mu(x,y), \bar{Y}_\mu(x,y) \) is used in (5.9,5.11), I recover the expressions for the phase factor derivatives given in step (iii) of the LF Green function transformation procedure [54].

The kinetic equations (5.40,5.41) together with the definitions (5.42,5.43) give the full gauge-invariant description of the dynamics of the system for a choice of a straight-line path in space-time. For the case of no periodic potential assumed in the LF approach, the Wigner transformation of the kinetic equations (5.40,5.41) involves simply a Fourier transform and agrees with the equations (30,31) of [55].

### 5.4 The link to PZW

Next I identify the path in space-time that leads to Green function dynamics associated with the usual PZW results [43, 45]. I choose a special point \( R \), which for an atom could be the position of a nucleus, considered fixed in this simple example.

Parts of the paths in space-time I consider here will be paths in space at fixed time. I introduce continuous, time-independent functions \( p(u; x_1, x_2) \) to describe paths in three-space going from \( x_2 \) to \( x_1 \) as \( u \) ranges from \( u_{\text{initial}} \) to \( u_{\text{final}} \). I also introduce the corresponding functions \( p_{\text{inv}}(u; x_2, x_1) \), related to \( p(u; x_1, x_2) \) by

\[
p_{\text{inv}}^k(u; x_2, x_1) \equiv p^k(u_{\text{final}} - u + u_{\text{initial}}; x_1, x_2). \tag{5.44}
\]

That is, the paths \( p^k(u; x_1, x_2) \) and \( p_{\text{inv}}^k(u; x_2, x_1) \) go through the same points in space, but in the reversed order. I now specify a path going through point \( R \), with a corresponding phase factor that can be written as three contributions of the form (5.1,5.2),

\[
\Gamma(x,y) = \Gamma_3((x^0, x), (x^0, R)) + \Gamma_2((x^0, R), (y^0, R)) + \Gamma_1((y^0, R), (y^0, y)).
\]

The paths of the three \( \Gamma_i(x,y) \) are as follows (see Figure 5.2):

- In \( \Gamma_1((y^0, R), (y^0, y)) \) the path goes from \((y^0, y)\) to \((y^0, R)\) at fixed time \( y^0 \), along an arbitrary but fixed path in space \( z^k = p_{\text{inv}}^k(u; R, y) \).

- In \( \Gamma_2((x^0, R), (y^0, R)) \) the path goes from \((y^0, R)\) to \((x^0, R)\) keeping the position \( R \) in space fixed.

- In \( \Gamma_3((x^0, x), (x^0, R)) \) the path goes from \((x^0, R)\) to \((x^0, x)\) at fixed time \( x^0 \), along the reverse path \( z^k = p^k(u; x, R) \) than that used for \( \Gamma_1 \).

I am interested in the form that the kinetic equations (5.29), (5.30) take for the full path described above. To find the functions \( X(x,y) \) and \( Y(x,y) \), I work out the expressions (5.14) and (5.15) for the
Chapter 5. Gauge-invariant Green function dynamics

Figure 5.2: A path in space-time going through a special point in space

tensors $K_0^\alpha\nu$ and $L_0^\alpha\nu$. I find that $K_0^0\nu$, $L_0^\alpha\nu$ yield

\begin{align*}
K_0(w; x, y) &= \delta(w_0 - x_0)M_0(w; x, R), \\
L_0(w; x, y) &= -\delta(w_0 - y_0)M_0(w; y, R),
\end{align*}

(5.45)

where the matrix $M_0^{\alpha\nu}$ has only space-time components

\[
M_0(w; x, R) = \begin{pmatrix}
0 & s^1(w, x; R) & s^2(w, x; R) & s^3(w, x; R) \\
-s^1(w, x; R) & 0 & 0 & 0 \\
-s^2(w, x; R) & 0 & 0 & 0 \\
-s^3(w, x; R) & 0 & 0 & 0
\end{pmatrix},
\]

and the function $s(w, x; R)$ is the line integral (2.56) identified in the PZW theory. Tensors $K_i^\alpha\nu$ and $L_i^\alpha\nu$ have only space-space components,

\begin{align*}
K_i(w; x, y) &= \delta(w_0 - x_0)M_i(w; x, R), \\
L_i(w; x, y) &= -\delta(w_0 - y_0)M_i(w; y, R),
\end{align*}

(5.47)

(5.48)

where the matrix $M_i^{\alpha\nu}$ is defined as

\[
M_i(w; x, R) = \begin{pmatrix}
0 & 0 & -\alpha_{3i}(w, x; R) & \alpha_{2i}(w, x; R) \\
0 & 0 & \alpha_{3i}(w, x; R) & 0 \\
0 & -\alpha_{2i}(w, x; R) & \alpha_{1i}(w, x; R) & 0 \\
0 & -\alpha_{2i}(w, x; R) & 0 & \alpha_{1i}(w, x; R)
\end{pmatrix},
\]

where $\alpha_{ji}(w, x; R)$ is the line integral (2.57) of the PZW theory.

The tensors $K_\mu^\alpha\nu(w; x, y)$ and $L_\mu^\alpha\nu(w; x, y)$ are much simpler in this example than they are in general, for $K_\mu^\alpha\nu(w; x, y)$ is in fact independent of $y$, and $L_\mu^\alpha\nu(w; x, y)$ is independent of $x$. So when I use the results (5.45) and (5.47) in equation (5.16) for $X_\mu(x, y)$, and (5.46) and (5.48) in (5.17) for $Y_\mu(x, y)$, and then use (5.27, 5.28) for $\bar{X}_\mu(x, y)$ and $\bar{Y}_\mu(x, y)$, I find that

\begin{align*}
\bar{X}_\mu^R(x, y) &= X_\mu^R(x, y) = -\Omega_\mu(x), \\
\bar{Y}_\mu^R(x, y) &= Y_\mu^R(x, y) = \Omega_\mu(y),
\end{align*}

(5.49)
where I use a superscript $R$ to indicate this is the form when a special point is introduced as I have here, and I introduce a notation

$$\Omega_0(x) = \int d^3w \, s(w, x; R) \cdot E(w, x),$$

$$\Omega_j(x) = \int d^3w \, B_k(w, x) \alpha_{kj}(w; x; R).$$  \hfill (5.50, 5.51)

In line with the nonrelativistic notation used here for the electric and magnetic fields I have put $t_x = x^0/c$.

I can now write the kinetic equations (5.29), (5.30) as

$$-c \left( \frac{\hbar}{i} \frac{\partial}{\partial x^0} - \frac{e}{c} \Omega_0(x) \right) \bar{G}^t(x, y) = \left[ \frac{1}{2m} \left( \frac{\hbar}{i} \frac{\partial}{\partial x^j} - \frac{e}{c} \Omega_j(x) \right)^2 + V(x) \right] \bar{G}^t(x, y) + \hbar \delta(x - y),$$

$$c \left( \frac{\hbar}{i} \frac{\partial}{\partial y^0} + \frac{e}{c} \Omega_0(y) \right) \bar{G}^t(x, y) = \left[ \frac{1}{2m} \left( \frac{\hbar}{i} \frac{\partial}{\partial y^j} + \frac{e}{c} \Omega_j(y) \right)^2 + V(y) \right] \bar{G}^t(x, y) + \hbar \delta(x - y),$$

and the charge and current densities (5.31), (5.32) as

$$\langle j^0(x) \rangle = -iec \bar{G}^t(x, x_+),$$

$$\langle j^k(x) \rangle = -ie \frac{2m}{\hbar} \left( \frac{\hbar}{i} \frac{\partial}{\partial x^k} - \frac{e}{c} \Omega_k(x) \right) \bar{G}^t(x, y) + \frac{ie^2}{mc} \Omega_k(x) \bar{G}^t(x, x_+).$$ \hfill (5.52, 5.53, 5.54, 5.55)

The expressions (5.52) - (5.55) are derived with the use of the general formalism introduced in section 5.2. Now I will show that the Green function dynamics (5.52, 5.53) can be also seen as following from an effective gauge-invariant Hamiltonian. This arises because the generalized Peierls phase appearing here can be rewritten in a form in which $x$ and $y$ coordinates separate. I have

$$\Gamma^R(x, y) = \gamma(x) - \gamma(y),$$ \hfill (5.56)

where $\gamma(x)$ contains an integral of the four-potential done along a path consisting of the following two parts (see Figure 5.3):

- The first part goes from some fixed time $cT$ to $x^0$ at a fixed position $R$
- The second part goes from the point $R$ to $x$ at a fixed time $x^0$

In nonrelativistic notation I thus have

$$\gamma(x) = -c \int_T^t \phi(R, t')dt' + \int_R^x A(p, t) \cdot dp.$$ \hfill (5.57)

I refer to the general form (5.56), where the generalized Peierls phase is given by the difference of the values a function of one space-time variable takes at the two space-time points in the Green function, as a generalized Peierls phase of the “first type;” generalized Peierls phases that cannot be written in this way are said to be of the “second type.” For generalized Peierls phases of the first type it is possible to understand the gauge-invariant Green function as a Green function formed from transformed field
operators, $\bar{\psi}(x)$ and $\bar{\psi}^\dagger(y)$,

$$G^i(x,y) = -i \langle T[\bar{\psi}(x)\bar{\psi}^\dagger(y)] \rangle,$$

(5.58)

with the new operators given by

$$\bar{\psi}(x) = e^{-\frac{i\hbar}{\epsilon}\gamma(x)}\psi(x) = e^{-i\Phi(t)}e^{-i\Phi_{PZW}(x,t)}\psi(x,t),$$

(5.59)

where the second equality is the special result for $\gamma(x)$ given by (5.57). Here $\Phi_{PZW}(x,t)$ is precisely the phase associated with the Power-Zienau-Wooley transformation [45], see eq. (2.55), and

$$\bar{\Phi}(t) = -\frac{e}{\hbar} \int_t^T \phi(R,t')dt'$$

(5.60)

generates an additional transformation that depends on time only.

Since for a generalized Peierls phase of the first type there exist transformed field variables, I can ask what effective Hamiltonian governs their evolution. For the phase (5.56), using (5.18) I find that the dynamics of the transformed field operators (5.59) is governed by an effective Hamiltonian $H_{\text{eff}}(t)$,

$$i\hbar \frac{\partial \bar{\psi}(x,t)}{\partial t} = [\bar{\psi}(x,t), H_{\text{eff}}(t)],$$

(5.61)

where

$$H_{\text{eff}}(t) = \frac{1}{2m} \int d^3x \frac{\hbar^2}{\epsilon} \left( \frac{\hbar}{\epsilon} \nabla - \frac{e}{c} \Omega(x,t) \right)^2 \bar{\psi}(x,t)$$

$$- e \int d^3x \Omega_0(x,t) \bar{\psi}^\dagger(x,t)\bar{\psi}(x,t)$$

$$+ \int d^3x V(x)\bar{\psi}^\dagger(x,t)\bar{\psi}(x,t).$$

(5.62)

I note that were the transformation from $\psi(x,t)$ to $\bar{\psi}(x,t)$ done using only the Power-Zienau-Wooley phase $\Phi_{PZW}(x,t)$ rather than $\Phi_{PZW}(x,t) + \bar{\Phi}(t)$, the result would differ from (5.62) by $Q\phi(R,t)$, a function only of time since $Q$ (recall (2.70)) is a constant of motion. Thus that extra term has no effect on the dynamics.

Since the transformation (5.59) is in essence the PZW transformation, I can expect that the effective Hamiltonian (5.62) is the transformed Hamiltonian associated with that approach. This is easily confirmed. Expanding the square of the first term of (5.62) and recalling the results (5.50,5.51) for $\Omega_0(x,t)$
and \( \Omega(x, t) \), on collecting terms I can write

\[
H_{\text{eff}}(t) = H_0 + H_{\text{int}}(t),
\]

with

\[
H_0 = -\frac{\hbar^2}{2m} \int d^3x \, \bar{\psi}(x, t) \nabla^2 \psi(x, t) + \int d^3x \, V(x) \bar{\psi}(x, t) \psi(x, t),
\]

and

\[
H_{\text{int}}(t) = -\int d^3x \, p(x, t) \cdot E(x, t) - \int d^3x \, m^{(P)}(x, t) \cdot B(x, t) - \frac{1}{2} \int d^3x \, m^{(D)}(x, t) \cdot B(x, t),
\]

where the operators \( p(x, t) \) and \( m^{(P,D)}(x, t) \) are the polarization (2.66) and magnetization (2.67) fields of the PZW theory; the para- and dia-magnetization are defined in terms of the para-magnetic current (2.64) and dia-magnetic current (2.65) respectively. Equation (5.65) is precisely the transformed interaction Hamiltonian of PZW (2.80) up to a term that has no effect on the dynamics.

In summary, when I use the general approach of section 5.2 with a generalized Peierls phase of the first type, the Green function that results can be understood as following from transformed field operators; in the special case of (5.57) the effective Hamiltonian (5.63) governing the evolution of those transformed field operators describes the interaction of the charge-current distribution with the electromagnetic field through terms involving polarization and magnetization fields and the electric and magnetic fields. Further, as has been stressed by Healy, any path \( p(u; x_1, x_2) \) in space can be used to generate a gauge-invariant Hamiltonian (5.63), with the gauge freedom in the minimal coupling Hamiltonian replaced by the freedom in the choice of this path. Within the framework of section 5.2, we would say that even when a path of the form shown in Figure 5.2 is adopted there is a significant freedom that remains in a choice of path. Indeed, this remaining freedom is precisely that which has been identified by Healy. And so the strategy of section 5.2 can be understood not only as a generalization of the work of Levanda and Fleurov, but also as a generalization of Healy’s approach – which involves a special point and an arbitrary path in space to generate a gauge-invariant Hamiltonian – to the use of an arbitrary path in space-time that might not employ a special point.

### 5.5 Hamiltonians and Green functions

In their applications to the physics of atoms and small molecules, the polarization and magnetization fields (2.66,2.67) are characterized by the property that they are nonzero only on the order of a few atomic spacings from the charge and current densities that are responsible for them. I will call polarization and magnetization fields characterized by this property ideal.

