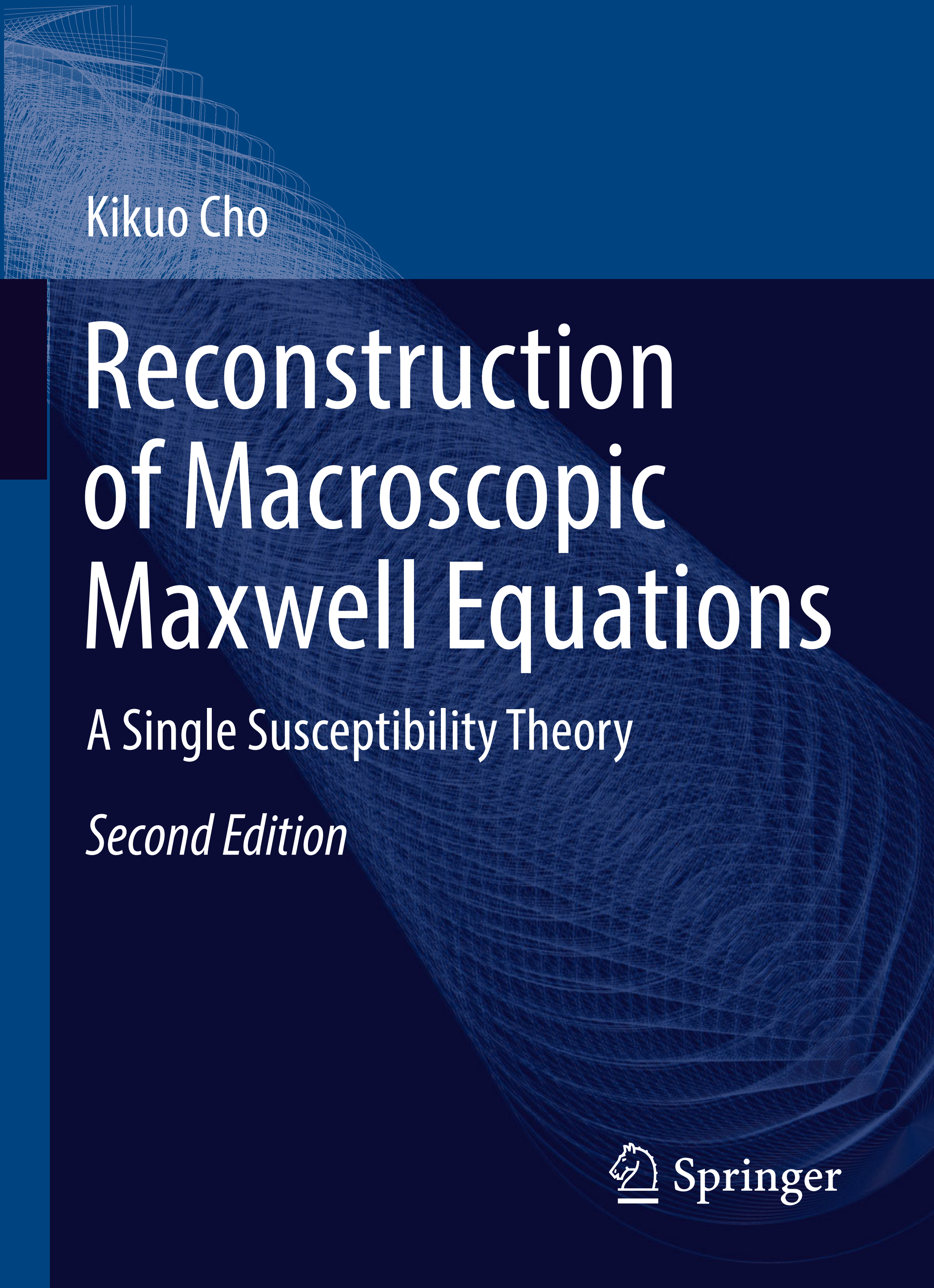
Springer Tracts in Modern Physics 237

Kikuo Cho

Reconstruction of Macroscopic Maxwell Equations

A Single Susceptibility Theory

Second Edition



Springer

Springer Tracts in Modern Physics

Volume 237

Springer Tracts in Modern Physics provides comprehensive and critical reviews of topics of current interest in physics. The following fields are emphasized:

- Elementary Particle Physics
- Condensed Matter Physics
- Light Matter Interaction
- Atomic and Molecular Physics
- Complex Systems
- Fundamental Astrophysics

Suitable reviews of other fields can also be accepted. The Editors encourage prospective authors to correspond with them in advance of submitting a manuscript. For reviews of topics belonging to the above mentioned fields, they should address the responsible Editor as listed in “Contact the Editors”.

More information about this series at http://www.springer.com/series/426

Kikuo Cho

Reconstruction of Macroscopic Maxwell Equations

A Single Susceptibility Theory

Second Edition

Springer

Kikuo Cho
Toyota Physical and Chemical
Research Institute
Aichi, Japan

ISSN 0081-3869 ISSN 1615-0430 (electronic)
Springer Tracts in Modern Physics
ISBN 978-3-662-58423-1 ISBN 978-3-662-58424-8 (eBook)
https://doi.org/10.1007/978-3-662-58424-8

Library of Congress Control Number: 2018962386

1st edition: © Springer-Verlag Berlin Heidelberg 2010
2nd edition: © Springer-Verlag GmbH Germany, part of Springer Nature 2018
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer-Verlag GmbH, DE part of Springer Nature
The registered company address is: Heidelberger Platz 3, 14197 Berlin, Germany

Preface to the Second Edition

Since the publication of the first edition, the author has felt it necessary, on various occasions of giving lectures and having discussions, to add several arguments reinforcing the general background of the theory. On writing the revision, the general scheme is kept unchanged. The addition is made on the following points.

(1) Gauge Invariance of Many-Body Schrödinger Equation (Sect. 2.2.5)

In the first edition, the micro- and macroscopic susceptibilities are given by explicitly assuming Coulomb gauge. Though this is a commonly used procedure, the gauge invariance character, which exists at the level of the minimal coupling Lagrangian for starting the argument, is not manifest in the form of many-body Schrödinger equation to perform perturbation calculation. This might give a (wrong) impression that the expressions of the susceptibilities are valid only in Coulomb gauge. The revised argument shows that the minimal coupling Lagrangian for an arbitrary gauge can be rewritten, through a variable transformation of the Lagrangian, into a form where the gauge dependent components of EM field appear only in a total time-derivative term, which can safely be omitted according to the least-action principle. The remaining Lagrangian is that of Coulomb gauge. Since this is the result of more general rule of least-action principle than choosing a gauge, its use is guaranteed to lead to gauge invariant results.

(2) Consequences of Relativistic Corrections (Sect. 2.2.6)

Typical aspects of theoretical description of EM response are discussed in the presence of spin degree of freedom with examples from atomic spectroscopy, impurity spectroscopy in crystals, and excitons in semiconductors. As a special topic of this category, a brief introduction is given about emergent electromagnetism, which is a new fundamental approach to the Pauli equation from the viewpoint of general gauge theory. Its relevance to the present theory of micro- and macroscopic EM response has not yet been studied, but, in view of the general nature of the two formulations, the inclusion of this subject is expected to be a seed of future studies.

(3) Comparison of Three Single-Susceptibility Theories (Sect. 3.5.2)

In this book, the constitutive equation to couple with Maxwell equation is the induced current density as a function of source EM field. Since current density consists of electric polarization and magnetization, one may define other types of constitutive equations. Landau-Lifshitz proposed to use (i) "generalized polarization" including $\boldsymbol{M}$ instead of current density. Similarly, (ii) "generalized magnetization" including $\boldsymbol{P}$ is possible. Examining their mutual relationship, we can show their mutual transformation properties, common dispersion equation, and the quantum mechanical susceptibility tensor in each case.

(4) Dispersion Relation in Chiral Media (Sect. 3.4 Latter Half and Sect. 4.1.5)

A comparison of chiral constitutive (ChC) equations of this book with the conventional Drude-Born-Fedorov (DBF) constitutive equations is made in terms of the dispersion curves near a chiral resonance. It shows a qualitative difference in the analytical form of dispersion equation, and, for a particular model, in the ability and inability of reproducing linear crossing at $k = 0$, indicating the superiority of ChC over DBF equations.

(5) Velocity Gauge Versus Length Gauge (Sect. 5.3.2)

In the EM response theory of matter, there has been a long debate about the right choice of interaction term, $-\boldsymbol{P} \cdot \boldsymbol{E}$ or $(e/c)\boldsymbol{p} \cdot \boldsymbol{A}$, which are often called length gauge and velocity gauge, respectively. These forms are known to be mutually transformed via the Power-Zienau-Woolley transformation in the Lagrangian, as discussed in Sect. 5.3.1 (Sect. 5.3 of original version). The additional argument in Sect. 5.3.2 is that the generating functional of PZW transformation is similar but not equal to that of gauge transformation. This means that the naming of length and velocity gauge is not appropriate.

(6) Static Polarization in the Present Framework (Sect. 5.7.2, Latter Half)

The present scheme relies on the constitutive equation between current density $\boldsymbol{J}$ and EM field. In the familiar relation $\boldsymbol{J} = \partial \boldsymbol{P}/\partial t + c\nabla \times \boldsymbol{M}$, the first term vanishes in static case. This might suggest that static polarization cannot be described by this theory. But we can show that the $\omega \to 0$ limit of the constitutive equation for $\boldsymbol{J}$ leads to that of static $\boldsymbol{P}$ correctly.

The remaining parts of the book are essentially same as the first edition, except for the minor changes related to the revision mentioned above.

Some of the research works of the added contents were supported in part financially by the Grant-in-Aid (No. 22540338) of the Ministry of Education, Sports, Culture, Science and Technology of Japan. Also, the supports by Yamada

Science Foundation (in 2014) and Toyota Physical and Chemical Research Institute (in 2016) are deeply appreciated for allowing the author to organize two workshops on themes closely related to fundamental EM responses, which is one of the motivations for this revision.

Kobe, Japan
September 2018

Kikuo Cho

Preface to the First Edition

Throughout my whole career including student time, I have had a feeling that leaning and teaching electromagnetism, especially macroscopic Maxwell equations (M-eqs) is difficult. In order to make a good use of these equations, it seemed necessary to be able to use certain empirical knowledges and model-dependent concepts, rather than pure logics. Many of my friends, colleagues and the physicists I met on various occasions have expressed similar impressions. This is not the case with microscopic M-eqs and quantum mechanics, which do not make us reluctant to teach because of the clear logical structure.

What makes us hesitate to teach is probably because we have to explain what we ourselves do not completely understand. Logic is an essential element in physics, as well as in mathematics, so that it does not matter for physicists to experience difficulties at the initial phase, as far as the logical structure is clear. As the well-known principles of physics say, "a good theory should be logically consistent and explain relevant experiments". The reason why we feel reluctant to teach macroscopic M-eqs may be related with some incompleteness of their logical structure.

There seem to have been explicit and implicit arguments about the problematic points of macroscopic M-eqs with respect to the uniqueness and consistency. A most frequent question I heard was how to uniquely separate total current density into the true and polarization charge densities. A similar problem of non-uniqueness exists when we divide transverse current density into the contributions of electric and magnetic polarizations. Also, there has been no answer to the question, "why do we need two susceptibility tensors in macroscopic M-eqs, while we need only one in microscopic response?". Further, it is strange that no general expression of magnetic permeability, except for the case of spin resonance, is known, while there are many general descriptions of dielectric function.

The present author has devoted himself to the studies of light-matter interaction and optical science, where M-eqs play an essential role. The main effort has been spent for the construction of microscopic nonlocal response theory. The result is published in a book "Optical Response of Nanostructures: Microscopic Nonlocal Theory" (Springer Verlag, 2003), where I intended to give a clear description of a well-founded microscopic semi-classical theory of light-matter interaction. Through the construction of this microscopic nonlocal response theory, we have established a deeper understanding of the hierarchical structure of the electromagnetic response theories as (i) quantum electrodynamics (QED), (ii) microscopic nonlocal response theory, and (iii) "macroscopic local response theory" with the descending accuracy in this order. The main application of this theory has been to the studies of nanostructures, which have sensitive dependence on the size, shape and internal structure of matter. An entirely new direction of its application, as the basis of deriving macroscopic M-eqs in a logically more complete fashion, was born, when I heard a talk on metamaterials in a research meeting some years ago. More specifically, I thought it feasible, as a new method of derivation of macroscopic M-eqs, to apply long wavelength approximation to the fundamental equations of the microscopic nonlocal response theory. The result was expected to be more reliable than the conventional ones because the microscopic theory is built from the first principles.

What is the "derivation" of the macroscopic from the microscopic M-eqs? A reasonable answer would be to extract the relations among the macroscopic (long wavelength) components of the relevant dynamical variables from the microscopic motions of charged particles and the microscopic M-eqs. The logically correct way to do so is to apply the approximation for macroscopic averaging to reliable microscopic equations. Thereby, it is important not to fix the goal of the argument beforehand. In many textbooks dealing with the derivation of macroscopic M-eqs, it is argued how one derives the "known" form of macroscopic M-eqs from the microscopic equations of matter and electromagnetic (EM) field. To fix the result of argument from the beginning is logically dangerous, because it may lead to an insufficient check of the validity condition of each step of the argument. In fact, the macroscopic M-eqs obtained in this book by a new method of derivation has a more general form than the conventional ones, and the former reduces to the latter only under a certain limited condition, which has nothing to do with macroscopic averaging. The new form of macroscopic M-eqs is free from all the problematic points of the conventional form with respect to the uniqueness and consistency. This is a relief of the long standing discomfort.

Although I believe that the logical structure of the new derivation is more complete than many previous arguments, I would still need to fight with a big pile of historical facts and arguments before the new result is widely accepted in the physics communities. Since the initial phase of this study, I have had a plenty of chances to discuss with experts personally and to give talks in various seminars and conferences for domestic and international audience. On such occasions, I did not encounter any embarrassing questions and comments, which require a fundamental change in my theory. Some gave me very positive comments and advices, but many others remained silent. This reaction is understandable, if we consider the rebelling aspect of this work against the well accepted knowledge of physics community. Some of my friends and colleagues made comments, with the tone of warning, such as "Isn't it bold ?", "You are brave" or "Retired professors tend to be interested in such a problem".

In order to make this theory acceptable to the physics community, the study of historical aspects would certainly be important because there are long accumulated results of the very successful conventional macroscopic M-eqs, with which the new theory must coexist. Since the author's knowledge on such historical aspects is limited, I would very much like to have readers collaboration. If a reader knows or finds a past argument which might be in conflict with the present theory, please bring it to my attention for further considerations.

In constructing this theory, I have been indebted to Profs. K. Shimoda, K. Ohtaka, F. Bassani, G. La Rocca, W. Brenig, M. Saitoh, and M.-A. Dupertuis for useful discussions. Especially, the very positive comment of Prof. Bassani, who passed away last year to my great regret, was quite encouraging. This work started almost at the same time when the author moved into Toyota Physical and Chemical Research Institute (TPCRI) in 2006. Its unique founding policy since 1940, allowing a very wide range of research works of fundamental and applicational nature, has been a great spiritual support of this work. Financially, this work was supported in part by TPCRI, and by the Grant-in-Aid (No.18510092) of the Ministry of Education, Sports, Culture, Science and Technology of Japan. Finally, I would like to thank my wife Satsuki for her continual support of my life as a physicist.

Nagakute Kikuo Cho
December 2009

Contents

Abbreviations

ABC	Additional boundary condition
AG	Agranovich–Ginzburg
BC	Boundary condition
ChC	Chiral constitutive
DBF-eqs	Drude-Born-Fedorov equations
E1	Electric dipole
E2	Electric quadrupole
EE	Emergent electromagnetism
EM	Electromagnetic
IK	Il'inskii-Keldysh
LHM	Left handed materials
L	Longitudinal
l.h.s.	Left hand side
LL	Landau–Lifshitz
LO	Longitudinal optical
LW	Long wavelength
LWA	Long wavelength approximation
M1	Magnetic dipole
M-eqs	Maxwell equations
PZW	Power-Zienau-Woolley
QED	Quantum electrodynamics
r.h.s.	Right hand side
SPR	Smith–Purcell radiation
SRR	Split ring resonator
SS	Self-sustaining
SU(2)	Special unitary (group of order) 2
T	Transverse

Notations

The meaning of the notations of physical quantities is tabulated. The choice is made for the frequently appearing ones in the text. Those limited only to a particular section are omitted.

1. Current densities

    - $\boldsymbol{J}$: (orbital) current density
    - $\boldsymbol{J}_{\mathrm{orb}}$: orbital current density
    - $\boldsymbol{J}_{\mathrm{s}}$: current density due to spin magnetization
    - $\boldsymbol{J}_0$: $\boldsymbol{A}$-independent part of current density
    - $\boldsymbol{I}_{\mathrm{t}}$: total current density, $= \boldsymbol{J}_{\mathrm{orb}} + \boldsymbol{J}_{\mathrm{s}}$
    - $\boldsymbol{I}$: $\boldsymbol{A}$-independent part of $\boldsymbol{I}_{\mathrm{t}}$
    - $\boldsymbol{I}_{\mathrm{t}}^{(\mathrm{T})}$: transverse component of $\boldsymbol{I}_{\mathrm{t}}$
    - $\boldsymbol{I}_{\mathrm{t}}^{(\mathrm{L})}$: longitudinal component of $\boldsymbol{I}_{\mathrm{t}}$
    - $\boldsymbol{I}_{\mathrm{eE}}$: $\boldsymbol{P}$ component of current density induced by electric field $\boldsymbol{E}$
    - $\boldsymbol{I}_{\mathrm{eB}}$: $\boldsymbol{P}$ component of current density induced by magnetic field $\boldsymbol{B}$
    - $\boldsymbol{I}_{\mathrm{mB}}$: $\boldsymbol{M}$ component of current density induced by magnetic field $\boldsymbol{B}$
    - $\boldsymbol{I}_{\mathrm{mE}}$: $\boldsymbol{M}$ component of current density induced by electric field $\boldsymbol{E}$

2. Polarizations

    - $\boldsymbol{P}_{\mathrm{E}}$: electric polarization induced by electric field $\boldsymbol{E}$
    - $\boldsymbol{P}_{\mathrm{B}}$: electric polarization induced by magnetic field $\boldsymbol{B}$
    - $\boldsymbol{M}_{\mathrm{E}}$: magnetic polarization induced by electric field $\boldsymbol{E}$
    - $\boldsymbol{M}_{\mathrm{B}}$: magnetic polarization induced by magnetic field $\boldsymbol{B}$
    - $\boldsymbol{P}_{\mathrm{ET}}$: electric polarization induced by transverse electric field $\boldsymbol{E}^{(\mathrm{T})}$
    - $\boldsymbol{P}_{\mathrm{EL}}$: electric polarization induced by longitudinal electric field $\boldsymbol{E}^{(\mathrm{L})}$
    - $\boldsymbol{M}_{\mathrm{ET}}$: magnetic polarization induced by transverse electric field $\boldsymbol{E}^{(\mathrm{T})}$
    - $\boldsymbol{M}_{\mathrm{EL}}$: magnetic polarization induced by longitudinal electric field $\boldsymbol{E}^{(\mathrm{L})}$

3. Susceptibilities

- $\chi_{\rm cd}$: microscopic nonlocal susceptibility
- $\chi_{\rm em}$: macroscopic susceptibility derived from $\chi_{\rm cd}$
- $\chi_{\rm em}^{\rm (T)}$: the component of $\chi_{\rm em}$ producing transverse field
- $\chi_{\rm em0}$: $O(k^0)$ term of $\chi_{\rm em}$
- $\chi_{\rm em1}$: $O(k^1)$ term of $\chi_{\rm em}$
- $\chi_{\rm em2}$: $O(k^2)$ term of $\chi_{\rm em}$
- $\chi_{\rm eE}$: electric susceptibility due to $\boldsymbol{E}$
- $\chi_{\rm eB}$: electric susceptibility due to $\boldsymbol{B}$
- $\chi_{\rm mE}$: magnetic susceptibility due to $\boldsymbol{E}$
- $\chi_{\rm mB}$: magnetic susceptibility due to $\boldsymbol{B}$
- $\chi_{\rm JEL}$: susceptibility due to external longitudinal field

4. Hamiltonians

- $\boldsymbol{H}_{\rm EM}$: Hamiltonian of vacuum EM field
- $\boldsymbol{H}_{\rm Mem}$: Hamiltonian of charged particles in a EM field
- $\boldsymbol{H}_0$: matter Hamiltonian without spin part
- $\boldsymbol{H}^{(0)}$: matter Hamiltonian with relativistic corrections
- $\boldsymbol{H}_{\rm int1}$: matter-EM field interaction, linear part
- $\boldsymbol{H}_{\rm int2}$: matter-EM field interaction, quadrutic part
- $\boldsymbol{H}_{\rm int}$: linear matter-EM field interaction including spin
- $\boldsymbol{H}_{\rm sZ}$: spin Zeeman Hamiltonian

5. Energies

- $\boldsymbol{E}_{\rm int}$: interaction energy of two current densities via EM field
- $\boldsymbol{E}_{\rm int}^{\rm (T)}$: $E_{\rm int}$ via transverse EM field
- $\boldsymbol{E}_{\rm int}^{\rm (L)}$: $E_{\rm int}$ via longitudinal EM field
- $\mathscr{A}_{\mu\sigma,\nu\tau}$: radiative interaction between two current densities

6. Green functions

- $\boldsymbol{G}_q$: scalar Green function of vacuum EM field, $q = \omega/c$
- $\mathbf{G}_q$: tensor Green function of vacuum EM field, $q = \omega/c$
- $\mathbf{G}_q^{\rm (T)}$: the part of $\mathbf{G}_q$ producing transverse field
- $\mathbf{G}_q^{\rm (L)}$: the part of $\mathbf{G}_q$ producing longitudinal field

7. Others

- q: light wave number in vacuum, $q = \omega/c$
- $v_{\rm g}$: group velocity

Chapter 1
Introduction

Abstract On the basis of traditional form of macroscopic Maxwell equations, several inherent problems are pointed out. In order to find its logically more complete form, it is argued to derive it from a reliable higher-rank theory through long wavelength approximation, i.e., to put it in an appropriate position within the single hierarchy of all the EM response theories interrelated via a series of valid approximations.

1.1 Purpose of the Book

Maxwell equations (M-eqs) are the essence of electromagnetic theory, consisting of a set of Gauss laws for electricity and magnetism, Ampère law and Faraday law. They have played one of the main roles in the tremendous development of physics in the last century. There are two sets of M-eqs, i.e., microscopic and macroscopic M-eqs. Historically, the latter appeared first and the former was derived from the latter according to the particle picture of matter. The former is used as one of the basic set of equations to construct quantum electrodynamics (QED). The agreement between the prediction of QED and related experiment is quite high in accuracy, which guarantees the reliability of its constituent theories, quantum mechanics, relativity, and microscopic M-eqs.

The macroscopic M-eqs, an approximate form of the microscopic M-eqs, have been quite successfully applied to a vast range of macroscopic phenomena including both fundamental and applicational problems, so that they have been well accepted by most research people. Still today they are indispensable as an essential tool in various research fields such as metamaterials, left-handed systems, near field optics, photonic crystals, etc., and they are also of basic importance as a curriculum in physics.

Since M-eqs describe the relationship between electromagnetic (EM) field and the dynamical variables of matter, i.e., charged particles, all the EM phenomena are governed by, not only M-eqs, but also Schrödinger (or Dirac, Newton) equations. The diversity of the EM phenomena is endless through that of matter. Since the proposal of the M-eqs in the latter half of the 19th century, various aspect of matter-

K. Cho, *Reconstruction of Macroscopic Maxwell Equations*,
Springer Tracts in Modern Physics 237,
https://doi.org/10.1007/978-3-662-58424-8_1

EM field coupled systems have been studied, but we still find new problems in both fundamental and applicational phenomena.

The research subjects mentioned above (metamaterials, etc.) are those for macroscopic M-eqs. A common central feature of metamaterials, left-handed systems and multi-ferroic systems is the coexistence of electric and magnetic polarizations of matter. Though such an aspect existed before as individual problems, its appearance as a central feature of a group of macroscopic phenomena seems to be a new trend. This gives a motivation to re-investigate whether the macroscopic M-eqs are good enough for the study of such problems.

There are a number of attitudes toward the macroscopic M-eqs. The easiest one is to regard them as a phenomenology, as they were proposed in the 19th century, when there were neither quantum mechanics, nor relativity theory, and also the particle picture of matter was not yet well established. From this viewpoint, one does not pursue the rigorous logic and consistency of the macroscopic scheme, regarding dielectric constant ε and magnetic permeability μ just as free material parameters. The second one is to accept the conventional macroscopic M-eqs as semi-quantitatively correct scheme, admitting the standard ways of deriving macroscopic from microscopic M-eqs found in many textbooks. The third one, though seemingly a minority, finds the standard derivation logically incomplete, and requires a new one which will give solutions to a number of questionable points in the conventional macroscopic M-eqs.

The present author belongs to the third group of mind mentioned above, and therefore the aim of this book is [a] to discuss the incompleteness in the derivation of the conventional macroscopic M-eqs from the microscopic ones based on the particle picture of matter, [b] to show an alternative method of derivation and its result, and [c] to discuss the conventional form in the light of the new result. The new result allows us to find out the conditions to derive the conventional scheme, and various questionable points inherent to the conventional scheme, mentioned in Sect. 1.5, are answered in the light of the new result.

Another motivation of this work is to stress the importance of using microscopic description of matter-EM field systems as a basis of arguments from both logical and practical points of view. This is because such a microscopic theory with a sufficiently general applicability has been established rather recently (due to the popularity of nanostructure studies), so that most of the previous derivations of macroscopic M-eqs had no chance to make use of it, neither an intension to do so because all the measurements in previous time were macroscopic (see, for example, p. 1 (footnote) of [1]). As will be understood later, the use of such a microscopic theory as the basis of derivation allows us to establish a better scheme of macroscopic M-eqs in a mathematically well-defined form without loss of logical generality.

Since Galilei's time, physics has made a firm and extensive progress on the two fundamental principles, "logical consistency of theory" and "agreement between theory and experiment". In particular, the requirement of logical consistency applies to every step of any theoretical frameworks from very fundamental to applicational levels. The existing EM response theories constitute the hierarchy shown in Table 1.1, where the accuracy of each theory decreases from the top to the bottom. If we have

Table 1.1 Classification of EM response theories

(A) Main hierarchy

EM field	Matter	Theory	Applied mainly to
Quantized	Rel. q-mechanics	Rel. QED	Elementary particles
Quantized	Non-rel. q-mechanics	Non-rel. QED	Atoms
Classical	Non-rel. q-mechanics	Semiclassical	

rel. = relativistic, q-mechanics = quantum mechanics

(B) Substructure of semi-classical theory

Theory	Applied mainly to
Microscopic nonlocal theory	Atoms ∼ nanostructures
Macroscopic local theory	Macroscopic media

two theoretical frameworks $\mathscr{T}_1$ and $\mathscr{T}_2$, where $\mathscr{T}_1$ is derived from $\mathscr{T}_2$ under an approximation $\mathscr{C}$, the reliability of $\mathscr{T}_1$ depends, not only on that of $\mathscr{T}_2$, but also on a clear knowledge (including the validity condition) of $\mathscr{C}$. In the case of macro- and microscopic M-eqs, $\mathscr{T}_1$ is the conventional macroscopic M-eqs, $\mathscr{T}_2$ is the microscopic M-eqs plus the (classical or quantum) mechanics of charged particles, and $\mathscr{C}$ is the "macroscopic averaging". From the requirement of logical consistency for theories of physics, any EM response theory should belong to the hierarchy of Table 1.1(A), (B). Table 1.1(A) gives the main hierarchy and Table 1.1(B) gives the substructure inside the semiclassical theory.

In the conventional way of derivation, one looks for the arguments which reproduce the known form of macroscopic M-eqs without considering the possibility of finding a more general scheme than the known one. Another frustrating point, which will be mentioned in more detail in Sect. 1.6, is the lack of generality and unambiguous definition of $\mathscr{T}_2$ and $\mathscr{C}$. In the new derivation in this book, on the other hand, we take the fundamental equations of microscopic EM response theory [2] for $\mathscr{T}_2$, and LWA for $\mathscr{C}$, which are all physically and mathematically well defined concept and procedure without empirical knowledge and model-dependence. This leads in fact to a new macroscopic scheme with more general character than the conventional one. Figure 1.1 shows the historical developments, from the author's viewpoint, about the micro- and macroscopic M-eqs including the present one.

The remarkable simplicity and generality of the new derivation arise from the form of the constitutive equation in the microscopic nonlocal response theory, where the nonlocal susceptibility is written as a separable integral kernel in general. This feature has been utilized in the microscopic nonlocal response theory to reduce the integral equations into simultaneous polynomial (linear in the case of linear response) equations, but it is also useful in performing LWA in the microscopic constitutive equation to obtain the macroscopically averaged constitutive equation. It is not an exaggeration to say that without this separability we could not construct a general scheme of the new macroscopic M-eqs.

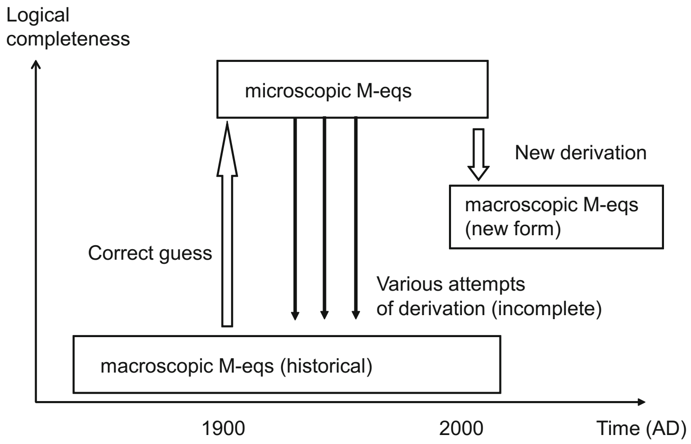


Fig. 1.1 Historical development of Maxwell equations

In the rest of this book, the author explains all the details of the background, the motivation of this study, the formulation of a new scheme, the results, the comparison with the conventional theories, and the consequences to various researches and teaching. Mathematical details and subsidiary physical aspects are given in Chap. 5, where each section is devoted to an independent subject.

As for the units system to be used in this book, we give all the numbered equations in two forms with cgs Gauss units and SI units. The latter is given in braces $\left[\cdots\right]_{\rm SI}$. It is often omitted, when it is same as, or easily derived from, those given previously. Short equations in the text are given in cgs Gauss units only to avoid congestion. The different dimension of field variables and susceptibilities in SI units, a tedious aspect of SI units in comparison with cgs Gauss units, is explicitly considered in Sect. 5.8 for the new susceptibilities defined in Sect. 3.1.