The problem is qualitatively different if a crystal or other extended system is considered, where the charges are not assumed to be confined to a particular, identifiable unit, and in principle can range over macroscopic distances. Here the approach of employing a special point to introduce polarization and magnetization fields, so natural in atomic and molecular physics, would be physically suspect. Nonetheless, for arbitrary fields and charges that are free to move beyond a single unit polarization and magnetization fields have been introduced, without the use of a special point [117, 118]. But these
strategies lead to polarization and magnetization that are not ideal, in the sense defined above.

A natural question arises: For charges that are not confined to particular units, is it possible to construct ideal polarization and magnetization fields so that the dynamics can be described by a Hamiltonian of the form (5.63)?

The answer is no, as can be confirmed by examining a magnetic Aharonov-Bohm scenario, as shown in Figure 5.4. A solenoid carries a magnetic flux $\Phi = \pi R_s^2 B$, where $R_s$ is the radius of the solenoid and $B$ the magnetic field; electrons are confined to the region where $x$ is such that $|x_\perp| \geq r$, where $x_\perp$ is the projection of $x$ onto the plane perpendicular to the solenoid, and $r > R_s$.

From the magnetic Aharonov-Bohm effect I know the electrons are sensitive to the presence of the flux in the solenoid; in the minimal-coupling Hamiltonian the effect arises through the presence of the vector potential, which is nonzero in the region to which the electrons are confined. Yet were ideal polarization and magnetization fields possible, they would only exist in the region where there is no magnetic field and the Hamiltonian (5.63) could not lead to the Aharonov-Bohm effect. Thus such ideal polarization and magnetization fields do not exist.

Besides the nonideal polarization and magnetization fields mentioned above [117, 118], which do not make use of a special point, Healy [119] showed that nonideal polarization and magnetization fields associated with a special point can also be introduced to treat Aharonov-Bohm scenarios: For the scenario depicted in Figure 5.4, the center of the solenoid can be taken to be the special point (and the origin of the coordinate system) and the path in space-time to be that of Figure 5.2, with a straight-line path chosen for $p(u; x, y)$. Although the current $j(y, t)$ exists only in the region $|y_\perp| \geq r$, the resulting magnetization

$$m_j(x, t) = \frac{\epsilon_{jk}}{c} \int d^3 y \int_0^1 y^j \delta(x - u y) j_k(y, t) u d u$$

will exist for all $|x_\perp|$ smaller than the furthest extent $|y_\perp|$ of the current distribution. Although mathematically satisfying relations such as (2.68,2.69), such nonideal magnetization fields (and nonideal polarization fields) seem to have little physical significance, and do not allow for easy approximation strategies.

Since only such problematic, nonideal magnetization and polarization fields are available for extended systems, it is interesting to contrast the description of the dynamics they provide, via (5.63), with that following from Green function kinetic equations employing a Peierls phase of the second type, where no such effective Hamiltonian with magnetization and polarization fields is introduced. There are of course a wide range of such phases that could be considered; I only give a ”toy calculation” here to illustrate...
their possible usefulness. For the system sketched in Figure 5.4, I consider a path that consists of the following parts (see Figure 5.5):

- I move from \((y^0, y)\) to \((x^0, y)\), changing only the time variable.
- I move from \((x^0, y)\) to \((x^0, x)\) at a fixed time along a straight-line path in space.

![Figure 5.5: The path not making use of a special point](image)

Consider first \(\bar{X}_\mu(x, y)\) which here equals \(X_\mu(x, y)\) (recall (5.27)). I find \(X_0(x, y) = 0\) and, writing the magnetic field as \(B(x) = B \Theta(R_s - |x_\perp|)\), where \(\Theta\) is the Heaviside function, I have

\[
X(x, y) = (x - y) \times B \mathcal{J}(x, y),
\]

where

\[
\mathcal{J}(x, y) = \int_0^1 u du \Theta(R_s - |y_\perp + u(x_\perp - y_\perp)|).
\]

This can be evaluated to yield

\[
\mathcal{J}(x, y) = \Theta \left( x_\perp^2 - x_\perp \cdot y_\perp \right) \Theta \left( y_\perp^2 - x_\perp \cdot y_\perp \right) \Theta \left( R_s |x_\perp - y_\perp| - |x_\perp \times y_\perp| \right) \\
\times \frac{2 \left[ y_\perp^2 - x_\perp \cdot y_\perp \right] \sqrt{R_s^2(x_\perp - y_\perp)^2 - (x_\perp \times y_\perp)^2}}{(x_\perp - y_\perp)^4}.
\]

If the particle moves on a cylinder with a radius \(r_c\),

\[
x_\perp = r_c (\cos \phi_2, \sin \phi_2),
\]

\[
y_\perp = r_c (\cos \phi_1, \sin \phi_1),
\]

then (5.69) reduces to

\[
\mathcal{J}(x, y) = \Theta \left( R_s \left| \frac{x_\perp}{r_c} \right| - \cos \left( \frac{\Delta \phi}{2} \right) \right) \sqrt{2 \left( \frac{R_s}{r_c} \right)^2 \left( 1 - \cos \Delta \phi \right) - \sin^2 \Delta \phi} \frac{\sin \Delta \phi}{2(1 - \cos \Delta \phi)},
\]

where \(\Delta \phi = \phi_1 - \phi_2\); see plot in Figure 5.6. The characterization of \(\bar{Y}_\mu(x, y) = Y_\mu(x, y)\) follows in the same manner.

The results for \(\bar{X}_\mu(x, y)\) (and \(\bar{Y}_\mu(x, y)\)) for the path chosen in the paragraph above – a path that does not lead to a Green function \(\bar{G}_t(x, y)\) that can be written in terms of transformed field operators (5.58) –
leads to kinetic equations so complicated that it might seem hard to imagine recommending the strategy for addressing problems involving the scenario pictured in Figure 5.4. But recall that the Aharonov-Bohm magnetic effects arise because of the coherence of the wave function around the solenoid. Suppose that because of electron-phonon scattering, or other decoherence mechanisms, the Green function \( \hat{G}(x, y) \) (and thus \( \hat{\bar{G}}(x, y) \)) is nonvanishing only for \( |x - y| \) sufficiently small; this is typical in “normal” materials where even in the absence of decoherence mechanisms \( \hat{G}(x, y) \) is limited to \( |x - y| \) on the order of a lattice spacing for \( t_x \) close enough to \( t_y \). For \( r_e \) sufficiently large, from the kinetic equations following from the path for \( \Gamma(x, y) \) chosen in the paragraph above it follows immediately that the presence of the magnetic field will be irrelevant to the electron motion, since \( J(x, y) \) will vanish at the small (or large) \( \Delta \phi \) for which \( \hat{\bar{G}}(x, y) \) is nonzero.

Of course, this result would follow from a full treatment of the problem using either the usual minimal coupling Hamiltonian, or a Hamiltonian (5.63) that would necessarily involve nonideal polarization and magnetization fields – whether associated with a special point [119] or not [117], [118]. But such a Hamiltonian would include either a vector potential or magnetic fields “where the electron is not,” and thus it is impossible to see from the beginning if the electron motion will in practice be susceptible to Aharonov-Bohm effects. Basing a description on kinetic equations involving a \( \hat{\bar{G}}(x, y) \) where the associated generalized Peierls phase \( \Gamma(x, y) \) depends only on the electromagnetic fields at points “between” \( y \) and \( x \), of which the choice made above – a generalization of the path chosen by Levanda and Fleurov – is an example, allows one to see more readily how Aharonov-Bohm effects can vanish in the presence of small correlation lengths; this would be included naturally and automatically in any calculations.

Moving away from “toy calculations” such as that associated with Figure 5.6, the GLF approach leads to new strategies for describing the response of matter to electromagnetic fields, in particular in normal materials where the correlation lengths are small compared to the wavelength of light, and approximations can be brought to bear to simplify the terms in the kinetic equations involving fields in the region between \( y \) and \( x \).

### 5.6 Summary

I have shown that two seemingly distinct methods for introducing gauge-invariant forms of the interaction of the electromagnetic field with matter can be seen as special cases of a general formalism. The first method is the approach of Levanda and Fleurov for the introduction of gauge-invariant Green function dynamics.
dynamics. Here the gauge-invariant Green function is introduced using a generalized Peierls phase involving a straight-line path in space-time. The second method, due to Power, Zienau, and Wooley, follows from a generalized Peierls phase that involves a special point. In fact a third method [117], in which polarization and magnetization fields are introduced apparently without the aid of a special point, also falls in this category, for it can be seen to follow from the averaging over all possible special points associated with the function of a space-time variable [118]. Thus all these strategies can be associated with generalized Peierls phases. The original gauge freedom in the minimal coupling Hamiltonian is replaced by a freedom in the choice of a path in space-time, a freedom which has not been fully exploited.

The generalized Peierls phases are of two types. The PZW results follow from a generalized phase of the first type, where the phase can be written as the difference of the values a function of one space-time variable takes at the two space-time points in the Green function. Here an effective Hamiltonian can be introduced, expressing the interaction of the system with the electromagnetic field through the products of the polarization and magnetization fields with respectively the electric and magnetic fields. In generalized Peierls phases of the second type there is no such effective Hamiltonian. For systems where electrons are not considered bound to identifiable units such as atoms or molecules, I have argued on the basis of the magnetic Aharonov-Bohm effect that for phases of the first type any such polarization and magnetization fields must in general be nonideal, in a sense that they do not vanish outside the region where the charge-current distribution they describe is nonzero. Such polarization and magnetization fields are of dubious physical significance, and do not allow for easy approximation. For this reason, phases of the second type are more useful for the description of the interaction of the electromagnetic field with normal condensed matter. Here there is no gauge-invariant interaction Hamiltonian, but within a Green function framework simplifications can be introduced based on the small variation of the electromagnetic field over the short correlation lengths that characterize the system. The latter approach I use in the next chapter to treat a linear response of a clean cold semiconductor.
Chapter 6

Microscopic charge-current densities, polarizations and magnetizations


6.1 Introduction

In the previous chapter I derived dynamical equations for the gauge-invariant Green functions, which describe electrons in the independent particle approximation interacting with arbitrary electromagnetic fields. Green functions can be linked to the expectation values of the microscopic charge and current density operators, and thus they characterize the microscopic state of the material medium. However, a full knowledge of even those densities is not necessary to find the optical response of the medium: while the response is affected by changes in the charge distributions at a macroscopic scale, the atomic-scale variations are not important. Therefore, spatially averaged charge and current distributions are of more interest than the exact microscopic distributions. The averaged distributions are described in the macroscopic Maxwell equations in terms of polarization and magnetization fields. Thus a microscopic theory of polarization and magnetization is needed in order to make the connection between the optical response coefficients that enter the usual phenomenological theories and the microscopic properties of the material medium.

In elementary treatments of the electrodynamics of molecular fluids, the macroscopic polarization and magnetization fields are associated with the electric and magnetic dipole moments of the molecules. Systematic generalizations to include higher order moments were introduced almost half a century ago [1], where the whole analysis can be drastically simplified by introducing microscopic polarization and magnetization fields via the Power-Zienau-Wooley (PZW) gauge transformation [120, 121]. An expansion of electromagnetic fields around the center of a molecule within expressions for polarization and magnetization fields leads in a natural way to molecular multipole moments, the effects of which can be derived up to any order.
Such an approach cannot be immediately extended to crystalline solids, where electrons cannot be assumed to be confined to particular molecules or unit cells. It was only with the advent of the “modern theory of polarization” [48], anticipated by the early work of Adams and Blount [122], that progress was made in developing a theory for the macroscopic polarization of crystals. It was argued [48, 51], that only the change in any alleged macroscopic polarization has physical significance, since this corresponds to the current density. Nonetheless, that polarization could then be linked with the dipole moment of the Wannier functions, with a quantum of ambiguity in its definition linked to the ambiguity in how the Wannier functions are associated with the lattice sites. Recent work [52] has addressed the more difficult problem of determining the orbital magnetoelectric polarizability.

An additional complication arises when magnetization is considered. For an insulator in the ground state [49, 114], the magnetization was shown to contain a contribution due to the magnetic dipole moment of the Wannier functions, but also a contribution due to the moment of a current involving Wannier functions at different lattice sites. The formulation of magnetization in terms of these moments was subsequently extended to systems in a uniform applied electric field [53], while the susceptibility governing the response of the orbital magnetization to the magnetic field has been extracted from the evaluation of the energy in the presence of the magnetic field [123, 124, 125, 126].

Approaches based on the energy of systems in the presence of applied fields cannot immediately be generalized in a rigorous way to take into account the response to time-dependent fields, which is necessary for the treatment of problems in optics. Here a number of the processes one would like to describe, including forbidden second harmonic generation [127] and the optical activity of solids [128, 129], depend on the variation of electromagnetic fields through the crystal. Phenomenological treatments [130] of such processes, for which an underlying microscopic theory should provide expressions for the response coefficients that appear, are often based on the introduction of multipole moments per unit volume.

All this motivates the calculation of the response to electromagnetic fields from an electrodynamic perspective rather than from energy considerations, and a modern introduction of microscopic polarization and magnetization fields in extended solids, as an approach to the construction of macroscopic multipole moments per unit volume.

One might object to this strategy on the grounds that the freedom in choosing Wannier functions would lead to an ambiguity in the higher order moments more drastic than the quantized ambiguity that arises, for example, in the polarization of the ground state. Ambiguities in higher order moments are not just a problem for the treatment of extended solids, of course. Even in the description of a molecular fluid, if the dipole moment of a molecule is nonvanishing, then its quadrupole moment will depend on the choice of the origin. This might lead one to suspect the validity – or at least the uniqueness – of a usual description of the optical activity of such a medium, where the response of the quadrupole moment per unit volume to the electric field enters, as well as the response of the magnetic dipole moment per unit volume to the electric field and the response of the dipole moment per unit volume to the magnetic field and the variation of the electric field. However, in that theory it is discovered that only combinations of these different responses that are insensitive to the choice of origin enter in the description of physical phenomena such as the rotation of the plane of polarization of light [3, 131]. While the situation is more complicated for extended solids, the result for molecular fluid shows that higher order multipoles need not be eschewed simply because there is some ambiguity in their definition, and that examining the possibility of constructing microscopic polarization and magnetization, and using them as the basis
of multipole expansions, is worth investigating. This is the program I initiate in this chapter.

In this investigation I show how the approach can be used to treat an insulator responding linearly to applied uniform and static electric and magnetic fields. With a view towards ultimate applications in the optical regime, in section 6.2 I consider the linear response of an insulator of Chern class zero to an arbitrary applied electromagnetic field. In a manner that allows for a natural expansion of an electromagnetic field varying little over the distance between lattice sites, I construct the nonequilibrium Green function that results. In this chapter I work within the independent particle approximation, but the Green function approach is generally a powerful one because it need not be restricted to considering only the Hamiltonian evolution that is often assumed in calculating the response to applied fields; any appropriate self-energy can be considered. Thus it connects naturally to modern approaches for the computation of material properties based on density functional theory and its generalizations [132]. For simplicity and to more clearly demonstrate the strategy I also neglect electron spin, assume that the bands are nondegenerate, and neglect local field corrections and any induced lattice distortion. Once the Green function is determined, the microscopic charge and current densities immediately follow. Earlier density matrix calculations in the presence of static and uniform fields [52] can be seen as precursors to this approach.

While microscopic polarization and magnetization fields do not uniquely follow from the charge and current densities they describe, I show in section 6.3 that there is a natural way to construct the microscopic polarization and magnetization fields associated with each site by employing the perturbed Wannier functions. There is some ambiguity left in the definition of the microscopic fields, which is related to the non-uniqueness of Wannier functions. However, the first moments of these fields are either gauge-invariant, or have a quantum of ambiguity that simply reflects the well-known ambiguity at the macroscopic level. The higher moments do not contribute to the macroscopic response in the case of static and uniform electric and magnetic fields, nevertheless these are non-unique; this is expected, as even in the extreme case of a molecular crystal, to which a more general PZW theory of a solid should reduce in the appropriate limit, the higher moments depend on the choice of the expansion point with respect to which the moments are taken.

In section 6.4, I verify that the spatial averaging of the generalized PZW microscopic fields introduced in section 6.3 results in the macroscopic polarization and magnetization in agreement with earlier calculations. In the absence of applied fields I recover the established expressions for the polarization and magnetization in the ground state [48, 49]. In the presence of the perturbation I recover the established response coefficients of the polarization to the applied electric field [133, 134] and of the magnetization to the magnetic field [124], what might be called the “diagonal susceptibilities”. While the second of these has been calculated before, the approach proposed in this chapter is the first that does not begin from the evaluation of the energy in a magnetic field. I also recover the response coefficients that describe the response of the polarization to the magnetic field and of the magnetization to the electric field [53], which might be called the “mixed susceptibilities”.

This is the first time that the ground state polarization and magnetization, and both the diagonal and mixed susceptibilities, have been derived from a common framework. It is based on electrodynamics, rather than energy considerations or thermodynamic potentials, and it is the first time microscopic polarization and magnetization fields have been introduced associated with individual lattice sites. While only the first moments of these fields arise in the response to static and uniform fields, the theory I present provides a first step towards a more general description of systems in spatially varying and
Chapter 6. Microscopic charge-current densities, polarizations and magnetizations

I present the conclusions in section 6.5.

6.2 Linear response to time and spatially dependent fields

I consider electrons interacting with an external classical electromagnetic field and with the periodic potential of a lattice, in the independent particle approximation. Within the usual approach the evolution of the system would be described with the minimal coupling Hamiltonian identified by eq. (5.18). Instead, I use the gauge-invariant formulation of the dynamics introduced in chapter 5. I describe the system with the lesser gauge-invariant Green function, taken in the equal time limit. In this chapter I find it convenient to use the non-relativistic notation. I use a short-hand notation

\[ G(x, y; T) = \lim_{x^0 = y^0 = cT} \bar{G}^<(x, y), \]

for the Green function in the equal-time limit, where \( \bar{G}^<(x, y) \) is the gauge-invariant lesser Green function identified by eq. (5.33). The dynamics of the Green function (6.1) follows from the kinetic equations (5.34, 5.35), and is given by

\[ i\hbar \frac{\partial G(x, y; T)}{\partial T} = [\mathcal{H}_0(x) - \mathcal{H}_0(y)] G(x, y; T) + U(x, y; T)G(x, y; T). \]  

(6.2)

Here the differential operator on the right hand side is split into a free Hamiltonian term, \( \mathcal{H}_0(x) = -\frac{\hbar^2}{2m} \nabla^2 + V(x) \), and the differential operator that depends on the externally applied fields,

\[ U(x, y; T) = \frac{1}{m} \left( p(x) + \frac{e}{c} X(x, y; T) \right)^2 \frac{p^2(x)}{2m} - \frac{1}{2m} \left( p(y) + \frac{e}{c} Y(x, y; T) \right)^2 \frac{p^2(y)}{2m} + e T(x, y; T). \]  

(6.3)

In eq. (6.3) I put \( p_j(x) = -i\hbar \partial / \partial x^j \) for the momentum operator and I identified the interaction functions,

\[ X(x, y; T) = \lim_{x^0 = y^0 = cT} X(x, y), \]

(6.4)

\[ Y(x, y; T) = \lim_{x^0 = y^0 = cT} Y(x, y), \]

(6.5)

\[ T(x, y; T) = \lim_{x^0 = y^0 = cT} [X_0(x, y) - Y_0(x, y)], \]

(6.6)

where \( X_\mu(x, y) \) and \( Y_\mu(x, y) \) on the right-hand side of (6.4-6.6) are identified by eqs. (5.10, 5.12). When writing (6.2), I assumed for simplicity that the gauge-invariant Green function (6.1) is introduced by means of a Peierls phase (5.1) that vanishes in the limit when the two coordinates of the Green function overlap, \( \Gamma(x, x) = 0 \); in this case \( \bar{X}_\mu(x, y) = X_\mu(x, y) \) and \( \bar{Y}_\mu(x, y) = Y_\mu(x, y) \) and the bars over the interaction functions can be dropped in (6.3). The functions (6.4-6.6) can be found directly from the
Chapter 6. Microscopic charge-current densities, polarizations and magnetizations

72

definitions (5.10, 5.12). Written explicitly in terms of the electromagnetic fields, these are given by

\[ X_j(x, y; T) = \int_y^x \frac{\partial z^k}{\partial x^j} [dz \times B(z, T)]_k, \quad (6.7) \]

\[ Y_j(x, y; T) = \int_y^x \frac{\partial z^k}{\partial y^j} [dz \times B(z, T)]_k, \quad (6.8) \]

\[ T(x, y; T) = -\int_y^x dz \cdot E(z, T). \quad (6.9) \]

For convenience, I also give here the expressions for the microscopic charge and current densities defined in terms of the equal-time Green function (6.1). These expressions follow immediately from (5.31-5.32) and in the non-relativistic notation take the form

\[ \rho(x, T) = -ieG(x, x; T), \quad (6.10) \]

\[ j(x, T) = -\frac{ie}{2m} \lim_{y \to x} [p(x) - p(y)]G(x, y; T) - \frac{ie^2}{2mc} [X(x, x; T) - Y(x, x; T)]G(x, x; T). \quad (6.11) \]

The equations given in the previous paragraph are valid for Green functions introduced with an arbitrary Peierls phase that vanishes in the limit when the two coordinates of the Green function overlap. An important special case corresponds to the choice of the Peierls phase defined with the use of the special point in space. When the path in the Peierls phase goes through a special point \( R \), and in addition to that the path from \( y \) to \( R \) is the inverse of the path from \( R \) to \( x \) (see Fig. 6.1), the description of the system reduces to the usual PZW theory as discussed in chapter 5. The Green function dynamics (6.2) is in this special case is described by eqs. (5.52, 5.53). For the equal-time Green function (6.1) the dynamical equation that results takes the form

\[ i\hbar \frac{\partial G(x, y; T)}{\partial T} = [\mathcal{H}_R(x, T) - \mathcal{H}_R^*(y, T)]G(x, y; T), \]

where \( \mathcal{H}_R(x, T) \) is the PZW Hamiltonian density,

\[ \mathcal{H}_R(x, T) = \frac{1}{2m} \left( p(x) - \frac{e}{c} \Omega^R(x, T) \right)^2 - e\Omega^0(x, T) + V(x). \quad (6.12) \]

Here I put \( \Omega^R(x, T) = \lim_{x'=cT} \Omega_\mu(x) \) where \( \Omega_\mu(x) \) is identified by eqs. (5.50, 5.51) and I explicitly
indicated the dependence of the PZW Hamiltonian density on the choice of the special point of the transformation.

### 6.2.1 Construction of the multipole Green function

The PZW formalism, although very powerful, can be usefully applied only to systems small compared with the wavelength of light. In an extended system the Peierls transformation (5.33) cannot globally depend on one special point in space, and the charge transfer between atoms in a solid prohibits carrying out a transformation independently for each atom. For these reasons I consider a Peierls phase that depends on \( x \) and \( y \) coordinates only, and for simplicity choose the path in the Peierls phase (5.1) to be a straight line connecting the two coordinates; I will refer to the phase resulting from this choice of a path as the “global Peierls phase”, and denote it by \( \Gamma_{GL}(x, y; T) \). For the Green function (6.1) defined with the phase \( \Gamma_{GL}(x, y; T) \) I seek a solution of (6.2) in a form such that a multipole expansion around lattice sites of the induced charges and currents and their interaction with the applied fields can easily be generated; I will refer to this form of the Green function as the “multipole Green function”.

The strategy I adopt resembles the PZW treatment, with the difference that I account for the motion of charge between lattice sites. Thus instead of evaluating the interaction terms \( X(x, y; T) \), \( Y(x, y; T) \), \( T(x, y; T) \) directly from eqs. (6.7-6.9) for the straight line path, I locally rewrite this path in the neighborhood of each lattice site as a path going through a lattice site together with a path forming a closed loop (see Fig. 6.2). The path going through \( R \) defines a phase \( \Gamma_{PZW}^R(x, y; T) \) leading to the usual PZW transformation with the multipole expansion point at \( R \). The closed loop defines a phase \( \phi(x, y, R; T) \) that is a flux of the magnetic field through a triangle spanned by \( x \), \( y \) and \( R \), which I denote by \( \Delta_{xyR} \). For any site \( R \) I can thus write

\[
\Gamma_{GL}(x, y; T) = \Gamma_{PZW}^R(x, y; T) + \phi(x, y, R; T),
\]

where

\[
\phi(x, y, R; T) = \int_{\Delta_{xyR}} B(s, T) \cdot n,
\]

with

\[
n = -\frac{(x - R) \times (y - R)}{|(x - R) \times (y - R)|}.
\]
I now use (6.13) to evaluate the interaction terms (6.7-6.9). Differentiating (6.13) with respect to \( T \) I arrive at a relation

\[
{T}^{GL}(x, y, T) = -\Omega_{0}^{R}(x, T) + \Omega_{0}^{R}(y, T) + \frac{1}{c} \frac{\partial}{\partial T} \phi(x, y, R, T). \tag{6.16}
\]

Similarly, taking space derivatives of (6.13) I arrive at

\[
X_{j}^{GL}(x, y, T) = -\Omega_{j}^{R}(x, T) + \frac{\partial}{\partial x^{j}} \phi(x, y, R, T), \tag{6.17}
\]

\[
Y_{j}^{GL}(x, y, T) = \Omega_{j}^{R}(y, T) + \frac{\partial}{\partial y^{j}} \phi(x, y, R, T). \tag{6.18}
\]

Relations (6.16-6.18) are exact and in principle hold even for the points \( x, y \) and \( R \) far apart. Physically however, the relation (6.13) is only useful when carried out locally, and will be applied to the projections of \( G(x, y; T) \) on lattice sites rather than for the Green function itself.