1.2 Macro- and Microscopic Maxwell Equations

The fundamental equations of electromagnetism are a set of equations to determine the electric field $\boldsymbol{E}$ and magnetic field $\boldsymbol{B}$ from a given set of charge and current densities. Their macroscopic form, established by Maxwell, Heaviside, and Hertz is the collection of Ampère law

$$\nabla \times \boldsymbol{H} = \frac{4\pi}{c}\boldsymbol{J}_{\rm c} + \frac{1}{c}\frac{\partial \boldsymbol{D}}{\partial t}, \quad \left[\nabla \times \boldsymbol{H} = \boldsymbol{J}_{\rm c} + \frac{\partial \boldsymbol{D}}{\partial t}\right]_{\rm SI}, \tag{1.1}$$

Faraday law

$$\nabla \times \boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t}, \quad \left[\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}\right]_{\rm SI}, \tag{1.2}$$

and Gauss laws of electricity and magnetism

$$\nabla \cdot \boldsymbol{D} = 4\pi\rho_{\mathrm{t}} \,, \quad \left[\nabla \cdot \boldsymbol{D} = \rho_{\mathrm{t}}\right]_{\mathrm{SI}} \,, \tag{1.3}$$

$$\nabla \cdot \boldsymbol{B} = 0 \,, \quad \left[\nabla \cdot \boldsymbol{B} = 0\right]_{\mathrm{SI}} \,, \tag{1.4}$$

respectively, where $\boldsymbol{J}_{\mathrm{c}}$ and ρ_{t} are the conduction current density and true charge density satisfying the continuity equation

$$\nabla \cdot \boldsymbol{J}_{\mathrm{c}} + \frac{\partial \rho_{\mathrm{t}}}{\partial t} = 0 \,, \tag{1.5}$$

and the field amplitudes $\boldsymbol{D}$ and $\boldsymbol{H}$ are defined as

$$\boldsymbol{D} = \boldsymbol{E} + 4\pi\boldsymbol{P} \,, \quad \left[\boldsymbol{D} = \varepsilon_0\boldsymbol{E} + \boldsymbol{P}\right]_{\mathrm{SI}} \,, \tag{1.6}$$

$$\boldsymbol{H} = \boldsymbol{B} - 4\pi\boldsymbol{M} \,, \quad \left[\boldsymbol{H} = \frac{1}{\mu_0}\boldsymbol{B} - \boldsymbol{M}\right]_{\mathrm{SI}} \,, \tag{1.7}$$

in terms of the electric polarization $\boldsymbol{P}$ and magnetization $\boldsymbol{M}$. The electric permittivity (or dielectric constant) of vacuum ε_0 and the magnetic permeability of vacuum μ_0 satisfy the relation $\varepsilon_0\mu_0 = 1/c^2$, where c is the light velocity in vacuum.

Both $\boldsymbol{P}$ and $\boldsymbol{M}$ represent the response of matter to an applied EM field, so that they have characteristic behavior of each material. For a weak EM perturbation, they are usually treated as linear functions of EM field as

$$\boldsymbol{P} = \chi_{\mathrm{e}}\boldsymbol{E} \,, \quad \left[\boldsymbol{P} = \varepsilon_0\chi_{\mathrm{e}}\boldsymbol{E}\right]_{\mathrm{SI}}, \tag{1.8}$$

$$\boldsymbol{M} = \chi_{\mathrm{m}}\boldsymbol{H} \,, \quad \left[\boldsymbol{M} = \chi_{\mathrm{m}}\boldsymbol{H}\right]_{\mathrm{SI}}, \tag{1.9}$$

where electric and magnetic susceptibilities, χ_{e} and χ_{m}, respectively, are considered to be material parameters. This kind of additional relationships to solve the M-eqs are called constitutive equations.

On the other hand, the microscopic form of M-eqs is again the collection of Ampère law

$$\nabla \times \boldsymbol{B} = \frac{4\pi}{c}\boldsymbol{J} + \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t}, \quad \left[\frac{1}{\mu_0}\nabla \times \boldsymbol{B} = \boldsymbol{J} + \varepsilon_0\frac{\partial \boldsymbol{E}}{\partial t}\right]_{\mathrm{SI}} \,, \tag{1.10}$$

Faraday law

$$\nabla \times \boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t}, \quad \left[\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}\right]_{\mathrm{SI}} \,, \tag{1.11}$$

and Gauss laws of electricity and magnetism

$$\nabla \cdot \boldsymbol{E} = 4\pi\rho \,, \quad \left[\nabla \cdot \boldsymbol{E} = \frac{1}{\varepsilon_0}\rho\right]_{\mathrm{SI}} \,, \tag{1.12}$$

$$\nabla \cdot \boldsymbol{B} = 0 \,, \quad \left[\nabla \cdot \boldsymbol{B} = 0\right]_{\mathrm{SI}} \,. \tag{1.13}$$

The charge and current densities, ρ and $\boldsymbol{J}$, respectively, are now no more continuous variables, but the summations over discrete variables corresponding to all the charged particles of matter, i.e.,

$$\rho(\boldsymbol{r}) = \sum_{\ell} \mathrm{e}_{\ell}\, \delta(\boldsymbol{r} - \boldsymbol{r}_{\ell}) \tag{1.14}$$

$$\boldsymbol{J}(\boldsymbol{r}) = \sum_{\ell} \mathrm{e}_{\ell}\boldsymbol{v}_{\ell}\, \delta(\boldsymbol{r} - \boldsymbol{r}_{\ell}), \tag{1.15}$$

where $\mathrm{e}_{\ell}, \boldsymbol{r}_{\ell}, \boldsymbol{v}_{\ell}$ are the electric charge, coordinate, and velocity, respectively, of the ℓth particle. These expressions can be used as the operator forms of these quantities in quantum mechanics of charged particles, and as operators, they satisfy the continuity equation (see Sect. 5.1)

$$\nabla \cdot \boldsymbol{J} + \frac{\partial \rho}{\partial t} = 0\,. \tag{1.16}$$

The microscopic M-eqs need also to be supplemented with a constitutive equation, which relates the induced current density $\boldsymbol{J}(\boldsymbol{r}, \omega)$ with source EM fields, which in this book are chosen transverse vector potential $\boldsymbol{A}(\boldsymbol{r}, \omega)$ and longitudinal external electric field $\boldsymbol{E}_{\mathrm{extL}}(\boldsymbol{r}, \omega)$. The characteristic point in this case is the nonlocal relationship between $\boldsymbol{J}(\boldsymbol{r})$ and source fields, through which the microscopic spatial variation is correctly taken into account reflecting the details of quantum mechanical excited states of matter. This is the core part of the microscopic nonlocal response theory, and will be described in Chap. 2.

Historically the microscopic form of M-eqs was obtained, or correctly guessed, from the macroscopic one [3], but from the hierarchical viewpoint the macroscopic one is an approximate form of the microscopic one. This recognition makes us try to derive the latter from the former via an appropriate approximation for macroscopic average. Thereby we need to consider the macroscopic forms of both the M-eqs and the constitutive equations as a set for determining EM response correctly. Using an appropriate procedure for "macroscopic averaging", we should be able to rewrite the microscopic forms into the macroscopic ones.

The M-eqs can be simplified by the use of vector and scalar potentials. The Gauss law for magnetism $\nabla \cdot \boldsymbol{B} = 0$ describes the transverse (T) nature of the vector field $\boldsymbol{B}$, so that we may introduce a vector potential $\boldsymbol{A}$ as

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}, \quad [\boldsymbol{B} = \nabla \times \boldsymbol{A}]_{\mathrm{SI}} \tag{1.17}$$

which always satisfies $\nabla \cdot \boldsymbol{B} = 0$. Inserting this into the Faraday law, we have

$$\nabla \times \left(\boldsymbol{E} + \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}\right) = 0\,, \quad \left[\nabla \times \left(\boldsymbol{E} + \frac{\partial \boldsymbol{A}}{\partial t}\right) = 0\right]_{\mathrm{SI}} \tag{1.18}$$

Since this relation claims the longitudinal (L) nature of $\boldsymbol{E} + (1/c)(\partial \boldsymbol{A}/\partial t)$, we may introduce a scalar potential ϕ to write $\boldsymbol{E}$ as

$$\boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} - \nabla\phi \;, \quad \left[\boldsymbol{E} = -\frac{\partial \boldsymbol{A}}{\partial t} - \nabla\phi\right]_{\mathrm{SI}} \;, \tag{1.19}$$

by using the identity $\nabla \times \nabla\phi = 0$.

The definition of the T and L character of a vector field $\boldsymbol{C}$ with translational symmetry is usually made in terms of its Fourier components as $\boldsymbol{k} \cdot \boldsymbol{C}_{\boldsymbol{k}} = 0$ for T, and $\boldsymbol{k} \times \boldsymbol{C}_{\boldsymbol{k}} = 0$ for L field. For a general case without translational symmetry, "$\nabla \cdot \boldsymbol{C} = 0$ for T, and $\nabla \times \boldsymbol{C} = 0$ for L field, at all points" is the generalized condition, which reduces to the usual one for translational symmetry by taking Fourier transform.

The relation between $(\boldsymbol{E}, \boldsymbol{B})$ and $(\boldsymbol{A}, \phi)$ is not unique, since the new set $(\boldsymbol{A}', \phi')$

$$\boldsymbol{A}' = \boldsymbol{A} + c\nabla\bar{\chi} \;, \quad \left[\boldsymbol{A}' = \boldsymbol{A} + \nabla\bar{\chi}\right]_{\mathrm{SI}} \;, \tag{1.20}$$

$$\phi' = \phi - \frac{\partial \bar{\chi}}{\partial t} \;, \quad \left[\phi' = \phi - \frac{\partial \bar{\chi}}{\partial t}\right]_{\mathrm{SI}} \tag{1.21}$$

in terms of an arbitrary analytic function $\bar{\chi}(\boldsymbol{r}, t)$ gives the same set of $(\boldsymbol{E}, \boldsymbol{B})$. This is called gauge transformation, and each choice of $\bar{\chi}$ defines a new gauge.

Among various choices, there are two frequently chosen cases, i.e., Coulomb gauge

$$\nabla \cdot \boldsymbol{A} = 0 \;, \quad \left[\nabla \cdot \boldsymbol{A} = 0\right]_{\mathrm{SI}} \tag{1.22}$$

and Lorentz gauge

$$\nabla \cdot \boldsymbol{A} + \frac{1}{c}\frac{\partial \phi}{\partial t} = 0 \;, \quad \left[\nabla \cdot \boldsymbol{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t} = 0\right]_{\mathrm{SI}} . \tag{1.23}$$

The M-eqs in the Coulomb gauge are given as

$$-\nabla^2\phi = 4\pi\rho \;, \quad \left[-\nabla^2\phi = \frac{1}{\varepsilon_0}\rho\right]_{\mathrm{SI}} \;, \tag{1.24}$$

$$-\nabla^2\boldsymbol{A} + \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} = \frac{4\pi}{c}\boldsymbol{J} - \frac{1}{c}\frac{\partial \nabla\phi}{\partial t} \;,$$
$$\left[-\nabla^2\boldsymbol{A} + \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} = \mu_0\boldsymbol{J} - \frac{1}{c^2}\frac{\partial \nabla\phi}{\partial t}\right]_{\mathrm{SI}} \tag{1.25}$$

and in the Lorentz gauge as

$$-\nabla^2\phi + \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} = 4\pi\rho \;, \quad \left[-\nabla^2\phi + \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} = \frac{1}{\varepsilon_0}\rho\right]_{\mathrm{SI}} \;, \tag{1.26}$$

$$-\nabla^2\boldsymbol{A} + \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} = \frac{4\pi}{c}\boldsymbol{J} \;, \quad \left[-\nabla^2\boldsymbol{A} + \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} = \mu_0\boldsymbol{J}\right]_{\mathrm{SI}} . \tag{1.27}$$

The symmetric form in Lorentz gauge is useful for the description of relativistic regime because of its apparently invariant form for Lorentz transformation.

The M-eqs in the Coulomb gauge can be split into T and L components, i.e., (1.24) is $\nabla \boldsymbol{E}^{(\mathrm{L})} = 4\pi\rho$ for the L field, and the T component of (1.25) is

$$-\nabla^2 \boldsymbol{A}^{(\mathrm{T})} + \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}^{(\mathrm{T})}}{\partial t^2} = \frac{4\pi}{c}\boldsymbol{J}^{(\mathrm{T})} \,, \quad \left[-\nabla^2 \boldsymbol{A}^{(\mathrm{T})} + \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}^{(\mathrm{T})}}{\partial t^2} = \mu_0 \boldsymbol{J}^{(\mathrm{T})}\right]_{\mathrm{SI}} \tag{1.28}$$

where the suffix T is deliberately attached to $\boldsymbol{A}$ to stress the T character. The L component of (1.25) leads, by taking its divergence, to

$$\frac{4\pi}{c}\nabla \cdot \boldsymbol{J}^{(\mathrm{L})} - \frac{1}{c}\frac{\partial \nabla^2 \phi}{\partial t} = 0 \,, \quad \left[\mu_0 \nabla \cdot \boldsymbol{J}^{(\mathrm{L})} - \frac{1}{c^2}\frac{\partial \nabla^2 \phi}{\partial t} = 0\right]_{\mathrm{SI}} \tag{1.29}$$

which is equivalent to the continuity equation (1.16) by the use of the Poisson equation (1.24).

It is noteworthy that the gauge transformation affects only the way to split $\boldsymbol{E}^{(\mathrm{L})}$ into $-\nabla\phi$ and $-(1/c)\partial \boldsymbol{A}^{(\mathrm{L})}/\partial t$, while $\boldsymbol{E}^{(\mathrm{T})}$, i.e., $\boldsymbol{A}^{(\mathrm{T})}$, remains intact. Thus the T components of M-eqs, (1.28), is not affected by the gauge transformation. It suggests the usefulness of the separate consideration of the T and L components of EM response. An additional support of this viewpoint is obtained from rewriting the self-energy of L field

$$\mathrm{H}_{\mathrm{EM}}^{(\mathrm{L})} = \frac{1}{8\pi}\int \mathrm{d}\boldsymbol{r}\, \{\boldsymbol{E}^{(\mathrm{L})}\}^2 \,, \tag{1.30}$$

$$\left[= \frac{\varepsilon_0}{2}\int \mathrm{d}\boldsymbol{r}\, \{\boldsymbol{E}^{(\mathrm{L})}\}^2\right]_{\mathrm{SI}} \tag{1.31}$$

produced by all the charged particles. Using the Gauss law $\nabla \cdot \boldsymbol{E}^{(\mathrm{L})} = 4\pi\rho$ and its solution

$$\boldsymbol{E}^{(\mathrm{L})}(\boldsymbol{r}) = -\nabla \int \mathrm{d}\boldsymbol{r}'\, \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} \quad \left[= -\frac{1}{4\pi\varepsilon_0}\nabla \int \mathrm{d}\boldsymbol{r}'\, \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\right]_{\mathrm{SI}} \tag{1.32}$$

we can rewrite $\mathrm{H}_{\mathrm{EM}}^{(\mathrm{L})}$ into the well-known form of Coulomb potential as

$$\mathrm{H}_{\mathrm{EM}}^{(\mathrm{L})} = = -\frac{1}{8\pi}\int \mathrm{d}\boldsymbol{r}\, \nabla \int \mathrm{d}\boldsymbol{r}'\, \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} \cdot \boldsymbol{E}^{(\mathrm{L})}(\boldsymbol{r}) \tag{1.33}$$

$$= \frac{1}{2}\int\int \mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{r}'\, \frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} \tag{1.34}$$

$$= \sum_{\ell}\sum_{\ell'>\ell} \frac{e_\ell e_{\ell'}}{|\boldsymbol{r}_\ell - \boldsymbol{r}_{\ell'}|}$$

$$\left[= \frac{1}{4\pi\varepsilon_0} \sum_{\ell} \sum_{\ell' > \ell} \frac{e_\ell e_{\ell'}}{|\boldsymbol{r}_\ell - \boldsymbol{r}_{\ell'}|} \right]_{\mathrm{SI}} . \tag{1.35}$$

This rewriting is gauge independent, because we use only the Gauss law and the charge density in particle picture. As discussed in Sect. 5.2.3, the choice of Coulomb gauge removes the L field from the kinetic energy term of the Hamiltonian $\sum\{\boldsymbol{p} - (e/c)\boldsymbol{A}\}^2/2m$. In this way, the T field is represented by the vector potential, and the L field is included in the Coulomb potential. The external L field is described by an external charge, and the external T field by a solution of (1.28) for $\boldsymbol{J}^{(\mathrm{T})} = 0$. In the absence of external L field, all the charges are included in the "matter", and the matter-EM field interaction is described by the (T) vector potential alone. This scheme introduces different forms of interaction term, i.e., $-\int \mathrm{d}\boldsymbol{r}\, \boldsymbol{E} \cdot \boldsymbol{P}$ for the L-field, and $(-1/c) \int \mathrm{d}\boldsymbol{r}\, \boldsymbol{J} \cdot \boldsymbol{A}$ for the T field. The main part of the microscopic response theory is constructed for the T field response with all the L component of $\boldsymbol{E}$ is incorporated in the internal field of matter. The M-eqs in this case are represented only by the single equation, (1.28), which is gauge independent. The case of excitation by external L field is described in Sect. 5.7.

1.3 Standard Derivation of Macroscopic Maxwell Equations

The standard argument to derive the macroscopic form from the microscopic one is as follows. The charge and current densities after macroscopic averaging are considered to have several components, according to which the charge density ρ consists of true and polarization charge densities, ρ_{t} and ρ_{p} as

$$\rho = \rho_{\mathrm{t}} + \rho_{\mathrm{p}} . \tag{1.36}$$

The latter represents the distortion of a neutral charge density perturbed by an electric field, and the former the remaining part of ρ in the case with net charges. Since the distortion of a neutral charge density causes an unbalance of charges, it should produce an electric polarization $\boldsymbol{P}$ in such a way as

$$\nabla \cdot \boldsymbol{P} = -\rho_{\mathrm{p}} . \tag{1.37}$$

On the other hand, the current density is caused by the motion of the charge density, which again consists of several components. One is the motion of ρ_{t} which causes $\boldsymbol{J}_{\mathrm{c}}$ $(\nabla \cdot \boldsymbol{J}_{\mathrm{c}} + \partial\rho_{\mathrm{t}}/\partial t = 0)$, and the other is the motion of $\boldsymbol{P}$ which causes polarization current density

$$\boldsymbol{J}_{\mathrm{p}} = \frac{\partial \boldsymbol{P}}{\partial t} . \tag{1.38}$$

Similarly, it is known that magnetization $\boldsymbol{M}$ with rotational structure produces a current density

$$\boldsymbol{J}_{\mathrm{M}} = c\nabla \times \boldsymbol{M} \ , \quad \left[\boldsymbol{J}_{\mathrm{M}} = \nabla \times \boldsymbol{M}\right]_{\mathrm{SI}} \ . \tag{1.39}$$

Altogether, $\boldsymbol{J}$ is the sum of the three components as

$$\boldsymbol{J} = \boldsymbol{J}_{\mathrm{c}} + \frac{\partial \boldsymbol{P}}{\partial t} + c\nabla \times \boldsymbol{M} \ , \quad \left[\boldsymbol{J} = \boldsymbol{J}_{\mathrm{c}} + \frac{\partial \boldsymbol{P}}{\partial t} + \nabla \times \boldsymbol{M}\right]_{\mathrm{SI}} \ . \tag{1.40}$$

This decomposition is consistent with the two continuity equations (1.5) and (1.16), since

$$\nabla \cdot \boldsymbol{J} = \nabla \cdot \boldsymbol{J}_{\mathrm{c}} - \frac{\partial \rho_{\mathrm{p}}}{\partial t} = -\frac{\partial(\rho_{\mathrm{t}} + \rho_{\mathrm{p}})}{\partial t} = -\frac{\partial \rho}{\partial t} \tag{1.41}$$

where we used the T character of the vector field $\nabla \times \boldsymbol{M}$, i.e., $\nabla \cdot \nabla \times \boldsymbol{M} = 0$.

The microscopic Gauss law, (1.12), after substitution of (1.36) and (1.37) becomes

$$\nabla \cdot \boldsymbol{E} = 4\pi\rho_{\mathrm{t}} - 4\pi\nabla \cdot \boldsymbol{P} \ , \quad \left[\nabla \cdot \boldsymbol{E} = \frac{1}{\varepsilon_0}(\rho_{\mathrm{t}} - \nabla \cdot \boldsymbol{P})\right]_{\mathrm{SI}} \ , \tag{1.42}$$

which is equivalent to the macroscopic Gauss law (1.3), and the microscopic Ampére law (1.10) inserted with (1.40) is

$$\begin{gathered} \nabla \times \boldsymbol{B} = \frac{4\pi}{c}(\boldsymbol{J}_{\mathrm{c}} + c\nabla \times \boldsymbol{M}) + \frac{1}{c}\frac{\partial(\boldsymbol{E} + 4\pi\boldsymbol{P})}{\partial t} \ , \\ \left[\frac{1}{\mu_0}\nabla \times \boldsymbol{B} = \boldsymbol{J}_{\mathrm{c}} + \nabla \times \boldsymbol{M} + \varepsilon_0\frac{\partial(\boldsymbol{E} + \boldsymbol{P})}{\partial t}\right]_{\mathrm{SI}} \ , \end{gathered} \tag{1.43}$$

which is equivalent to the macroscopic Ampère law (1.1).

The macroscopic variables $\boldsymbol{J}_{\mathrm{c}}$, $\boldsymbol{P}$, $\boldsymbol{M}$ represent the conduction current density due to the motion of true charge density, electric polarization, and magnetization, respectively, of the matter in consideration. They are dependent on the EM field in the matter. It is usual to introduce electric and magnetic susceptibilities, χ_{e} and χ_{m}, respectively, dielectric constant ε, and magnetic permeability μ in the regime of linear response as

$$\boldsymbol{P} = \chi_{\mathrm{e}}\boldsymbol{E} \ \left[\boldsymbol{P} = \varepsilon_0\chi_{\mathrm{e}}\boldsymbol{E}\right]_{\mathrm{SI}} \ , \tag{1.44}$$

$$\boldsymbol{M} = \chi_{\mathrm{m}}\boldsymbol{H} \ \left[\boldsymbol{M} = \chi_{\mathrm{m}}\boldsymbol{H}\right]_{\mathrm{SI}} \ , \tag{1.45}$$

$$\boldsymbol{D} = \varepsilon\boldsymbol{E} \ \left[\boldsymbol{D} = \varepsilon\boldsymbol{E}\right]_{\mathrm{SI}} \ , \tag{1.46}$$

$$\boldsymbol{B} = \mu\boldsymbol{H} \ \left[\boldsymbol{B} = \mu\boldsymbol{H}\right]_{\mathrm{SI}} \tag{1.47}$$

with additional relationship as

$$\varepsilon = 1 + 4\pi\chi_{\mathrm{e}} \ \left[\varepsilon = \varepsilon_0(1 + \chi_{\mathrm{e}})\right]_{\mathrm{SI}} \ , \tag{1.48}$$

$$\mu = 1 + 4\pi\chi_{\mathrm{m}} \ \left[\mu = \mu_0(1 + \chi_{\mathrm{m}})\right]_{\mathrm{SI}} \ . \tag{1.49}$$

As macroscopic material constants, they describe the response of individual material samples. Later we raise a question about the appropriateness of these linear response coefficients from the viewpoint of the new single susceptibility theory.

1.4 Hierarchy of EM Response Theories

There are several different theoretical schemes to describe the light-matter interaction. They are classified in the hierarchy:

(1) Relativistic QED (quantum electrodynamics)
(2) Non-relativistic QED
(3) microscopic nonlocal response theory (non-relativistic and semi-classical)
(4) macroscopic local response theory (non-relativistic and semi-classical)

The schemes (3) and (4) are semi-classical theories, where EM field is treated as classical variables. The matter variables are treated quantum mechanically in (3), and as macroscopically averaged quantities in (4). The conventional macroscopic M-eqs correspond to (4). Table 1.1 summarizes the relationship of these different schemes.

The scheme (1) is the fully quantum mechanical treatment of coupled matter-EM field system in the relativistic regime. The matter part, e.g. electrons, should be described by Dirac equation. The scheme (2) is the non-relativistic version of the scheme (1), treating the matter motion in terms of Schrödinger equation. If we treat the EM field as classical dynamical variables without introducing quantized photons, the schemes (3) and (4) arise. While the quantum mechanical motions of charged particles (in the non-relativistic regime) are precisely taken into account in the scheme (3), all the dynamical variables in the scheme (4) are treated as classical or macroscopically averaged quantities. From this sketch of the different schemes, it is obvious that the accuracy decreases according to the order from (1) to (4).

Since there is only one EM theory in physics, these four schemes are, or should be, logically related. A lower rank scheme is derived from the upper one by a certain approximation. Namely, we derive (2) from (1) by assuming that the velocity of matter particles is generally much smaller than the light velocity c. For the step from (2) to (3), we neglect the commutation relations of photon operators and describe the state of each mode in terms of a complex c-number, i.e., we replace the statistical distribution of photon states in amplitude and phase with a complex c-number for each mode. In these two cases, the logics are clear. One uses a reliable scheme as a starting point, then applies a well defined approximation to the starting scheme without presetting the resulting form. As a consequence, we find a less exact but often simpler form of theory.

As for the derivation of (4) from (3), the treatments in various textbooks of electromagnetic theory do not seem to be so logical as the cases from (1) to (2), and from (2) to (3). As we discussed in Sect. 1.3, the past derivations aimed at rewriting the microscopic M-eqs into the already known macroscopic M-eqs which had been historically established. This "derivation" was motivated by the belief that the microscopic scheme based on the particle picture and quantum mechanics, established in the 20th C., is more fundamental than the macroscopic one known since the 19th C. It is understandable to be lured to set the aim of the derivation to the search of a reasoning somehow to reproduce the known form of macroscopic M-eqs. From the logical point of view, however, it is not appropriate to fix the goal of argument from the beginning. The goal should be the result of an argument, not the aim to be fixed beforehand. If one fixes the goal at the onset, the arguments in the intermediate stage tend to be oriented toward the fixed goal. This contains a risk to miss the proper check of logical steps to be taken, e.g., whether or not the separation of $\boldsymbol{J}$ into the sum of the contributions from two independent variables $\boldsymbol{P}$ and $\boldsymbol{M}$ can be done without restrictions.

The attempt of this book to reconstruct the macroscopic Maxwell eqs from the microscopic ones is motivated by the observation mentioned above. The proposed logic to derive the scheme (4) from (3) is very simple, i.e., to apply LWA to the fundamental equations of (3). This will establish the deeper understanding of the hierarchy.

1.5 "Problems" of the Conventional Maxwell Equations

One of the problems about the standard "derivation" of the macroscopic M-eqs, mentioned in Sect. 1.3, is the non-uniqueness of splitting ρ into ρ_{t} and ρ_{p}, (1.36), and $\boldsymbol{J}$ into $\boldsymbol{J}_{\mathrm{c}}$, $\partial \boldsymbol{P}/\partial t$, and $c\nabla \times \boldsymbol{M}$, (1.40). It is possible to raise examples how these split components arise on a given model. For example, a neutral atom affected by an external electric field gives rise to a dipole moment, which contributes to $\boldsymbol{P}$. However, for a given vector field of induced current density $\boldsymbol{J}(\boldsymbol{r}, t)$, there is no general recipe, to the author's knowledge, to split it into $\boldsymbol{J}_{\mathrm{c}}$, $\partial \boldsymbol{P}/\partial t$, and $c\nabla \times \boldsymbol{M}$, and $\rho(\boldsymbol{r}, \omega)$ $\{= (-i/\omega)\nabla \cdot \boldsymbol{J}(\boldsymbol{r}, \omega)\}$ into ρ_{t} and ρ_{p}. In order for the standard derivation of macroscopic M-eqs to be logically acceptable, there should be a general recipe to make the decomposition uniquely.

The second problem is related to the first one. As a result of the splitting, we have $\boldsymbol{P}$ and $\boldsymbol{M}$, which represent the electric and magnetic properties of matter via susceptibilities χ_{e} and χ_{m}, defined in (1.44). These susceptibilities are tensors in general, and independent of each other. As functions of frequency, χ_{e} has poles at electric excitation energies, and χ_{m} at magnetic excitation energies. We need the two susceptibility tensors to describe the linear response of a matter macroscopically. However, in the microscopic M-eqs, we need only one susceptibility between $\boldsymbol{J}$ and $\boldsymbol{E}$ (or $\boldsymbol{B}$). (The charge density is related with $\nabla \cdot \boldsymbol{J}$, i.e., the L component $\boldsymbol{J}^{(\mathrm{L})}$, so that it does not require a new susceptibility.) This susceptibility has poles at all the excitation energies of matter system, and it is sufficient to have this susceptibility to describe the microscopic linear response. Thus, a question arises, why the macroscopic averaging in deriving macroscopic M-eqs increases the number of necessary susceptibility tensors? The answer to this question, like the first one, cannot be found in textbooks.

This problem is the core part of this book, about which the author claims a single susceptibility scheme of macroscopic M-eqs in contrast to the conventional scheme with two susceptibilities. It will be discussed later if we can reduce the single susceptibility of the new scheme into the two components corresponding to χ_e and χ_m. This requires us to check two points: (i) chiral and non-chiral symmetry condition, and (ii) the rewriting of interaction Hamiltonian $-(1/c)\int \mathrm{d}\boldsymbol{r}\boldsymbol{J}\cdot\boldsymbol{A}$ into the form linear in electric and magnetic fields as a legitimate procedure of analytic mechanics, which leads to the preference of $\boldsymbol{B}$ to $\boldsymbol{H}$ as source magnetic field.

The third problem is about the form of dispersion equation. In the charge neutral system ($\rho_\mathrm{t} = 0,\ \boldsymbol{J}_\mathrm{c} = 0$), the dispersion equation for a plane wave with frequency ω and wave vector $\boldsymbol{k}$ is given as

$$\frac{c^2k^2}{\omega^2} = \varepsilon\mu\ , \quad \left[\frac{k^2}{\omega^2} = \varepsilon\mu\right]_\mathrm{SI} \tag{1.50}$$

This is easily obtained by eliminating $\boldsymbol{E}$ or $\boldsymbol{B}$ from (1.2) and (1.1). The ω dependence of ε and μ is generally a superposition of single poles, according to the time dependent perturbation theory, and the poles correspond to the matter excitation energies of the electric dipole (E1) and magnetic dipole (M1) characters, respectively.

If the symmetry of matter is high, the E1 and M1 characters of excitations do not mix with each other from a symmetry ground, i.e., they belong to different irreducible representations of the group of a given symmetry. In this case, the excitation energies of E1 and M1 transitions are generally different, so that the product $\varepsilon\mu$ is a superposition of single poles. On the other hand, if the symmetry of matter mixes the E1 and M1 characters of transitions, the excitation energies of the mixed transitions contribute to both ε and μ. In this case the product $\varepsilon\mu$ would be no more a superposition of single poles, but contains second order poles. This change of the pole structure is bizarre in the linear response regime. Any excitation of matter should contribute to the response function of matter as a single poles in the lowest order time dependent perturbation theory as shown in Sect. 2.2. This is so even in the presence of an additional term of matter Hamiltonian corresponding to the lower symmetry, because after diagonalizing the total Hamiltonian we again have a series of eigenvalues which give the poles of the response function. Thus the change in the pole structure of the dispersion equation depending on the coupling or decoupling of the E1 and M1 transitions is physically unacceptable. (Later in Chap. 3 we discuss this problem from two points of view. One is the validity condition and possible extension of this dispersion equation, and the other is the correct definition of μ when this form of dispersion equation is allowed.)

One could raise another problem as an example showing the incompleteness of the conventional treatment of macroscopic M-eqs. The magnetic permeability μ represents the effect of magnetic transitions of matter. There are two well-known examples of M1 transitions. One is the spin resonances of electron, nucleus, etc., and the other is the orbital M1 transitions which induce orbital magnetic moments. The latter occurs at large variety of transition energies in various systems, such as atoms, excitons in solids, and nuclei. The famous Mössbauer line at 54.7 keV of ^{57}Fe nucleus is M1 transition, and this is why it is so sharp (10^{-7} eV width). However, the conventional ways to connect these transitions to μ (or to spectral intensity) are different.