With relations (6.16-6.18) in hand I can now construct the multipole Green function. Within the independent particle approximation, the dynamical equation (6.2) is exact and completely general. I now assume the insulator to be at zero temperature and I neglect the spin degrees of freedom; then before any perturbing fields are applied, the lesser Green function is given by

\[
G^{(0)}(x, y) = i \sum_{n} \int_{BZ} d^{3}k \psi_{nk}(x) \psi_{nk}^{*}(y), \tag{6.19}
\]

where as indicated the integral ranges over the Brillouin zone, and I write \( \psi_{nk}(x) \) for the unperturbed Bloch function for band index \( n \) and crystal momentum \( k \),

\[
\mathcal{H}_{0}(x) \psi_{nk}(x) = E_{nk} \psi_{nk}(x), \tag{6.20}
\]

normalized according to

\[
\int d^{3}x \psi_{nk}^{*}(x) \psi_{nk'}(x) = \delta_{nn'} \delta(k - k'), \tag{6.21}
\]

where \( v \) denotes valence bands. Introducing

\[
I^{(1)}(x, y; T) = U(x, y; T)G^{(0)}(x, y; T), \tag{6.22}
\]

where I restrict myself to terms in \( U(x, y; T) \) linear in \( X(x, y; T), Y(x, y; T) \) and \( T(x, y; T) \), I solve (6.2) for \( G(x, y; T) \) to linear order in the electromagnetic field, writing the first order correction to \( G^{(0)}(x, y) \) as \( G^{(1)}(x, y; T) \). I find

\[
G^{(1)}_{n_{1}k_{1}; n_{2}k_{2}}(T) = \lim_{\eta \to 0^{+}} \frac{1}{\hbar} \int_{-\infty}^{T} dt' \int_{-\infty}^{T} dt' I^{(1)}_{n_{1}k_{1}; n_{2}k_{2}}(t') \text{exp} \left[ \frac{i}{\hbar} (E_{n_{1}k_{1}} - E_{n_{2}k_{2}} - i\eta)(t' - T) \right], \tag{6.23}
\]

where the Bloch representation of a general function of the form \( g(x, y; T) \) is given in terms of coefficients \( g_{n_{1}k_{1}; n_{2}k_{2}}(T) \) by

\[
g(x, y; T) = \sum_{n_{1}n_{2}} \int_{BZ} d^{3}k_{1} d^{3}k_{2} \psi_{n_{1}k_{1}}(x) \psi_{n_{2}k_{2}}^{*}(y) g_{n_{1}k_{1}; n_{2}k_{2}}(T). \tag{6.24}
\]
To make use of (6.16-6.18) I now recast the Green function into the Wannier function representation. I assume that for the material in the ground state a complete, orthonormal set of well localized Wannier functions can be introduced. For the case of non-degenerate bands I consider here, the Wannier and Bloch functions are related through transformations [46]

\[ W_{nR}(x) = \sqrt{\frac{V_c}{(2\pi)^3}} \int_{BZ} d^3k e^{-i k \cdot R} \psi_{nk}(x), \]

(6.25)

\[ \psi_{nk}(x) = \sqrt{\frac{V_c}{(2\pi)^3}} \sum_R e^{i k \cdot R} W_{nR}(x), \]

(6.26)

where \(V_c\) is the volume of a unit cell and \(R\) labels lattice sites. Introducing a Wannier representation of a general function of the form \(g(x, y; T)\),

\[ g(x, y; T) = \sum_{n_1, n_2} W_{n_1 R}(x) W^{*}_{n_2 R'}(y) g_{n_1 R; n_2 R'}(T), \]

(6.27)

from (6.23) I find

\[ G^{(1)}_{n_1 R; n_2 R'}(T) = \frac{1}{\hbar} \lim_{\eta \to 0^+} \left[ \frac{V_c}{(2\pi)^3} \right]^2 \int_{BZ} d^3k_1 d^3k_2 \int_0^T dt \left( \frac{i}{\hbar} (E_{n_1 k_1} - E_{n_2 k_2} - i\eta) (t' - T) \right) \times \sum_{R_1, R_2} e^{ik_1 \cdot (R - R_1)} e^{i k_2 \cdot (R_2 - R')} I^{(1)}_{n_1 R_1; n_2 R_2}(t'). \]

(6.28)

In each term \(G^{(1)}_{n_1 R; n_2 R'}(T)\) I now employ the relations (6.16-6.18) to arrive at the multipole expansions of the interaction with the applied electromagnetic fields around sites \(R\) and \(R'\). I do this in a symmetric way, taking the average of the expansions around sites \(R\) and \(R'\),

\[ I^{(1)}_{n_1 R_1; n_2 R_2}(t') = \frac{1}{2} I^{(1)}_{n_1 R_1; n_2 R_2}(t', R) + \frac{1}{2} I^{(1)}_{n_1 R_1; n_2 R_2}(t', R'), \]

(6.29)

where the second argument in the bracket indicates an expansion point. Assuming the applied electromagnetic fields to be of the form

\[ E(x, T) = E(x, \omega)e^{-i \omega T} + E(x, -\omega)e^{i \omega T}, \]

\[ B(x, T) = B(x, \omega)e^{-i \omega T} + B(x, -\omega)e^{i \omega T}, \]

I find

\[ I^{(1)}_{n_1 R_1; n_2 R_2}(t', R) = I^{(1)}_{n_1 R_1; n_2 R_2}(\omega, R)e^{-i \omega t'} + I^{(1)}_{n_1 R_1; n_2 R_2}(-\omega, R)e^{i \omega t'}, \]

(6.30)

with

\[ I^{(1)}_{n_1 R_1; n_2 R_2}(\omega, R) = \langle n_1 R_1 | \hat{H}^{(1)}(\omega) | n_2 R_2 \rangle \]

\[ + \frac{ie}{\hbar c} \int d^3z_1 d^3z_2 g^{(0)}(z_1, z_2) \phi(z_1, z_2, R; \omega) [\mathcal{H}_0(z_1) - \mathcal{H}_0(z_2) - i\omega] W^{*}_{n_1 R_1}(z_1) W_{n_2 R_2}(z_2). \]
Here I put
\[\langle n_1 \mathbf{R}_1 | \hat{H}^{(1)}_R (\omega) | n_2 \mathbf{R}_2 \rangle = \int d^3 z W^*_{n_1 \mathbf{R}_1} (z) \mathcal{H}^{(1)}_R (z, \omega) W_{n_2 \mathbf{R}_2} (z),\]
where \(\mathcal{H}^{(1)}_R (z, \omega)\) is the frequency component of the PZW Hamiltonian density (6.12) that is linear in fields,
\[H^{(1)}_R (x, \omega) = \frac{i e \hbar}{2 mc} \left[ 2 \Omega_j^R (x, \omega) \frac{\partial}{\partial x^j} + \frac{\partial \Omega_j^R (x, \omega)}{\partial x^j} \right] - e \Omega_0^R (x, \omega),\]
and in writing (6.30) I have employed \(\sum_v [\delta_{n_1v} - \delta_{n_2v}] = f_{n_1n_2} - f_{n_1} - f_{n_2}\), where the occupation factor \(f_n = \sum_v \delta_{nv}\). Carrying out a similar calculation for \(I_{n_1 \mathbf{R}_1, n_2 \mathbf{R}_2} (t', \mathbf{R}'), \) using (6.29) in (6.28), employing
\[\phi(z_1, z_2, \mathbf{R}_1; \omega) = \phi(z_1, z_2, \mathbf{R}_2; \omega) - \phi(\mathbf{R}_1, z_2, \mathbf{R}_2; \omega) - \phi(z_1, \mathbf{R}_1, \mathbf{R}_2; \omega),\]
and invoking the completeness of Wannier functions I arrive at the final result,
\[G^{(1)}(x, y; \omega) = G^{(1)}(x, y; \omega)e^{-i\omega T} + G^{(1)}(x, y; -\omega)e^{i\omega T}.\]

The frequency component of the Green function is given by
\[G^{(1)}(x, y; \omega) = \frac{e}{\hbar c} \sum_{v \mathbf{R}} \phi(x, y, \mathbf{R}; \omega) W_v \phi(x) W_v^* (y)\]
\[-\frac{e}{2\hbar c} \sum_{n \mathbf{R} \mathbf{R}'} W_{n \mathbf{R}} (x) W_{v \mathbf{R}}^* (y) \langle n \mathbf{R} | \phi(\mathbf{R}, \mathbf{R}'; \omega) | v \mathbf{R}' \rangle\]
\[-\frac{e}{2\hbar c} \sum_{n \mathbf{R} \mathbf{R}'} W_{v \mathbf{R}} (x) W_{n \mathbf{R}'}^* (y) \langle v \mathbf{R} | \phi(\mathbf{R}, \mathbf{R}'; \omega) | n \mathbf{R}' \rangle\]
\[+ \sum_{\mathbf{R} \mathbf{R}'} \sum_{n_1n_2} W_{n_1 \mathbf{R}} (x) W_{n_2 \mathbf{R}'}^* (y) f_{n_1n_2} \mathcal{O}_{n_1 \mathbf{R}_1, n_2 \mathbf{R}_2} (\omega),\]
where I define
\[\langle n_1 \mathbf{R} | \phi(\mathbf{R}, \mathbf{R}'; \omega) | n_2 \mathbf{R} \rangle = \int d^3 z W^*_{n_1 \mathbf{R}_1} (z) \phi(z, \mathbf{R}, \mathbf{R}'; \omega) W_{n_2 \mathbf{R}_2} (z),\]
and I put
\[\mathcal{O}_{n_1 \mathbf{R}_1, n_2 \mathbf{R}_2} (\omega) = \frac{1}{i} \sum_{\mathbf{R}_1 \mathbf{R}_2} \left[ \frac{V_e}{(2\pi)^3} \right]^2 \mathcal{F}_{n_1 \mathbf{R}_1, n_2 \mathbf{R}_2} (\omega) \lim_{\eta \to 0^+} \int_{BZ} d^3 k_1 d^3 k_2 e^{ik_1 \cdot (\mathbf{R} - \mathbf{R}_1)} e^{-ik_2 \cdot (\mathbf{R}' - \mathbf{R}_2)} \frac{e^{-i\eta E_{n_1 k_1} - E_{n_2 k_2} - \hbar \omega - i\eta}}{E_{n_1 k_1} - E_{n_2 k_2} - \hbar \omega - i\eta},\]
with
\[\mathcal{F}_{n_1 \mathbf{R}_1, n_2 \mathbf{R}_2} (\omega) = \frac{1}{2} \langle n_1 \mathbf{R}_1 | \hat{H}^{(1)}_R (\omega) | n_2 \mathbf{R}_2 \rangle + \frac{1}{2} \langle n_1 \mathbf{R}_1 | \hat{H}^{(1)}_R (\omega) | n_2 \mathbf{R}_2 \rangle.\]

The Green function (6.32) is the main result of this section and gives an exact, microscopic information about the charge and current density response to the linear order in EM fields. Although no approximations regarding the spatial variation of the applied fields were made in deriving this result, the multipole
expansion of the interaction with the applied fields in (6.32) is most useful in situations where the
electron correlation length is much smaller than the distance over which the electromagnetic field varies
significantly; in treating the optical response of materials, this is identified by the wavelength of light.

The response (6.32) takes a particularly simple form for a static, homogeneous electromagnetic field
to which I restrict the considerations in the remaining part of the chapter. In this special case I note
that I have
\[ \Omega^R_0(x) = E \cdot (x - R), \]  
\[ \Omega^R(x) = \frac{1}{2} B \times (x - R), \]  
\[ \phi(x, y, R) = \frac{1}{2} B \cdot [(y - R) \times (x - R)], \]
where \( E = 2E(\omega = 0) \), etc. I now use (6.34-6.36) in equation (6.32), and evaluate moments of Wannier
functions as \( k \)-space integrals over the Brillouin zone. I find
\[ G^{(1)}_{DC}(x, y) = \sum_{n_1 n_2} \int_{BZ} d^3 k \psi_{n_1 k}(x) \psi_{n_2 k}^*(y) \beta_{n_1 n_2}(k) + \sum_{n v} \int_{BZ} d^3 k \psi_{n k}(x) \psi_{n k}^*(y) \gamma_{n v}(k) \]
\[ + \sum_{n v} \int_{BZ} d^3 k \psi_{n k}(x) \psi_{n k}^*(y) \gamma_{n v}(k) + \frac{e}{c \hbar} \sum_{v R} \phi(x, y, R) W_v R(x) W_v^*(y), \]
where I use a notation
\[ \beta_{n_1 n_2}(k) = i e \frac{f_{n_2 n_1}}{E_{n_1 n_2}(k)} \left( E \cdot \zeta_{n_1 n_2}(k) + \frac{1}{2m c} B \cdot [L_{n_1 n_2}(k) + X_{n_1 n_2}(k)] \right), \]
\[ \gamma_{n_1 n_2}(k) = \frac{i e}{4c \hbar} B \cdot D_{n_1 n_2}(k). \]
Here I define vectors
\[ L_{n_1 n_2}(k) = \frac{1}{2} \sum_s \zeta_{s n_2}(k) \times p_{n_1 s}(k) + \frac{1}{2} \sum_s \zeta_{n_1 s}(k) \times p_{s n_2}(k), \]
\[ X_{n_1 n_2}(k) = m \frac{2 \hbar}{i} \left[ \zeta_{n_1 n_2}(k) \times \frac{\partial (E_{n_1 k} + E_{n_2 k})}{\partial k} \right], \]
\[ D_{n_1 n_2}(k) = \frac{\partial}{\partial k} \times \zeta_{n_1 n_2}(k), \]
and write \( \zeta_{n_1 n_2}(k) \) for Berry connection
\[ \zeta_{n_1 n_2}(k) = \frac{1}{V_c} \int_{V_c} d^3 x u_{n_1 k}^*(x) i \frac{\partial}{\partial k} u_{n_2 k}(x), \]
with \( u_{n k}(x) \) being the periodic part of the Bloch function, \( \psi_{n k}(x) = (2\pi)^{-3/2} u_{n k}(x) e^{i k \cdot x} \), and for
the momentum matrix element I put \( \langle n_1 | \hat{p} | n_2 k' \rangle = p_{n_1 n_2}(k) \delta(k - k') \). The first term in (6.37),
which involves the difference of Fermi factors \( f_{n_2 n_1} \), is of the usual type expected in perturbation
theory [135], with \( \beta_{n_1 n_2}(k) \) (6.38) indicating a perturbation energy divided by the energy difference
\[ E_{n_1 n_2}(k) = E_{n_1}(k) - E_{n_2}(k) \] between two Bloch states, which are eigenstates of the unperturbed Hamiltonian. Recalling that the electric and (para)-magnetic dipole moment matrix elements between Wannier
functions at the same site, found using (6.25) and a partial integration, are given by [49]

\[
\int d^3x W_{n0}^*(x) x W_{m0}(x) = \frac{V_e}{(2\pi)^3} \int_{BZ} d^3k \zeta_{nm}(k),
\]

(6.43)

\[
\int d^3x W_{n0}^*(x) x \times \hat{p}(x) W_{m0}(x) = \frac{V_e}{(2\pi)^3} \int_{BZ} d^3k \mathbf{L}_{nm}(k),
\]

(6.44)

we can see that the first two terms on the right-hand side of (6.38) are what one would expect were appropriate \(k\) components of the atomic-like matrix elements (6.43,6.44) associated with each transition energy \(E_{m,n}(k)\). Since the group velocity associated with a given band \(n\) and wavevector \(k\) is given by \(\partial E_{n,k}/\partial (\hbar \mathbf{k})\), the last term on the right-hand side of (6.38) can be identified as a perturbation term involving the magnetic energy associated with the motion of electrons through the bands. The terms involving \(\gamma_{n_1n_2}(k)\) in (6.37) are more complicated, and describe effects due to the variation of the Berry connections throughout the Brillouin zone. The last term in (6.37) I leave in a real space representation, in which it has a simple form; in the next section I indicate the physical significance of this term.

### 6.3 Polarization and magnetization fields associated with lattice sites

The microscopic charge and current distribution driven by an applied electromagnetic field can be found using the expression (6.32) for the Green function. However, when the variation of the applied field is over a length scale much larger than the lattice constant in a normal material, a coarse-grained description with the charge and current densities averaged over the microscopic degrees of freedom is of more interest. For a fluid, this can be an ensemble average [1]; in a solid it can be a spatial average [2]. In the PZW treatment, a coarse-grained description is formulated by introducing the microscopic polarization and magnetization fields (2.66,2.67) in an electrodynamic approach that is also valid for systems driven out of equilibrium. The definitions (2.66,2.67) however, although useful in the treatment of isolated atoms and molecules, are not immediately applicable to extended systems like solids. In this section I formulate a generalized description in the spirit of the PZW treatment, where I introduce microscopic polarization and magnetization fields taking into account the motion of electrons between lattice sites. I restrict the consideration here to the case of static, uniform fields.

The system I consider in this section is described by the Green function \(G(x,y) = G^{(0)}(x,y) + G^{(1)}_{DC}(x,y)\), with the ground-state contribution given by (6.19), and the linear order correction given by (6.37). The difficulty in introducing polarization and magnetization fields is related to the fact that the charge and current densities that follow from \(G(x,y)\) are not localized. In a material with short electron correlation lengths I can avoid these difficulties by searching for a decomposition of the Green function into contributions associated with lattice sites, \(G_R(x,y)\), each of which is nonzero only if \(x\) and \(y\) are within a few lattice constants of \(R\). I construct such a decomposition, ensuring that in the isolated atom limit \(G_R(x,y)\) reduces to a Green function formed from the transformed field operators (5.59) introduced in the PZW approach.

The PZW treatment can be applied to a molecular crystal where Wannier functions associated with different lattice sites are assumed to have no common support, and it is equivalent to introducing a gauge-invariant Hamiltonian (5.63) at each lattice site using the Peierls phase \(\Gamma_{PZW}(x,y;T)\). In this limit of a molecular crystal, the global Peierls phase \(\Gamma_{GL}(x,y;T)\) used to define \(G(x,y;T)\) in section
6.2 differs from the PZW phase by a flux term $\phi(x, y; R; T) = \Gamma_{\text{GL}}^R (x, y; T) - \Gamma_{\text{PZW}}^R (x, y; T)$ at each $R$. So in the more general problem I seek a decomposition of the form

$$G(x, y) = \sum_R G_R(x, y) e^{-\frac{i\pi}{\hbar} \phi(x, y; R)}$$

(6.45)

in the limit of static, uniform fields (see (6.14,6.36)). Here equation (6.45) does not yet define $G_R(x, y)$ uniquely, rather a sum $\sum_R G_R^{(n)}(x, y)$ at each order $n$ in electromagnetic fields. From (6.19) I see that in the zeroth order I have

$$\sum_R G_R^{(0)}(x, y) = i \sum_y W_{vR}(x) W_{vR}^*(y),$$

(6.46)

and the obvious physical choice of $G_R(x, y)$ before an electromagnetic field is applied is

$$G_R^{(0)}(x, y) = i \sum_y W_{vR}(x) W_{vR}^*(y).$$

(6.47)

Here I consider Wannier functions (6.25) for an insulator with non-degenerate bands, as the formalism of section 6.2 is restricted to that special case. For the first order term I use (6.37,6.47) to find

$$\sum_R G_R^{(1)}(x, y) = \sum_{n_1 n_2} \int_{BZ} d^3 k \psi_{n_1 k}(x) \psi_{n_2 k}^*(y) \beta_{n_1 n_2}(k)$$

$$+ \sum_{n v} \int_{BZ} d^3 k \psi_{n k}(x) \psi_{v k}^*(y) \gamma_{n v}(k)$$

$$+ \sum_{n v} \int_{BZ} d^3 k \psi_{n k}(x) \psi_{v k}^*(y) \gamma_{v n}(k),$$

(6.48)

showing that the last term in (6.37) arises from using a global rather than a local Peierls phase for each site. Note that in (6.46) the Wannier functions provide a Schmidt (biorthogonal) decomposition of the function $\sum_R G_R^{(0)}(x, y)$. Even with the application of an electromagnetic field, I can always seek a Schmidt decomposition of $\sum_R G_R(x, y)$. Of course, if there is degeneracy in the amplitudes of the product functions (as there is in (6.46)), that Schmidt decomposition will not be unique; for example, both Eqs. (6.19) and (6.46) provide equivalent Schmidt decompositions of $G_R^{(0)}(x, y)$. Nonetheless, at least within a perturbative approach a natural strategy would be to seek a Schmidt decomposition of the form

$$G_R(x, y) = i \sum_v \tilde{W}_{vR}(x) \tilde{W}_{vR}^*(y),$$

(6.49)

where the $\tilde{W}_{vR}(x)$ are Wannier functions modified by the applied electromagnetic field but still satisfying

$$\int \tilde{W}_{vR}^*(x) \tilde{W}_{v'R}(x) d^3x = \delta_{v v'}. $$

(6.50)

Further, in the special case I consider here, of a perturbative response to uniform fields, I would expect the Wannier functions associated with each lattice site to be identical except for translation; that is $\tilde{W}_{vR}(x-R) = \tilde{W}_{vR}(x-R')$. With this assumption I can introduce Bloch functions $\tilde{\psi}_{v k}(x)$ associated with $\tilde{W}_{vR}(x)$,

$$\tilde{\psi}_{v k}(x) = \sqrt{\frac{V_c}{(2\pi)^3}} \sum_R e^{ikR} \tilde{W}_{vR}(x),$$

(6.51)
Chapter 6. Microscopic charge-current densities, polarizations and magnetizations

satisfying $\tilde{\psi}_{vk}(x + R) = e^{ik \cdot R} \tilde{\psi}_{vk}(x)$, and write

$$\sum_R G_R(x, y) = i \sum_v \int_{BZ} d^3 k \tilde{\psi}_{vk}(x) \tilde{\psi}^*_{vk}(y). \quad (6.52)$$

These natural assumptions, together with (6.48), are enough to significantly constrain the possible $\tilde{W}_{vR}(x)$, as I now show. Noting that in the absence of an applied field $\tilde{\psi}^{(0)}_{vk}(x) = \psi_{vk}(x)$ (see (6.47,6.51)), from (6.52, 6.48) I immediately see than one of the solutions is given by

$$\tilde{\psi}_{vk}(x) = \phi_{vk}(x) \equiv \psi_{vk}(x) + \frac{1}{i} \sum_c \psi_{ck}(x) \beta_{cv} + \frac{1}{i} \sum_n \psi_{nk}(x) \gamma_{nv}. \quad (6.53)$$

The general solution is a unitary transformation of (6.53),

$$\tilde{\psi}_{vk}(x) = \sum_{v'} U_{v'v}(k) \phi_{v'k}(x), \quad U_{v'v}(0) = \delta_{v'v}, \quad (6.54)$$

which however does not affect the dipole moment contributions to polarization and magnetization, while the higher order moments do not contribute in the case of uniform electromagnetic fields, as I will show in section 6.4. Thus in the simple case I consider in this chapter I can choose $\tilde{\psi}_{vk}(x) = \phi_{vk}(x)$. More generally, for non-uniform fields or degenerate energy bands a unitary transformation of the Bloch functions could be considered to minimize the spread of the Wannier function. The modified Wannier function follows from the inverse of (6.51),

$$\tilde{W}_{vR}(x) = \sqrt{\frac{V_c}{(2\pi)^3}} \int_{BZ} d^3 k e^{-ik \cdot R} \phi_{vk}(x), \quad (6.55)$$

and the lattice-site Green function (6.49) can be constructed. The charge and current densities that correspond to it,

$$\rho_R(x) = -ieG_R(x, x), \quad (6.56)$$

$$j^k_R(x) = \frac{ie}{2m} \lim_{y \to x} \left[ \frac{\hbar}{i} \frac{\partial}{\partial x^k} - \frac{\hbar}{i} \frac{\partial}{\partial y^k} \right] G_R(x, y) + \frac{ie^2}{mc} \Omega_{k}^{R}(x)G_R(x, x), \quad (6.57)$$

when summed over lattice sites lead to the total charge and current densities

$$\rho(x) = \sum_R \rho_R(x), \quad (6.58)$$

$$j^k(x) = \sum_R j^k_R(x). \quad (6.59)$$

The densities $\rho_R(x)$ and $j_R(x)$ (6.56,6.57) are localized about $R$ and their moments are well defined. I can thus use them to introduce polarization and local circulation magnetization fields, in an approach that follows the PZW definitions (2.66,2.67). I write

$$p^i_R(x) = \int d^3 y s^i(x, y; R) \rho_R(y), \quad (6.60)$$

$$m^i_R(x) = \frac{1}{c} \int d^3 y a^{ik}(x, y; R) j^k_R(y). \quad (6.61)$$
where the functions $s_i(x, y; R)$ and $\alpha^{ik}(x, y; R)$ are given by (2.56,2.57). Using relations (2.59,2.58), I can rewrite the charge and current densities as

$$\rho(x) = -\sum_R \nabla \cdot p_R(x) + \rho_{ref}(x), \quad (6.62)$$

and

$$j(x) = e\sum_R \nabla \times m_R(x) + \tilde{j}(x), \quad (6.63)$$

where I define the reference charge density,

$$\rho_{ref}(x) = \sum_R \delta(x - R) Q_R, \quad Q_R = \int d^3y \rho_R(y), \quad (6.64)$$

and an additional current density,

$$\tilde{j}(x) = -\sum_R \int d^3y s(x, y; R) \Theta_R(y). \quad (6.65)$$

A term $\partial p_R/\partial T$ would normally be present on the right-hand-side of (6.62), but is absent because I have assumed that all fields are time-independent. Were that time-dependence reintroduced, both $\Theta_R(x)$ and $\tilde{j}(x)$ would also become time dependent, with $\Theta_R(x, T) = \partial \rho_R(x, T)/\partial T + \nabla \cdot j_R(x, T)$; in the time-independent problem I consider here, $\Theta_R(x) = \nabla \cdot j_R(x)$. This term thus describes the fact that although the total charge-current density $(\rho(x), j(x))$ naturally satisfies the continuity equation, the individual charge-current densities I have introduced, $(\rho_R(x), j_R(x))$, need not. And so I can identify $\tilde{j}(x)$ as a current associated with the motion of electrons between sites, or a "hopping current" for short. Using (6.49,6.57) I find

$$\Theta_R(x) = \frac{ie}{\hbar} \sum_v \hat{W}^*_v R(x) \mathcal{H}_R(x) \hat{W}_v R(x) + c.c, \quad (6.66)$$

where $\mathcal{H}_R(x)$ is the static limit of the PZW Hamiltonian density (6.12) associated with site $R$,

$$\mathcal{H}_R(x) = \frac{1}{2m} \left[ p(x) - \frac{e}{2c} B \times (x - R) \right]^2 + V(x) - eE \cdot (x - R).$$

The equation (6.66) is valid in general for the Green function decomposition of the form (6.45,6.49), with the densities defined through (6.56,6.57). However, since I identified the modified Wannier functions (6.55) only up to linear order in the applied electromagnetic fields, when using Eq. (6.55) in (6.66) only terms up to linear order should be kept; this applies also to the remaining equations in this section, for which I drop the subscripts indicating the order of the perturbation theory for a clearer notation.

The equation (6.63) is not yet in its final form. For a collection of isolated molecules $\Theta_R(x) = \tilde{j}(x) = 0$, and then $m_R(x)$ given by (6.61) is the total contribution to the magnetization per lattice site; but this is not the case for a solid where the hopping current $\tilde{j}(x)$ gives rise to the itinerant part of magnetization. I note that physically the change in the charge-current densities leading to the violation of the "site continuity equation" $\Theta_R(x)$ should be understood as the influx of charge from neighbouring sites diminished by the outflow to other sites. To identify the hopping matrix elements in the presence of external electromagnetic fields I first need to establish the dynamical equation satisfied by the lattice-site Green function. Using the decomposition (6.45) in (6.2) together with relations (6.16-6.18) I have in the
The current (6.71) is divergenceless and thus can be represented as the curl of an itinerant magnetization up to linear order. I note here that the relation (6.68) guarantees indempotency of the usual gauge-satisfying functions (6.55), rather than being orthonormal, satisfy

\[ \int \! d^3x \tilde{W}^*_{v_1R_1}(x) \tilde{W}_{v_2R_2}(x) e^{- \frac{i}{\hbar} \phi(x,R_1R_2)} = \delta_{v_1v_2}\delta_{R_1R_2} \quad (6.68) \]

up to linear order. I note here that the relation (6.68) guarantees indempotency of the usual gauge-dependent density matrix \(-iG^{\text{GD}}(x,y;T) = -iG(x,y;T)e^{i\phi(x+y;T)} \) with \( G(x,y;T) \) given by (6.45,6.49). Multiplying now both sides of equation (6.67) by \( e^{i\phi(x,y;T)} \tilde{W}_{v'R'}(y) \), using (6.31,6.68) and integrating over space, I arrive at

\[ \mathcal{H}_R(x) \tilde{W}_{vR}(x) = \sum_{v'R'} t_{v'R}' \tilde{W}_{v'R'}(x) e^{- \frac{i}{\hbar} \phi(x,R,R')} \]

with the hopping matrix elements

\[ t_{v'R}' = \int \! d^3x \tilde{W}^*_{v'R}(x) e^{- \frac{i}{\hbar} \phi(x,R',R)} \mathcal{H}_R(x) \tilde{W}_{vR}(x) = \int \! d^3x \tilde{W}^*_{v'R}(x) H_R(x) e^{- \frac{i}{\hbar} \phi(x,R',R)} \tilde{W}_{vR}(x) \]

satisfying \([t_{v'R}']^* = t_{v'R}'\). Having identified the hopping matrix elements it is now straightforward to construct a decomposition

\[ \Theta_R(x) = \sum_{R'} [\Theta_{R'R}(x) - \Theta_{RR}(x)] \quad (6.69) \]

of the function (6.66), where

\[ \Theta_{R'R}(x) = \frac{i\epsilon}{\hbar} \sum_{v,v'} t_{v'R}^* \tilde{W}^*_{v'R}(x) \tilde{W}_{v'R'}(x) e^{- \frac{i}{\hbar} \phi(x,R,R')} \]

describes the motion of charge from \( R' \) to \( R \). From (6.69) I now see that the hopping current (6.65) is an inherently two-site quantity, and I can identify the part of \( \tilde{j}(x) \) associated with hopping between sites \( R \) and \( R' \),

\[ \tilde{j}_{RR'}(x) = -\frac{1}{2} \int \! d^3y [\Theta_{R'R}(y) - \Theta_{RR}(y)] [s(x,y;R) - s(x,y;R')] \quad (6.70) \]