In the case of spin resonance [4], one writes the resonant part of magnetic susceptibility as

$$\chi_{\rm m} = \frac{\beta_{\rm m}}{\omega_0 - \omega - i\gamma} \tag{1.51}$$

where $\hbar\omega_0$ is the spin flip energy, $\beta_{\rm m}$ the intensity of the magnetic transition and γ is the width of the transition energy. From this expression, μ is obtained as a ω-dependent but wave vector ($\boldsymbol{k}$) independent quantity. Such μ together with ε of matter leads to the response spectra, from which we get the resonance energy, intensity, and width.

On the other hand, the intensity of (orbital M1 + E2) transitions is calculated by expanding the matrix element of the light-matter interaction $\boldsymbol{p} \cdot \boldsymbol{A}$ under the LWA of $\boldsymbol{A} = \boldsymbol{A}_0 \exp(i\boldsymbol{k} \cdot \boldsymbol{r}) = \boldsymbol{A}_0(1 + i\boldsymbol{k} \cdot \boldsymbol{r} + \cdots)$. Omitting the first E1-active term, one gets the (M1 + E2) term as the matrix element of

$$i(\boldsymbol{p} \cdot \boldsymbol{A}_0)(\boldsymbol{k} \cdot \boldsymbol{r}) \, . \tag{1.52}$$

The matrix element of the dyadic $\boldsymbol{pr}$ becomes non-zero for (M1 + E2) transitions [5, 6], while that of $\boldsymbol{p}$ is non-vanishing for E1 transitions. Since the matrix element of (1.52) is linear in k, the intensity of this transition is $O(k^2)$. Thus, μ is proportional to k^2 in this case.

In spite of the same M1 character, the above two treatments lead to different k-dependence, $O(k^0)$ and $O(k^2)$. This difference seems to have been overlooked for a long time. However, if we consider the popularity of meta-materials or left handed systems, where the coexistence of E1 and M1 transitions leads to new subjects, we need to have a general expression of μ including both spin and orbital M1 transitions. In Sects. 3.1 and 3.2, this problem will be solved by rewriting the single susceptibility tensor of the new scheme.

There is a related question, the historical truth of which the present author has been asking to many of his friends, colleagues and teachers including experts without getting a satisfying answer. In the early days, the microscopic magnetic field was written as, not $\boldsymbol{B}$, but $\boldsymbol{H}$. The microscopic Ampère law had the form

$$\nabla \times \boldsymbol{h} = \frac{4\pi}{c}\boldsymbol{j} + \frac{1}{c}\frac{\partial \boldsymbol{e}}{\partial t} \, , \quad \left[= \boldsymbol{j} + \varepsilon_0 \frac{\partial \boldsymbol{e}}{\partial t} \right]_{\rm SI} \tag{1.53}$$

where the dynamical variables for $\boldsymbol{E}, \boldsymbol{H}, \boldsymbol{J}$ are written in the lower case letters to stress their microscopic character. In taking a macroscopic average of this equation, we often see a statement "macroscopic average of microscopic magnetic field $\boldsymbol{h}$ is usually written as $\boldsymbol{B}$" in various textbooks [1, 7, 8]. It means that, by the macroscopic averaging, we should, not only extract the LW component of $\boldsymbol{h}$, but also rewrite $\boldsymbol{H}$ into $\boldsymbol{B}$, a different physical quantity including magnetization. Without rewriting $\boldsymbol{H}$ into $\boldsymbol{B}$, we cannot obtain the macroscopic Ampère law, because of (1.7). This requirement is understandable as a mean to derive the macroscopic Ampère law, but logically not acceptable. Macroscopic averaging of a physical quantity should be the elimination of the short wavelength and the preservation of the LW components of the quantity. The rewriting of $\boldsymbol{h}$ into $\boldsymbol{B}$ contains an idea outside the macroscopic average.

This is not just a problem of semantic. It is related with the definition of magnetic susceptibility. Since the microscopic form of magnetic interaction should be

$$\mathrm{H}_{\mathrm{mag}} = -\int \mathrm{d}\boldsymbol{r}\, \boldsymbol{m} \cdot \boldsymbol{h} \,, \tag{1.54}$$

in the same way of using the lower case letters for microscopic quantities, the linear response calculation would give an induced magnetization proportional to $\boldsymbol{h}$. Its macroscopic average should lead to macroscopic constitutive equation for magnetization. In the conventional definition of macroscopic magnetic susceptibility χ_{m}, (1.9), it seems that $\boldsymbol{h}$ is simply replaced by $\boldsymbol{H}$. If one should rewrite $\boldsymbol{h}$ into $\boldsymbol{B}$ in the process of macroscopic averaging, the same replacement in the interaction Hamiltonian $\mathrm{H}_{\mathrm{mag}}$ and in the calculated result of linear response would give a constitutive equation $\boldsymbol{M} = \chi_{\mathrm{B}}\boldsymbol{B}$. This definition of magnetic susceptibility leads, by the use of $\boldsymbol{B} = \boldsymbol{H} + 4\pi\boldsymbol{M}$, to $\mu = 1/(1 - 4\pi\chi_{\mathrm{B}})$. This means that the magnetic excitation energies in χ_{B} correspond to the zeros of μ, while those in χ_{m} correspond to the poles of μ. One would ask, which is the correct excitation energy of a given matter? This difference in physical picture needs to be clarified together with the form of corresponding matter Hamiltonian defining the excitation energies.

Thus, the last question is “When and how was the earlier way of writing vacuum magnetic field as $\boldsymbol{H}$ changed to the today’s form $\boldsymbol{B}$, and is it done consistently with the definition of magnetic susceptibility and μ?”. If the use of χ_{B} instead of χ_{m} were the general understanding today, this problem would not bother us very much. But χ_{m} seems to be widely used still today in various fields using macroscopic M-eqs. Most textbooks use χ_{m}, though a rare case using χ_{B} does exist [9].[1] In view of the fact that susceptibilities are not just parameters, but the quantities to be calculated quantum mechanically with their poles at the excitation energies of a well-defined matter Hamiltonian, this mismatch would lead to an essential error in the resonant region of EM response. Although the rewriting of field variables is allowed within the framework of analytical mechanics (Sect. 5.3), which leads to various sets of “matter Hamiltonian and interaction term”, the use of $\boldsymbol{E}$ and $\boldsymbol{H}$ does not fit to the

---

[1] The official documents of IUPAP and IUPAC on this subject go as follows. In the IUPAC-2007 document [Quantities, Units and Symbols in Physical Chemistry, 3rd edition, IUPAC-2007 RSC Publishing], we find $\chi = \mu_{\mathrm{r}} - 1$ in the table of Sect. 2.3, which is also given in the IUPAP-1987 document [Table 12 of Physica 146A (1987) 1–67]. This corresponds to $\chi_{\mathrm{m}} = \mu/\mu_0 - 1$ according to the notations of this book (in SI units). Since the IUPAP document has not been revised since 1987, it is the valid recommendation today by IUPAP and IUPAC to use the definition of χ_{m} as $M = \chi_{\mathrm{m}} H$. In addition to this, there is no description in these documents about the chiral susceptibility (or admittance). From the viewpoint of the present author, these documents need to be revised by taking the microscopic consideration of susceptibility into account.

well-accepted matter Hamiltonian, i.e., the sum of the kinetic and Coulomb potential energies of charged particles.

All the problems in the conventional macroscopic M-eqs seem to arise from the lack of simple logical step, i.e., the preparation of the object to be averaged in an explicit mathematical form. The arguments in Chap. 2 will show how to fill this gap from the first-principles approach.

1.6 Meaning of Macroscopic Averaging

In order to derive the macroscopic from microscopic M-eqs, we need to take a macroscopic average of the latter. However, what does a macroscopic average mean in practice? There should be a clear mathematical definition of what is the object to be averaged and how it is done. In view of the fact that the microscopic response is obtained from the solution of "microscopic M-eqs and constitutive eqs", a straightforward logic with a clear mathematical meaning would be to extract the LW components of these fundamental equations of microscopic response.

But the past derivations do not seem to follow this line of argument. The main point is how to write the constitutive equations for macroscopic variables, and for this purpose, we need to have the general form of the microscopic constitutive equations containing all the wavelength components. But it is rather recently that this kind of microscopic constitutive equations for a general matter-EM field system has become in practical use. In the former days where most of these derivations were made, one had rather used empirical or model-dependent treatments. The main stress was, not on the general nature of a model, but on the technical point of macroscopic averaging. A typical expression for this procedure was "to take an average over a distance much larger than atomic scale but smaller than the relevant wavelength of EM field". From this statement, we can guess that the coherence of matter excitations were assumed to be of atomic (or molecular) scale. The use of appropriate models for the matter excitations leads to the electric and magnetic polarizations, respectively, which can be used to derive the typical expressions of χ_e and χ_m. Within a given model, this is an acceptable treatment. What the present author believes to be the origin of the various incompleteness of the conventional M-eqs is the lack of arguments about the general, model-independent criterion to split current density into an independent sum of the components due to $\boldsymbol{P}$ and $\boldsymbol{M}$.

The microscopic M-eqs are the equations to determine $\boldsymbol{E}$ and $\boldsymbol{B}$ from the given dynamical variables of matter, ρ and $\boldsymbol{J}$, which are determined by the quantum mechanics of charged particles. Since the motion of ρ and $\boldsymbol{J}$ is affected by EM field, we have to determine $\boldsymbol{E}$, $\boldsymbol{B}$, ρ and $\boldsymbol{J}$ selfconsistently. The auxiliary equation to allow this self-consistent determination is "microscopic constitutive equation", which together with the M-eqs gives a unique solution for a given initial condition of the dynamical variables.

All of the variables $\boldsymbol{E}$, $\boldsymbol{B}$, ρ and $\boldsymbol{J}$ are generally functions of position and time, and their position dependences contain all the wavelength components if one solves the set of M-eqs and constitutive equation selfconsistently. To extract the equations for the LW components alone, should we apply LWA to all the M-eqs, (1.24) and (1.25), and the constitutive equations relating $\{\rho, \boldsymbol{J}\}$ with $\{\phi, \boldsymbol{A}\}$?

Though there is a proposal by Nelson [10] to apply LWA to the Hamiltonian of matter, which corresponds to making LWA of (1.24), we do not take this viewpoint. Since (1.24) gives the (microscopic) Coulomb potential due to a charge density, it is directly related with the quantum mechanical energy eigenvalues and eigenfunctions. Application of LWA to the microscopic Coulomb potential would make a drastic change in the eigenvalues and eigenfunctions. Then, the poles of the macroscopic response functions do not represent the quantum mechanical excitation energies.

What we actually want to have is the susceptibilities with poles corresponding to the quantum mechanically correct eigen energies of matter, and with LWA averaged spatial structure. In this sense, we apply LWA only to the matrix elements of current density operator in the constitutive equation relating $\boldsymbol{J}$ and $\boldsymbol{A}$. The concrete form of this constitutive equation is given in the next chapter.

In some cases, macroscopic average is meant to contain also a statistical average, when we consider the statistical distribution of (A) the initial ensemble of matter states in calculating its susceptibilities and/or (B) randomly located defect or impurity centers with given transition energies. In carrying out the time dependent perturbation theory to calculate the susceptibilities of matter, we need to define the initial states of matter. This can be given as an ensemble of microscopic matter states, e.g., a canonical ensemble for a given macroscopic temperature T, which leads to an expression of susceptibilities with a weighted average via the initial ensemble. The explicit expression in the next chapter is a special case of $T = 0°$ K, and a general case is described in Sect. 5.4. It is explicitly shown that this part of weighted average via the initial ensemble is not affected by LWA, i.e., the same weighting factor remains in the macroscopic susceptibility. Therefore, the macroscopic average should not contain this kind of ensemble average.

In the case of the randomly distributed impurities or defects, however, statistical average has a meaning of macroscopic average. Though the system has no translational symmetry in a microscopic sense, it can be regarded as homogeneous after taking its macroscopic average, if the distribution is uniform. Even in this case, however, one could observe scattered light due to the absence of translational symmetry, especially near the resonance, which may be ascribed to the invalid situation of macroscopic averaging.

When we make LWA to a given microscopic system, we may introduce an intermediate step of LWA in addition to the full LWA regime. An example is the resonant Bragg scattering of the inner-core excitations of a crystal. The induced current density due to the excitation of an inner core level of an atom is well-localized in comparison with the wavelength of a resonant X ray, which allows us to use LWA to the description of the induced current density at each atom site. If we assume the full LWA regime, the crystal is described as a uniform assembly of the E1, M1, E2 etc.

multipoles of the resonant inner core transitions, which does not give any Bragg scattering. Bragg scattering becomes possible when we admit that the resonant X ray has a wavelength λ comparable to the lattice constant $d_{\rm L}$ of the crystal. This corresponds to describing the crystal as a periodic array of the E1, M1, E2 etc. multipoles, where we introduce two typical lengths of matter, the size of atom $a_{\rm A}$ and the lattice constant $d_{\rm L}$, where LWA is applicable only for $\lambda \gg a_{\rm A}$. This should be called "intermediate LWA", in contrast with "full LWA" where the conditions $\lambda \gg a_{\rm A}$ and $\lambda \gg d_{\rm L}$ do not allow Bragg scattering. This will be discussed more in detail in Sect. 4.3.

A similar situation arises in metamaterials made, e.g., of an array of split ring resonators (SRR). Though the conventional treatment of such an array is done in the full LWA regime, called homogenization, it is conceivable that an intermediate LWA may bring about a new regime where the interaction among SRR's due the induced charge densities on them introduces the nonlocality of coherent excitation made of the induced current densities. This may lead to a new category of the field, "nonlocal metamaterials", an intermediate entity between nano- and macroscopic materials. More discussion will be given in Sect. 4.1.4.

References

1. van Vleck, J.H.: Theory of Electric and Magnetic Susceptibilities. Oxford University Press, New York (1932)
2. Cho, K.: Optical Response of Nanostructures: Microscopic Nonlocal Theory. Springer, Heidelberg (2003)
3. Lorentz, H.A.: The Theory of Electrons. Teubner, Leipzig (1916)
4. Slichter, C.P.: Principles of Magnetic Resonance. Harper & Row, New York (1963)
5. Oppenheimer, J.R.: Lectures on Electrodynamics. Gordon and Breach, New York (1970)
6. Elliott, R.J.: Phys. Rev. **124**, 340 (1961)
7. Landau, L.D., Lifshitz, E.M.: Electromagnetics of Continuous Media. Pergamon Press, Oxford (1960)
8. Il'inskii, Yu.A., Keldysh, L.V.: Electromagnetic Response of Material Media. Plenum Press, New York (1994)
9. Brenig, W.: Statistical Theory of Heat, p. 152. Springer, Berlin (1989)
10. Nelson, D.F.: Electric, Optic, and Acoustic Interactions in Dielectrics. Wiley, New York (1979)

Chapter 2
New Form of Macroscopic Maxwell Equations

Abstract As a one-step higher EM response theory of macroscopic M-eqs, we give a detailed description of "microscopic nonlocal response theory", allowing semiclassical treatment of quantum mechanical matter systems from macroscopic to nano (or atomic) scale. Applying long wavelength approximation to this theory, we derive first-principles macroscopic M-eqs and constitutive equation, and also the dispersion equation for the macroscopically uniform system under LWA. The new feature of susceptibility is that it contains all the effects of electric, magnetic, and chiral polarizations in a single 3×3 tensor. We newly add (i) an argument to guarantee the gauge invariance even if one uses the Lagrangian in Coulomb gauge, and (ii) some new aspects brought about by the consideration of relativistic corrections.

2.1 New Strategy for Derivation

As mentioned in the introduction, a proper derivation of macroscopic M-eqs would need a new strategy to make the whole processes of the derivation logically and mathematically well-defined, and to avoid the problems described in Sect. 1.5.

Since a reliable approximate theory can generally be obtained from a higher rank theory by applying a valid approximation, we need to choose such a theory and an approximation for the present problem of deriving macroscopic M-eqs. In the conventional derivation, this process is described as "to derive macroscopic M-eqs from the microscopic M-eqs by applying macroscopic averaging", but the mathematical procedure to do it was not quite clear in the sense mentioned above. The main point of derivation was to derive the constitutive equations for macroscopic variables from appropriate models of matter. In such a derivation, it was admitted that the induced current density $\boldsymbol{J}$ consists of the contributions of the induced electric and magnetic polarizations as $(\partial \boldsymbol{P}/\partial t) + c\nabla \times \boldsymbol{M}$. A frequently used model to calculate polarizations is an assembly of molecules, which gives a detailed description of susceptibilities through the quantum mechanical properties of molecules.

This type of argument is acceptable as an example, but may contain a risk to miss something important about its generality. In fact, when a material system has a low

K. Cho, *Reconstruction of Macroscopic Maxwell Equations*,
Springer Tracts in Modern Physics 237,
https://doi.org/10.1007/978-3-662-58424-8_2

symmetry, which does not allow to distinguish axial and polar vectors, we cannot define electric and magnetic polarizations independently. (The symmetry condition of matter has nothing to do with the macroscopic averaging. If it affects the final result, we should consider it separately.) In this case, we have to go back to the microscopic description and see how it is possible to introduce $\boldsymbol{P}$ and $\boldsymbol{M}$ from the microscopic $\boldsymbol{J}$. For this purpose, the microscopic scheme needs to be general enough to enable us to judge it. Though this kind of general scheme had not been established during the time where most of the conventional derivations of macroscopic M-eqs were made, one could have derived it via standard time dependent perturbation theory, as shown in the next sections. The lack of the motivation to do it was, to the author's opinion, the origin of the problems discussed in Sect. 1.5.

As a new strategy, we employ the recently established scheme of microscopic optical response [1] as the basic theory, to which we apply LWA and derive the new macroscopic M-eqs and constitutive eqs [2]. The formulation of this microscopic optical response theory is made in a model independent way, so that one can choose any cases of symmetry through the choice of eigen functions of quantum mechanical matter states. The merit of this scheme is that we can start with a very general form of matter Hamiltonian and matter-EM field interaction within the semi-classical and non-relativistic regime, which however allows the inclusion of relativistic correction terms, such as spin Zeeman interaction, spin-orbit interaction and so on. Therefore this scheme can cover most of the matter Hamiltonians used for materials studies in non-relativistic regime, including various effective Hamiltonians applicable to a restricted energy range used for certain selected purposes.

It should be stressed that all we use here are the well-known principles and methods of physics and mathematics, such as analytic and quantum mechanics, time dependent perturbation theory, Taylor expansion, Fourier transform, etc. No exotic or fancy method is employed. The only new aspect is the attempt to make the logics as firmly consistent as possible.

The fundamental equations of this microscopic nonlocal response theory consist of the microscopic M-eqs and a microscopic constitutive equation between current density and source EM field. All the detailed information about the material is included in the microscopic nonlocal susceptibility, containing the symmetry related aspects, which is a sufficiently general basis to answer the problems of Sect. 1.5. Another practical merit of this theory is the separability of the microscopic nonlocal susceptibility as an integral kernel, which is quite useful both for the microscopic calculation and for the application of LWA in deriving the macroscopic M-eqs.

2.2 Microscopic Nonlocal Response Theory

In this section, we give a detailed description of microscopic nonlocal response theory, from which we derive the new form of macroscopic M-eqs later by applying LWA. Though a similar description is found in [1], we give it here because it is the core part of the present theory. We try to present the description as general as possible, so that the final result can be applied to a broader range of problems. This is done by choosing the matter Hamiltonian and matter-EM field interaction in a model independent form, and their explicit spin dependence is taken into account via relativistic correction terms (spin-orbit interaction, spin Zeeman interaction, etc.). By preparing the matter Hamiltonian and matter-EM field interaction in such a general form, we can cover a wide range of problems of EM response of matter. We will be mostly concerned with linear response, since it is the main field of interest in comparing the conventional and new schemes of macroscopic M-eqs. Extension to nonlinear response will be mentioned in Sect. 4.5.

The interaction of matter and EM field may be divided into two categories according to the T and L characters of the vector fields involved. Though there can be mixing between two cases, the T-field interaction is essentially related with optical response, and the L-field interaction with the response of matter to the excitation by external charges. For this reason, and for an additional one mentioned just below, we split the formulation into two parts, and give the T-field part in this section, and the L-field part in Sect. 5.7. The second reason to split the description into two parts is that the interaction Hamiltonians for the T and L fields appear in different forms. In Coulomb gauge, the L and T components of interaction arise from the Coulomb potential and the $\boldsymbol{A}$ dependent terms of the kinetic energy, respectively. The former can be rewritten as $-\int \boldsymbol{P}^{(\mathrm{L})} \cdot \boldsymbol{E}^{(\mathrm{L})}\, \mathrm{d}\boldsymbol{r}$. The standard form of the latter is $-(1/c) \int \boldsymbol{J}^{(\mathrm{T})} \cdot \boldsymbol{A}^{(\mathrm{T})}\, \mathrm{d}\boldsymbol{r}$, and it is not possible to rewrite it into $-\int \boldsymbol{P}^{(\mathrm{T})} \cdot \boldsymbol{E}^{(\mathrm{T})}\, \mathrm{d}\boldsymbol{r}$ without distorting the matter Hamiltonian in an unusual way (see Sect. 5.3 for details).

There is an another aspect of L field to be mentioned at this point. When charged particles are excited by some external field, they induce L, as well as T, field by the change induced in their quantum mechanical states. The problem is whether we treat this L field as an external field or not. Concerning this point, we have two choices, either (I) consider it as a part of matter Hamiltonian, or (II) regard it as a component of external EM field. The interaction of this L field with the polarization of matter is generally written as the interaction energy H_{C} among the induced charge densities of matter (p. 8 of [1]). Therefore, the choice (I) or (II) means whether we keep this interaction energy as a part of matter Hamiltonian or add it to matter-EM field interaction. Depending on this choice, the energy levels of the states containing L-mode character change by the amount caused by H_{C}. Historically, this energy difference has been called by various names, such as LT splitting, depolarization shift, or electron-hole exchange energy. Their unified description in terms of the induced charge densities of the relevant modes has been given rather recently [3]. This scheme, applicable to any type of L field caused by electron or phonon systems, localized or extended states, etc., is used in Sect. 5.7. In the following section, we consider $\boldsymbol{E}^{(\mathrm{L})}$ as the internal field of matter, i.e., we take the full Coulomb interaction energy into the matter Hamiltonian.

2.2.1 *Precise Definition of "Matter, EM Field and Interaction"*

In order to answer the questions raised in Sect. 1.5, we need to define "matter, EM Field, and their interaction" as precisely as possible, because some of the problems are related with the definition of the starting Hamiltonian. For that purpose, it will be most appropriate to take the minimal coupling Lagrangian for interacting matter-EM field systems

$$\begin{aligned} L = \sum_\ell &\left\{ \frac{1}{2} m_\ell v_\ell^2 - e_\ell \phi(\boldsymbol{r}_\ell) + \frac{e_\ell}{c} \boldsymbol{v}_\ell \cdot \boldsymbol{A}(\boldsymbol{r}_\ell) \right\} \\ &+ \frac{1}{8\pi} \int \mathrm{d}\boldsymbol{r} \left\{ \left(\frac{1}{c} \frac{\partial \boldsymbol{A}}{\partial t} + \nabla\phi \right)^2 - (\nabla \times \boldsymbol{A})^2 \right\} \\ \Bigg[= \sum_\ell &\left\{ \frac{1}{2} m_\ell v_\ell^2 - e_\ell \phi(\boldsymbol{r}_\ell) + e_\ell \boldsymbol{v}_\ell \cdot \boldsymbol{A}(\boldsymbol{r}_\ell) \right\} \\ &+ \frac{\varepsilon_0}{2} \int \mathrm{d}\boldsymbol{r} \left\{ \left(\frac{\partial \boldsymbol{A}}{\partial t} + \nabla\phi \right)^2 - c^2 (\nabla \times \boldsymbol{A})^2 \right\} \Bigg]_{\mathrm{SI}} \end{aligned} \tag{2.1}$$

where $\boldsymbol{A}$ and ϕ are the vector and scalar potentials, respectively, and the integral part on the r.h.s. is the Lagrangian of vacuum EM field. It is noteworthy that the interaction part of the Lagrangian can be rewritten in the following integral form

$$\begin{aligned} &\sum_\ell \left\{ -e_\ell \phi(\boldsymbol{r}_\ell) + \frac{e_\ell}{c} \boldsymbol{v}_\ell \cdot \boldsymbol{A}(\boldsymbol{r}_\ell) \right\} = \int \mathrm{d}\boldsymbol{r} \left\{ -\rho(\boldsymbol{r})\phi(\boldsymbol{r}) + \frac{1}{c} \boldsymbol{J}(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r}) \right\} \\ &\left[\sum_\ell \{ -e_\ell \phi(\boldsymbol{r}_\ell) + e_\ell \boldsymbol{v}_\ell \cdot \boldsymbol{A}(\boldsymbol{r}_\ell) \} = \int \mathrm{d}\boldsymbol{r} \{ -\rho(\boldsymbol{r})\phi(\boldsymbol{r}) + \boldsymbol{J}(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r}) \} \right]_{\mathrm{SI}} \end{aligned} \tag{2.2}$$

where charge density ρ and current density $\boldsymbol{J}$ are defined as (1.14) and (1.15), respectively. This integral expression is useful in carrying out the least action principle for $\boldsymbol{A}$ and ϕ.

As a Lagrangian, this contains three kinds of generalized coordinates, $\boldsymbol{r}_\ell$, $\boldsymbol{A}(\boldsymbol{r})$, $\phi(\boldsymbol{r})$ and the corresponding generalized velocities $\boldsymbol{v}_\ell$, $\partial \boldsymbol{A}/\partial t$ (the time derivative of ϕ is not contained). The least action principle of Lagrangian, or Lagrange equation for each set of generalized coordinate and velocity, gives the Newton equation of motion, and microscopic M-eqs for ϕ (Poisson eq) and $\boldsymbol{A}$ (wave equation) (see Sect. 5.2). The Newton equation is

$$m_\ell \frac{\mathrm{d}\boldsymbol{v}_\ell}{\mathrm{d}t} = e_\ell \left\{ \boldsymbol{E}(\boldsymbol{r}_\ell) + \frac{1}{c}\boldsymbol{v}_\ell \times \boldsymbol{B}(\boldsymbol{r}_\ell) \right\} \tag{2.3}$$

$$= \left[e_\ell \{ \boldsymbol{E}(\boldsymbol{r}_\ell) + \boldsymbol{v}_\ell \times \boldsymbol{B}(\boldsymbol{r}_\ell) \} \right]_{\mathrm{SI}} \tag{2.4}$$

where $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ and $\boldsymbol{E} = -\nabla\phi - (1/c)\partial \boldsymbol{A}/\partial t$. The r.h.s. is the Lorentz force due to the EM field acting on the charged particle. The Poisson equation is the same as (1.24), and the wave equation for $\boldsymbol{A}$ is (1.25) (see Sect. 5.2).

The fact that this Lagrangian gives the well established equations of motion for charged particles and EM field, as mentioned above, guarantees the soundness of this Lagrangian as a basis of further developments in various directions. In fact, it is used for the (non-relativistic) QED, and now we are going to use it for the semiclassical arguments.

Hamiltonian is obtained by the standard procedure of Lagrangian formalism. Defining the generalized momentum p_ℓ for a generalized coordinate q_ℓ via $p_\ell = \partial L/\partial \dot{q}_\ell$, where $\dot{q}_\ell$ is the time derivative of q_ℓ, we derive Hamiltonian as $H = \sum_\ell p_\ell \dot{q}_\ell - L$. The details of this argument applied to the present Lagrangian are given in Sect. 5.2.3, and the Hamiltonian is given as

$$\begin{aligned} H = &\sum_\ell \frac{1}{2m_\ell} \left\{ \boldsymbol{p}_\ell - \frac{e_\ell}{c}\boldsymbol{A}(\boldsymbol{r}_\ell) \right\}^2 + U_{\mathrm{C}} \\ &+ \frac{1}{8\pi} \int \mathrm{d}\boldsymbol{r} \left[\left\{ \frac{1}{c} \frac{\partial \boldsymbol{A}^{(T)}}{\partial t} \right\}^2 + \left\{ \nabla \times \boldsymbol{A}^{(T)} \right\}^2 \right], \\ \Bigg[= &\sum_\ell \frac{1}{2m_\ell} \{ \boldsymbol{p}_\ell - e_\ell \boldsymbol{A}(\boldsymbol{r}_\ell) \}^2 + U_{\mathrm{C}} \\ &+ \frac{\varepsilon_0}{2} \int \mathrm{d}\boldsymbol{r} \left[\left\{ \frac{\partial \boldsymbol{A}^{(T)}}{\partial t} \right\}^2 + c^2 \{ \nabla \times \boldsymbol{A}^{(T)} \}^2 \right] \Bigg]_{\mathrm{SI}}, \end{aligned} \tag{2.5}$$

where U_{C} is the Coulomb potential among the particles

$$U_{\mathrm{C}} = \frac{1}{2} \sum_\ell \sum_{\ell' \neq \ell} \frac{e_\ell e_{\ell'}}{|\boldsymbol{r}_\ell - \boldsymbol{r}_{\ell'}|} \quad \left[= \frac{1}{8\pi\varepsilon_0} \sum_\ell \sum_{\ell' \neq \ell} \frac{e_\ell e_{\ell'}}{|\boldsymbol{r}_\ell - \boldsymbol{r}_{\ell'}|} \right]_{\mathrm{SI}}. \tag{2.6}$$

In this expression, we have rewritten the self-energy of the longitudinal EM field into the Coulomb potential by using the Gauss law $\nabla \cdot \boldsymbol{E} = 4\pi\rho$ and the definition of ρ in the particle picture. This form of Hamiltonian applies to any gauge, i.e., the vector potential in the first term on the r.h.s. can have the L, as well as T, component.

According to the argument in newly added Sect. 2.2.5, any Lagrangian for EM response of matter in arbitrary gauge can be reduced to that in Coulomb gauge. This is not through the choice of gauge, but through the use of general rule of variational principle. In this sense, we may always use the Schrödinger equation in Coulomb gauge for the calculation of gauge invariant EM response.