Assigning half of the current (6.70) to site \( R \) and half to \( R' \), I put

\[ \tilde{j}_R(x) = \frac{1}{2} \sum_{R'} \left[ \tilde{j}_{RR'}(x) + \tilde{j}_{RR'}(x) \right] \quad (6.71) \]

The current (6.71) is divergenceless and thus can be represented as the curl of an itinerant magnetization

\[ \tilde{m}_R(x) = -\frac{1}{c} \int \! d^3y \alpha^k(x,y;R) \tilde{j}_R^k(y) \quad (6.72) \]
I now arrive at the final representation of the charge and current densities,

\[
\rho(x) = - \sum_R \nabla \cdot p_R(x),
\]

\[
\mathbf{j}(x) = c \sum_R \nabla \times \left[ \mathbf{m}_R(x) + \tilde{m}_R(x) \right],
\]

(6.73, 6.74)

where to write (6.73) I use the fact that for linear response the reference charge density (6.64) vanishes after the ionic background is included. The polarization and magnetization fields in equations (6.73,6.74) serve as a basis for constructing a multipole expansion of the charge and current distributions that when spatially averaged leads to a macroscopic description of a system.

### 6.4 Multipole expansion of polarization and magnetization fields

I am interested in applying the formalism of section 6.3 and its generalizations to problems where normal materials are subject to radiation with wavelengths in the optical regime or longer. In general any global polarization and magnetization fields that describe charge-current densities not confined in space must be linked with those densities in a nonlocal way as discussed in chapter 5. But for normal materials the Green function \( G(x, y; T) \) can be expected to vanish as the distance between \( x \) and \( y \) is increased beyond a few lattice spacings, and the \( G_R(x, y) \) (6.49) constructed perturbatively from (6.47) can be expected to vanish for \( x \) and \( y \) removed from \( R \) by a few lattice spacings, a distance over which the electromagnetic field varies little. A multipole expansion of the charge and current densities associated with each site \( R \) within the expressions for polarization and magnetization fields can then be successfully introduced which, when complemented with a spatial averaging procedure to define macroscopic fields, gives a macroscopic description of the material response. In this section I construct the multipole expansion for the static case of section 6.3. I confirm that within my approach I recover the earlier results for the ground-state polarization and magnetization [48, 49], the electric susceptibility [133, 134], the magnetoelectric response [53], and the magnetic susceptibility [124].

I start by using the Dirac delta expansion (2.83) in the expressions (6.60,6.61) to find

\[
p^i_R(x) = \mu^i_R \delta(x - R) - q^{ij}_R \frac{\partial}{\partial x^j} \delta(x - R) + \ldots,
\]

\[
m^i_R(x) = v^i_R \delta(x - R) + \ldots,
\]

(6.75, 6.76)

where the dipole, quadrupole and the local circulation magnetic dipole moment are given by

\[
\mu^i_R = \int d^3 y (y - R) \rho_R(y),
\]

\[
q^{ij}_R = \frac{1}{2} \int d^3 y (y^i - R^i) (y^j - R^j) \rho_R(y),
\]

\[
v_R = \frac{1}{2c} \int d^3 y (y - R) \times \mathbf{j}_R(y)
\]

(6.77, 6.78, 6.79)

For the itinerant magnetization (6.72) I first do the integral \( \int d^3 y \alpha^{ik}(x, y; R_1)s^k(y, z; R_2) \) and then keeping the lowest term in the expansions around \( u(R_2 - R_1) \) and \( u'(z - R_2) \) I get

\[
\tilde{m}_R(x) = \delta(x - R)\tilde{v}_R + \ldots,
\]

(6.80)
with the itinerant magnetic dipole moment given by

$$\vec{v}_R = \frac{1}{4c} \sum_{R'} (R' - R) \times \int \, \mathrm{d}^3z (z - R) \left[ \Theta_{RR'}(z) - \Theta_{RR}(z) \right]. \quad (6.81)$$

The expansions (6.75,6.76,6.80) need to be complemented with an appropriate coarse-graining procedure. I do this in the standard way [2] by defining macroscopic fields $F^{\text{mac}}(x)$ in terms of the corresponding microscopic fields $F(x)$ by spatial averaging, see eq. (2.10). From Eqs. (6.73,6.74) I then have

$$\rho^{\text{mac}}(x) = -\nabla \cdot \mathbf{P}(x),$$
$$j^{\text{mac}}(x) = c\nabla \times \mathbf{M}(x),$$

where using (6.75,6.76,6.80) I have for the macroscopic polarization and magnetization fields

$$\mathbf{P}(x) = \mathbf{P}(x) - \nabla \cdot \mathbf{Q}(x) + \ldots, \quad (6.82)$$
$$\mathbf{M}(x) = \mathbf{M}(x) + \ldots, \quad (6.83)$$

with

$$\mathbf{P}(x) = \sum_R f(x - R) \mu_R,$$
$$\mathbf{Q}^{ij}(x) = \sum_R f(x - R) q^{ij}_R,$$
$$\mathbf{M}(x) = \sum_R f(x - R) (v_R + \tilde{v}_R),$$

where $f(x)$ is the averaging function. The terms surviving in (6.82,6.83) result from the terms kept in the use of the expansion (2.83); in general more could be included. Expansions such as (6.82,6.83) are common in the discussion of the electrodynamics of fluids, where the macroscopic fields are traditionally defined via ensemble averages [1]. In the case I consider here, with uniform electric and magnetic fields, the multipole moments are independent of lattice site, $\mu_R = \mu$, etc., and taking the sum over all lattice sites of $f(x - R)$ to be essentially independent of $x$ by virtue of (2.12), $\mathbf{P}(x) = \mathbf{P}$, etc., I have

$$\mathbf{P} = n\mu, \quad (6.84)$$
$$\mathbf{M} = n(v + \tilde{v}), \quad (6.85)$$

where $n$ is the number density of lattice sites.

I now explicitly evaluate the moments in (6.84,6.85) to find the ground state contributions, the electric and magnetic susceptibilities and the magnetoelectric tensor, defined as coefficients in the expansion

$$\mathbf{P}^i = \mathbf{P}^i_{(0)} + \mathbf{\chi}^{ij} E_j + \tilde{\mathbf{\alpha}}^{ij} B_j,$$
$$\mathbf{M}^i = \mathbf{M}^i_{(0)} + \mathbf{\chi}^{ij} B_j + \tilde{\mathbf{\alpha}}^{ij} E_j.$$
For the ground state contribution to polarization and magnetization I find

\[ \mathcal{P}_{(0)} = \epsilon \sum_v \int_{BZ} \frac{d^3k}{(2\pi)^3} \zeta_{vv}(k), \]  

(6.86) \]

\[ \mathcal{M}_{(0)} = \epsilon \sum_v \int_{BZ} \frac{d^3k}{(2\pi)^3} J_{vv}(k), \]  

(6.87) \]

in agreement with King-Smith et al. [48] and Thonhauser et al. [49] and where I have introduced a shorthand

\[ J_{n_1n_2}(k) = \frac{1}{2mc} [L_{n_1n_2}(k) + X_{n_1n_2}(k)], \]  

(6.88) \]

Identifying the term in the expression for the polarization proportional to the electric field I find the electric susceptibility

\[ \chi_{ij}^{(E)} = \sum_{n_1n_2} \int_{BZ} \frac{d^3k}{(2\pi)^3} \frac{e^2 f_{n_1n_2}}{E_{n_1n_2}(k)} \zeta_{n_2n_1}^{i}(k) \zeta_{n_1n_2}^{j}(k), \]  

(6.89) \]

in agreement with Sipe et al. [133, 134]. The magnetoelastic tensor I calculate both as a response of the electric dipole moment to the magnetic field and the magnetic dipole with respect to electric field; the result is

\[ \tilde{\alpha}_{ij} = \sum_{n_1n_2} \int_{BZ} \frac{d^3k}{(2\pi)^3} \frac{e^2 f_{n_1n_2}}{E_{n_1n_2}(k)} \zeta_{n_2n_1}^{i}(k) J_{n_1n_2}^{j}(k) - \frac{e^2}{2\hbar} \text{Re} \sum_{nv} \int_{BZ} \frac{d^3k}{(2\pi)^3} \zeta_{vn}^{i}(k) D_{vn}^{j}(k), \]  

(6.90) \]

and agrees with that of Malashevitch et al. [53] for the simple case of nondegenerate bands that I consider in this chapter and in the limit where screening effects are neglected. The magnetic susceptibility tensor I write as a sum of four terms,

\[ \chi_{ij}^{(B)} = \chi_{a}^{ij} + \chi_{b}^{ij} + \chi_{c}^{ij} + \chi_{d}^{ij}. \]  

(6.91) \]

The first two terms are given by

\[ \chi_{a}^{ij} = \sum_{n_1n_2} \int_{BZ} \frac{d^3k}{(2\pi)^3} \frac{e^2 f_{n_1n_2}}{E_{n_1n_2}(k)} J_{n_1n_2}^{i}(k) J_{n_1n_2}^{j}(k), \]  

(6.92) \]

\[ \chi_{b}^{ij} = - \frac{e^2}{4mc^2} \sum_{nv} \int_{BZ} \frac{d^3k}{(2\pi)^3} \left[ \delta^{ij} \zeta_{vn}(k) \cdot \zeta_{vn}(k) - \frac{1}{2} \left( \zeta_{vn}^{i}(k) \zeta_{vn}^{i}(k) + \zeta_{vn}^{j}(k) \zeta_{vn}^{j}(k) \right) \right], \]  

(6.93) \]

and in the limit of isolated atoms they reduce respectively to the usual paramagnetic and diamagnetic atomic susceptibilities. The remaining two terms involve itinerant currents, in that they would disappear were the Wannier functions at different sites to have no common support. They are given by

\[ \chi_{c}^{ij} = - \frac{e^2}{\hbar c} \text{Re} \sum_{nv} \int_{BZ} \frac{d^3k}{(2\pi)^3} \left\{ J_{vn}^{i}(k) D_{vn}^{j}(k) + \frac{1}{8\hbar c} E_{vn}(k) D_{vn}^{i}(k) D_{vn}^{j}(k) \right\}, \]  

(6.94) \]

\[ \chi_{d}^{ij} = \frac{e^2}{8\hbar^2 c^2} \sum_{nv} \epsilon^{iml} \epsilon^{jus} \int_{BZ} \frac{d^3k}{(2\pi)^3} \frac{\partial^2 E_{vk}(k)}{\partial k^i \partial k^j} \left[ \zeta_{vn}^{m}(k) \zeta_{vn}^{u}(k) + \zeta_{vn}^{m}(k) \zeta_{vn}^{u}(k) \right], \]  

(6.95) \]

where the curly bracket indicates a symmetrization in \( i, j \) indices,

\[ \{ A^i B^j \} = \{ A^j B^i \} = \frac{1}{2} \left( A^i B^j + A^j B^i \right). \]  

(6.96) \]
Eq. (6.91) is in agreement with the earlier result found for crystals assumed to have a cubic symmetry and derived using energy considerations [124].

The gauge transformation of equations (6.86-6.90) was discussed in earlier works [48, 49, 53]. I now show that the form (6.91) of the magnetic susceptibility for a crystal of arbitrary symmetry is explicitly gauge-invariant; to this end I introduce ”gauge-covariant” matrix elements $J_{n_1n_2}(k)$ and $D_{n_1n_2}(k)$,

$$
J_{n_1n_2}(k) = \frac{1 - \delta_{n_1n_2}}{4\hbar} \zeta_{n_1n_2}(k) \times \frac{\partial (E_{n_1k} + E_{n_2k})}{\partial k} + \frac{1}{4mc} \sum_{s \neq n_2} \zeta_{sn_2}(k) \times p_{n_2s}(k) + \frac{1}{4mc} \sum_{s \neq n_1} \zeta_{ns}(k) \times p_{sn_1}(k),
$$

(6.97)

$$
D_{n_1n_2}(k) = i \sum_{s \neq n_1, n_2} \zeta_{ns}(k) \times \zeta_{sn_2}(k),
$$

(6.98)

that differ from $J_{n_1n_2}(k), D_{n_1n_2}(k)$ (6.88,6.42) by the absence of diagonal Berry connections. The susceptibility tensor (6.91) I now write as

$$
\chi^{ij} = \tilde{\chi}^{ij}_a + \tilde{\chi}^{ij}_b + \tilde{\chi}^{ij}_c + \tilde{\chi}^{ij}_d + \tilde{\chi}^{ij}_e,
$$

(6.99)

where the first four terms

$$
\tilde{\chi}^{ij}_a = \sum_{n_1n_2} \int_{BZ} \frac{d^3k}{(2\pi)^3} \frac{E_{n_1n_2}(k)}{E_{n_1n_2}(k)} J^j_{n_2n_1}(k) J^i_{n_1n_2}(k),
$$

(6.100)

$$
\tilde{\chi}^{ij}_b = -\frac{e^2}{4mc^2} \sum_{n_1n_2} \int_{BZ} \frac{d^3k}{(2\pi)^3} \left( \delta^{ij} \zeta_{vn}(k) \cdot \zeta_{vn}(k) - \frac{1}{2} (\zeta^i_{vn}(k) \zeta^i_{vn}(k) + \zeta^j_{vn}(k) \zeta^j_{vn}(k)) \right),
$$

(6.101)

$$
\tilde{\chi}^{ij}_c = -\frac{e^2}{c} \mathrm{Re} \sum_{n_1n_2} \int_{BZ} \frac{d^3k}{(2\pi)^3} \left( J^i_{vn}(k) D^j_{vn}(k) + \frac{1}{8\hbar} E_{vn}(k) D^i_{vn}(k) D^j_{vn}(k) \right),
$$

(6.102)

$$
\tilde{\chi}^{ij}_d = \frac{e^2}{8\hbar^2c^2} \sum_{n_1n_2} \sum_{u \neq v} \epsilon^{ilm} \epsilon^{jus} \int_{BZ} \frac{d^3k}{(2\pi)^3} \frac{\partial^2 E_{vk}(k)}{\partial k^l \partial k^s} \left[ \zeta^m_{vn}(k) \zeta^u_{vn}(k) + \zeta^u_{vn}(k) \zeta^m_{vn}(k) \right],
$$

(6.103)

differ from (6.92-6.95) in that the diagonal Berry connections are now excluded; the excluded terms I collect and using appropriate sum rules outlined in appendix B I rewrite as

$$
\tilde{\chi}^{ij}_e = -\frac{e^2}{2\hbar} \sum_{n_1n_2} \int_{BZ} \frac{d^3k}{(2\pi)^3} \left( J^i_{vn}(k) D^j_{vn}(k) \right).
$$

(6.104)

6.5 Discussion and future outlook

I have presented a microscopic description of the linear response of a crystal to electromagnetic fields that is based entirely on electrodynamic considerations. The starting point was a gauge-invariant Green function with a Peierls phase defined only in terms of the coordinates of the Green function. In the neighbourhood of each lattice site the phase was rewritten as the Power-Zienau-Wooley (PZW) phase modified by a flux of the magnetic field, with the lattice site being the special point of the transformation. An expansion of applied fields around the lattice sites within the Green function followed in a natural way, and microscopic charge and current densities could be extracted for the response of the system exposed to time and spatially varying electromagnetic fields.
Restricting the description to static and uniform electric and magnetic fields, I introduced microscopic polarization and magnetization fields associated with each site, taking into account the itinerant current "between" lattice sites that gives an additional contribution to the magnetization. Macroscopic polarization and magnetization fields were introduced as spatial averages of the microscopic fields. The results that followed for the macroscopic polarization and magnetization, both before any field is applied as well as in linear response, are all in agreement with accepted expressions. This is the first time that expressions for the ground state polarization and magnetization, and the diagonal and mixed susceptibilities, have all been derived within a single framework and based on purely electrodynamic considerations.

The description of the response of solids to time-varying and spatially-varying electromagnetic fields – where for transverse fields the former necessarily implies the latter through the Maxwell equations – is an outstanding problem in the optical response of materials. A formal expansion of the minimal-coupling Hamiltonian in terms of the wave vector of the applied fields is an obvious strategy and always a possible approach, but it offers little physical insight, and since the calculations are made for a particular choice of the scalar and vector potentials they are generally not manifestly gauge invariant. In contrast, only the electromagnetic fields appear in a PZW approach. But even in the usual PZW treatment of molecular fluids, nonunique polarization and magnetization fields arise once spatially varying applied fields are introduced, and moments beyond the first of the molecular charge and current densities appear in the description. Nonetheless, an unambiguous account results for both the charge and current densities in the bulk of the material, and the physics of interfaces involving such calculations as the intensity of reflected light [3, 136].

In the extended PZW approach of a solid I initiate the situation is even more complicated, but the nonuniqueness of the higher order moments of the Wannier functions does mirror that of the nonuniqueness of the higher order moments in molecular fluids. Whether the generalization of the formalism to treat optical response could provide an unambiguous account of both the charge and current densities in the bulk, and the optical properties of interfaces, is an open question and beyond the scope of this thesis. But the success of such an extended PZW approach in describing all the linear response properties in the limit of static and uniform fields is promising, and its demonstration here provides an introduction to a strategy for addressing the outstanding problem of the optical response of solids to time-dependent and spatially varying electric and magnetic fields.
Chapter 7

Conclusions

In this thesis I analyzed the response of nanoparticle arrays and regular solids to electromagnetic fields. The analysis is based on a multipolar expansion around lattice sites of charge-current distributions and the electromagnetic fields radiated by them, and on extensions of the usual multipolar expansion to account for motion of charge between lattice sites. I used those methods to study optical processes in nanoparticle arrays that are associated with higher order multipole moments of nanoparticles, and to describe at the microscopic level the response of solids to electromagnetic fields.

In the first part of this thesis I analyzed the optical response of 2d nanoparticle arrays. Nanoparticles in an array are spatially separated so that there is no transfer of charge between them, and thus they are modeled within a multipole expansion of charge-current distributions carried out independently for each nanoparticle. I treated the optical response of the array at the level of including the electric dipole moments, the magnetic dipole moments, and the electric quadrupole moments of the nanoparticles. I took into account the response of each of those moments to the electric field, the magnetic field, and to the gradient of the electric field at the position of the multipole.

In chapter 3, I evaluated all the periodic Green function dyadics that describe radiative interactions between multipole moments in arrays of arbitrary geometry. The dyadics take the form of slowly convergent summations over lattice sites, and they need to be evaluated with the use of acceleration techniques. I accelerated the summations using a hybrid method, which involves a dimensionality reduction, Poisson transformation, and the dominant part extraction. This method gives exact analytic results for the radiative contributions to periodic Green functions. These are the contributions that affect the energy balance on the lattice. I found those radiative contributions in a simple analytic form, such that the radiation reaction terms associated with each diffracted s- and p-polarized beam can be easily identified. The remaining non-radiative contributions I found in a form of rapidly converging series, and wrote them as linear combinations of a few functions. An identification of radiation reaction terms in an exact analytic form facilitates the analysis of the optical response of the system: these terms identify energy flow in a system and they describe the scaling of the radiative linewidth with lattice parameters. Furthermore, these analytic expressions can be used to construct simple approximate expressions for the dyadics that satisfy the optical theorem exactly. By way of an example, I proposed an approximate scheme to evaluate periodic Green functions dyadics in the long wavelength limit. The radiative contributions to the dyadics are treated exactly. The non-radiative contributions are approximated by their static limit and leading dynamic corrections, which significantly extend the range of validity of the
static approximation. This chapter thus provides all the tools necessary to analyze 2d magneto-electric quadrupolar arrays within one formalism that leads to easy calculations, and identifies restrictions on the approximate treatment of those systems that follow from the optical theorem.

In chapter 4, I used the formalism developed in chapter 3 to analyze the multipolar structure of radiation from a 2d array of gold spheres of moderate sizes, and its dependence on frequency and illumination conditions. I carried out the analysis within the full multipolar model of the previous chapter, including multipole moments that give non-resonant contributions to radiation when the sphere is isolated. I identified surface-lattice resonances (SLRs) that are due to a dominant dipolar response coupled to higher multipoles, and SLRs that are associated with higher multipole moments only. I showed that the coupling between the resonant and non-resonant multipole moments strongly affects the position and profile of SLRs of both types, as well as the very existence of the multipolar SLRs, resulting in its extreme sensitivity to illumination conditions. Thus even though the non-resonant multipole moments are negligible in a description of an isolated sphere, they become important in describing the collective effects in an array of spheres. However, even within the multipolar model the description of the system is still complex, and further approximations were needed to gain additional insight into the electrodynamics of the system. To gain more insight I linked SLRs to an excitation of normal modes of the array, discussing in detail the direct and in-direct excitation mechanisms with a careful consideration of the energy balance. I then linked the properties of SLRs – their position, width, and sensitivity to illumination conditions – to the mode dispersion relations. Normal modes are identified by the poles in the response functions, and their dispersion relations are easier to model than the full multipolar response of the system. I approximated the mode dispersion relations with a simplified model, which I then used to explain the extreme sensitivity of the multipolar SLRs to illumination conditions and the sudden termination at low frequencies of the mode dispersion relations associated with SLRs of both types. These characteristics of the optical response of the array are not easily understood without the consideration of its normal mode structure. While in this chapter I considered only one lattice and one size of spheres, the methods and concepts that were applied here should be useful more generally to other systems as well.

A description of the optical response of a solid at the microscopic level is more complicated than that of an array of nanoparticles. In a solid, electrons cannot be considered to be bound to individual unit cells, and thus a straightforward application of the multipolar expansion with respect to the lattice sites is not possible. In the second part of the thesis I developed strategies to describe charges and currents and polarization and magnetization fields induced in an insulator responding to electromagnetic fields, using an approach based on a gauge-invariant formulation of electron dynamics.

In chapter 5, I presented a gauge-invariant Green function formalism introduced by means of a Peierls phase involving an arbitrary choice of a path in space-time. I found the dynamical equations satisfied by the gauge-invariant Green function that results, extending the earlier formalism where only straight-line paths were considered. I use those generalized equations to establish a connection between the gauge-invariant Green function description and the Power-Zienau-Wooley (PZW) transformation. The PZW transformation is a highly successful theory used in atomic physics, but it relies on a choice of a special point in space. The gauge-invariant Green function description can be introduced without the use of a special point, and thus is promising for the applications to extended systems. Thus establishing a connection between those formalisms is useful in exploring the strategies to treat the optical response of solids. I found that all the gauge-invariant formulations of the system dynamics that can be introduced
within the approach of chapter 5 can be seen to belong to one of the two categories: those for which an effective Hamiltonian can be introduced, and those for which this is not possible. The PZW transformation is an example of the former. Peierls transformations that do not make use of the special point in space belong to the latter. For Peierls phases such that a global effective Hamiltonian with polarization and magnetization fields can be introduced, I discuss the limitations on these fields by considering an Aharonov-Bohm scenario. I show that in extended systems, in contrast with isolated atoms, it is not possible to construct an effective Hamiltonian with polarization and magnetization fields which would be nonzero only over atomic distances from the charge distributions responsible for them. Since only “nonlocal” polarization and magnetization fields are possible, and they are of dubious physical significance, a description based on Peierls phases that do not lead to an effective global Hamiltonian is more promising for the description of extended systems.

In chapter 6, I used the formalism developed in chapter 5 to treat the linear response of an insulator to electromagnetic fields. I chose a Peierls phase that corresponds to a straight-line integration path in space. For this choice of the path a global effective Hamiltonian cannot be introduced. However, projecting the gauge-invariant Green function on lattice sites with the use of Wannier functions, I could locally reformulate the global Peierls phase as the phase corresponding to the PZW transformation carried out with respect to those lattice sites, plus corrections. The microscopic charge and current densities that result are given by expressions such that their expansion around the lattice sites follows in a natural way, and in the limit of non-overlapping molecules they immediately reduce to the standard expressions obtained with the PZW theory. The charge and current densities were found for arbitrary electromagnetic fields. Restricting then my attention to static and uniform electromagnetic fields, I showed that there is a natural way to introduce microscopic polarization and magnetization fields associated with each lattice site, accounting for the contributions to magnetization that are associated with the motion of charge between the sites. When these expressions are complemented with the usual spatial averaging procedure, I recovered within one formalism all the response coefficients found earlier with other methods: I find the electric susceptibility, magnetic susceptibility, and the “mixed” susceptibilities that describe the response of polarization to magnetic field and the response of magnetization to electric field. While these were found earlier, previous approaches were based on macroscopic and often thermodynamic considerations. The approach introduced in chapter 6, however, is fully dynamical and makes no reference to energy or thermodynamic potentials.

A generalization of polarization and magnetization fields to a description of solids interacting with time- and spatially-varying electromagnetic fields is beyond the scope of this thesis. Nevertheless, in chapters 5 and 6 I initiated a first step towards such a generalized description of solids, by formulating the description of electron dynamics in arbitrary fields and identifying the electron correlation function that results up to linear order in those fields. Understanding the dynamics of charges is important, as for time-varying fields polarization and magnetization fields need to be introduced based on electrodynamic rather than on thermodynamic considerations. However, identifying the electron correlation function and the microscopic charges and current that result is not sufficient to find the expressions for susceptibility tensors, which enter the usual phenomenological theories of the optical response. It remains an open question whether or not microscopic polarization and magnetization fields can be identified that would generalize the polarization and magnetization introduced in chapter 6 to a description of a solid in arbitrary electromagnetic fields, and whether or not these would provide a useful description of a solid in this more general case. Here one of the challenges is related to the fact that polarization and
magnetization fields that can be introduced are in general non-unique. Even in a molecule, the multipole 
moments beyond the lowest non-vanishing ones depend on the choice of origin with respect to which 
the multipole expansion is introduced. Nevertheless, in theories of molecular crystals that rely on the 
introduction of higher order multipole moments it is found that the measurable quantities are expressed 
by combinations of multipole moments that are origin independent. The question of whether a similar 
scenario would be repeated when formulating a multipolar theory of regular solid is not addressed in 
this thesis. However, the success of the extended PZW approach in describing all the linear response 
properties for static and uniform fields, presented in chapter 6, is promising, and shows that building the 
description of a solid based on polarization and magnetization fields introduced from an electrodynamic 
perspective, even for more complicated excitation scenarios, is worth investigation.
Appendix A

Free-space Green functions

I use $g_0(r)$ to denote the scalar free-space Green function,

$$g_0(r) = \frac{1}{4\pi r} e^{i\omega n r},$$

(A.1)

which is a solution to the free-space Helmholtz equation,

$$[\nabla^2 + (\omega n)^2] g_0(r) = -\delta(r).$$

(A.2)

In terms of the scalar Green function, the four independent dyadic Green functions are given by the usual expressions

$$g^{Ep}_{ij}(r) = \frac{\mu}{\epsilon_0 n^2} \delta_{ij} g_0(r),$$

(A.3)

$$g^{Em}_{ij}(r) = \frac{i\omega \mu}{\epsilon_0 c} \epsilon_{isj} \partial_s g_0(r),$$

(A.4)

$$g^{Eq}_{ijk}(r) = -\frac{\mu}{\epsilon_0 n^2} \frac{1}{2} (\partial_k \delta_{ij} + \partial_j \delta_{ik}) g_0(r),$$

(A.5)

$$g^{Fq}_{ijkl}(r) = -\frac{\mu}{\epsilon_0 n^2} \frac{1}{4} (\partial_j \delta_{ik} + \partial_i \delta_{jk}) g_0(r) - \frac{\mu}{\epsilon_0 n^2} \frac{1}{4} (\partial_j \delta_{il} + \partial_i \delta_{jl}) g_0(r),$$

(A.6)

where I introduce a notation

$$\delta_{ij} = \partial_i \partial_j + (\omega n)^2 \delta_{ij},$$

(A.7)

and the remaining dyadics follow from (A.3-A.6) in an obvious manner.