In Coulomb gauge we have $\nabla \cdot \boldsymbol{A} = 0$, i.e., $\boldsymbol{A} = \boldsymbol{A}^{(\mathrm{T})}$, $\boldsymbol{A}^{(\mathrm{L})} = 0$. Thus the vector potential appearing in the following arguments has pure T character. We omit the superscript T from $\boldsymbol{A}^{(\mathrm{T})} = 0$ hereafter, unless it is better to stress it. The Hamiltonian is a sum of two contributions. One is the Hamiltonian of vacuum EM field

$$\begin{aligned} H_{\mathrm{EM}} &= \frac{1}{8\pi}\int \mathrm{d}\boldsymbol{r}\,\left\{\left(\frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}\right)^2 + (\nabla \times \boldsymbol{A})^2\right\} \\ &\left[= \frac{\varepsilon_0}{2}\int \mathrm{d}\boldsymbol{r}\,\left\{\left(\frac{\partial \boldsymbol{A}}{\partial t}\right)^2 + c^2(\nabla \times \boldsymbol{A})^2\right\}\right]_{\mathrm{SI}} \end{aligned} \tag{2.7}$$

and the other is the Hamiltonian of the charged particles in a given EM field in the Coulomb gauge

$$\begin{aligned} H_{\mathrm{Mem}} &= \sum_\ell \frac{1}{2m_\ell}\left\{p_\ell - \frac{e_\ell}{c}\boldsymbol{A}(\boldsymbol{r}_\ell)\right\}^2 + U_{\mathrm{C}}\,. \\ &\left[= \sum_\ell \frac{1}{2m_\ell}\{p_\ell - e_\ell \boldsymbol{A}(\boldsymbol{r}_\ell)\}^2 + U_{\mathrm{C}}\,.\right]_{\mathrm{SI}} \end{aligned} \tag{2.8}$$

The matter Hamiltonian H_0 is defined as H_{Mem} for $\boldsymbol{A} = 0$, i.e.,

$$\begin{aligned} H_0 &= \sum_\ell \frac{p_\ell^2}{2m_\ell} + U_{\mathrm{C}} \\ &\left[= \sum_\ell \frac{p_\ell^2}{2m_\ell} + U_{\mathrm{C}}\right]_{\mathrm{SI}} \end{aligned} \tag{2.9}$$

which is the sum of the kinetic energy and potential energy of particles, and the matter-EM field interaction is the $\boldsymbol{A}$-dependent terms of H_{Mem}, which is the sum of the two terms $H_{\mathrm{int1}} + H_{\mathrm{int2}}$. The $\boldsymbol{A}$-linear term is

$$\begin{aligned} H_{\mathrm{int1}} &= -\frac{1}{c}\int \mathrm{d}\boldsymbol{r}\;\boldsymbol{J}_0(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r}) \\ &\left[= -\int \mathrm{d}\boldsymbol{r}\,\boldsymbol{J}_0(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r})\right]_{\mathrm{SI}} \end{aligned} \tag{2.10}$$

and the $\boldsymbol{A}$-quadratic term is

$$\begin{aligned} H_{\mathrm{int2}} &= \frac{1}{2c^2}\int \mathrm{d}\boldsymbol{r}\;\hat{N}(\boldsymbol{r})A(\boldsymbol{r})^2\,, \\ &\left[= \frac{1}{2}\int \mathrm{d}\boldsymbol{r}\;\hat{N}(\boldsymbol{r})A(\boldsymbol{r})^2\right]_{\mathrm{SI}}\,, \end{aligned} \tag{2.11}$$

where $\hat{N}$ is defined as

$$\hat{N}(\boldsymbol{r}) = \sum_{\ell} \frac{e_\ell^2}{m_\ell} \, \delta(\boldsymbol{r} - \boldsymbol{r}_\ell) \, , \tag{2.12}$$

and $\boldsymbol{J}_0$ is the $\boldsymbol{A}$ independent part of current density, i.e., (1.15) with $\boldsymbol{v}_\ell$ replaced by $\boldsymbol{p}_\ell / m_\ell$,

$$\boldsymbol{J}_0(\boldsymbol{r}) = \sum_{\ell} \frac{e_\ell}{2m_\ell} [\boldsymbol{p}_\ell \delta(\boldsymbol{r} - \boldsymbol{r}_\ell) + \delta(\boldsymbol{r} - \boldsymbol{r}_\ell) \boldsymbol{p}_\ell] \, . \tag{2.13}$$

The (orbital) current density operator (1.15) is the sum of $O(A^0)$ and $O(A^1)$ terms

$$\begin{aligned} \boldsymbol{J}_{\text{orb}}(\boldsymbol{r}) &= \boldsymbol{J}_0(\boldsymbol{r}) - \frac{1}{c} \hat{N}(\boldsymbol{r}) \boldsymbol{A}(\boldsymbol{r}) \, , \\ \Big[&= \boldsymbol{J}_0(\boldsymbol{r}) - \hat{N}(\boldsymbol{r}) \boldsymbol{A}(\boldsymbol{r}) \, \Big]_{\text{SI}} \, . \end{aligned} \tag{2.14}$$

We write a suffix "orb" to stress its orbital character and to distinguish it from the current density induced by spin magnetization to be discussed later.

When an external L field exists, it should be ascribed to an external charge density ρ_{ext} in the Coulomb gauge. This means that the charge density in U_{C} contains the internal and external parts, ρ_{int} and ρ_{ext}. In this case, there arises a new term of interaction due to the Coulomb interaction between ρ_{int} and ρ_{ext}. As discussed in Sect. 5.7, the natural form of the interaction Hamiltonian is $-\int \mathrm{d}\boldsymbol{r} \, \boldsymbol{J_0} \cdot \boldsymbol{A}$ for T field and $-\int \mathrm{d}\boldsymbol{r} \, \boldsymbol{P} \cdot \boldsymbol{E}^{(\text{L})}$ for L field, and it does not seem possible to write them in one unified form $-\int \mathrm{d}\boldsymbol{r} \, \boldsymbol{P} \cdot \boldsymbol{E}$ without distorting the matter Hamiltonian (Sect. 5.3). This point is often overlooked in the conventional EM response theories, so that it is appropriate to stress it at this stage of the present formulation. Since the calculation of the susceptibilities goes similarly in both cases, we describe the case of the T field in detail in the main text and leave the case of L field in Sect. 5.7. (For the discussion of T-field response, the relevant ρ is ρ_{int} alone and the suffix "int" will be omitted.)

For the linear response to the T field excitation, we need only H_{int1}. As an operator to be used in quantum mechanical calculation, we have symmetrized the non-commutative quantities $\{\boldsymbol{p}_\ell, \boldsymbol{r}_\ell\}$ in $\boldsymbol{J}_0(\boldsymbol{r})$. The operator $\hat{N}(\boldsymbol{r})$ has contributions from various charged particles, but, because of the factor e_ℓ^2/m_ℓ, lighter electrons make much more contribution than heavier ions. The electron term is written as $(e^2/m_0)\hat{\rho}_{\text{el}}(\boldsymbol{r})$, where $e, m_0, \hat{\rho}_{\text{el}}(\boldsymbol{r})$ are the charge, mass, and the density of electron(s), respectively.

For a given set of matter Hamiltonian and matter-EM field interaction, we can calculate the induced current density, which gives the microscopic constitutive equation. The forms of H_0 and H_{int1} given above are model independent and have a rather general character. However, in order to increase the range of their applicability, we would like to include their explicit spin dependence, which is important for the magnetic properties of matter. Paramagnetism is typically caused by localized spin states due to the spin Zeeman interaction with static magnetic field. The resonance transition between these spin levels can be induced by a microwave with corresponding frequency. This transition is caused by the (spin Zeeman) interaction between spin magnetization and microwave EM field. This interaction is also necessary to analyze the intra- and interband magneto-optics in semiconductors. In addition, the spin-orbit interaction gives rise to spin-dependent energy level structure for matter systems containing heavy atoms. These examples show the necessity of introducing the explicit spin dependence of the matter Hamiltonian and matter-EM field interaction, which leads to the realistic resonant structure of susceptibilities of spin related systems.

The explicit spin dependence of H_0 and H_{int1} arises from the relativistic correction to the non-relativistic Hamiltonian [4]. In the Dirac equation dealing with an electron in the relativistic regime, there emerges the entity "spin" by the requirement of relativistic invariance of the equation consisting of the linear terms of time and space derivatives. The expansion of the positive eigenvalue $E = E' + m_0c^2$ of the Dirac equation with respect to $(E' - V)/2m_0c^2$, where V is the potential energy of the electron, gives various correction terms. Among them we have spin-orbit interaction

$$\frac{\hbar}{2m_0^2c^2}\boldsymbol{\sigma} \cdot [(\nabla V) \times \boldsymbol{p}] \,. \tag{2.15}$$

Similar expansion for an electron in an EM potential gives the spin-Zeeman term

$$-\frac{e\hbar}{m_0c}\boldsymbol{\sigma} \cdot \boldsymbol{B}$$

$$\left[-\frac{e\hbar}{m_0}\boldsymbol{\sigma} \cdot \boldsymbol{B} \right]_{\mathrm{SI}} \tag{2.16}$$

as an additional term of the Hamiltonian in the (non-relativistic) Schrödinger equation. Here, $\hbar\boldsymbol{\sigma}$ is the spin angular momentum of an electron. (The magnetic field $\boldsymbol{H}$ in [4] is rewritten into $\boldsymbol{B}$ in accordance with our definition of magnetic field in microscopic M-eqs.)

In addition to these spin dependent correction terms, there are spin independent correction terms, such as mass velocity term (due to the velocity dependent mass correction) and Darwin term (due to the velocity-induced nonlocality of the potential V) [5].

From our viewpoint to put L field into matter Hamiltonian, the spin-orbit interaction, H_{so}, mass velocity term, and Darwin term should be included in the matter Hamiltonian H_0, and the spin Zeeman term into H_{int1}. For many electron systems, we should take a sum over all the electrons for (2.16) and (2.15). Thus, the matter Hamiltonian is now

$$H^{(0)} = H_0 + H_{\mathrm{rel-corr}} \tag{2.17}$$

where $H_{\mathrm{rel-corr}} = H_{\mathrm{so}} + H_{\mathrm{mass-v}} + H_{\mathrm{Darwin}}$.

For the new form of H_{int1}, we can rewrite the spin Zeeman term as

$$H_{\mathrm{sZ}} = -\sum_{\ell} \frac{e\hbar}{m_0c}\,\boldsymbol{\sigma}_\ell \cdot \boldsymbol{B} \,, \tag{2.18}$$

$$= -\int \mathrm{d}\boldsymbol{r}\; \boldsymbol{M}_{\mathrm{s}}(\boldsymbol{r}) \cdot (\nabla \times \boldsymbol{A}(\boldsymbol{r}))\;, \tag{2.19}$$

$$= -\int \mathrm{d}\boldsymbol{r}\; \nabla \times \boldsymbol{M}_{\mathrm{s}}(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r})\;, \tag{2.20}$$

where $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ and partial integration are used, and the spin magnetic polarization is

$$\boldsymbol{M}_{\mathrm{s}}(\boldsymbol{r}) = \sum_{\ell} \frac{e\hbar}{m_0 c}\, \boldsymbol{\sigma}_{\ell}\, \delta(\boldsymbol{r} - \boldsymbol{r}_{\ell})\;. \tag{2.21}$$

Defining spin induced current density as

$$\begin{aligned} \boldsymbol{J}_{\mathrm{s}}(\boldsymbol{r}) &= c\nabla \times \boldsymbol{M}_{\mathrm{s}}(\boldsymbol{r})\;, \\ &\Big[= \nabla \times \boldsymbol{M}_{\mathrm{s}}(\boldsymbol{r})\;, \Big]_{\mathrm{SI}} \end{aligned} \tag{2.22}$$

we can rewrite the last line of the equations for H_{sZ} as

$$\begin{aligned} H_{\mathrm{sZ}} &= -\frac{1}{c} \int \mathrm{d}\boldsymbol{r}\; \boldsymbol{J}_{\mathrm{s}}(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r})\;. \\ &\left[= -\int \mathrm{d}\boldsymbol{r}\; \boldsymbol{J}_{\mathrm{s}}(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r})\;. \right]_{\mathrm{SI}} \end{aligned} \tag{2.23}$$

Adding this term to H_{int1}, we generalize linear matter-EM field interaction as

$$\begin{aligned} H_{\mathrm{int}} &= -\frac{1}{c} \int \mathrm{d}\boldsymbol{r}\; \boldsymbol{I}(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r})\;, \\ &\left[= -\int \mathrm{d}\boldsymbol{r}\; \boldsymbol{I}(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r})\;, \right]_{\mathrm{SI}} \end{aligned} \tag{2.24}$$

where the generalized current density $\boldsymbol{I}$ is the $\boldsymbol{A}$-independent part of the total current density (the sum of orbital and spin-induced current densities)

$$\boldsymbol{I}_{\mathrm{t}}(\boldsymbol{r}) = \boldsymbol{J}_{\mathrm{orb}}(\boldsymbol{r}) + \boldsymbol{J}_{\mathrm{s}}(\boldsymbol{r})\;, \tag{2.25}$$

$$= \boldsymbol{I}(\boldsymbol{r}) - (1/c)\hat{N}(\boldsymbol{r})\boldsymbol{A}(\boldsymbol{r})\;, \tag{2.26}$$

$$= \big[\boldsymbol{I}(\boldsymbol{r}) - \hat{N}(\boldsymbol{r})\boldsymbol{A}(\boldsymbol{r})\big]_{\mathrm{SI}}\;,$$

i.e.,

$$\boldsymbol{I}(\boldsymbol{r}) = \boldsymbol{J}_0(\boldsymbol{r}) + \boldsymbol{J}_{\mathrm{s}}(\boldsymbol{r})\;, \tag{2.27}$$

where $\boldsymbol{J}_0$ is defined in (2.13). In terms of these generalized Hamiltonians with explicit spin dependence, $H^{(0)}$ and H_{int}, we can treat a broader range of problems of matter-EM field coupled systems.

It may be worth mentioning that most of the effective Hamiltonians used for various specialized purposes are in fact derived via certain approximation from the first-principles Hamiltonian discussed above. A typical example is a spin Hamiltonian for the analysis of spin resonance [6], where one looks at a very small energy range corresponding to the energy levels of the spin system in consideration, and derives an effective Hamiltonian of spin operators. Thereby, one adds a consideration on symmetry to restrict the possible invariant forms of the combinations of spin operators. The coefficients of such allowed terms are usually taken as free parameters, but one could estimate them by the perturbational calculation using the basis set of states including the abandoned ones. Besides the effective spin Hamiltonians, there are many examples of effective Hamiltonians to describe a particular properties of matter states and various interactions, such as energy band Hamiltonian with an effective one-particle potential, Heisenberg model of ferro- and antiferromagnetism, Hubbard Hamiltonian to study the electron correlation, BCS Hamiltonian for superconductivity, Frölich Hamiltonian for electron-LO phonon coupling, etc. All of them should be derivable from the first principles form of Hamiltonians $H^{(0)}$ and $H_{\rm int}$, as far as one stays in the weakly relativistic regime of charged particle systems.

2.2.2 Calculation of Microscopic Nonlocal Susceptibility

We now calculate the current density $\boldsymbol{I}(\boldsymbol{r})$ of a system of charged particles induced by the application of a T field $\boldsymbol{A}(\boldsymbol{r}, t)$ to the lowest order of $\boldsymbol{A}$. For this calculation, we only need the matter Hamiltonian $H^{(0)}$, (2.17), and the matter-EM field interaction $H_{\rm int}$, (2.24). (At this stage, $\boldsymbol{A}$ is just a T field interacting with the matter system. Later, on looking for a selfconsistent solution, it turns out to be the sum of an incident field and the one induced by the induced current density.) The induced current density is written in terms of the eigen values and eigen functions of $H^{(0)}$. In this sense, our result is model independent. Model dependence arises when we evaluate the energy eigen values and the matrix elements of current density operator for a particular system. The expression of induced current density is given in a general form, so that it can be applied to any model systems. The necessity of relativistic correction should also be judged at the stage of such an evaluation.

Let us consider the Schrödinger equation of a system of charged particles in a EM field $\boldsymbol{A}(\boldsymbol{r}, t)$

$$i\hbar \frac{\partial \Psi}{\partial t} = (H^{(0)} + H_{\rm int})\ \Psi\ . \tag{2.28}$$

Using the interaction representation $\Psi(t) = \exp(-iH^{(0)}t/\hbar)\ \tilde{\Psi}(t)$, we rewrite the Schrödinger equation as

$$i\hbar \frac{\partial \tilde{\Psi}}{\partial t} = H_{\rm int}(t)\ \tilde{\Psi} \tag{2.29}$$

where

$$H_{\text{int}}(t) = \exp(iH^{(0)}t/\hbar)\ H_{\text{int}}\ \exp(-iH^{(0)}t/\hbar)\ . \tag{2.30}$$

Assuming that the matter state was initially in its ground state of $H^{(0)}$ and the interaction was switched on adiabatically from the remote past, we can solve this equation by iteration as

$$\tilde{\Psi}(t) = \Psi_0 + \frac{-i}{\hbar}\int_{-\infty}^{t} \mathrm{d}t_1\ H_{\text{int}}(t_1)\ \mathrm{e}^{\gamma t_1}\ \Psi_0 + \cdots\ , \tag{2.31}$$

where the wave function $\tilde{\Psi}(-\infty)$ is written as Ψ_0, the ground state wave function of $H^{(0)}$. (The case of more general initial state described by an ensemble will be treated in Sect. 5.4.) The factor $\gamma = 0^+$ is a positive infinitesimal quantity, representing the adiabatic switching of the interaction at the remote past.

The induced current density is the expectation value of the total current density operator $\boldsymbol{I}_{\mathrm{t}}(\boldsymbol{r}) = \boldsymbol{J}_{\text{orb}}(\boldsymbol{r}) + \boldsymbol{J}_{\mathrm{s}}(\boldsymbol{r})(= \boldsymbol{I}(\boldsymbol{r}) - (1/c)\hat{N}(\boldsymbol{r})\boldsymbol{A}(\boldsymbol{r}))$ with respect to the wave function $\Psi(t)(= \exp(-iH^{(0)}t/\hbar)\tilde{\Psi}(t))$. Let us expand $\Psi(t)$ as

$$\Psi(t) = \sum_{\nu} a_{\nu}(t)|\nu\rangle \tag{2.32}$$

where $|\nu\rangle$ is the eigenstate of $H^{(0)}$, i.e., $H^{(0)}|\nu\rangle = E_{\nu}|\nu\rangle$. Then, for $\nu = 0$, we have

$$a_0(t) = \exp(-i\omega_0 t)\ , \quad (\omega_0 = E_0/\hbar) \tag{2.33}$$

to the lowest order of $\boldsymbol{A}$, and for $\nu \neq 0$, we have

$$a_{\nu}(t) = \frac{i}{\hbar}\ \langle\nu| \int_{-\infty}^{t} \mathrm{d}t_1 \exp(-iH^{(0)}t/\hbar)H_{\text{int}}(t_1)\exp(\gamma t_1)|0\rangle\ . \tag{2.34}$$

Using the Fourier expansion of $\boldsymbol{A}(\boldsymbol{r}, t)$

$$\boldsymbol{A}(\boldsymbol{r}, t) = \sum_{\omega} \boldsymbol{A}(\boldsymbol{r}, \omega)\mathrm{e}^{-i\omega t}\ , \tag{2.35}$$

we can calculate the integral over t_1 as

$$\begin{aligned}
&\int_{-\infty}^{t} \mathrm{d}t_1 \langle\nu|H_{\text{int}}(t_1)\exp(\gamma t_1)|0\rangle \\
&= -\frac{1}{c}\sum_{\omega}\int \mathrm{d}\boldsymbol{r}\langle\nu|\boldsymbol{I}_{\mathrm{t}}(\boldsymbol{r})|0\rangle \cdot \boldsymbol{A}(\boldsymbol{r}, \omega)\int_{-\infty}^{t} \mathrm{d}t_1 \exp[i(\omega_{\nu 0} - \omega - i\gamma)t_1] \\
&= \frac{i}{c}\sum_{\omega}\frac{\exp[i(\omega_{\nu 0} - \omega - i\gamma)t]}{\omega_{\nu 0} - \omega - i\gamma}\int \mathrm{d}\boldsymbol{r}\langle\nu|\boldsymbol{I}_{\mathrm{t}}(\boldsymbol{r})|0\rangle \cdot \boldsymbol{A}(\boldsymbol{r}, \omega)\ , \qquad (2.36) \\
&= \left[\text{same expression without } 1/c\right]_{\text{SI}}\ ,
\end{aligned}$$

where $\hbar\omega_{\nu 0} = E_\nu - E_0$ is the excitation energy of matter. This leads, for $\nu \neq 0$, to

$$\begin{aligned} a_\nu(t) &= -\frac{1}{\hbar c} \sum_\omega \frac{\exp[-i(\omega_0 + \omega + i\gamma)t]}{\omega_{\nu 0} - \omega - i\gamma} \int \mathrm{d}\boldsymbol{r} \langle \nu | \boldsymbol{I}_\mathrm{t}(\boldsymbol{r}) | 0 \rangle \cdot \boldsymbol{A}(\boldsymbol{r}, \omega) \;, \\ &= \big[\text{same expression without } 1/c\big]_\mathrm{SI} \;. \end{aligned} \tag{2.37}$$

The $\boldsymbol{A}$-linear part of the induced current density $\langle \Psi(t) | \boldsymbol{I}_\mathrm{t}(\boldsymbol{r}) | \Psi(t) \rangle$ arises in two different ways. One is from the first term $\langle \Psi_0 | \boldsymbol{I}_\mathrm{t}(\boldsymbol{r}) | \Psi_0 \rangle$, proportional to $|a_0|^2$, through the $\boldsymbol{A}$-linear term of $\boldsymbol{J}_\mathrm{orb}$, (2.14), and the other is from the terms proportional to $\{a_0 a_\nu^*\}$'s or $\{a_0^* a_\nu\}$'s through the $\boldsymbol{A}$-linear dependence of a_ν with the $\boldsymbol{A}$-independent part of $\boldsymbol{I}_\mathrm{t}$. The sum of these two terms gives the full expression of the $\boldsymbol{A}$-linear part of the induced current density of frequency ω as

$$\boldsymbol{I}(\boldsymbol{r}, \omega) = \int \mathrm{d}\boldsymbol{r}' \; \chi_\mathrm{cd}(\boldsymbol{r}, \boldsymbol{r}'; \omega) \cdot \boldsymbol{A}(\boldsymbol{r}', \omega) \tag{2.38}$$

where the microscopic susceptibility χ_cd is given as

$$\begin{aligned} \chi_\mathrm{cd}(\boldsymbol{r}, \boldsymbol{r}'; \omega) = &- \frac{1}{c} \langle 0 | \hat{N}(\boldsymbol{r}) | 0 \rangle \; \delta(\boldsymbol{r} - \boldsymbol{r}') \\ &+ \frac{1}{c} \sum_\nu \big[g_\nu(\omega) \boldsymbol{I}_{0\nu}(\boldsymbol{r}) \boldsymbol{I}_{\nu 0}(\boldsymbol{r}') + h_\nu(\omega) \boldsymbol{I}_{\nu 0}(\boldsymbol{r}) \boldsymbol{I}_{0\nu}(\boldsymbol{r}') \big] \\ &= \big[\text{same expression without } 1/c\big]_\mathrm{SI} \end{aligned} \tag{2.39}$$

in terms of

$$\boldsymbol{I}_{\mu\nu}(\boldsymbol{r}) = \langle \mu | \boldsymbol{I}(\boldsymbol{r}) | \nu \rangle \;, \tag{2.40}$$

$$g_\nu(\omega) = \frac{1}{\hbar(\omega_{\nu 0} - \omega - i\gamma)} \;, \tag{2.41}$$

$$h_\nu(\omega) = \frac{1}{\hbar(\omega_{\nu 0} + \omega + i\gamma)} \;. \tag{2.42}$$

As mentioned in the previous subsection, the first term on the r.h.s. of (2.39) is mainly contributed from the electron density in the ground state (times e^2/mc), and the second term represents the contribution from all the excited states, where the factor $g_\nu(\omega)$ and $h_\nu(\omega)$ give the resonance condition, and the product of two matrix elements of current density works as position-dependent weighting factors of each resonance.

The $(\boldsymbol{r}, \boldsymbol{r}')$ dependence of $\chi_\mathrm{cd}(\boldsymbol{r}, \boldsymbol{r}')$ shows the nonlocal character of the response, i.e., an EM field applied to the position $\boldsymbol{r}'$ can induce current density at a different position $\boldsymbol{r}$. This nonlocal response occurs within the spatial extension of relevant wave functions $\{|\nu\rangle, |0\rangle\}$. It should be stressed that this nonlocal character arises from the quantum mechanical extension of the wave functions. The matter-EM field interaction itself is local, as explicitly given in (2.10), i.e., they interact only at the same positions in space. Therefore, we should strictly distinguish between the "nonlocal response" and "nonlocal interaction".

The nonlocal response is the characteristic feature of microscopic response. In the macroscopic response, we generally use a local relationship between polarization(s) and source EM field, e.g., $\boldsymbol{P}(\boldsymbol{r}, \omega) = \chi_{\mathrm{e}}(\omega)\boldsymbol{E}(\boldsymbol{r}, \omega)$. Thus, the macroscopic averaging should contain a recipe to reduce the nonlocal response to a local one. For this purpose, the expression of χ_{cd} given above has a very convenient general form with respect to the $(\boldsymbol{r}, \boldsymbol{r}')$ dependence, i.e., it is a sum of the products of a function of $\boldsymbol{r}$ and that of $\boldsymbol{r}'$. As an integral kernel, this behavior is called separable, and greatly serves to simplify the solution of the integral equations, as shown below.

2.2.3 *Fundamental Equations to Determine Microscopic Response*

From the arguments of the previous sections, the fundamental equations to determine the set of microscopic variables $\{\boldsymbol{A}$ and $\boldsymbol{I}_{\mathrm{t}}\}$ in the linear response regime are the microscopic M-eqs

$$-\nabla^2 \boldsymbol{A} + \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} = \frac{4\pi}{c}\boldsymbol{I}_{\mathrm{t}}^{(\mathrm{T})} \quad \left[= \mu_0 \boldsymbol{I}_{\mathrm{t}}^{(\mathrm{T})}\right]_{\mathrm{SI}} \tag{2.43}$$

and the constitutive equation (2.38). These are the coupled equations to determine the T components of the two vector fields $\boldsymbol{A}$ and $\boldsymbol{I}_{\mathrm{t}}$ for a given initial condition. In the M-eqs $\boldsymbol{I}_{\mathrm{t}}^{(\mathrm{T})}$ is the source term of $\boldsymbol{A}$, and in the constitutive equation $\boldsymbol{A}$ induces $\boldsymbol{I}_{\mathrm{t}}^{(\mathrm{T})}$ (and also $\boldsymbol{I}_{\mathrm{t}}^{(\mathrm{L})}$ if symmetry allows), and the solution of the coupled equations gives us a self-consistent set of $\boldsymbol{A}$ and $\boldsymbol{I}_{\mathrm{t}}^{(\mathrm{T})}$. The L component of $\boldsymbol{I}_{\mathrm{t}}$ is obtained from the selfconsistently determined $\boldsymbol{A}$ via the constitutive equation, and $\boldsymbol{A}$ has no L component in Coulomb gauge. The case of exciting matter via external charge source, which introduces an L electric field as initial condition, will be treated in Sect. 5.7.

The initial condition of matter is already taken into account in calculating the induced current density by choosing the ground state of $H^{(0)}$ as the matter state at the remote past as mentioned in the previous subsection. (Its extension to the more generalized case of density matrix description will be given in Sect. 5.4.) The initial condition for the vector potential corresponds to the choice of incident EM field inducing matter polarization, which is contained in the solution of the M-eqs (2.43). The solution is a sum of the general solution for the homogeneous equation for $\boldsymbol{I}_{\mathrm{t}}^{(\mathrm{T})} = 0$ and a special solution in the presence of finite $\boldsymbol{I}_{\mathrm{t}}^{(\mathrm{T})}$. The general solution contains two free parameters corresponding to the two independent solutions of the second order differential equation. The values of these parameters are chosen to fit the asymptotic situation, e.g., in the remote past in accordance with the incident field.

In order to solve (2.43) and (2.38) in a neat way, we renormalize the $\langle 0|\hat{N}(\boldsymbol{r})|0\rangle$ term of $\chi_{\rm cd}(\boldsymbol{r}, \boldsymbol{r}'; \omega)$ into the resonant terms as given in the Sect. 5.5. This approximation is valid in LWA and in the non-relativistic regime. In this case, the microscopic susceptibility is written as

$$\chi_{\rm cd}(\boldsymbol{r}, \boldsymbol{r}'; \omega) = \frac{1}{c}\sum_{\nu}\left[\bar{g}_{\nu}(\omega)\boldsymbol{I}_{0\nu}(\boldsymbol{r})\boldsymbol{I}_{\nu 0}(\boldsymbol{r}') + \bar{h}_{\nu}(\omega)\boldsymbol{I}_{\nu 0}(\boldsymbol{r})\boldsymbol{I}_{0\nu}(\boldsymbol{r}')\right] \tag{2.44}$$
$$= \left[\text{same expression without } 1/c\right]_{\rm SI}$$

where

$$\bar{g}_{\nu}(\omega) = g_{\nu}(\omega) - \frac{1}{\hbar\omega_{\nu 0}}\,, \tag{2.45}$$
$$\bar{h}_{\nu}(\omega) = h_{\nu}(\omega) - \frac{1}{\hbar\omega_{\nu 0}}\,. \tag{2.46}$$

Using this form in the susceptibility $\chi_{\rm cd}$, we can rewrite the ω-Fourier component of (2.43) and (2.38) into a set of linear equations for new variables $F_{\mu\nu}(\omega)$ defined as

$$F_{\mu\nu}(\omega) = \int {\rm d}\boldsymbol{r}\ \langle\mu|\hat{\boldsymbol{I}}(\boldsymbol{r})|\nu\rangle \cdot \boldsymbol{A}(\boldsymbol{r}, \omega)\,. \tag{2.47}$$

In terms of $\{F_{\mu\nu}\}$, the induced current density $\boldsymbol{I}_{\rm t}(\boldsymbol{r}, \omega)$ is written as

$$\boldsymbol{I}_{\rm t}(\boldsymbol{r}, \omega) = \frac{1}{c}\sum_{\nu}\left[\bar{g}_{\nu}(\omega)\boldsymbol{I}_{0\nu}(\boldsymbol{r})F_{\nu 0}(\omega) + \bar{h}_{\nu}(\omega)\boldsymbol{I}_{\nu 0}(\boldsymbol{r})F_{0\nu}((\omega)\right] \tag{2.48}$$
$$= \left[\text{same expression without } 1/c\right]_{\rm SI}$$

The variables $\{F_{\mu\nu}\}$ depend on the quantum numbers μ, ν and frequency ω, but not on the coordinate $\boldsymbol{r}$, and as shown just above, they are the expansion coefficients of the induced current density in terms of the basis set $\{\boldsymbol{I}_{0\nu}(\boldsymbol{r}), \boldsymbol{I}_{\nu 0}(\boldsymbol{r})\}$. Namely, $\boldsymbol{I}_{\rm t}(\boldsymbol{r}, \omega)$ is a linear combination of $\{F_{\mu\nu}\}$. Since the basis set $\{\boldsymbol{I}_{0\nu}(\boldsymbol{r}), \boldsymbol{I}_{\nu 0}(\boldsymbol{r})\}$ should be given for any fixed model of matter, we only need to determine the expansion coefficients $\{F_{\nu 0}, F_{0\nu}\}$.

Taking the ω-Fourier component of the M-eqs (2.43), we obtain its general solution in the form

$$\boldsymbol{A}(\boldsymbol{r}, \omega) = \boldsymbol{A}_0(\boldsymbol{r}, \omega) + \frac{1}{c}\int {\rm d}\boldsymbol{r}'\ G_q(\boldsymbol{r}, \boldsymbol{r}')\ \boldsymbol{I}_{\rm t}^{\rm (T)}(\boldsymbol{r}', \omega)\,, \tag{2.49}$$
$$= \left[\text{same expression with} 1/c \text{ replaced by } \mu_0/4\pi\right]_{\rm SI}\,,$$

where $\boldsymbol{A}_0$ is the incident field satisfying the homogeneous equation (for $\boldsymbol{I}_{\rm t}^{\rm (T)} = 0$). The field $\boldsymbol{A}_0$ is a linear combination of two independent solutions of the homogeneous equation, with their coefficients to be chosen according to the initial condition, and $\boldsymbol{I}_{\mathrm{t}}^{(\mathrm{T})}(\boldsymbol{r},\omega)$ is given as (2.48) with the vector fields $\boldsymbol{I}_{\mu\nu}(\boldsymbol{r})$ replaced by their T components $\boldsymbol{I}_{\mu\nu}^{(T)}(\boldsymbol{r})$. The EM Green function is defined as

$$-\nabla^2 G_q(\boldsymbol{r},\boldsymbol{r}') - q^2 G_q(\boldsymbol{r},\boldsymbol{r}') = 4\pi\delta(\boldsymbol{r}-\boldsymbol{r}')\,, \tag{2.50}$$

where $q = \omega/c$ is the wave number in vacuum of the EM field with frequency ω, and a special solution of G_q is given as

$$G_q(\boldsymbol{r},\boldsymbol{r}') = \frac{\mathrm{e}^{iq|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\,. \tag{2.51}$$

By applying the operation $-\nabla^2 - q^2$ from the left to (2.49), we can assure that it is the general solution of (2.43) for frequency ω. The scattered field, the integral part, of (2.49) is also a T-field, which can be seen by taking its divergence and carrying out partial integration.