The expressions for the Fourier transformed free-space Green function dyadics can be found using the Green function formalism for planar structures [72], generalized to include the quadrupole moments. At points in space such that $z \neq 0$ all the Green functions are of the form,

$$g(\kappa, z) = \tilde{g}_+(\kappa) e^{i\omega z} \theta(z) + \tilde{g}_-(\kappa) e^{-i\omega z} \theta(-z),$$

(A.8)

where $\tilde{g}_\pm(\kappa)$ are the Green function amplitudes of the upward and downward propagating electromag-
netic fields. The amplitudes of the dipole Green functions are given by [72],
\[ g_{Ep}^{±}(κ) = \frac{i\tilde{ω}^2 n μ}{2ε_0 w} (\hat{s}\hat{s} + \hat{p}_±\hat{p}_±), \]  
(A.9)  
\[ g_{Em}^{±}(κ) = \frac{i\tilde{ω}^2 n μ}{2ε_0 cw} (\hat{p}_±\hat{s} - \hat{s}\hat{p}_±). \]  
(A.10)  
The amplitudes of the quadrupole Green functions are given by
\[ g_{Eq}^{±}(κ) = \frac{ω^3 n μ}{2ε_0 w} (\hat{s}\{\hat{s}, \hat{v}_±\} + \hat{p}_±\{\hat{p}_±, \hat{v}_±\}), \]  
(A.11)  
\[ g_{Fq}^{±}(κ) = \frac{iω^4 n^2 μ}{2ε_0 w} \{\hat{s}, \hat{v}_±\} \{\hat{s}, \hat{v}_±\} + \frac{iω^4 n^2 μ}{2ε_0 w} \{\hat{p}_±, \hat{v}_±\} \{\hat{p}_±, \hat{v}_±\}. \]  
(A.12)
Appendix B

Sum rules

In this appendix I prove the equivalence of Eqs. (6.91) and (6.99). I first establish relations between the vectors $J_{n_1n_2}(k)$, $D_{n_1n_2}(k)$ (6.88,6.42), and the gauge-covariant vectors $\mathcal{J}_{n_1n_2}(k)$, $\mathcal{D}_{n_1n_2}(k)$ (6.97,6.98),

$$J_{n_1n_2}(k) = \mathcal{J}_{n_1n_2}(k) + \frac{\delta_{n_1n_2}}{2\hbar} \zeta_{n_1n_1}(k) \times \frac{\partial E_{n_1}}{\partial k} + (\zeta_{n_1n_1}(k) + \zeta_{n_2n_2}(k)) \times \frac{p_{n_1n_2}(k)}{4mc}, \quad (B.1)$$

$$D_{n_1n_2}(k) = \mathcal{D}_{n_1n_2}(k) + i\zeta_{n_1n_1}(k) \times \zeta_{n_1n_2}(k) - i\zeta_{n_2n_2}(k) \times \zeta_{n_1n_2}(k). \quad (B.2)$$

Next I use (B.1,B.2) to decompose (6.92-6.95),

$$\chi^{ij}_p = \tilde{\chi}^{ij}_p + g^{ij}_p, \quad p = a,b,c,d,$$

where $\tilde{\chi}^{ij}_p$ are the explicitly gauge-invariant contributions given by (6.100-6.103), and $g^{ij}_p$ are the remaining contributions involving the diagonal Berry connections. The sum of $g^{ij}_b$ and $g^{ij}_c$ after some manipulations can be written as

$$g^{ij}_{a+c} = \text{Re} \sum_v \int_{BZ} \frac{d^3k}{(2\pi)^3} \left\{ G^{ij}_{uv}(k) [\zeta_{uv}(k) \times \zeta_{uv}(k)] \right\}$$

$$- \frac{e^2}{m\hbar c^2} \sum_v \int_{BZ} \frac{d^3k}{(2\pi)^3} \left\{ D^{ij}_{uv}(k) \zeta_{uv}(k) \right\}, \quad (B.3)$$

where I define

$$G^{ij}_{uv}(k) = \frac{ie^2}{\hbar c^2 m} \left( L^{ij}_{uv}(k) - \frac{i}{2} \nabla \times p^{ij}_{uv}(k) \right) + \frac{ie^2}{\hbar^2 c^2} \zeta^{ij}_{uv}(k) \times \frac{\partial E_{uv}}{\partial k} + \frac{e^2}{2\hbar^2 c^2} E_{uv}(k) \zeta_{uv}(k) \times \zeta_{uv}(k), \quad (B.4)$$

and the remaining two terms are

$$g^{ij}_b = - \frac{e^2}{4mc^2} \sum_v \int_{BZ} \frac{d^3k}{(2\pi)^3} \left[ - \zeta^{ij}_{uv}(k) \zeta^{ij}_{uv}(k) + \delta^{ij} \zeta^{ij}_{uv}(k) \cdot \zeta^{ij}_{uv}(k) \right],$$

$$g^{ij}_d = e^2 \frac{i \epsilon \epsilon^{ijus}}{4\hbar^2 c^2} \sum_v \int_{BZ} \frac{d^3k}{(2\pi)^3} \frac{\partial^2 E_{uv}}{\partial k^j \partial k^s} \zeta^{ij}_{uv}(k) \epsilon^{us}_{uv}(k).$$
I now simplify expression (B.3) by noting that, although formally involving sums over all bands, in fact it really depends on the properties of the valence band manifold. To show that this is the case, I rewrite the first line of (B.3) using sum rules for the three contributions coming from each line in the definition (B.4). For the contributions from lines two and three of (B.4) I use respectively the expression for the Berry curvature tensor

\[ \partial_l \zeta_{mn}^i(k) - \partial_m \zeta_{nl}^i(k) = i \sum_n \zeta_{vn}^l(k) \zeta_{mn}^r(k) - i \sum_n \zeta_{nv}^l(k) \zeta_{mn}^r(k), \]  
\[ \text{(B.5)} \]

and the effective mass tensor sum rule

\[ m \frac{\partial^2 E_{vk}}{\partial k^m \partial k^l} = \hbar^2 \delta^{im} + i \hbar \sum_n \zeta_{vn}^m(k) p_{nv}^l(k) - i \hbar \sum_n p_{vn}^l(k) \zeta_{nv}^m(k). \]  
\[ \text{(B.6)} \]

I write the contribution from the first line of (B.4) in real space in terms of the Wannier function moments of the position and orbital angular momentum operators,

\[ Z_R(z) = z - R, \]
\[ L_R(z) = (z - R) \times p(z). \]

When the first line of (B.4) is inserted in (B.3) I find

\[ \text{Re} \sum_{nv} i \int_{BZ} \frac{d^3k}{(2\pi)^3} \left[ L_{vn}(k) - \frac{i}{2} \nabla \times p_{vn}(k) \right] i \left[ \zeta_{vn}^i(k) \times \zeta_{nv}^i(k) \right]^j \]
\[ = \text{Re} \sum_{R' R''} \frac{i}{V_c} \sum_{nv} \langle v R | \hat{L}_R^i | n R' \rangle \left( \langle v R' | \hat{Z}_R | v R \rangle \times \langle n R' | \hat{Z}_{R''} | v R'' \rangle \right)^j. \]  
\[ \text{(B.7)} \]

Recalling now that Wannier functions are complete, and summing over \( n \) and \( R' \) I find

\[ \text{Re} \sum_{nv} i \int_{BZ} \frac{d^3k}{(2\pi)^3} \left[ L_{vn}(k) - \frac{i}{2} \nabla \times p_{vn}(k) \right] i \left[ \zeta_{vn}^i(k) \times \zeta_{nv}^i(k) \right]^j \]
\[ = \frac{i}{V_c} \sum_{R''} \epsilon^{ilm} \left[ \langle v R' | \hat{Z}_{R''}^i | v R \rangle \langle v R | \hat{Z}_R \hat{L}_R^i | v R'' \rangle - R \leftrightarrow R'' \right] \]
\[ - \frac{2mc^2 \hbar}{e^2} g_{ib} - \frac{1}{2} \sum_v \int_{BZ} \frac{d^3k}{(2\pi)^3} \left[ L_{vn}^i(k) D_{vn}^j(k) + \epsilon^{ilm} \epsilon^{irs} \frac{\partial^2 \zeta_{vn}^s(k)}{\partial k^r} L_{vn}^m(k) \right], \]  
\[ \text{(B.8)} \]

where I introduced the orbital angular momentum tensor

\[ L_{vn}^{ms}(k) = \frac{1}{2} \sum_n \left[ \zeta_{vn}^m(k) p_{vn}^s(k) + \zeta_{nv}^m(k) p_{vn}^s(k) \right], \]  
\[ \text{(B.9)} \]

which can be expressed in terms of the orbital angular momentum vector (6.40) through

\[ L_{vn}^{ms}(k) = \frac{1}{2} \epsilon^{ims} L_{vn}^i(k) + \frac{m}{2}\hbar \gamma_{vn}^s(k) \frac{\partial E_{vk}}{\partial k^m} + \frac{m}{2}\hbar \gamma_{vn}^s(k) \frac{\partial E_{vk}}{\partial k^s}. \]  
\[ \text{(B.10)} \]

The terms in the first line on the right-hand side of (B.8) cancel out, which can be seen either by explicitly evaluating the Wannier function moments as \( k \)-space integrals over Berry connections and
Appendix B. Sum rules

momentum matrix elements, or more simply by expanding the periodic part of the wavefunction $u_{v\mathbf{k}}(\mathbf{x}) = \sum_{\mathbf{G}} u_{v\mathbf{k}}(\mathbf{G}) e^{i\mathbf{G} \cdot \mathbf{x}}$, and formally writing

$$1 \sum_{\mathbf{R}'} \epsilon_{jlm} \langle v| \hat{Z}_R^* | v\mathbf{R}' \rangle \langle v\mathbf{R}' | \hat{Z}_{R'}^* \hat{L}_{R'} | v\mathbf{R} \rangle = i \hbar \int_{BZ} \frac{d^3k}{(2\pi)^3} \sum_{\mathbf{G}' \mathbf{G}} \left[ \frac{\partial u_{v\mathbf{k}}^*(\mathbf{G})}{\partial \mathbf{k}} \times \frac{\partial u_{v\mathbf{k}}(\mathbf{G}')}{\partial \mathbf{k}'} \right]^j \times \left[ u_{v\mathbf{k}}(\mathbf{G})(\mathbf{k} + \mathbf{G}') \times \frac{\partial u_{v\mathbf{k}}(\mathbf{G}')}{\partial \mathbf{k}'} \right]^i.$$  

Finally I establish a relation

$$\int_{BZ} \frac{d^3k}{(2\pi)^3} \left\{ \epsilon_{iml} \epsilon_{irs} \frac{\partial \zeta_{vv}(\mathbf{k})}{\partial \mathbf{k}'} \times \zeta_{nv}(\mathbf{k}) \right\}^j T_{vn}^i(\mathbf{k})$$

$$= \int_{BZ} \frac{d^3k}{(2\pi)^3} \left\{ mc J_{vv}(\mathbf{k}) D_{vv}(\mathbf{k}) - D_{vv}(\mathbf{k}) X_{vv}(\mathbf{k}) + \frac{m}{\hbar} \epsilon_{iml} \epsilon_{irs} \zeta_{vv}(\mathbf{k}) \zeta_{vv}(\mathbf{k}) \frac{\partial^2 E_{v\mathbf{k}}}{\partial \mathbf{k}' \partial \mathbf{k}''} \right\},$$  

(B.11)

by using (B.10) in the first line. Using now (B.11) together with (B.8) I arrive at the third “sum rule”

$$\text{Re} \sum_{nv} i \int_{BZ} \frac{d^3k}{(2\pi)^3} \left\{ [\zeta_{vv}(\mathbf{k}) \times \zeta_{nv}(\mathbf{k})]^j T_{vn}^i(\mathbf{k}) \right\}$$

$$= \frac{1}{2} \sum_{v} \int_{BZ} \frac{d^3k}{(2\pi)^3} \left\{ D_{vv}(\mathbf{k}) X_{vv}(\mathbf{k}) \right\} + \frac{mc^2 \hbar}{\epsilon^2} \left[ \chi_{ei}^i - 2g_{ij}^l - 2g_{ij}^d \right],$$  

(B.12)

with $T_{vn}(\mathbf{k}) = L_{vn}(\mathbf{k}) - \frac{i}{2} \nabla \times p_{vn}(\mathbf{k})$. Using now (B.5,B.6,B.12) in (B.3) I arrive at

$$g_{ij}^{l+c} = \tilde{\chi}_{ei}^i - g_{ij}^l - g_{ij}^d.$$  

(B.13)

Adding now all the contributions to susceptibility, I get (6.99).
Bibliography