There is an another useful expression of the same quantity, where the T character is carried by the tensor EM Green function $\mathbf{G}_q(\boldsymbol{r},\boldsymbol{r}')$ as

$$\begin{aligned}\boldsymbol{A}(\boldsymbol{r},\omega) &= \boldsymbol{A}_0(\boldsymbol{r},\omega) + \frac{1}{c}\int \mathrm{d}\boldsymbol{r}'\,\mathbf{G}_q^{(T)}(\boldsymbol{r},\boldsymbol{r}')\cdot\boldsymbol{I}(\boldsymbol{r}',\omega)\,, \\ &= \left[\text{same expression with } 1/c \text{ replaced by } \mu_0/4\pi\right]_{\mathrm{SI}}\,,\end{aligned} \tag{2.52}$$

where

$$\mathbf{G}_q^{(T)}(\boldsymbol{r},\boldsymbol{r}') = G_q(\boldsymbol{r}-\boldsymbol{r}')\mathbf{1} + \frac{1}{q^2}\big[G_q(\boldsymbol{r}-\boldsymbol{r}') - G_0(\boldsymbol{r}-\boldsymbol{r}')\big]\nabla'\nabla'\,. \tag{2.53}$$

For details, see Sect. 5.7.1.

In terms of $\{F_{\nu0}, F_{0\nu}\}$, (2.49) can be rewritten as

$$\begin{aligned}\boldsymbol{A}(\boldsymbol{r},\omega) &= \boldsymbol{A}_0(\boldsymbol{r},\omega) + \frac{1}{c}\sum_{\nu}\left[\bar{g}_\nu F_{\nu0}\boldsymbol{A}_{0\nu}(\boldsymbol{r},\omega) + \bar{h}_\nu F_{0\nu}\boldsymbol{A}_{\nu0}(\boldsymbol{r},\omega)\right], \\ &= \left[\text{same expression with} 1/c \text{ replaced by } \mu_0/4\pi\right]_{\mathrm{SI}}\,,\end{aligned} \tag{2.54}$$

where

$$\begin{aligned}\boldsymbol{A}_{\mu\nu}(\boldsymbol{r},\omega) &= \frac{1}{c}\int \mathrm{d}\boldsymbol{r}'\,\mathbf{G}_q^{(T)}(\boldsymbol{r},\boldsymbol{r}';\omega)\cdot\boldsymbol{I}_{\mu\nu}(\boldsymbol{r}',\omega) \\ &= \left[\text{same expression with} 1/c \text{ replaced by } \mu_0/4\pi\right]_{\mathrm{SI}}\,,\end{aligned} \tag{2.55}$$

is the vector potential produced by the current density $\boldsymbol{I}_{\mu\nu}^{(T)}(\boldsymbol{r}')$.

If we further insert this result into the definition of $\{F_{\mu\nu}\}$, (2.47), it gives us a set of linear equations for $\{F_{\nu0}, F_{0\nu}\}$. In doing so, let us note that we can replace the current density $\boldsymbol{I}_{\mu\nu}(\boldsymbol{r})$ in (2.47) with its T component $\boldsymbol{I}_{\mu\nu}^{(T)}(\boldsymbol{r})$. This is because a

L-field can be written as the gradient of a scalar function ($\nabla f(\boldsymbol{r})$), and because the integral of the inner product of a L-field and $\boldsymbol{A}$ (T-field) turns out to be zero as

$$\int \mathrm{d}\boldsymbol{r}\ \nabla f(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r}) = -\int \mathrm{d}\boldsymbol{r}\ f(\boldsymbol{r})\ \nabla \cdot \boldsymbol{A}(\boldsymbol{r}) = 0\ . \tag{2.56}$$

where we have made partial integration and used $\nabla \cdot \boldsymbol{A}(\boldsymbol{r}) = 0$. This leads to

$$F_{\mu\nu}(\omega) = \int \mathrm{d}\boldsymbol{r}\ \boldsymbol{I}^{(T)}_{\mu\nu}(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r}, \omega)\ . \tag{2.57}$$

Inserting (2.54) into the definitions of $F_{\nu 0}$ and $F_{0\nu}$, (2.57), we obtain

$$F_{\nu 0} = F^{(0)}_{\nu 0} - \sum_{\mu} \left[\bar{g}_{\mu} \mathscr{A}_{\nu 0, 0\mu} F_{\mu 0} + \bar{h}_{\mu} \mathscr{A}_{\nu 0, \mu 0} F_{0\mu}\right] \tag{2.58}$$

$$F_{0\nu} = F^{(0)}_{0\nu} - \sum_{\mu} \left[\bar{g}_{\mu} \mathscr{A}_{0\nu, 0\mu} F_{\mu 0} + \bar{h}_{\mu} \mathscr{A}_{0\nu, \mu 0} F_{0\mu}\right] \tag{2.59}$$

where

$$F^{(0)}_{\mu\nu} = \int \mathrm{d}\boldsymbol{r}\ \boldsymbol{I}^{(T)}_{\mu\nu}(\boldsymbol{r}) \cdot \boldsymbol{A}_0(\boldsymbol{r}, \omega)\ , \tag{2.60}$$

and

$$\begin{aligned} \mathscr{A}_{\nu\sigma, \mu\tau} &= -\frac{1}{c} \int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r}' \boldsymbol{I}^{(T)}_{\nu\sigma}(\boldsymbol{r})\ G_q(\boldsymbol{r}, \boldsymbol{r}', \omega)\ \boldsymbol{I}^{(T)}_{\mu\tau}(\boldsymbol{r}') \\ &= \left[\text{same expression with} 1/c \text{ replaced by } \mu_0/4\pi\right]_{\mathrm{SI}} \end{aligned} \tag{2.61}$$

represents the radiative (radiation mediated) interaction energy between the two current densities associated with the transitions $\{\nu \leftrightarrow \sigma\}$ and $\{\mu \leftrightarrow \tau\}$. The real and imaginary part of the diagonal element $\mathscr{A}_{\nu 0, 0\nu}$ gives the shift and radiative width of the transition energy $E_{\nu 0}$ (see Sect. 3 of [1]). As mentioned above in connection with the tensor Green function (Sect. 5.7.1), it is possible to write the radiative correction in terms of $\mathbf{G}^{(T)}_q(\boldsymbol{r}, \boldsymbol{r}')$ as

$$\begin{aligned} \mathscr{A}_{\nu\sigma, \mu\tau} &= -\frac{1}{c} \int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r}' \boldsymbol{I}_{\nu\sigma}(\boldsymbol{r}) \cdot \mathbf{G}^{(T)}_q(\boldsymbol{r}, \boldsymbol{r}', \omega) \cdot \boldsymbol{I}_{\mu\tau}(\boldsymbol{r}')\ , \\ &= \left[\text{same expression with} 1/c \text{ replaced by } \mu_0/4\pi\right]_{\mathrm{SI}}\ . \end{aligned} \tag{2.62}$$

This rewriting is a rather general feature in describing T (L) field propagation, i.e., one ascribes the T (L) nature either to the source or to the propagator (Green function).

For a given incident field, $\{F^{(0)}_{\mu\nu}\}$ is a known set of quantities, so that it is straightforward to solve the simultaneous linear equations (2.58) and (2.59). The solution $\{F_{\nu 0}, F_{0\nu}\}$, directly determines the response fields, $\boldsymbol{A}$ and $\boldsymbol{I}$. Originally, this scheme was developed to describe the microscopic variation correctly, as given in [1] in detail. It has been used mainly for the study of nanostructures, but it can be used also as a starting theory to derive the macroscopic M-eqs, because it describes both microscopic and macroscopic spatial variations correctly. This is what we are now going to do in the following.

Before we proceed to derive the new macroscopic M-eqs and the corresponding constitutive equation, we give some characteristic aspects of the microscopic nonlocal response theory to show that our derivation of macroscopic M-eqs is based on a reliable foundation.

2.2.4 Characteristics of Microscopic Nonlocal Response Theory

Though the contents of this subsection will not be used directly in the derivation of the new macroscopic M-eqs, they will show the nature of the higher rank theory from which we are going to derive the macroscopic theory.

Microscopic Spatial Variation

The response fields $\boldsymbol{A}$ and $\boldsymbol{I}$ are expanded in terms of $\boldsymbol{A}_{\mu\nu}(\boldsymbol{r}, \omega)$ and $\boldsymbol{I}_{\nu\mu}(\boldsymbol{r})$, respectively. Since the basis for the matrix representation consists of the eigen functions of $H^{(0)}$, they have microscopic spatial variations like atomic wave functions, which is reflected in the spatial structure of $\boldsymbol{A}$ and $\boldsymbol{I}$. But, at the same time, they also contain rather smooth, or macroscopic, spatial variation, as a superposition of the contributions from infinitely many excited states. The relative weights of the contributions of individual excitations depend on the frequency range of interest. In the neighborhood of a particular resonance, the spatial structure of the induced current density of the resonant transition is dominant. The microscopic nonlocal theory is a scheme enabling us to treat these effects correctly in principle, and also in practice within the limit of numerical calculation.

Resonant Enhancement of Microscopic Spatial Structure

The selfconsistent solution of response is obtained by solving the equations of the variables $F_{\mu\nu}(\omega)$, which is the expansion coefficients of $\boldsymbol{I}_{\mathrm{t}}$. Since the solution has resonance effect at each excitation energy of matter, the response fields $\boldsymbol{A}$ and $\boldsymbol{I}_{\mathrm{t}}$ show corresponding resonant behavior. Each resonance is accompanied by a characteristic spatial structure of the resonance fields, so that this structure is enhanced at the resonance. The microscopic structures accompanying various resonances are different from one another.

This feature does not exist in the macroscopic case, because the resonances in macroscopic response are specified by the resonant frequencies and the corresponding residues given by the (mainly first order) moments of the matrix element of the induced current density. Thus, at any resonance, the spatial structure is specified only by a wave number, which becomes infinitely large, in the absence of non-radiative damping, as ω approaches the resonant frequency. This is an unphysical behavior introduced by the macroscopic averaging. The spatial structure of the resonance before carrying out LWA is quite different from the one described by a wave number. Therefore, unless it is smeared out by a non-radiative damping, we should be aware of its unphysical nature.

The spatial coherence is closely related with the applicability of LWA. If the extension of the spatial coherence is comparable to or larger than the wavelength of interest, LWA is not a good approximation, and we should keep the microscopic description of the resonance. Bulk excitons are typical examples of this kind. Since any quantum mechanical excitations has its own coherence, we should judge the applicability of LWA for each resonance in the frequency region of interest. In the case of impurity transitions, we can usually neglect the coherence over different impurities, if the density is low. For a high density case, we need to consider that a large number of degenerate transitions occur at various positions in a medium. Generally, there exists a coupling between the excitations at different positions via the Coulomb interaction among electrons, working even if the overlap of wave functions is negligible. Its main term is the dipole-dipole interaction, which is proportional to $1/R^3$ ($R =$distance between two impurities), and its strength reflects the dipole moments of the transitions. For a small density case, average R is large, so that this energy is negligible in comparison with the non-radiative width of each excitation or the fluctuation of site energy. Then, all the excitations can be treated independently, and the spatial coherence has an extension of a single impurity transition. If, however, the dipole-dipole interaction is not negligible (due, for example, to the high density of impurities, or to the large E1 moment of the transition), the eigenstates of the impurity transitions need to be diagonalized with the inclusion of the dipole-dipole interaction, which greatly changes the coherence character of the transitions. By this rearrangement, some of the eigenstates may have a large spatial extension. (If the impurities are regularly positioned, all the rearranged states are specified by some wave vector, so that all of them are extended infinitely.) There is a possibility that metamaterials might have a situation of this kind. Then, we need to describe the response of such metamaterials microscopically, i.e., with the nonlocal character kept explicitly (see Sect. 4.1.4).

Self-sustaining Modes

The formalism of the microscopic nonlocal response is applicable to a large variety of matter systems from individual atoms to bulk materials. As discussed in detail in Sect. 2.2.3, its fundamental equations for linear response are the simultaneous linear equations $\mathbf{S}\boldsymbol{X} = \boldsymbol{F}^{(0)}$ (in a matrix notation rewritten for variables $X_{\nu 0} = \bar{g}_\nu F_{\nu 0},\ X_{0\nu} = \bar{h}_\nu F_{0\nu}$), where the incident field is included in $\boldsymbol{F}^{(0)}$, and the solution $\boldsymbol{X}$ gives the amplitudes of selfconsistently determined current densities. The response EM field is obtained by solving the M-eqs with this current density as the source term. The coefficient matrix $\mathbf{S}$ of the equations consist of the eigenvalues of matter Hamiltonian $H^{(0)}$ and the matrix elements of induced current density with respect to the eigenfunctions of $H^{(0)}$.

The condition for the existence of non-trivial solution in the absence of an incident field, i.e., the vanishing of the determinant of the coefficient matrix, $\det|\mathbf{S}| = 0$ has a particular physical meaning. It gives the finite amplitude solution in the absence of incident field, representing the eigen mode of coupled matter excitation and EM field, which are sustaining each other without the help of incident field. In this sense, they may be called the "self-sustaining (SS) modes" of the interacting matter-EM field system.

The eigen-frequency of a SS mode is generally complex, the difference of which from the matter excitation energy represents the radiative shift and width of the relevant matter excitation. Since $\det|\mathbf{S}|$ occurs in the denominator of the solution, $\boldsymbol{X} = \mathbf{S}^{(-1)}\boldsymbol{F}^{(0)}$, the (real) ω dependence of $\boldsymbol{X}$ is resonant at the real part of the SS mode with a width given by its imaginary part. In this way the SS mode frequencies describe the resonant behavior of the response spectrum.

In the presence of several resonances, they affect each other, so that the exact positions of resonances (peaks and/or dips) are shifted from the isolated resonances. For an isolated resonance, the complex frequency of the self-sustaining mode exactly gives the peak position and, in the absence of non-radiative damping, its half width.

The realistic picture of the SS modes takes various form. In the case of an isolated atom, it is the atomic excitation energy with radiative correction. In the case of an exciton in a non-metallic crystal of infinite size, the condition $\det|\mathbf{S}| = 0$ gives the dispersion equation of exciton-polariton without radiative width. Similarly, various surface polaritons, such as surface (exciton, phonon, or plasmon) polaritons, are also SS modes. More exotic examples are the cavity modes of the dielectrics with a particular size and shape, and the dynamically scattered X rays in a crystal. See Sect. 3.1 of [1] for more details.

Radiative Correction

It is one of the characteristic points of the microscopic nonlocal response theory that it contains the interaction among the induced current densities via the T and L components of EM field. The L component represents the Coulomb field due to the current density (or charge density), and its interaction with matter polarization is included in the matter Hamiltonian as the Coulomb potential among charged particles. The interaction among the induced current densities mediated by the T component of induced EM field plays an important role in this framework as "radiative correction". This is the interaction energy of the T field produced by a current density with another current density (or with itself), as defined in (2.61). Since the EM Green function connecting the two current densities is generally a complex quantity, the resultant interaction energy is complex, too. Its physical meaning is that the continuum of the EM field energy works as a bath for the decay of a matter excitation energy, giving a finite lifetime to the matter excitation.

This kind of interaction via T field is also taken into account in the macroscopic M-eqs, if we solve them selfconsistently with the constitutive equation. The main difference is that the radiative correction is defined quantum mechanically with all the details of matter excitations in the microscopic theory, so that we can calculate the radiative correction from the first-principles. In fact, we can study the size and shape dependence of the radiative correction for a given finite matter system. See Sect. 4.1 of [1].

It is worth mentioning that the radiative width calculated from this general expression is exact in comparison with the result of QED. For a single atom in vacuum, the self interaction of an induced current density via emitted EM field of a given frequency produces radiative shift and width of the current density. The radiative width (FWHM) is exactly the same as the result of QED, which is usually given in LWA as

$$\Gamma = \frac{4}{3} q^3 |\boldsymbol{\mu}|^2 \tag{2.63}$$

where $\boldsymbol{\mu}$ is the electric dipole moment of the transition with energy E, and $q = E/\hbar c$. (The corresponding radiative shift depends on the details of the wave functions of the transition, so that it cannot be uniquely fixed by the value of $\boldsymbol{\mu}$ alone.)

The definition of radiative correction (2.61) is valid also outside LWA. Therefore, the size dependence mentioned above can be calculated smoothly across the validity limit of LWA, which allows to discuss the connection of two different regimes, i.e., within or beyond LWA. See Sect. 4.5 of [1].

Boundary Conditions

It is the most pronounced aspect of the microscopic nonlocal response theory, that it does not require the boundary conditions (BC's) to connect the EM fields in- and outside a matter system. It is a matter of fact for a microscopic theory not to use BC's, because no boundary can be drawn for a microscopic material distinguishing the in- and outside the matter, and because the fundamental equations $\mathbf{S}\boldsymbol{X} = \boldsymbol{F}^{(0)}$ to determine the response for a given incident field are complete without BC's. A given shape of matter in its ground state defines the BC's for electrons which govern the EM response of matter. This allows us to calculate the microscopic nonlocal susceptibility in a position dependent manner. As we described in Sects. 2.2.2 and 2.2.3 in detail, this knowledge of nonlocal susceptibility is enough to determine the selfconsistent response uniquely. The use of BC's in macroscopic M-eqs is a standard technique to solve problems, but it is a specialty only in macroscopic M-eqs. It should be remembered that no BC is required in the higher rank theories.

The necessity of BC arises when we approximate a part or all of the induced current density by a macroscopic one, which is the subject of Sect. 3.7.

2.2.5 *Gauge Invariance of Many-Body Schrödinger Equation*

From the viewpoint that all the EM response theories, from relativistic QED to semiclassical macroscopic Maxwell eqs, belong to a single hierarchy, we are going to present macroscopic constitutive equations to be derived from a one-rank higher theory, i.e., the non-relativistic quantum theory for the motion of charged particles of matter in an EM field. Since the quantum mechanical description of such a system requires the use of scalar and vector potentials for the EM field, there arises a question of gauge for the explicit calculation of constitutive equations.

In Sect. 2.2.1, we have chosen Coulomb gauge, and it is used throughout the book. The general expression of the susceptibility relating induced current density and source EM field is given in terms of the eigenvalues and eigenfunctions of the many-body Hamiltonian containing the Coulomb potential among the charged particles of matter. Though it is standard to use the Coulomb gauge, one might ask whether the result is in any way specific to the use of Coulomb gauge.

From the structure of minimal coupling Lagrangian equation (2.1), the motion of the system is seen to be independent on the choice of gauge. The change in the Lagrangian caused by the gauge transformation (1.20) and (1.21) can be written in a total time derivative form as

$$\delta L = \int \mathrm{d}\boldsymbol{r} \left[\rho \frac{\partial \psi}{\partial t} + \boldsymbol{J} \cdot \nabla \psi \right] = \frac{\mathrm{d}}{\mathrm{d}t} \int \mathrm{d}\boldsymbol{r} \, \rho \psi \tag{2.64}$$

where $\nabla \cdot \boldsymbol{J}^{(\mathrm{L})} = -\partial \rho / \partial t$ is used. Since a total time derivative term in Lagrangian does not contribute to the minimum action principle, this additional term δL does not change the Lagrange equations of the system. This argument shows that the description of interacting matter-EM field system via the minimal coupling Lagrangian is gauge invariant. This makes us expect the gauge invariance of the constitutive equation, i.e., the induced current density as a functional of EM field calculated from the many-body Schrödinger equation. However, the derivation of the many-body Schrödinger equation from the minimal coupling Lagrangian is usually made for a fixed gauge, which obscures the argument of gauge invariance at the level of Schrödinger equation.

The reason for it is the lack of the velocity term corresponding to scalar potential in L, which gives zero for the generalized momentum of scalar potential. This is considered to be a problem on introducing the commutation relation for this set of variables, when we treat EM field quantum mechanically in the (non-relativistic) QED. In the semiclassical treatment, EM field is not a quantum mechanical quantity, so that we may put aside the problem of commutation relation for the scalar potential. In this sense, the Hamiltonian (2.5) is all right for use. However, the many-body Schrödinger equation $i\hbar\partial\Psi/\partial t = H\psi$ does not seem to be invariant for the transformation $\{\boldsymbol{A}, \phi, \Psi\} \rightarrow \{\boldsymbol{A}', \phi', \Psi'\}$ mediated by an arbitrary scalar function $\bar{\chi}(\boldsymbol{r}, t)$

$$\boldsymbol{A}' = \boldsymbol{A} + c\nabla\bar{\chi} \ , \quad \phi' = \phi - \frac{\partial \bar{\chi}}{\partial t} \ , \quad \Psi' = \exp(i\Theta)\Psi \ , \tag{2.65}$$

$$\Theta = \sum_{\ell} \frac{e_\ell}{\hbar} \bar{\chi}(\boldsymbol{r}_\ell, t) \ = \frac{1}{\hbar} \int \mathrm{d}\boldsymbol{r} \ \rho_{\mathrm{int}}(\boldsymbol{r}) \bar{\chi}(\boldsymbol{r}, t) \ . \tag{2.66}$$

which is the many body version of gauge invariance. Then the next question is "how is the gauge invariance ensured?"

There is another point of view about the gauge invariance. In this book, we emphasize that the Hamiltonian (2.5) or its extended form with relativistic correction terms is used, not only for semiclassical treatment, but also for weakly-relativistic QED by just quantizing EM field. This forms the backbone of the single hierarchy nature of EM response theories mentioned in Table 1.1. To avoid the difficulty in defining the commutation relation for ϕ, we want to find a logically clear way to rewrite the minimal coupling Lagrangian into many-body Hamiltonian, which can be used for the gauge invariant calculation of the expectation value of physical quantities.

For this purpose, a key role is played by a transformation of dynamical variables in Lagrangian, where we eliminate the generalized coordinate having no corresponding velocity in the Lagrangian. The general scheme for this transformation is described by Cohen-Tannoudji et al. [7] (p. 84). Since this transformation plays an essential role in our logical development, we reproduce the relevant part to see its general nature.

We consider a Lagrangian L_0 where one of the generalized coordinates, x_0, does not have its corresponding velocity. Denoting the other coordinates as $\{x_\ell; \ell = 1, 2, \ldots\}$ and the velocities as $\{v_\ell; \ell = 1, 2, \ldots\}$, we write $L_0 = L_0(x_0, \{x_\ell, v_\ell\})$. In this case, a special type of variable transformation is possible without changing the set of Lagrange equations. The Lagrange equation derived from the variation of x_0 is

$$\frac{\partial L_0}{\partial x_0} = 0\,, \tag{2.67}$$

since v_0 is missing in L_0. This equation can be solved for x_0 as a function of $\{x_\ell, v_\ell\}$, i.e.,

$$x_0 = f(\{x_\ell, v_\ell\})\,. \tag{2.68}$$

Inserting this result in L_0, we obtain a new Lagrangian

$$\hat{L}_0(\{x_\ell, v_\ell\}) = L_0(f(\{x_\ell, v_\ell\}), \{x_\ell, v_\ell\})\,. \tag{2.69}$$

The two Lagrangians L_0 and $\hat{L}_0$ give the same Lagrange equations for the variables $\{x_\ell; \ell = 1, 2, \ldots\}$, because

$$\frac{\partial \hat{L}_0}{\partial x_\ell} = \frac{\partial L_0}{\partial x_\ell} + \frac{\partial L_0}{\partial x_0}\frac{\partial x_0}{\partial x_\ell} = \frac{\partial L_0}{\partial x_\ell}\,, \tag{2.70}$$

$$\frac{\partial \hat{L}_0}{\partial v_\ell} = \frac{\partial L_0}{\partial v_\ell} + \frac{\partial L_0}{\partial x_0}\frac{\partial x_0}{\partial v_\ell} = \frac{\partial L_0}{\partial v_\ell}\,, \tag{2.71}$$

by the use of the Lagrange equation $\partial L_0/\partial x_0 = 0$. This result also shows that the generalized momenta $\{\partial L_0/\partial v_\ell\}$ are not affected by the transformation. This allows us to introduce Hamiltonian and quantization condition for all the remaining variables without worrying about the vanishing commutation relation mentioned above. In this way, we can prepare a logically safe description of such a system in both classical and quantum mechanical regimes. Below we will see how this general scheme of variable transformation works in the system of interacting EM field and charged particles.

In the minimal coupling Lagrangian, the dynamical variables are the coordinate and velocity of the charged particles, $\{\boldsymbol{r}_\ell, \boldsymbol{v}_\ell\}$, scalar potential ϕ, and vector potential and its time derivative, $\{\boldsymbol{A}, \partial \boldsymbol{A}/\partial t\}$. The minimum action principle due to the variation of $\boldsymbol{r}_\ell$ and ϕ leads to the Newton equation under Lorentz force, (5.34) and Gauss law, (5.38), respectively. The variation of $\boldsymbol{A}$ leads to the EM wave equation, (5.43), which consists of the L component

$$\frac{\partial \boldsymbol{E}_\mathrm{L}}{\partial t} = -4\pi \boldsymbol{J}_\mathrm{L} \,, \tag{2.72}$$

$$\left[\frac{\partial \boldsymbol{E}_\mathrm{L}}{\partial t} = -\frac{1}{\varepsilon_0} \boldsymbol{J}_\mathrm{L} \right]_\mathrm{SI} \tag{2.73}$$

and the T component

$$\frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}_\mathrm{T}}{\partial t^2} - \nabla^2 \boldsymbol{A}_\mathrm{T} = \frac{4\pi}{c} \boldsymbol{J}_\mathrm{T} \,, \tag{2.74}$$

$$\left[\frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}_\mathrm{T}}{\partial t^2} - \nabla^2 \boldsymbol{A}_\mathrm{T} = \mu_0 \boldsymbol{J}_\mathrm{T} \right]_\mathrm{SI} \tag{2.75}$$

Divergence of the L component leads, with the help of Gauss law, to the charge conservation law $\nabla \cdot \boldsymbol{J} + \partial \rho/\partial t = 0$. By using $\boldsymbol{J}_\mathrm{L} = \partial \boldsymbol{P}_\mathrm{L}/\partial t$, it can be rewritten also as $\partial(\nabla \cdot \boldsymbol{P}_\mathrm{L} + \rho)/\partial t = 0$ or $\partial(\boldsymbol{E}_\mathrm{L} + 4\pi \boldsymbol{P}_\mathrm{L})/\partial t = 0$.

The variable ϕ plays the role of x_0 mentioned above. The Lagrange equation for ϕ is $\partial L/\partial \phi = 0$, which leads to Gauss law. Its solution is

$$\boldsymbol{E}_\mathrm{L}(\boldsymbol{r}) = -\nabla \phi_\mathrm{c}(\boldsymbol{r}) \,, \tag{2.76}$$

$$\phi_\mathrm{c}(\boldsymbol{r}) = \int \mathrm{d}\boldsymbol{r}' \, \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \quad \left[= \frac{1}{4\pi\varepsilon_0} \int \mathrm{d}\boldsymbol{r}' \, \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \right]_\mathrm{SI} \tag{2.77}$$

Using $\boldsymbol{E}_\mathrm{L} = -(1/c)(\partial \boldsymbol{A}_\mathrm{L}/\partial t) - \nabla \phi$, we can express ϕ in terms of the other variables as

$$\nabla \phi = -\frac{1}{c}\frac{\partial \boldsymbol{A}_\mathrm{L}}{\partial t} + \nabla \phi_\mathrm{c} \; = \left[-\frac{\partial \boldsymbol{A}_\mathrm{L}}{\partial t} + \nabla \phi_\mathrm{c} \right]_\mathrm{SI} \tag{2.78}$$

The variable ϕ appears in two terms of L. One is the L component of EM field energy, which can be rewritten into Coulomb potential as

$$\frac{1}{8\pi} \int \mathrm{d}\boldsymbol{r} \, \boldsymbol{E}_\mathrm{L}^2 = \frac{1}{2} \sum_\ell \sum_{\ell'} \frac{e_\ell e_{\ell'}}{|\boldsymbol{r}_\ell - \boldsymbol{r}_{\ell'}|} = U_\mathrm{C} \,, \tag{2.79}$$

$$\left[\frac{\varepsilon_0}{2} \int \mathrm{d}\boldsymbol{r} \, \boldsymbol{E}_\mathrm{L}^2 = \frac{1}{8\pi\varepsilon_0} \sum_\ell \sum_{\ell'} \frac{e_\ell e_{\ell'}}{|\boldsymbol{r}_\ell - \boldsymbol{r}_{\ell'}|} = U_\mathrm{C} \right]_\mathrm{SI} \tag{2.80}$$

and the other is the interaction term $-\int \mathrm{d}\boldsymbol{r}\,\rho\,\phi$. By using $\nabla\cdot\boldsymbol{P}_\mathrm{L} = -\rho$, this term takes the form

$$-\int \mathrm{d}\boldsymbol{r}\rho\phi = \int \mathrm{d}\boldsymbol{r}(\nabla\cdot\boldsymbol{P}_\mathrm{L})\phi = \int \mathrm{d}\boldsymbol{r}\ \boldsymbol{P}_\mathrm{L}\cdot\left(\frac{1}{c}\frac{\partial\boldsymbol{A}_\mathrm{L}}{\partial t} - \nabla\phi_\mathrm{c}\right), \tag{2.81}$$

$$\left[= \int \mathrm{d}\boldsymbol{r}\ \boldsymbol{P}_\mathrm{L}\cdot\left(\frac{\partial\boldsymbol{A}_\mathrm{L}}{\partial t} - \nabla\phi_\mathrm{c}\right)\right]_\mathrm{SI}. \tag{2.82}$$

Noting $\int \mathrm{d}\boldsymbol{r}\boldsymbol{P}_\mathrm{L}\cdot\nabla\phi_\mathrm{c} = 2U_\mathrm{C}$ and $\boldsymbol{J}_\mathrm{L} = \partial\boldsymbol{P}_\mathrm{L}/\partial t$, we obtain the new Lagrangian after elimination of ϕ as

$$L' = \sum_\ell \frac{1}{2}m_\ell v_\ell^2 - U_\mathrm{C} + \frac{1}{c}\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{J}_\mathrm{T}\cdot\boldsymbol{A}_\mathrm{T} - \frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}t}\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{P}_\mathrm{L}\cdot\boldsymbol{A}_\mathrm{L}$$
$$+\frac{1}{8\pi}\int \mathrm{d}\boldsymbol{r}\ \left\{\left(\frac{1}{c}\frac{\partial\boldsymbol{A}_\mathrm{T}}{\partial t}\right)^2 - (\nabla\times\boldsymbol{A}_\mathrm{T})^2\right\}, \tag{2.83}$$

$$\Bigg[\ L' = \sum_\ell \frac{1}{2}m_\ell v_\ell^2 - U_\mathrm{C} + \int \mathrm{d}\boldsymbol{r}\ \boldsymbol{J}_\mathrm{T}\cdot\boldsymbol{A}_\mathrm{T} - \frac{\mathrm{d}}{\mathrm{d}t}\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{P}_\mathrm{L}\cdot\boldsymbol{A}_\mathrm{L}$$
$$+\frac{\varepsilon_0}{2}\int \mathrm{d}\boldsymbol{r}\left\{\left(\frac{\partial\boldsymbol{A}_\mathrm{T}}{\partial t}\right)^2 - c^2(\nabla\times\boldsymbol{A}_\mathrm{T})^2\right\}\Bigg]_\mathrm{SI} \tag{2.84}$$

The only gauge dependent quantity $\boldsymbol{A}_\mathrm{L}$ appears in the total time derivative term of L'. This term is the time derivative of a function of general coordinates, as seen from (5.7) giving $\boldsymbol{P}(\boldsymbol{r})$ as a function of $\{\boldsymbol{r}_\ell\}$ alone. Analytical mechanics generally allows to omit such a term from Lagrangian without affecting the minimum action principle. Thus, irrespective of any starting gauge, we can use

$$L_\mathrm{T} = \sum_\ell \frac{1}{2}m_\ell v_\ell^2 - U_\mathrm{C} + \frac{1}{c}\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{J}\cdot\boldsymbol{A}_\mathrm{T}$$
$$+\frac{1}{8\pi}\int \mathrm{d}\boldsymbol{r}\ \left\{\left(\frac{\partial\boldsymbol{A}_\mathrm{T}}{\partial t}\right)^2 - (\nabla\times\boldsymbol{A}_\mathrm{T})^2\right\}, \tag{2.85}$$

$$\Bigg[\ L_\mathrm{T} = \sum_\ell \frac{1}{2}m_\ell v_\ell^2 - U_\mathrm{C} + \int \mathrm{d}\boldsymbol{r}\ \boldsymbol{J}\cdot\boldsymbol{A}_\mathrm{T}$$
$$+\frac{\varepsilon_0}{2}\int \mathrm{d}\boldsymbol{r}\ \left\{\left(\frac{\partial\boldsymbol{A}_\mathrm{T}}{\partial t}\right)^2 - c^2(\nabla\times\boldsymbol{A}_\mathrm{T})^2\right\}\Bigg]_\mathrm{SI} \tag{2.86}$$

to determine the dynamical motions of the system. In the interaction term, $\boldsymbol{J}_\mathrm{T}$ can be written as $\boldsymbol{J}$ because $\nabla\cdot\boldsymbol{A}_\mathrm{T} = 0$ eliminates the contribution of $\boldsymbol{J}_\mathrm{L}$ in the integral.

This Lagrangian has the same form as that of Coulomb gauge. It should be stressed, however, that this result does not arise from choosing Coulomb gauge, but from the use of higher principle of Lagrange formalism, i.e., the independence of minimum action principle on the total time derivative term in Lagrangian, which is valid, not only for electromagnetism, but also for more general cases. **This is the logical basis allowing the use of Lagrangian in Coulomb gauge to calculate gauge invariant EM response**.

The Lagrange equations derived from L_{T} are the wave equation (2.74) for $\boldsymbol{A}_{\mathrm{T}}$, and the Newton equation of motion for the ℓth particle

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(m_\ell \boldsymbol{v}_\ell + \frac{e_\ell}{c}\boldsymbol{A}_{\mathrm{T}}\right) = -\nabla_\ell U_{\mathrm{C}} + \frac{e_\ell}{c}\nabla_\ell[\boldsymbol{v}_\ell \cdot \boldsymbol{A}_{\mathrm{T}}(\boldsymbol{r}_\ell)]\,, \tag{2.87}$$

$$\left[\frac{\mathrm{d}}{\mathrm{d}t}(m_\ell \boldsymbol{v}_\ell + e_\ell \boldsymbol{A}_{\mathrm{T}}) = -\nabla_\ell U_{\mathrm{C}} + e_\ell \nabla_\ell[\boldsymbol{v}_\ell \cdot \boldsymbol{A}_{\mathrm{T}}(\boldsymbol{r}_\ell)]\right]_{\mathrm{SI}} \tag{2.88}$$

By noting $\nabla_\ell U_{\mathrm{C}} = \nabla\phi_{\mathrm{c}}(\boldsymbol{r}_\ell) = -e_\ell \boldsymbol{E}_{\mathrm{L}}(\boldsymbol{r}_\ell)$ and (5.27), this equation turns out to be the Newton equation of motion, (5.34). Together with Gauss law and charge conservation law, the set of Lagrange equations derived from L_{T} is exactly the same as that from L. In order to determine the motions of all the dynamical variables consistently, we first solve (2.74) and (5.34) for a given initial condition, obtaining $\boldsymbol{J}$ and $\boldsymbol{A}_{\mathrm{T}}$, which are then used to determine ρ via charge conservation law and $\boldsymbol{E}_{\mathrm{L}}$ via Gauss law.

For quantum mechanical treatment, we derive the generalized momenta of $\boldsymbol{r}_\ell$ and $\boldsymbol{A}_{\mathrm{T}}$ as

$$\boldsymbol{p}_\ell = m_\ell \boldsymbol{v}_\ell + \frac{e_\ell}{c}\boldsymbol{A}_{\mathrm{T}}(\boldsymbol{r}_\ell)\,, \quad \boldsymbol{\Pi}_{\mathrm{AT}} = \frac{1}{4\pi c^2}\frac{\partial \boldsymbol{A}_{\mathrm{T}}}{\partial t}\,, \tag{2.89}$$

$$\left[\boldsymbol{p}_\ell = m_\ell \boldsymbol{v}_\ell + e_\ell \boldsymbol{A}_{\mathrm{T}}(\boldsymbol{r}_\ell)\,, \quad \boldsymbol{\Pi}_{\mathrm{AT}} = \varepsilon_0 \frac{\partial \boldsymbol{A}_{\mathrm{T}}}{\partial t}\right]_{\mathrm{SI}} \tag{2.90}$$

respectively, and the corresponding Hamiltonian as

$$H_{\mathrm{T}} = \sum_\ell \frac{1}{2m_\ell}\left[\boldsymbol{p}_\ell - \frac{e_\ell}{c}\boldsymbol{A}_{\mathrm{T}}(\boldsymbol{r}_\ell)\right]^2 + U_{\mathrm{C}} + \frac{1}{8\pi}\int \mathrm{d}\boldsymbol{r}\left\{\left(\frac{1}{c}\frac{\partial \boldsymbol{A}_{\mathrm{T}}}{\partial t}\right)^2 + (\nabla \times \boldsymbol{A}_{\mathrm{T}})^2\right\}, \tag{2.91}$$

$$\left[H_{\mathrm{T}} = \sum_\ell \frac{1}{2m_\ell}[\boldsymbol{p}_\ell - e_\ell \boldsymbol{A}_{\mathrm{T}}(\boldsymbol{r}_\ell)]^2 + U_{\mathrm{C}} + \frac{\varepsilon_0}{2}\int \mathrm{d}\boldsymbol{r}\left\{\left(\frac{\partial \boldsymbol{A}_{\mathrm{T}}}{\partial t}\right)^2 + c^2(\nabla \times \boldsymbol{A}_{\mathrm{T}})^2\right\}\right]_{\mathrm{SI}} \tag{2.92}$$

Introducing the quantization condition for each pair of coordinate and momentum, $\{\boldsymbol{r}_\ell, \boldsymbol{p}_{\ell'}\}$ and $\{\boldsymbol{A}_{\mathrm{T}}, \Pi_{\mathrm{AT}}\}$, we can study the dynamical motion of the system quantum mechanically. The quantization of $\boldsymbol{A}_{\mathrm{T}}$ leads to the (transverse) photons characterized by wave number and polarization, which is the basis of QED.

In QED, EM fields $\boldsymbol{E}$, $\boldsymbol{B}$ are operators, and the equations of motion for particles and EM field are operator equations, whose solution needs to be converted to ensemble average over initial ensemble of the EM field in order to be compared with observable quantities. For semiclassical description, however, EM field is a classical variable without quantization, so that the particle parts in the equations of motion need to be expressed as expectation values determined by the initial condition of particle states.

Though the main interest of this book is the coupled system of EM field and charged particles in non-relativistic regime, we sometimes need to consider the relativistic correction term, such as spin-orbit interaction, spin Zeeman interaction, mass velocity term, and Darwin terms. These terms are derived from Dirac equation as the correction to electron energy next to the rest energy mc^2 for weakly relativistic situation. The addition of these correction terms to Hamiltonian does not change the argument about gauge invariance, since they are written in terms of $\boldsymbol{E}$ and $\boldsymbol{B}$ [4, 8]. The explicitly spin dependent terms, spin-orbit interaction and spin Zeeman term, are especially important for material science related with magnetism. As sketched in Sect. 2.2.6, these additional terms could be rewritten into a new invariant form, which are the combinations of new matter variables (such as spin current) and emergent EM fields. This rewriting is claimed to leads to the new point of view in the theory of electromagnetism called "emergent electromagnetism".

The Hamiltonian H_{T} does not contain the L component of electric field. One might wonder how one treat the problems of an isolated charged particles excited by external charges, such as a matter sample in a condenser or the case of electron energy loss spectroscopy, etc. Such a problem can be formulated by dividing the total Coulomb potential into the contributions of internal and external charges and the interaction between them. This will be discussed in Sect. 5.7.2.

2.2.6 Relativistic Correction Terms

Consideration of relativistic correction in EM response theory adds a new element "spin" to the description of matter-EM field interaction. It appears through spin Zeeman and spin-orbit interaction, (2.15), (2.16), which should be added to the Hamiltonian H_{T}, (2.91). This is the basic Hamiltonian not only for non-relativistic QED, but also for both semiclassical microscopic response theory and semiclassical macroscopic response theory according to the single hierarchy viewpoint.

Spin Zeeman term is the interaction energy between EM field and spin magnetization as shown in (2.19). Similar expression appears from the orbital contribution (2.10) by taking the orbital magnetization part $c\nabla \times \boldsymbol{M}$ of $\boldsymbol{J}_0$. Their sum gives the interaction energy between the total magnetization and magnetic field $\boldsymbol{B}$. In the calculation of constitutive equation, its zero frequency part is often included in unperturbed Hamiltonian, while its dynamical part is used as perturbation Hamiltonian, This is the case for spin resonance, cyclotron resonance, various types of magneto-optics, etc., where spin flip energy in a static magnetic field is regarded as a matter excitation energy. This is reflected in the presence of spin flip energy $\hbar\omega_0$ as a pole of magnetic susceptibility of spin resonance, (1.51). In such a case, the effect of static field is contained, not perturbationally, but rigorously. Similar consideration is also given in the problems of cyclotron resonance, where the quantized energies of orbital motion in a static magnetic field is included in unperturbed energy, so that the susceptibility of cyclotron resonance has poles at the frequencies corresponding to the transition energies between Landau levels.

There is no definite rule to separate zero and finite frequency components of H_{sZ} into unperturbed and perturbed parts of Hamiltonian for the calculation of susceptibilities, though zero frequency part can be naturally contained in the unperturbed part due to its time independent character. Only for the calculation of static susceptibility, it should be the perturbation Hamiltonian by definition. Further, in nonlinear processes, there arise sum and difference components of incident frequencies, among which there can be a zero frequency component. In such a case, it does not make much sense to regard the interaction energy with the zero frequency component as a part of unperturbed Hamiltonian.

The electronic states described by the weakly (or non-) relativistic Hamiltonian are $s = 1/2$ spinors with positive energy. The quantum mechanical states of matter are described by wave functions for orbital motions and spin states. Thus the quantum number of matter states, μ, ν, appearing in the expressions of susceptibilities in this book contains the quantum numbers of all the variables of matter.

Spin orbit interaction contains the slope of potential $\nabla V(\mathrm{r})$. Since the charge density is stronger around each nucleus, its effect becomes pronounced in materials containing heavy atomic elements. In the neighborhood of each nucleus we may expect centrally symmetric charge distribution as

$$\nabla V(\boldsymbol{r}) \cong \boldsymbol{r} f_{\mathrm{so}}(|\boldsymbol{r}|) \tag{2.93}$$

which leads to the more familiar form of spin-orbit interaction

$$H_{\mathrm{SO}} = \frac{\hbar}{2m_0^2c^2} f_{\mathrm{so}}(|\boldsymbol{r}|)\boldsymbol{L} \cdot \sigma \tag{2.94}$$

where $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$ is orbital angular momentum vector. This explicitly shows the mixing of spin and orbital angular momenta. Even in a spherical symmetric situation as in atoms, these two angular momenta are not separately conserved, but their sum, total angular momentum, is the conserved quantity. This is one of the important indices to determine the level scheme of matter.

The description of matter quantum states in the presence of spin-orbit interaction is diverse corresponding to the variety of matter states, such as nuclei, atoms, molecules, crystals, impurities and defects in a crystal, etc. For each system appropriate methods have been developed. We sketch some of them as representative examples, and finally a new interpretation is mentioned about the relativistic correction appearing in the studies of transport problems in the matter systems with spontaneous magnetic and/or electric polarizations, spin current, topological defects, etc. , where a new formulation of EM response theory might be useful in terms of “emergent electromagnetic field” [9].

Atomic Spectroscopy

In this case, one considers a single atom consisting of many electrons (except for hydrogen atom) around a given nucleus. Due to the Fermion nature of electrons, the wave functions need to be antisymmetric for the exchange of any two electrons. The one-electron basis functions are the eigenfunctions of single particle Hamiltonian consisting of “kinetic energy + central symmetric attractive potential’, and are classified according to the principal and angular quantum numbers [10]. The single particle potential is due to the Coulomb interaction with nucleus and inner core electrons and also to spin-orbit interaction. The basis of many electron system is defined as “configuration”, specified by a set of occupied one-electron states, each of which contains a spin part. The antisymmetric nature is taken into account by arranging the one-electron functions in the form of “Slater determinant”, To obtain the energy eigenvalues to compare measured atomic spectra, we further need to consider the Coulomb interaction for a set of configurations expected to form an atomic state. In the early days of active atomic spectroscopy, there was no big computers to diagonalize even a Hamiltonian matrix of reasonable size, and two tractable cases were considered as good approximations. They are the well-known schemes of “LS coupling” and “jj coupling”, which correspond, respectively, to the case “spin-orbit interaction is negligible compared with Coulomb interaction” and “spin-orbit interaction is dominant for a certain configuration”. In these limiting cases, the evaluation of the energy eigenvalues and wave functions were feasible without big computers.

Effective Hamiltonian Method

In spectroscopic studies of matter, we are usually interested in the matter excitations causing resonant structure in measured spectra. Fine structures in a resonant region and possibly its dependence on external field (e.g., magnetic field) gives us a detailed information about the quantum mechanical states in that energy region. In the presence of spin-orbit interaction which mixes various spin states via orbital motion, the descriptions of these processes is complicated due to the dependence on the details of quantum mechanical states. The case of atomic spectroscopy mentioned above is a rather straightforward treatment of quantum mechanical argument.

For atomic impurities in solids, the corresponding theoretical problem is rather different, because of the influence from surrounding atoms and/ions. This problem was an important issue at the initial stage of laser physics, where transition metal ions in insulating crystals are used as lasing medium, and their accurate level scheme is important to know their laser frequencies. The quantum mechanical behavior of the impurity ion is strongly influenced by its surroundings, and the main issue was how to integrate the effect of the surroundings into the level scheme of the ions.

A general recipe for this problem is to prepare an “effective Hamiltonian” or “spin Hamiltonian” for the relevant, functional subspace of transition metal ion, which describes the perturbation due to its surroundings and the effect of external magnetic field [11]. This can be viewed as cutting-out procedure of a finite size matrix from the Hamiltonian matrix of infinite size. The abandoned subspace contains the influence of the surroundings, so that its effect needs to be somehow included in the effective Hamiltonian. The standard way would be to calculate such additional contributions by lower order perturbation theory, which would be quite complicated due to the presence of infinitely many degrees of freedom.

An alternative approach is to construct effective Hamiltonian from a symmetry consideration. Hamiltonian matrix for a finite subspace of a transition metal ion can be described by a linear combination of effective tensor operators with known symmetry character. A detailed theoretical framework is constructed on the rigid basis of group theory [11], which of course consists of a lot of mathematical details. A simplified version of effective tensor operator can be seen in the case where the finite subspace in question consists of eigenfunctions of angular momentum operator $\boldsymbol{L}_{\mathrm{et}}$. Then, any element of the Hamiltonian matrix in this subspace can be written as a linear combination of the operators produced by the products of the components of $\boldsymbol{L}_{\mathrm{et}}$. Due to the finite size of the subspace, the number of the products required to describe the matrix is finite. This set of products is the effective tensor operators for this example. Their linear combination with an arbitrary coefficient for each term becomes the most general form of Hamiltonian for the assumed functional subspace.

The coefficients of the effective tensor operators are regarded as parameters, and can be determined by appropriate experiments. Since the necessary types and numbers of the tensor operators are determined uniquely by the symmetry of the system, this is a convenient formulation to analyze experiments. A real theoretical thrill arises when one tries to calculate the coefficients of the tensor operators and compare them with experiments. But the existence of many elements in the perturbation calculation will hinder an accurate comparison. This would be an unsatisfactory point of this method, but it offers a very useful tool to analyze experiments. Similar method of analysis in terms of spin Hamiltonian was developed in the field of spin resonance [6].

Electrons in a Crystal

Spatial periodicity of crystals allows us a useful concept, band picture, in describing the quantum mechanics of electrons. It provides, not only a simple criterion distinguishing metals and insulators, but also the idea of holes and effective masses, etc., which contribute to the detailed understandings of the various properties of crystal electrons. Especially in semiconductor physics band picture has played an essential role. The detailed understanding of the materials has lead to the great progress of semiconductor technologies in the fields of computers, lasers, nanomaterials, etc.

For the arguments of various properties of crystal electrons, one introduces single electron states for an electron moving in a periodic potential. The eigenstates are written in the form of Bloch functions as

$$\phi_{\mathrm{k}\lambda}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}} u_{\mathrm{k}\lambda}(\boldsymbol{r}) \tag{2.95}$$

where $\boldsymbol{k}$ is a wave vector, λ band index, and $u_{\mathrm{k}\lambda}(\boldsymbol{r})$ a function with lattice periodicity. The quantum number $\boldsymbol{k}$ takes continuous values in a Brillouin zone, so that the energy eigenvalues form a band structure, i.e., multi-branch dispersion curves as functions of $\boldsymbol{k}$. The band index λ corresponds to the branch index, but it also contains spin part σ. The spin part for a given $\boldsymbol{k}$ is not pure up or down spin state, but a mixture of them. Thus $\sigma = (\uparrow, \downarrow)$ indicates the dominant part of spin direction. The importance of H_{SO} arises for crystals containing heavy elements in general, or when we are interested in the degenerate bands in the absence of H_{SO}.

There is a simple symmetry rule about the form of energy eigenvalues. In terms of the single particle Schrödinger equation

$$H\phi_{\mathrm{k}\lambda} = \varepsilon_{\lambda}(\mathrm{k})\phi_{\mathrm{k}\lambda}, \tag{2.96}$$

the behavior of the eigenvalues for space inversion and time reversal is noteworthy. For space inversion, k changes its sign, while spin does not, and for time reversal, both k and spin change their signs. Therefore, if the system has space inversion symmetry, the states with (k, σ) and $(-k, \sigma)$ are degenerate, and if the system has time inversion symmetry, the states with (k, σ) and $(-k, -\sigma)$ are degenerate. If the system has both inversion and time reversal symmetry, all of the states with $(\pm k, \pm\sigma)$ are degenerate. If we apply this argument to a doubly degenerate band at $k = 0$ in a system lacking in inversion symmetry, the degeneracy is lifted linearly in k.

The band picture plays an important role in the development of semiconductor physics, where the detailed band structure near the band edges (highest occupied states, and lowest unoccupied states) affects many of the essential properties of semiconductors, such as effective masses of electron and hole, k-linear terms, donor and acceptor states, exciton states, etc. The precise understanding of these properties for various semiconductors has been the basis of the great development of semiconductor technology.

Semiconductors are the first group of materials providing well controlled nanostructures, i.e., quantum wells, wires and dots, and more complex structures. They were the materials showing "size, shape, and internal structure"-dependent EM responses, which opened a new field, "nano physics", after a long period of macroscopic response of condensed matter physics.

The success of the band picture in semiconductors is based on the effective mass theory. The one-electron eigen functions for a given $\boldsymbol{k}$ form a complete set, so that the eigen functions for a different $\boldsymbol{k}$ can be expanded in terms of them. A typical example is the case of the band edge occurring at $k = 0$. The energy band structure in the neighboring $\boldsymbol{k}(\neq 0)$ can be calculated by perturbation theory, where the unperturbed Hamiltonian is the one for $k = 0$, which gives the perturbation term $(\hbar/m_0)\boldsymbol{k} \cdot \boldsymbol{p}$. This scheme, called "$\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory", was used to calculate the details of semiconductor band edges (top of valence bands and bottom of conduction bands), and further extended to the studies of cyclotron resonance, donor and acceptor states, exciton states, etc. For effective mass theory, see [12].

Here also, effective Hamiltonian formalisms were developed. A most useful, nontrivial one is the Luttinger Hamiltonian for the top of valence band in typical cubic semiconductors, which consists mainly of the p-like ($\ell = 1$) wave functions of the constituent atoms. The three orbital components coupled with two spin parts split into two groups with total angular momentum $j = 3/2$ and $j = 1/2$, which consist of 4 and 2 components, respectively. Many semiconductors have $j = 3/2$ bands as the topmost valence bands at $k = 0$. Because of the four-fold degeneracy the band structure is quite complicated. There arise heavy and light effective masses, which depend on the direction of $\boldsymbol{k}$ (warping), and, in the absence of inversion symmetry as in zincblende type crystals, there appear $\boldsymbol{k}$-linear terms. Including the magnetic field dependence, which is necessary to study these details via cyclotron resonance, we can write the effective Hamiltonian of the $j = 3/2$ bands as a linear combination of independent products of the components of $j = 3/2$ angular momentum vectors [13]

$$\begin{aligned}
H_{\mathrm{L}} = \frac{1}{m_0}\Bigg[&\left(\gamma_1 + \frac{5}{2}\gamma_2\right)\frac{k^2}{2} \\
&+\gamma_2(k_x^2 j_x^2 + k_y^2 j_y^2 + k_z^2 j_z^2) \\
&-2\gamma_3(\{k_y k_z\}\{j_y j_z\} + \{k_z k_x\}\{j_z j_x\} + \{k_x k_y\}\{j_x j_y\})\Bigg] \\
&+2\mu_{\mathrm{B}}\kappa\, \boldsymbol{j}\cdot\boldsymbol{B} + 2\mu_{\mathrm{B}} q_{\mathrm{L}}(B_x j_x^3 + B_y j_y^3 + B_z j_z^3) \\
&-K_\ell\left[k_x\{(j_y^2 - j_z^2)j_x\} + k_y\{(j_z^2 - j_x^2)j_y\} + k_z\{(j_x^2 - j_y^2)j_z\}\right],
\end{aligned} \tag{2.97}$$

where $\gamma_1, \gamma_2, \gamma_3, \kappa, q_{\mathrm{L}}, K_\ell$ are parameters to be determined by measurements, $\boldsymbol{B}$ the static magnetic field, and a short hand notation $\{ab\} = (ab + ba)/2$ is used. This is the general form of the 4×4 Hermitean matrix spanned over the eigenstates of $j = 3/2$. The operators in the product forms of j_x, j_y, j_z are the "effective tensor operators" of this case. This is quite a useful theoretical invention for the study of semiconductors in the presence of spin-orbit interaction.

The effective mass theory is extended to the study of shallowly bound states of donors and acceptors, Landau levels, exciton states etc. and the effective Hamiltonian method are further developed to the study of symmetry breaking effects of exciton states [14].

Emergent Electromagnetic Field

Recently there is a new trend about the fundamental way of thinking on the electromagnetism of matter among the researchers working on spin current, spin Hall effect, multiferroics, topological insulators and defects, etc. [9]. For this purpose, it is proposed to reformulate EM response theory on a very fundamental level. One rewrites the weakly relativistic Hamiltonian, derived from Dirac Hamiltonian in a EM field, into a form consisting of invariant combinations of matter variables and "emergent EM field components".

Compared with the spinless case, which is governed by Schrödinger equation for scalar wave function with usual gauge field $\{\boldsymbol{A}, \phi\}$, the weakly relativistic case is described by Pauli equation for spinor wave function, where the additional terms "spin-orbit and spin Zeeman interactions" bring about a new gauge field for the spinor wave function. The proposed rewriting gives the new form of Pauli equation as [15][1]

$$i\hbar D_0\psi = -\frac{\hbar^2}{2m_0}D_{\mathrm{i}}^2\psi - \frac{1}{2m_0}\left(2e\bar{q}\frac{\tau^{\mathrm{a}}}{2}\boldsymbol{A}\cdot\boldsymbol{A}^{\mathrm{a}} + \frac{\bar{q}^2}{4}\boldsymbol{A}^{\mathrm{a}}\cdot\boldsymbol{A}^{\mathrm{a}}\right)\psi\ , \tag{2.98}$$

where ψ is $s = 1/2$ spinor wave function, τ^{a} Pauli matrices ($\boldsymbol{\sigma} = \boldsymbol{\tau}/2$), indices (i, a) represent space components (x, y, z). Constant $\bar{q}$ is Bohr magneton with relativistic correction. The operators D_0, D_{i} are covariant derivatives defined as

$$D_0 = \partial_0 - i\frac{\bar{q}}{\hbar}A_0^{\mathrm{a}}\frac{\tau^{\mathrm{a}}}{2} - i\frac{e}{\hbar}A_0\ , \tag{2.99}$$

$$D_{\mathrm{i}} = \partial_{\mathrm{i}} + i\frac{\bar{q}}{\hbar}A_{\mathrm{i}}^{a}\frac{\tau^{a}}{2} + i\frac{e}{\hbar}A_{\mathrm{i}}\ , \tag{2.100}$$

where $\partial_0 = \partial/\partial t$, $\partial_{\mathrm{i}} = \partial/\partial \mathrm{i}$. (A_0, A_{i}) are the time and space components of conventional gauge field, and those of the new gauge field are $(A_0^{\mathrm{a}}, A_{\mathrm{i}}^{\mathrm{a}})$ defined in terms of electric and magnetic fields $\boldsymbol{E}$, $\boldsymbol{B}$ and 3-dimensional Levi-Civita tensor as

$$A_0^{\mathrm{a}} = B^{\mathrm{a}}\ , \tag{2.101}$$

$$A_{\mathrm{i}}^{\mathrm{a}} = \varepsilon_{\mathrm{ia}\ell}E_\ell\ . \tag{2.102}$$

The case $\bar{q} = 0$ corresponds to the case of no relativistic correction, where Schrödinger equation for scalar field describes the system. In this case, the requirement of gauge theory is the invariance of this equation for the multiplication of a phase factor to the wave function and the corresponding change in the gauge potentials $(\boldsymbol{A}, \phi)$ in (2.65), (2.66), The phase function $\Theta(\boldsymbol{r}, t)$ is a real continuous function with arbitrary dependence on $\boldsymbol{r}$ and t, which means that the quantum mechanical behavior of matter does not depend on $\Theta(\boldsymbol{r}, t)$.

Similar requirement from the gauge theory to the Pauli equation for $s = 1/2$ spinor (SU(2)) field is that the wave function have the phase factor, not only of the type for scalar field, but also of the rotation of spin, $\exp(i\phi^{\mathrm{a}}\tau^{\mathrm{a}}/2)$. The gauge potentials in this case are $(A_0^{\mathrm{a}}, A_{\mathrm{i}}^{\mathrm{a}})$. They are coupled to "spin current" $(j_0^{\mathrm{a}}, j_i^{\mathrm{a}})$ defined as [9]

$$j_0^{\mathrm{a}} = \psi\sigma^{\mathrm{a}}\psi\ , \tag{2.103}$$

$$j_i^{\mathrm{a}} = \frac{1}{2m_0}[\psi^\dagger\sigma^{\mathrm{a}}D_{\mathrm{i}}\psi - D_{\mathrm{i}}\psi^\dagger\sigma^{\mathrm{a}}\psi] \tag{2.104}$$

where the indices a and i represent the directions of spin and its flow, respectively. This scheme is a case of general gauge theory for spin = 1/2 spinor field, sometimes called "emergent electromagnetism (EE)", and is used mainly for the analysis of spin related transport problems [8, 9, 15].

---

[1] On this subject we skip the [SI] form of equations, which seems to be rarely used.

It is not yet well examined how this EE theory (EE) is compared with the main scheme of this book, i.e., “microscopic nonlocal theory and its macroscopic form derived through LWA”. In the sense that they are both based on Pauli equation, they describe the same physics essentially. The aspect of SU(2) gauge theory is explicitly handled only in EE. Obviously it is a merit of EE to have an explicit parallelism with the general gauge theory of physics, which could provide a wide viewpoint in describing individual problems. This, however, would not restrict the other from handling a certain class of problems. Probably the difference may exist in how one formulate a problem. The theory of this book has mainly been used for spectroscopic problems, and transport problems have little been studied. For a better understanding of these two formulations, it would be useful, not only to discuss their formal structures, but also to find some concrete problems for testing the merit and demerit of them.

2.3 Long Wavelength Approximation (LWA)

We now proceed to make the macroscopic average of the fundamental equations of microscopic nonlocal response, i.e., the microscopic M-eqs (2.43) and the constitutive equation (2.38). As discussed previously, this means mathematically to take the LWA of these equations. Since it is an approximation, there is a validity condition which may be fulfilled or not according to the system in consideration. We will consider this problem later in Sect. 3.6. In this section, we just apply LWA, leaving the first few terms of the expansion. This means that we derive the expected form of macroscopic equations when LWA is a good approximation.

Application of LWA to a microscopic system does not necessarily lead to a uniform system in general. It is possible to result in a macroscopically non-uniform system, as, for example, in the case of an impurity system with macroscopically non-uniform distribution of density. Such a system would require an additional consideration after introducing a macroscopic description depending on the details of each problem. In this book, we omit these macroscopically non-uniform cases from our consideration. However, to complement this point, the case of resonant X-ray diffraction from a crystal will be discussed in Sect. 4.3.

The omission of macroscopically non-uniform systems after LWA allows us to work only in a uniform system, where wave vector $\boldsymbol{k}$ is a good quantum number. We use a space Fourier transform of a field $B(\boldsymbol{r})$ as

$$B(\boldsymbol{r}) = \frac{V}{8\pi^3}\int \mathrm{d}\boldsymbol{k}\ \tilde{B}(\boldsymbol{k})\exp(i\boldsymbol{k}\cdot\boldsymbol{r}), \quad \tilde{B}(\boldsymbol{k}) = \frac{1}{V}\int \mathrm{d}\boldsymbol{r}\ B(\boldsymbol{r})\exp(-i\boldsymbol{k}\cdot\boldsymbol{r})\ . \tag{2.105}$$

where V is a volume to define discrete $\boldsymbol{k}$ via periodic boundary condition, leading to the uniform density $V/8\pi^3$ of $\boldsymbol{k}$ in the continuum limit $V \to \infty$. The $(\boldsymbol{k}, \omega)$ Fourier component of microscopic M-eqs is ($q = \omega/c$)

$$(k^2 - q^2)\tilde{\boldsymbol{A}}(\boldsymbol{k}, \omega) = \frac{4\pi}{c}\tilde{\boldsymbol{I}}_{\rm t}^{\rm (T)}(\boldsymbol{k}, \omega) = \left[\mu_0 \tilde{\boldsymbol{I}}_{\rm t}^{\rm (T)}(\boldsymbol{k}, \omega)\right]_{\rm SI} . \tag{2.106}$$

This equation is not affected by LWA, except that both $\boldsymbol{A}$ and $\boldsymbol{I}^{\rm (T)}$ are appreciable only for small k. (If LWA is a good approximation, all the physical quantities should mainly consist of their LW components.)

The Fourier component of the constitutive equation (2.48) is

$$\begin{aligned}\tilde{\boldsymbol{I}}_{\rm t}(\boldsymbol{k}, \omega) &= \frac{1}{c}\sum_{\nu}\left[\bar{g}_{\nu}(\omega)\tilde{\boldsymbol{I}}_{0\nu}(\boldsymbol{k})F_{\nu 0}(\omega) + \bar{h}_{\nu}(\omega)\tilde{\boldsymbol{I}}_{\nu 0}(\boldsymbol{k})F_{0\nu}((\omega)\right] , \quad (2.107)\\ &= \left[\text{same expression without } 1/c\right]_{\rm SI} ,\end{aligned}$$

and the factor $F_{\mu\nu}$ is written in terms of the Fourier components as

$$F_{\mu\nu}(\omega) = \frac{V^2}{8\pi^3}\int {\rm d}\boldsymbol{k}\ \tilde{\boldsymbol{I}}_{\mu\nu}(-\boldsymbol{k}) \cdot \tilde{\boldsymbol{A}}(\boldsymbol{k}) . \tag{2.108}$$

Substituting (2.108) into (2.107), we have

$$\begin{aligned}\tilde{\boldsymbol{I}}_{\rm t}(\boldsymbol{k}, \omega) = \frac{V^2}{8\pi^3 c}\int {\rm d}\boldsymbol{k}' \sum_{\nu} &\left[\bar{g}_{\nu}(\omega)\tilde{\boldsymbol{I}}_{0\nu}(\boldsymbol{k})\tilde{\boldsymbol{I}}_{\nu 0}(-\boldsymbol{k}')\right.\\ &\left. + \bar{h}_{\nu}(\omega)\tilde{\boldsymbol{I}}_{\nu 0}(\boldsymbol{k})\tilde{\boldsymbol{I}}_{0\nu}(-\boldsymbol{k}')\right] \cdot \tilde{\boldsymbol{A}}(\boldsymbol{k}') \quad (2.109)\\ = \left[\text{same expression without } 1/c\right]_{\rm SI} .&\end{aligned}$$

In evaluating the matrix element of the current density in LWA, we begin with the operator form of the current density

$$\begin{aligned}\tilde{\boldsymbol{I}}(\boldsymbol{k}) &= \frac{1}{V}\int {\rm d}\boldsymbol{r}\ {\rm e}^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\ \{\boldsymbol{J}_0(\boldsymbol{r}) + c\nabla \times \boldsymbol{M}_{\rm s}(\boldsymbol{r})\} \quad (2.110)\\ &= \frac{1}{V}\int {\rm d}\boldsymbol{r}\ {\rm e}^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\ \boldsymbol{J}_0(\boldsymbol{r}) + i\frac{c}{V}\boldsymbol{k} \times \int {\rm d}\boldsymbol{r}\ {\rm e}^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\ \boldsymbol{M}_{\rm s}(\boldsymbol{r}) , \quad (2.111)\\ &= \left[\text{same expression without } c\right]_{\rm SI} ,\end{aligned}$$

where $\boldsymbol{J}_0$ is the $\boldsymbol{A}$ independent part of orbital current density, (2.13), and $\boldsymbol{M}_{\rm s}$ is the spin magnetization, (2.21), and the second equation is derived via partial integration. The $(\mu\nu)$ matrix element of this operator is given by the same expression with $\boldsymbol{J}_0(\boldsymbol{r})$ and $\boldsymbol{M}_{\rm s}(\boldsymbol{r})$ replaced by the matrix elements $\langle\mu|\boldsymbol{J}_0(\boldsymbol{r})|\nu\rangle$ and $\langle\mu|\boldsymbol{M}_{\rm s}(\boldsymbol{r})|\nu\rangle$.

If LWA is a good approximation, we can expand $\exp(i\boldsymbol{k} \cdot \boldsymbol{r})$ in Taylor series, and keep the first few terms. These terms are the various moments of $\langle\mu|\boldsymbol{J}_0(\boldsymbol{r})|\nu\rangle$ and $\langle\mu|\boldsymbol{M}_{\rm s}(\boldsymbol{r})|\nu\rangle$. Since the values of the moments depend on the center about which they are defined, we need to specify the center. In the situation where LWA is a good approximation, the transition $\mu \leftrightarrow \nu$ is usually localized, so that we can choose a "center" in the region where the wave functions Ψ_{μ} and Ψ_{ν} have appreciable amplitudes. Let us denote the center as $\bar{\boldsymbol{r}}$. Then, the matrix element of the current density in LWA is given as

$$\begin{aligned}\tilde{\boldsymbol{I}}_{\mu\nu}(\boldsymbol{k}) &= \frac{\exp(-i\boldsymbol{k}\cdot\bar{\boldsymbol{r}})}{V}\left[\bar{\boldsymbol{J}}_{\mu\nu} - i\boldsymbol{k}\cdot\bar{\mathbf{Q}}_{\mu\nu} + ic\boldsymbol{k}\times\bar{\boldsymbol{M}}^{(s)}_{\mu\nu} + \cdots\right] \\ &= \left[\text{same expression without } c\right]_{\text{SI}}\end{aligned} \tag{2.112}$$

up to the $O(k^1)$ terms, where

$$\bar{\boldsymbol{J}}_{\mu\nu} = \int \mathrm{d}\boldsymbol{r}\ \langle\mu|\boldsymbol{J}_0(\boldsymbol{r})|\nu\rangle\ , \tag{2.113}$$

$$\bar{\mathbf{Q}}_{\mu\nu} = \int \mathrm{d}\boldsymbol{r}\ (\boldsymbol{r}-\bar{\boldsymbol{r}})\langle\mu|\boldsymbol{J}_0(\boldsymbol{r})|\nu\rangle\ , \tag{2.114}$$

$$\bar{\boldsymbol{M}}^{(s)}_{\mu\nu} = \int \mathrm{d}\boldsymbol{r}\ \langle\mu|\boldsymbol{M}_{\text{spin}}(\boldsymbol{r})|\nu\rangle \tag{2.115}$$

From the form of the one particle operators included in $\boldsymbol{J}_0(\boldsymbol{r})$, the matrix element $\bar{\boldsymbol{J}}_{\mu\nu}$ is nonzero when the transition $\mu \leftrightarrow \nu$ contains electric dipole (E1) character, and $\bar{\mathbf{Q}}_{\mu\nu}$ is nonzero for the transition with magnetic dipole (M1) and electric quadrupole (E2) characters. We can explicitly separate $\bar{\mathbf{Q}}_{\mu\nu}$ into the M1 and E2 components as shown in Sect. 5.6. This allows us to rewrite (2.112) as

$$\begin{aligned}\tilde{\boldsymbol{I}}_{\mu\nu}(\boldsymbol{k}) &= \frac{\exp(-i\boldsymbol{k}\cdot\bar{\boldsymbol{r}})}{V}\left[\bar{\boldsymbol{J}}_{\mu\nu} - i\boldsymbol{k}\cdot\bar{\mathbf{Q}}^{(\text{e2})}_{\mu\nu} + ic\boldsymbol{k}\times\bar{\boldsymbol{M}}_{\mu\nu} + \cdots\right], \\ &= \left[\text{same expression without } c\right]_{\text{SI}}\ ,\end{aligned} \tag{2.116}$$

where $\bar{\boldsymbol{M}}_{\mu\nu}$ is the sum of spin and orbital magnetizations

$$\bar{\boldsymbol{M}}_{\mu\nu} = \bar{\boldsymbol{M}}^{(\text{s})}_{\mu\nu} + \bar{\boldsymbol{M}}^{(\text{orb})}_{\mu\nu}\ . \tag{2.117}$$

The explicit form of the orbital magnetization $\bar{\boldsymbol{M}}^{(\text{orb})}_{\mu\nu}$ is given in Sect. 5.6. Substituting this expression into (2.109), we can express the induced current density in terms of the separate contributions of E1, E2, and M1 transitions. Because of the assumption, at the beginning of this section, of neglecting the non-uniformity in the LWA averaged system, we should choose the $\boldsymbol{k}' = \boldsymbol{k}$ term in the integral over $\boldsymbol{k}'$ in (2.109). Supplying $(8\pi^3/V)\delta(\boldsymbol{k}-\boldsymbol{k}')$ (which corresponds to $\delta_{\boldsymbol{k},\boldsymbol{k}'}$ in discrete case) in the $\boldsymbol{k}'$-integral, we obtain

$$\begin{aligned}\tilde{\boldsymbol{I}}(\boldsymbol{k},\omega) &= \frac{V}{c}\sum_{\nu}\left[\bar{g}_{\nu}(\omega)\tilde{\boldsymbol{I}}_{0\nu}(\boldsymbol{k})\tilde{\boldsymbol{I}}_{\nu 0}(-\boldsymbol{k}) + \bar{h}_{\nu}(\omega)\tilde{\boldsymbol{I}}_{\nu 0}(\boldsymbol{k})\tilde{\boldsymbol{I}}_{0\nu}(-\boldsymbol{k})\right]\cdot\tilde{\boldsymbol{A}}(\boldsymbol{k})\ , \\ &= \left[\text{same expression without } 1/c\right]_{\text{SI}}\ ,\end{aligned} \tag{2.118}$$

i.e., the susceptibility is

$$\chi_{\mathrm{em}}(\boldsymbol{k},\omega) = \frac{V}{c}\sum_{\nu}\left[\bar{g}_{\nu}(\omega)\tilde{\boldsymbol{I}}_{0\nu}(\boldsymbol{k})\tilde{\boldsymbol{I}}_{\nu 0}(-\boldsymbol{k}) + \bar{h}_{\nu}(\omega)\tilde{\boldsymbol{I}}_{\nu 0}(\boldsymbol{k})\tilde{\boldsymbol{I}}_{0\nu}(-\boldsymbol{k})\right]$$
$$= \left[\text{same expression without } c\right]_{\mathrm{SI}} . \tag{2.119}$$

Note that the explicit $\bar{\boldsymbol{r}}$-dependence in (2.116) cancels out in this expression because of $\boldsymbol{k} = \boldsymbol{k}'$. When a same transition occurs at various positions with number density n_0, the factor $(1/V)\sum_{\nu}$ is replaced by $n_0\sum'_{\nu}$, where the prime on the summation sign means that a same transition (at different positions) is counted only once.

The T component of the induced current density, required in the M-eqs (2.106), is

$$\boldsymbol{I}_{\mathrm{t}}^{(\mathrm{T})} = \chi_{\mathrm{em}}^{(\mathrm{T})}\boldsymbol{A} \tag{2.120}$$

where

$$\chi_{\mathrm{em}}^{(\mathrm{T})}(\boldsymbol{k},\omega) = \left(\mathbf{1} - \hat{\boldsymbol{k}}\hat{\boldsymbol{k}}\right)\cdot\chi_{\mathrm{em}}(\boldsymbol{k},\omega)\ , \quad (\hat{\boldsymbol{k}} = \boldsymbol{k}/|\boldsymbol{k}|)\ . \tag{2.121}$$

The inner product $\hat{\boldsymbol{k}}\cdot(\hat{\boldsymbol{k}}\cdot\bar{\mathbf{Q}})$ appearing in this quantity is defined as

$$\hat{\boldsymbol{k}}\cdot(\hat{\boldsymbol{k}}\cdot\bar{\mathbf{Q}}) = \sum_{\xi}\sum_{\eta}\sum_{\zeta}\hat{k}_{\xi}\hat{k}_{\eta}\bar{Q}_{\eta\xi}\ . \tag{2.122}$$

This allows us to calculate the dispersion equation of the coupled waves of $\tilde{\boldsymbol{A}}$ and $\tilde{\boldsymbol{I}}_{\mathrm{t}}^{(\mathrm{T})}$ in a general form.

2.4 New Macroscopic Susceptibility

We have derived the LWA average of microscopic susceptibility to be used in the macroscopic constitutive equation. Combining this susceptibility with the M-eqs (2.106), we can selfconsistently determine the T components of vector potential and induced current density. Since a vector potential contains both electric and magnetic fields, and a current density is written in terms of the matrix elements of "E1, E2, M1 ..." characters, this selfconsistent solution describes the complete (linear) response of a coupled matter-EM field system for T field excitation. (The L components of $\boldsymbol{E}$ and $\boldsymbol{I}$, if they are allowed by symmetry, can be determined from the selfconsistent solution of the T components. See the last argument of this sub-section.)

The new macroscopic susceptibility relating $\tilde{\boldsymbol{I}}$ with $\tilde{\boldsymbol{A}}$ can be classified into $O(k^0)$, $O(k^1)$, $O(k^2)$ terms as

$$\chi_{\mathrm{em}}(\boldsymbol{k},\omega) = \chi_{\mathrm{em0}}(\omega) + k\ \chi_{\mathrm{em1}}(\hat{\boldsymbol{k}},\omega) + k^2\ \chi_{\mathrm{em2}}(\hat{\boldsymbol{k}},\omega) + \cdots \tag{2.123}$$

Each matrix element of current density consists of a sum of E1, E2, and M1 components, as seen from (2.112). Since the products of two matrix elements of current density are contained in the susceptibility χ_{em} in a dyadic form, the first one (on the l.h.s.) indicates the character (E1, E2, or M1) of the induced current density, and the second one represents the character of the interaction contributing to the term. In this sense, we can specify the contributions of (E1, E2, or M1) transitions in $\{\chi_{\text{emj}}(\omega);\ j = 1, 2, 3\}$.

The term $\chi_{\text{em0}}(\omega)$ has contribution only from the (E1, E1) transitions as

$$\begin{aligned}\chi_{\text{em0}}(\omega) &= \frac{1}{cV}\sum_{\nu}\left[\bar{g}(\omega)\bar{\boldsymbol{J}}_{0\nu}\ \bar{\boldsymbol{J}}_{\nu 0} + \bar{h}(\omega)\bar{\boldsymbol{J}}_{\nu 0}\ \bar{\boldsymbol{J}}_{0\nu}\right], \\ &= \left[\text{same expression without } 1/c\right]_{\text{SI}} .\end{aligned} \tag{2.124}$$

In a favorable symmetry condition, this term is related with the conventional electric susceptibility χ_{e}, as shown in Sect. 3.1

The term $\chi_{\text{em2}}(\omega)$ consists of the (M1 + E2, M1 + E2) transitions, i.e., the (M1, M1), (E2, E2) terms and their cross terms (M1, E2) and (E2, M1). The (M1, M1 + E2) terms contribute to the current density due to the induced magnetizations, and (E2, M1 + E2) terms contribute to the current density due to the induced electric quadrupole (E2) polarizations. The M1 and E2 characters can be distinguished, not by space inversion, but by time reversal. It will be shown in Sect. 3.1 that the (M1, M1) term can be used to derive the magnetic susceptibility χ_{B}, and $\mu = 1/(1 - 4\pi\chi_{\text{B}})$ in the case of non-chiral symmetry.

In contrast, the χ_{em1} term consists of the mixed transitions of (M1 + E2, E1) and (E1, M1 + E2) types. In order for this term to be non-vanishing, there must be the quantum mechanical excited states $\{|\nu\rangle\}$, which are active to both E1 and M1 (or E1 and E2) transitions. This is possible only when the system has no inversion symmetry, i.e., the case of chiral symmetry, or a system with optical activity. In this case, common poles appear among $\{\chi_{\text{emj}}(\omega);\ j = 1, 2, 3\}$.

If the system has inversion symmetry, all the excited states are classified according to the parity, so that the states contributing to E1 transitions and (M1 and E2) transitions belong to different irreducible representations. No excited state is active to both E1 and (M1 and E2) transitions, so that χ_{em1} is zero in this case. Then, the summation over the index ν can be divided into two groups, each of which contributes to either E1 or M1 (E2), i.e., the excitations contributing to χ_{em0} are different from those contributing to χ_{em2}.

Though the induced current density is given as a power series expansion about $\boldsymbol{k}$, it would be more physical to decompose it into the current densities due to electric field-induced electric polarization $\boldsymbol{I}_{\text{eE}}$, magnetic field-induced magnetic polarization $\boldsymbol{I}_{\text{mB}}$, magnetic field-induced electric polarization $\boldsymbol{I}_{\text{eB}}$, and electric field-induced magnetic polarization $\boldsymbol{I}_{\text{mE}}$. Their explicit forms are obtained by the substitution of (2.116) into (2.118):

$$\begin{aligned}\boldsymbol{I}_{\text{eE}} &= \frac{1}{cV}\sum_{\nu}\left[\bar{g}_{\nu}(\bar{\boldsymbol{J}}_{0\nu} - i\boldsymbol{k}\cdot\bar{\mathbf{Q}}^{(\text{e2})}_{0\nu})(\bar{\boldsymbol{J}}_{\nu 0} + i\boldsymbol{k}\cdot\bar{\mathbf{Q}}^{(\text{e2})}_{\nu 0})\right. \\ &\qquad \left.+\bar{h}_{\nu}(\bar{\boldsymbol{J}}_{\nu 0} - i\boldsymbol{k}\cdot\bar{\mathbf{Q}}^{(\text{e2})}_{\nu 0})(\bar{\boldsymbol{J}}_{0\nu} + i\boldsymbol{k}\cdot\bar{\mathbf{Q}}^{(\text{e2})}_{0\nu})\right]\cdot\tilde{\boldsymbol{A}}(\boldsymbol{k}) , \\ &= \left[\text{same expression without } 1/c\right]_{\text{SI}} ,\end{aligned} \tag{2.125}$$

$$\boldsymbol{I}_{\mathrm{mB}} = \frac{c}{V}\,\boldsymbol{k} \times \sum_{\nu} \left[\bar{g}_{\nu} \bar{\boldsymbol{M}}_{0\nu}(\boldsymbol{k} \times \bar{\boldsymbol{M}}_{\nu 0}) + \bar{h}_{\nu} \bar{\boldsymbol{M}}_{\nu 0}(\boldsymbol{k} \times \bar{\boldsymbol{M}}_{0\nu})\right] \cdot \tilde{\boldsymbol{A}}(\boldsymbol{k}) \,, \tag{2.126}$$

$$= \left[\text{same expression without } c\right]_{\mathrm{SI}} \,,$$

$$\begin{aligned} \boldsymbol{I}_{\mathrm{eB}} = \frac{-i}{V} \sum_{\nu} \Big[&\bar{g}_{\nu}(\bar{\boldsymbol{J}}_{0\nu} - i\boldsymbol{k} \cdot \bar{\mathbf{Q}}^{(\mathrm{e2})}_{0\nu})(\boldsymbol{k} \times \bar{\boldsymbol{M}}_{\nu 0}) \\ &+ \bar{h}_{\nu}(\bar{\boldsymbol{J}}_{\nu 0} - i\boldsymbol{k} \cdot \bar{\mathbf{Q}}^{(\mathrm{e2})}_{\nu 0})(\boldsymbol{k} \times \bar{\boldsymbol{M}}_{0\nu})\Big] \cdot \tilde{\boldsymbol{A}}(\boldsymbol{k}) \,, \end{aligned} \tag{2.127}$$

$$\begin{aligned} \boldsymbol{I}_{\mathrm{mE}} = \frac{i}{V}\,\boldsymbol{k} \times \sum_{\nu} \Big[&\bar{g}_{\nu} \bar{\boldsymbol{M}}_{0\nu}(\bar{\boldsymbol{J}}_{\nu 0} + i\boldsymbol{k} \cdot \bar{\mathbf{Q}}^{(\mathrm{e2})}_{\nu 0}) \\ &+ \bar{h}_{\nu} \bar{\boldsymbol{M}}_{\nu 0}(\bar{\boldsymbol{J}}_{0\nu} + i\boldsymbol{k} \cdot \bar{\mathbf{Q}}^{(\mathrm{e2})}_{0\nu})\Big] \cdot \tilde{\boldsymbol{A}}(\boldsymbol{k}) \,. \end{aligned} \tag{2.128}$$

In Sect. 3.1, this result is used to rewrite the constitutive equations in terms of electric and magnetic polarizations. This is an attempt to reproduce the conventional form of macroscopic M-eqs, but the result shows the difference in an essential way.

When the symmetry allows the mixing of the T and L components of response, there are non-zero elements in χ_{em} describing the L component of current density induced by the T field $\boldsymbol{A}$. In this case,

$$\tilde{\boldsymbol{I}}^{(\mathrm{L})}(\boldsymbol{k}, \omega) = \hat{\boldsymbol{k}}\hat{\boldsymbol{k}} \cdot \chi_{\mathrm{em}}(\boldsymbol{k}, \omega) \cdot \boldsymbol{A}(\boldsymbol{k}, \omega) \tag{2.129}$$

is a non-zero vector. The magnitude of this vector is determined, by substituting $\boldsymbol{A}$ of the selfconsistent solution ($\boldsymbol{A}$ and $\boldsymbol{I}^{(\mathrm{T})}$) into the r.h.s. of this equation. If, in this case, there exists also an external L-field, it induces the T, as well as L, components of current density. This case is treated in Sect. 5.7 and the result is shown to be neatly combined with that of the T field excitation in Sect. 3.2.

Thus, the single susceptibility tensor $\chi_{\mathrm{em}}(\boldsymbol{k}, \omega)$ describes all the possible situations, including electric and magnetic polarizations, and their mutual interference effect due to chiral symmetry. It should be stressed that this result is not a phenomenology, but a first-principles theory with explicit quantum mechanical expressions in a model-independent form, which allows both symmetry arguments, as given above, and numerical analysis of model systems.

2.5 Dispersion Equation

The coupled equations for $\boldsymbol{A}(\boldsymbol{k}, \omega)$ and $\tilde{\boldsymbol{I}}^{(T)}_{\mathrm{t}}(\boldsymbol{k}, \omega)$, (2.106) and (2.120), have a solution when a particular relation between $\boldsymbol{k}$ and ω, i.e., dispersion relation, is satisfied. Such an equation is obtained by substituting (2.120) into the M-eqs (2.106) as

$$(k^2 - q^2)\boldsymbol{A}(\boldsymbol{k}, \omega) = \frac{4\pi}{c}\ \chi_{\rm em}^{(\rm T)}(\boldsymbol{k}, \omega)\boldsymbol{A}(\boldsymbol{k}, \omega)\ , \tag{2.130}$$
$$= \left[\text{same expression with } 4\pi/c \text{ replacedby } \mu_0\right]_{\rm SI}\ ,$$

($q = \omega/c$). This is the homogeneous linear equations for the two T components of $\boldsymbol{A}$, and the condition for the existence of non-trivial solution is the vanishing of the determinant of the (2 × 2) coefficient matrix, i.e.,

$$\det|\frac{c^2k^2}{\omega^2}\mathbf{1} - \left\{\mathbf{1} + \frac{4\pi c}{\omega^2}\ \chi_{\rm em}^{(\rm T)}(\boldsymbol{k}, \omega)\right\}| = 0 \tag{2.131}$$
$$\left[\text{same expression with } 4\pi c \text{ replacedby } 1/\varepsilon_0\right]_{\rm SI}\ ,$$

In the conventional case, the dispersion equation is obtained from the M-eqs $\nabla \times \nabla \times \boldsymbol{E} = (\omega^2/c^2)\varepsilon\mu\boldsymbol{E}$, rewritten by eliminating magnetic field. The condition for the existence of non-trivial solution of T-character leads to

$$\det|\frac{c^2k^2}{\omega^2}\mathbf{1} - \{(\mathbf{1} + 4\pi\chi_{\rm e})(\mathbf{1} + 4\pi\chi_{\rm m})\}^{(\rm T)}| = 0\ , \tag{2.132}$$

$$\left[\det|\frac{k^2}{\omega^2}\mathbf{1} - \varepsilon_0\mu_0\{(\mathbf{1} + \chi_{\rm e})(\mathbf{1} + \chi_{\rm m})\}^{(\rm T)}| = 0\right]_{\rm SI}\ . \tag{2.133}$$

As mentioned in Sect. 1.5, it looks that the contributions of electric and magnetic transitions occur as a product in (2.132), while all the transitions in the new dispersion equation (2.131) occur as a sum of single poles of $\chi_{\rm em}$. Since the E1 and (M1 and E2) transitions are mutually mixed in chiral symmetry, $\chi_{\rm e}$ and $\chi_{\rm m}$ will have common poles, which will lead to the occurrence of second order poles in the $\varepsilon\mu$ part of (2.132). This is a clear distinction from the new result, and requires explanation. This apparent contradiction can be solved, in the case of non-chiral symmetry, by using the magnetic susceptibility defined as $\boldsymbol{M} = \chi_{\rm B}\boldsymbol{B}$, as shown in Sect. 3.3.

In the case of chiral symmetry, the dispersion equation of the conventional scheme needs a modification from (2.132). For this purpose, a phenomenology has been used with the name of Drude–Born–Fedorov constitutive equations, which generalize the relations $\boldsymbol{D} = \varepsilon\boldsymbol{E},\quad \boldsymbol{B} = \mu\boldsymbol{H}$ so as to include the effect of “magnetic field induced electric polarization” and “electric field induced magnetic polarization”. The dispersion equation in this case is also different from the new one (2.131), which will be discussed in Sect. 3.4. Though the form (2.130) is valid in both chiral and non-chiral symmetries, one could rewrite it in terms of the susceptibilities defined with respect to the electric and magnetic fields $\boldsymbol{E}$ and $\boldsymbol{B}$ (not $\boldsymbol{E}$ and $\boldsymbol{H}$), which leads to a set of constitutive equations somewhat similar to, but essentially different from, the DBF eqs, as shown in Sects. 3.1 and 3.4.

If the symmetry of matter is low, the excited states contributing to the poles of the susceptibilities $\chi_{\rm em}$, $\chi_{\rm e}$ and $\chi_{\rm m}$ may have LT-mixed character. This aspect is automatically taken care of by these susceptibilities through the pole positions and the residues of each term of the summands, though the T parts of the susceptibilities are selected in these dispersion equations. The contribution of pure L modes is excluded because of the vanishing interaction with ***A*** (T-field).

References

1. Cho, K.: Optical Response of Nanostructures: Microscopic Nonlocal Theory. Springer, Heidelberg (2003)
2. Cho, K.: J. Phys. Condens. Matter **20**, 175202 (2008)
3. Cho, K.: J. Phys. Soc. Jpn. **68**, 683 (1999)
4. Schiff, L.I.: Quantum Mechanics, 2nd edn. McGraw-Hill, New York (1955)
5. Cohen-Tannoudji, C., Diu, B., Laloë, F.: Quantum Mechanics, p. 1213. Hermann, Paris (1977)
6. Slichter, C.P.: Principles of Magnetic Resonance. Harper & Row, New York (1963)
7. Cohen-Tannoudji, C., Dupont-Roc, J., Grynberg, G.: Photons and Atoms. Wiley Interscience, New York (1989)
8. Fröhlich, J., Studer, U.M.: Rev. Mod. Phys. **65**, 733 (1993)
9. Nagaosa, N., Tokura, Y.: Phys. Scr. **T146**, 014020 (2012)
10. Condon, E.U., Shortley, G.H.: The Theory of Atomic Spectra. Cambridge University Press, Cambridge (1963)
11. Sugano, S., Tanabe, Y., Kamimura, H.: Multiplets of Transition Metal Ions in Crystals. Academic, New York (1970)
12. Kohn, W.: In: Seitz, F., Turnbull, D. (eds.) Solid State Physics, vol. 5, p. 258. Academic, New York (1957)
13. Luttinger, J.M.: Phys. Rev. **102**, 1030 (1956)
14. Cho, K.: Phys. Rev. B **14**, 4463 (1976)
15. Leurs, B.W.A., Nazario, Z., Santiago, D.I., Zaanen, J.: Ann. Phys. **323**, 907 (2008)

Chapter 3
Discussions of the New Results

Abstract New results are discussed from various angles. A reversible rewriting of the single susceptibility constitutive equation leads to a first-principles definition of $\boldsymbol{P}$ and $\boldsymbol{M}$ induced by both $\boldsymbol{E}$ and $\boldsymbol{B}$. This contains the microscopic definition of constitutive equations in chiral medium, more reliable than the phenomenological DBF eqs, and their comparison is made in details. A comparison with other types of single susceptibility theories of EM response, including that of Landau-Lifshitz, is made, which shows the advanced nature of the present theory. A short discussion about LWA is given, indicating its positive and negative meanings depending on the problem in consideration. As a special example of the application of this theory, dispersion curve and transmission window in a left-handed chiral medium are discussed. The aspects of L electric field is described.

3.1 Rewriting of the New Constitutive Equation

In Sect. 2.4, we have decomposed the new constitutive equation into the terms due to the electric field-induced electric polarization $\boldsymbol{I}_{\mathrm{eE}}$, magnetic field-induced magnetic polarization $\boldsymbol{I}_{\mathrm{mB}}$, magnetic field-induced electric polarization $\boldsymbol{I}_{\mathrm{eB}}$, and electric field-induced magnetic polarization $\boldsymbol{I}_{\mathrm{mE}}$, as in (2.125)–(2.128). In this form, there are two points of worth noting. The two terms due to induced magnetizations, $\boldsymbol{I}_{\mathrm{mB}}$ and $\boldsymbol{I}_{\mathrm{mE}}$, have the factor $\boldsymbol{k} \times \boldsymbol{M}_{0\nu}$ in front of their expressions. This means that these induced current densities correspond to the form "$\nabla\times$ magnetization".

The second noteworthy point is that $\boldsymbol{I}_{\mathrm{eB}}$ and $\boldsymbol{I}_{\mathrm{mB}}$ contain the factor $(\boldsymbol{k} \times \bar{\boldsymbol{M}}_{0\nu}) \cdot \tilde{\boldsymbol{A}}(\boldsymbol{k})$ at the end of their expressions. If we use the manipulation

$$(\boldsymbol{k} \times \bar{\boldsymbol{M}}_{0\nu}) \cdot \tilde{\boldsymbol{A}}(\boldsymbol{k}) = -(\boldsymbol{k} \times \tilde{\boldsymbol{A}}(\boldsymbol{k})) \cdot \bar{\boldsymbol{M}}_{0\nu} = i\boldsymbol{B}(\boldsymbol{k}) \cdot \bar{\boldsymbol{M}}_{0\nu} \ , \tag{3.1}$$

we understand that these components of the induced current density are the linear response of the system to the applied magnetic field $\boldsymbol{B}$. On the other hand, the induced electric polarization terms, $\boldsymbol{I}_{\mathrm{eE}}$ and $\boldsymbol{I}_{\mathrm{mE}}$, contain the factor $(\bar{\boldsymbol{J}}_{0\nu} - i\boldsymbol{k} \cdot \bar{\mathbf{Q}}^{(\mathrm{e2})}_{0\nu}) \cdot \tilde{\boldsymbol{A}}(\boldsymbol{k}, \omega)$ at the end. Rewriting $\tilde{\boldsymbol{A}}(\boldsymbol{k}, \omega)$ into $-i(c/\omega)\boldsymbol{E}^{(\mathrm{T})}(\boldsymbol{k}, \omega)$, we see that these terms

K. Cho, *Reconstruction of Macroscopic Maxwell Equations*,
Springer Tracts in Modern Physics 237,
https://doi.org/10.1007/978-3-662-58424-8_3

are caused by the interaction with (transverse) electric field. By using these relations, we can rewrite the $\boldsymbol{I}_{\mathrm{eE}}, \boldsymbol{I}_{\mathrm{mB}}, \boldsymbol{I}_{\mathrm{eB}}$, and $\boldsymbol{I}_{\mathrm{mE}}$ terms, i.e., (2.125)–(2.128), as

$$\begin{aligned} \boldsymbol{I}_{\mathrm{eE}} = \frac{-i}{\omega V} \sum_{\nu} \big[& \bar{g}_{\nu} (\bar{\boldsymbol{J}}_{0\nu} - i\boldsymbol{k} \cdot \bar{\mathbf{Q}}^{(\mathrm{e2})}_{0\nu})(\bar{\boldsymbol{J}}_{\nu 0} + i\boldsymbol{k} \cdot \bar{\mathbf{Q}}^{(\mathrm{e2})}_{\nu 0}) \\ & + \bar{h}_{\nu} (\bar{\boldsymbol{J}}_{\nu 0} - i\boldsymbol{k} \cdot \bar{\mathbf{Q}}^{(\mathrm{e2})}_{\nu 0})(\bar{\boldsymbol{J}}_{0\nu} + i\boldsymbol{k} \cdot \bar{\mathbf{Q}}^{(\mathrm{e2})}_{0\nu})\big] \cdot \tilde{\boldsymbol{E}}^{(\mathrm{T})}(\boldsymbol{k}) \ , \end{aligned} \tag{3.2}$$

$$\begin{aligned} \boldsymbol{I}_{\mathrm{mB}} &= \frac{ic}{V} \boldsymbol{k} \times \sum_{\nu} \big[\bar{g}_{\nu} \bar{\boldsymbol{M}}_{0\nu} \bar{\boldsymbol{M}}_{\nu 0} + \bar{h}_{\nu} \bar{\boldsymbol{M}}_{\nu 0} \bar{\boldsymbol{M}}_{0\nu}\big] \cdot \tilde{\boldsymbol{B}}(\boldsymbol{k}) \ , \qquad (3.3) \\ &= \big[\text{same expression without } c\big]_{\mathrm{SI}} \ , \end{aligned}$$

$$\begin{aligned} \boldsymbol{I}_{\mathrm{eB}} = \frac{1}{V} \sum_{\nu} \big[& \bar{g}_{\nu} (\bar{\boldsymbol{J}}_{0\nu} - i\boldsymbol{k} \cdot \bar{\mathbf{Q}}^{(\mathrm{e2})}_{0\nu}) \bar{\boldsymbol{M}}_{\nu 0} \\ & + \bar{h}_{\nu} (\bar{\boldsymbol{J}}_{\nu 0} - i\boldsymbol{k} \cdot \bar{\mathbf{Q}}^{(\mathrm{e2})}_{\nu 0}) \bar{\boldsymbol{M}}_{0\nu}\big] \cdot \tilde{\boldsymbol{B}}(\boldsymbol{k}) \ , \end{aligned} \tag{3.4}$$

$$\begin{aligned} \boldsymbol{I}_{\mathrm{mE}} = \frac{c}{\omega V} \boldsymbol{k} \times \sum_{\nu} \big[& \bar{g}_{\nu} \bar{\boldsymbol{M}}_{0\nu} (\bar{\boldsymbol{J}}_{\nu 0} + i\boldsymbol{k} \cdot \bar{\mathbf{Q}}^{(\mathrm{e2})}_{\nu 0}) \\ & + \bar{h}_{\nu} \bar{\boldsymbol{M}}_{\nu 0} (\bar{\boldsymbol{J}}_{0\nu} + i\boldsymbol{k} \cdot \bar{\mathbf{Q}}^{(\mathrm{e2})}_{0\nu})\big] \cdot \tilde{\boldsymbol{E}}^{(\mathrm{T})}(\boldsymbol{k}) \qquad (3.5) \\ = \big[\text{same expression without } c\big]_{\mathrm{SI}} \ . \end{aligned}$$

If we put

$$\begin{aligned} \boldsymbol{I}_{\mathrm{mB}} = ic\boldsymbol{k} \times \boldsymbol{M}_{\mathrm{B}} \ , \quad \boldsymbol{I}_{\mathrm{mE}} = ic\boldsymbol{k} \times \boldsymbol{M}_{\mathrm{E}} \ , \\ \big[\boldsymbol{I}_{\mathrm{mB}} = i\boldsymbol{k} \times \boldsymbol{M}_{\mathrm{B}} \ , \quad \boldsymbol{I}_{\mathrm{mE}} = i\boldsymbol{k} \times \boldsymbol{M}_{\mathrm{E}}\big]_{\mathrm{SI}} \ , \end{aligned} \tag{3.6}$$

the $\boldsymbol{B}$-induced and $\boldsymbol{E}$-induced magnetizations, $\boldsymbol{M}_{\mathrm{B}}$ and $\boldsymbol{M}_{\mathrm{E}}$, are given as

$$\boldsymbol{M}_{\mathrm{B}}(\boldsymbol{k}, \omega) = \frac{1}{V} \sum_{\nu} \big[\bar{g}_{\nu} \bar{\boldsymbol{M}}_{0\nu} \bar{\boldsymbol{M}}_{\nu 0} + \bar{h}_{\nu} \bar{\boldsymbol{M}}_{\nu 0} \bar{\boldsymbol{M}}_{0\nu}\big] \cdot \tilde{\boldsymbol{B}}(\boldsymbol{k}) \ , \tag{3.7}$$

$$\begin{aligned} \boldsymbol{M}_{\mathrm{E}}(\boldsymbol{k}, \omega) = \frac{-i}{\omega V} \sum_{\nu} \big[& \bar{g}_{\nu} \bar{\boldsymbol{M}}_{0\nu} (\bar{\boldsymbol{J}}_{\nu 0} + i\boldsymbol{k} \cdot \bar{\mathbf{Q}}^{(\mathrm{e2})}_{\nu 0}) \\ & + \bar{h}_{\nu} \bar{\boldsymbol{M}}_{\nu 0} (\bar{\boldsymbol{J}}_{0\nu} + i\boldsymbol{k} \cdot \bar{\mathbf{Q}}^{(\mathrm{e2})}_{0\nu})\big] \cdot \tilde{\boldsymbol{E}}^{(\mathrm{T})}(\boldsymbol{k}) \ . \end{aligned} \tag{3.8}$$

Similarly, by using the definition of $\boldsymbol{E}$- and $\boldsymbol{B}$-induced electric polarizations

$$\boldsymbol{I}_{\mathrm{eE}} = -i\omega \boldsymbol{P}_{\mathrm{E}} \ , \quad \boldsymbol{I}_{\mathrm{eB}} = -i\omega \boldsymbol{P}_{\mathrm{B}} \ , \tag{3.9}$$

we have

$$\boldsymbol{P}_{\mathrm{E}} = \frac{1}{\omega^2 V} \sum_{\nu} \big[\bar{g}_{\nu} (\bar{\boldsymbol{J}}_{0\nu} - i\boldsymbol{k} \cdot \bar{\mathbf{Q}}^{(\mathrm{e2})}_{0\nu})(\bar{\boldsymbol{J}}_{\nu 0} + i\boldsymbol{k} \cdot \bar{\mathbf{Q}}^{(\mathrm{e2})}_{\nu 0})$$

$$+\bar{h}_\nu(\bar{\boldsymbol{J}}_{\nu 0} - i\boldsymbol{k}\cdot\bar{\mathbf{Q}}^{(\mathrm{e2})}_{\nu 0})(\bar{\boldsymbol{J}}_{0\nu} + i\boldsymbol{k}\cdot\bar{\mathbf{Q}}^{(\mathrm{e2})}_{0\nu})\big]\cdot\tilde{\boldsymbol{E}}^{(\mathrm{T})}(\boldsymbol{k})\,, \qquad (3.10)$$

$$\boldsymbol{P}_{\mathrm{B}} = \frac{i}{\omega V}\sum_\nu \big[\bar{g}_\nu(\bar{\boldsymbol{J}}_{0\nu} - i\boldsymbol{k}\cdot\bar{\mathbf{Q}}^{(\mathrm{e2})}_{0\nu})\bar{\boldsymbol{M}}_{\nu 0} + \bar{h}_\nu(\bar{\boldsymbol{J}}_{\nu 0} - i\boldsymbol{k}\cdot\bar{\mathbf{Q}}^{(\mathrm{e2})}_{\nu 0})\bar{\boldsymbol{M}}_{0\nu}\big]\cdot\tilde{\boldsymbol{B}}(\boldsymbol{k})\,. \qquad (3.11)$$

In this way, we can redefine the induced magnetizations $\boldsymbol{M}_{\mathrm{B}}$ and $\boldsymbol{M}_{\mathrm{E}}$, and the induced electric polarizations $\boldsymbol{P}_{\mathrm{E}}$ and $\boldsymbol{P}_{\mathrm{B}}$ for general symmetry conditions. This allows the new definitions of the "electric, magnetic, and chiral" susceptibilities as

$$\boldsymbol{P}_{\mathrm{E}} = \chi_{\mathrm{eE}}\boldsymbol{E},\quad \boldsymbol{P}_{\mathrm{B}} = \chi_{\mathrm{eB}}\boldsymbol{B},\quad \boldsymbol{M}_{\mathrm{E}} = \chi_{\mathrm{mE}}\boldsymbol{E},\quad \boldsymbol{M}_{\mathrm{B}} = \chi_{\mathrm{mB}}\boldsymbol{B}. \qquad (3.12)$$

(Though we should write $\boldsymbol{E}$ as $\boldsymbol{E}^{(\mathrm{T})}$ more exactly, we may leave it as it is, because the argument in the next section allows the same form of χ_{em} as the susceptibility to relate induced current density and L source field $\boldsymbol{E}_{\mathrm{extL}}$.) The precise expressions of these susceptibilities are

$$\chi_{\mathrm{eE}} = \frac{1}{\omega^2 V}\sum_\nu \big[\bar{g}_\nu(\bar{\boldsymbol{J}}_{0\nu} - i\boldsymbol{k}\cdot\bar{\mathbf{Q}}^{(\mathrm{e2})}_{0\nu})(\bar{\boldsymbol{J}}_{\nu 0} + i\boldsymbol{k}\cdot\bar{\mathbf{Q}}^{(\mathrm{e2})}_{\nu 0}) + \bar{h}_\nu(\bar{\boldsymbol{J}}_{\nu 0} - i\boldsymbol{k}\cdot\bar{\mathbf{Q}}^{(\mathrm{e2})}_{\nu 0})(\bar{\boldsymbol{J}}_{0\nu} + i\boldsymbol{k}\cdot\bar{\mathbf{Q}}^{(\mathrm{e2})}_{0\nu})\big]\,, \qquad (3.13)$$

$$\chi_{\mathrm{eB}} = \frac{i}{\omega V}\sum_\nu \big[\bar{g}_\nu(\bar{\boldsymbol{J}}_{0\nu} - i\boldsymbol{k}\cdot\bar{\mathbf{Q}}^{(\mathrm{e2})}_{0\nu})\bar{\boldsymbol{M}}_{\nu 0} + \bar{h}_\nu(\bar{\boldsymbol{J}}_{\nu 0} - i\boldsymbol{k}\cdot\bar{\mathbf{Q}}^{(\mathrm{e2})}_{\nu 0})\bar{\boldsymbol{M}}_{0\nu}\big]\,, \qquad (3.14)$$

$$\chi_{\mathrm{mB}} = \frac{1}{V}\sum_\nu \big[\bar{g}_\nu\bar{\boldsymbol{M}}_{0\nu}\bar{\boldsymbol{M}}_{\nu 0} + \bar{h}_\nu\bar{\boldsymbol{M}}_{\nu 0}\bar{\boldsymbol{M}}_{0\nu}\big]\,, \qquad (3.15)$$

$$\chi_{\mathrm{mE}} = \frac{-i}{\omega V}\sum_\nu \big[\bar{g}_\nu\bar{\boldsymbol{M}}_{0\nu}(\bar{\boldsymbol{J}}_{\nu 0} + i\boldsymbol{k}\cdot\bar{\mathbf{Q}}^{(\mathrm{e2})}_{\nu 0}) + \bar{h}_\nu\bar{\boldsymbol{M}}_{\nu 0}(\bar{\boldsymbol{J}}_{0\nu} + i\boldsymbol{k}\cdot\bar{\mathbf{Q}}^{(\mathrm{e2})}_{0\nu})\big]\,. \qquad (3.16)$$

In terms of these new, quantum mechanical definitions of $\boldsymbol{P}$ and $\boldsymbol{M}$, the microscopic Ampère law, (1.10) can be rewritten as

$$i\boldsymbol{k}\times(\boldsymbol{B} - 4\pi\boldsymbol{M}) = -i\frac{\omega}{c}(\boldsymbol{E} + 4\pi\boldsymbol{P})\,, \qquad (3.17)$$

$$\left[i\boldsymbol{k}\times\left(\frac{1}{\mu_0}\boldsymbol{B} - \boldsymbol{M}\right) = -i\omega(\varepsilon_0\boldsymbol{E} + \boldsymbol{P})\right]_{\mathrm{SI}}\,, \qquad (3.18)$$

i.e., the same form as the conventional macroscopic one by writing $\boldsymbol{B} - 4\pi\boldsymbol{M} = \boldsymbol{H}$ and $\boldsymbol{E} + 4\pi\boldsymbol{P} = \boldsymbol{D}$. Combining this result with the usual definition $\boldsymbol{D} = \varepsilon\boldsymbol{E}$ and $\boldsymbol{B} = \mu\boldsymbol{H}$, we have

$$\varepsilon \boldsymbol{E} = (1 + 4\pi \chi_{\mathrm{eE}})\boldsymbol{E} + 4\pi \chi_{\mathrm{eB}} \boldsymbol{B} \ \left[= (\varepsilon_0 + \chi_{\mathrm{eE}})\boldsymbol{E} + \chi_{\mathrm{eB}} \boldsymbol{B} \right]_{\mathrm{SI}} , \tag{3.19}$$

$$\frac{1}{\mu}\boldsymbol{B} = (1 - 4\pi \chi_{\mathrm{mB}})\boldsymbol{B} - 4\pi \chi_{\mathrm{mE}} \boldsymbol{E} \ \left[= (\frac{1}{\mu_0} - \chi_{\mathrm{mB}})\boldsymbol{B} - \chi_{\mathrm{mE}} \boldsymbol{E} \right]_{\mathrm{SI}} . \tag{3.20}$$

This is the constitutive equations in terms of $\boldsymbol{E}$ and $\boldsymbol{B}$ in the general case including chiral symmetry. Though this looks like two vector equations, it is actually one, because it was derived from the single one $\boldsymbol{I} = \chi_{\mathrm{em}}\boldsymbol{A}$, or $\boldsymbol{I} = \chi_{\mathrm{em}}[\boldsymbol{A} + (c/i\omega)\boldsymbol{E}_{\mathrm{extL}}]$, including the content of the next section. Equations (3.19) and (3.20) are those to be compared with DBF constitutive equations, which is done in Sect. 3.4.

In the case of non-chiral symmetry, these equations reduce to

$$\varepsilon \boldsymbol{E} = (1 + 4\pi \chi_{\mathrm{eE}})\boldsymbol{E} \ \left[= (\varepsilon_0 + \chi_{\mathrm{eE}})\boldsymbol{E} \right]_{\mathrm{SI}} , \tag{3.21}$$

$$\boldsymbol{B} = \mu(1 - 4\pi \chi_{\mathrm{mB}})\boldsymbol{B} \ \left[= \frac{\mu}{\mu_0}(1 - \mu_0 \chi_{\mathrm{mB}})\boldsymbol{B} \right]_{\mathrm{SI}} , \tag{3.22}$$

i.e., $\varepsilon = 1 + 4\pi \chi_{\mathrm{E}}$ and $\mu = 1/(1 - 4\pi \chi_{\mathrm{mB}})$, or $\chi_{\mathrm{eE}} = \chi_{\mathrm{e}} = (c/\omega^2)\chi_{\mathrm{em0}}$, $\mu = 1 + 4\pi \chi_{\mathrm{m}} = 1/(1 - 4\pi \chi_{\mathrm{mB}})$, i.e., $\chi_{\mathrm{m}} = \chi_{\mathrm{mB}}/(1 - 4\pi \chi_{\mathrm{mB}})$.

The new scheme in terms of the single susceptibility χ_{em} can deal with the general situation including the chiral symmetry. But the popular trend in the study of metamaterials, near-field optics, photonic crystals etc. is to use ε and μ as independent free parameters. The criterion to allow this is the non-chiral symmetry of the system in consideration, as discussed above. This is particularly important in the resonant region of ε and μ. If this condition is not fulfilled, the description via "ε and μ" has no justification. Even when the use of independent ε and μ is allowed in the non-chiral case, one should use, not χ_{m}, but χ_{mB}, since the matter excitation energies to describe the resonance are correctly included as the poles of, not the former, but the latter.

The problem about the statement "the macroscopic average of microscopic magnetic field $\boldsymbol{h}$ is usually written as $\boldsymbol{B}$" described in Sect. 1.5, does not exist in the present scheme, because the magnetic field is always written as $\boldsymbol{B}$ both in the interaction Hamiltonian H_{int} and in the vacuum EM field Hamiltonian $H_{\mathrm{EM}}^{(\mathrm{T})}$ without any change before and after the application of LWA. The macroscopic average defined in this scheme does not logically allow such a change.

The expressions of the formulas in SI units system, especially in this section, will need a check from the dimensional point of view, which is done in Sect. 5.8.

3.2 Unified Susceptibility for T and L Source Fields

In the previous subsection, we have rewritten the new constitutive equation in a form similar to the conventional ones. However, this is limited to the response to the transverse field $\boldsymbol{A}$. The response to the longitudinal field is treated in Sect. 5.7, where the source field is an external L electric field $\boldsymbol{E}_{\text{extL}}$. In order to consider the general cases of EM response, we have only to combine these two formulations. In doing so, however, we find it awkward to have constitutive equations for T and L fields defined with respect to different kinds of field, $\boldsymbol{A}$ for T and $\boldsymbol{E}_{\text{extL}}$ for L field. In this subsection we will show how to unify them, i.e., how to rewrite the susceptibility χ_{JEL} in Sect. 5.7.2 in terms of χ_{em}. The result is quite simple, i.e., the whole macroscopic constitutive equation is given in the form

$$\begin{aligned} \boldsymbol{I}(\boldsymbol{k},\omega) &= \chi_{\text{em}}(\boldsymbol{k},\omega)\cdot[\boldsymbol{A}(\boldsymbol{k},\omega)+\frac{c}{i\omega}\boldsymbol{E}_{\text{extL}}(\boldsymbol{k},\omega)]\,, \\ &= \big[\text{same expression without } c\big]_{\text{SI}}\,. \end{aligned} \tag{3.23}$$

where the sum of $\boldsymbol{A}$ and $(c/i\omega)\boldsymbol{E}_{\text{extL}}$ represents the general form of source EM field with T- and L-components.

In order to rewrite the result of Sect. 5.7.2, let us take the (ξ,η,ζ) axes, as the Cartesian coordinate system to express L and T components, where ζ axis is parallel to $\boldsymbol{k}$. The induced current density by $\boldsymbol{E}_{\text{extL}}$ contains the matrix elements of the (L-component of) polarization operator $\boldsymbol{P}^{(\text{L})}$, as shown in Sect. 5.7.2. The operator equation $\boldsymbol{J}=\partial\boldsymbol{P}/\partial t+c\nabla\times\boldsymbol{M}$ discussed in Sect. 5.1 leads us to $\boldsymbol{J}^{(\text{L})}=\partial\boldsymbol{P}^{(\text{L})}/\partial t=(i/\hbar)(H\boldsymbol{P}^{(\text{L})}-\boldsymbol{P}^{(\text{L})}H)$, where the last equality is the Heisenberg equation of motion. Since we need the matrix elements of $\boldsymbol{P}^{(\text{L})}$ and $\boldsymbol{J}^{(\text{L})}$ with respect to the eigenstates of matter Hamiltonian $H^{(0)}$, we use $\boldsymbol{J}^{(\text{L})}=(i/\hbar)(H^{(0)}\boldsymbol{P}^{(\text{L})}-\boldsymbol{P}^{(\text{L})}H^{(0)})$. Then, we have

$$\langle\mu|\boldsymbol{J}(\boldsymbol{r})^{(\text{L})}|\nu\rangle=\frac{i}{\hbar}(E_\mu-E_\nu)\langle\mu|\boldsymbol{P}(\boldsymbol{r})^{(\text{L})}|\nu\rangle\,, \tag{3.24}$$

which allows us to rewrite the (τ,ζ) components $(\tau=\xi,\eta,\zeta)$ of χ_{JEL} as

$$\chi_{\text{JEL}}^{(\tau,\zeta)}=-i\hbar\sum_\nu\left[g_\nu(\omega)\boldsymbol{J}_{0\nu}^{(\tau)}(\boldsymbol{r})\boldsymbol{J}_{\nu0}^{(\zeta)}(\boldsymbol{r}')\frac{1}{E_{\nu0}}+h_\nu(\omega)\boldsymbol{J}_{\nu0}^{(\tau)}(\boldsymbol{r})\boldsymbol{J}_{0\nu}^{(\zeta)}(\boldsymbol{r}')\frac{-1}{E_{\nu0}}\right], \tag{3.25}$$

where $E_{\nu0}=E_\nu-E_0$. In this way we can rewrite the (τ,ζ) components of χ_{JEL} in terms of the matrix elements of $\boldsymbol{J}$ alone.

The manipulation

$$\left[\frac{1}{E_{\nu0}\mp z}\right]\frac{1}{E_{\nu0}}=\frac{\pm1}{z}\left\{\frac{1}{E_{\nu0}\mp z}-\frac{1}{E_{\nu0}}\right\},\quad(z=\hbar\omega+i0^+) \tag{3.26}$$

allows us to rewrite χ_{JEL} as

$$\begin{aligned} \chi_{\text{JEL}}^{(\tau,\zeta)}=-\frac{i}{\omega}\sum_\nu\Bigg[&\left\{g_\nu(\omega)-\frac{1}{E_{\nu0}}\right\}\boldsymbol{J}_{0\nu}^{(\tau)}(\boldsymbol{r})\boldsymbol{J}_{\nu0}^{(\zeta)}(\boldsymbol{r}') \\ &+\left\{h_\nu(\omega)-\frac{1}{E_{\nu0}}\right\}\boldsymbol{J}_{\nu0}^{(\tau)}(\boldsymbol{r})\boldsymbol{J}_{0\nu}^{(\zeta)}(\boldsymbol{r}')\Bigg] \end{aligned} \tag{3.27}$$

The r.h.s. of this expression is exactly same as the susceptibility $\chi_{\rm cd}$ (times $c/i\omega$) derived in Sect. 2.2.2, except for the difference in the assignment of tensor components. Thus, we may write

$$\chi_{\rm JEL}^{(\tau,\zeta)} = \frac{c}{i\omega}\chi_{\rm cd}^{(\tau,\zeta)} \quad \left[= \frac{1}{i\omega}\,\chi_{\rm cd}^{(\tau,\zeta)} \right]_{\rm SI} . \tag{3.28}$$

(The (ζ, ξ) and (ζ, η) components, describing the L current density induced by the T field, are already included in the susceptibility $\chi_{\rm cd}$ of Sect. 2.2.2.) Altogether, we have shown that the 3×3 matrix $\chi_{\rm cd}$ describes the linear response of matter generally for both the T field $\boldsymbol{A}$ and the L field $(c/i\omega)\boldsymbol{E}_{\rm extL}$.

To derive the macroscopic susceptibility for the $\boldsymbol{E}_{\rm extL}$-induced components, we can repeat the same calculation as that for the $\boldsymbol{A}$-induced components, which leads to the same form of $\chi_{\rm em}$, except for the assignment of the tensor components (τ, ζ). Hence, we obtain (3.23).

In Sect. 2.4, we decomposed the induced current density into a sum of the current densities $\boldsymbol{I}_{\rm eE}, \boldsymbol{I}_{\rm eB}, \boldsymbol{I}_{\rm mB}, \boldsymbol{I}_{\rm mE}$, and they are rewritten as the sum of $-i\omega(\boldsymbol{P}_{\rm E} + \boldsymbol{P}_{\rm B})$ and $ic\boldsymbol{k} \times (\boldsymbol{M}_{\rm B} + \boldsymbol{M}_{\rm E})$ in Sect. 3.1 Since this part of the current densities is caused by the transverse field $\boldsymbol{A}$, we rewrite them as $-i\omega(\boldsymbol{P}_{\rm ET} + \boldsymbol{P}_{\rm B})$ and $ic\boldsymbol{k} \times (\boldsymbol{M}_{\rm B} + \boldsymbol{M}_{\rm ET})$ to distinguish T and L electric fields. In the presence of external L field, we add the induced current densities, $\boldsymbol{I}_{\rm eEL}$ and $\boldsymbol{I}_{\rm mEL}$, due to the L electric field. These terms are defined by the same expression as (3.2) and (3.5) simply by replacing $\tilde{\boldsymbol{E}}(\boldsymbol{k})$ with $\tilde{\boldsymbol{E}}_{\rm extL}(\boldsymbol{k})$, and can be rewritten as $(-i\omega\boldsymbol{P}_{\rm EL} + ic\boldsymbol{k} \times \boldsymbol{M}_{\rm EL})$, where $\boldsymbol{P}_{\rm EL}$ and $\boldsymbol{M}_{\rm EL}$ are defined by (3.10) and (3.8), respectively, by replacing $\tilde{\boldsymbol{E}}(\boldsymbol{k})$ with $\tilde{\boldsymbol{E}}_{\rm extL}(\boldsymbol{k})$.

Thus, the general LWA form of the induced current density in the presence of $\boldsymbol{E}_{\rm extL}$ can be written as

$$\begin{aligned} \tilde{\boldsymbol{I}}(\boldsymbol{k}, \omega) &= -i\omega\tilde{\boldsymbol{P}}(\boldsymbol{k}, \omega) + ic\boldsymbol{k} \times \tilde{\boldsymbol{M}}(\boldsymbol{k}, \omega) , \\ &= \left[\text{same expression without } c\right]_{\rm SI} . \end{aligned} \tag{3.29}$$

where

$$\tilde{\boldsymbol{P}}(\boldsymbol{k}, \omega) = \tilde{\boldsymbol{P}}_{\rm ET}(\boldsymbol{k}, \omega) + \tilde{\boldsymbol{P}}_{\rm EL}(\boldsymbol{k}, \omega) + \tilde{\boldsymbol{P}}_{\rm B}(\boldsymbol{k}, \omega) , \tag{3.30}$$

$$\tilde{\boldsymbol{M}}(\boldsymbol{k}, \omega) = \tilde{\boldsymbol{M}}_{\rm ET}(\boldsymbol{k}, \omega) + \tilde{\boldsymbol{M}}_{\rm EL}(\boldsymbol{k}, \omega) + \tilde{\boldsymbol{M}}_{\rm B}(\boldsymbol{k}, \omega) . \tag{3.31}$$

Using this extended definition of $\tilde{\boldsymbol{P}}$ and $\tilde{\boldsymbol{M}}$ to define

$$\boldsymbol{D} = \boldsymbol{E} + 4\pi\tilde{\boldsymbol{P}} \quad \left[= \varepsilon_0\boldsymbol{E} + \boldsymbol{P} \right]_{\rm SI} , \tag{3.32}$$

$$\boldsymbol{H} = \boldsymbol{B} - 4\pi\tilde{\boldsymbol{M}} \quad \left[= \frac{1}{\mu_0}\boldsymbol{B} - \boldsymbol{M} \right]_{\rm SI} , \tag{3.33}$$

we can write the macroscopic Gauss law for electricity and Ampère law in the conventional form. It should be stressed that all this reformulation arises from the single susceptibility $\chi_{\rm em}$ as a full 3×3 matrix for the constitutive equation relating $\boldsymbol{I}$ and $[\boldsymbol{A} + (c/i\omega)\boldsymbol{E}_{\rm extL}]$. Since the T part of the source field $\boldsymbol{A}$ contains both electric and magnetic fields, this susceptibility describes both electric and magnetic responses. It should be stressed that this rewriting does not affect the polariton dispersion equation, (2.131).

3.3 New and Conventional Dispersion Equations

The new dispersion equation is $(ck/\omega)^2 = 1 + (4\pi c/\omega^2)\chi_{\rm em}(\boldsymbol{k}, \omega)$, while the conventional one is $(ck/\omega)^2 = \varepsilon\mu$ (more rigorously, (2.131) and (2.132)). Though it is usually not explicitly mentioned, the conventional form applies only to the case of non-chiral symmetry. In the case of chiral symmetry, where one cannot distinguish polar and axial vectors by their transformation properties with respect to mirror reflection and/or space inversion, electric field induces $\boldsymbol{M}$, as well as $\boldsymbol{P}$, and magnetic field induces $\boldsymbol{P}$, as well as $\boldsymbol{M}$. In order to describe such an extended situation, a phenomenology called Drude-Born-Fedorov (DBF) constitutive equations [1] has been used. As will be shown in Sect. 3.4, DBF eqs lead to a dispersion equation different from the new one. Thus, the new dispersion equation is different from the conventional one in the both cases of chiral and non-chiral symmetries.

The apparent difference between the two dispersion equations in non-chiral symmetry is in the pole structure of $\chi_{\rm em}$ and $\varepsilon\mu$ on the r.h.s. of the equations. In $\chi_{\rm em}$, the contribution of all the quantum mechanical transitions appears as a sum of single poles, which is a general result of the perturbation calculation given in Sect. 2.2.2. On the other hand, the poles of the product $\varepsilon\mu$ appear differently. Since ε is a sum of single poles of E1 (+ E2) character and μ that of M1 character, the contributions of E1 (+ E2) and M1 transitions appear as a product in $\varepsilon\mu$. This apparent controversy can be solved by using the magnetic susceptibility $\chi_{\rm B}(= \chi_{\rm mB}$ for cgs Gauss units, $= \mu_0\chi_{\rm mB}$ for SI units) rather than $\chi_{\rm m}$, where $\boldsymbol{M} = \chi_{\rm B}\boldsymbol{B} = \chi_{\rm m}\boldsymbol{H}$. This leads, together with $\boldsymbol{B} = \mu\boldsymbol{H}$, to $\mu = 1 + 4\pi\chi_{\rm m} = 1/(1 - 4\pi\chi_{\rm B})$, which allows us to rewrite the conventional equation $(ck/\omega)^2 = \varepsilon\mu$ into

$$\begin{aligned}\left(\frac{ck}{\omega}\right)^2 &= \frac{\varepsilon}{1 - 4\pi\chi_{\rm B}} \left[= \frac{1 + \chi_{\rm e}}{1 - \chi_{\rm B}}\right]_{\rm SI}, \\ &= \varepsilon + 4\pi\left(\frac{ck}{\omega}\right)^2\chi_{\rm B}\left[= 1 + \chi_{\rm e} + \left(\frac{ck}{\omega}\right)^2\chi_{\rm B}\right]_{\rm SI}. \qquad (3.34)\end{aligned}$$

In this form, the r.h.s. is the sum of the single poles due to E1 (+ E2) and M1 transitions, and the magnetic contribution appears with a multiplication factor of $O(k^2)$. This fact coincides with the derivation, in Sect. 2.4, of induced magnetization from the $O(k^2)$ term of the induced current density.

This solves also one of the problems of Sect. 1.5, i.e., the k-dependence of μ. The apparent difference in the k-dependence of μ between the two typical cases of M1 transition, spin resonance and optical (orbital) M1 transitions, is due to the different stages of theoretical description. In both types of experiment, a proper analysis would require the comparison of spectral peak position and intensity with those of theoretical prediction. For that purpose, we need to calculate the EM response of the medium based on the dispersion relation. The frequently used expression $\mu = 1 + 4\pi\chi_{\mathrm{m}}$ for spin resonance should be rewritten as $\mu = 1/(1 - 4\pi\chi_{\mathrm{B}})$, and the dispersion equation takes the form of (3.34). The argument for the intensity of orbital M1 transition given in Sect. 1.5 is made in accordance with (3.34). Thus, the apparent difference in the k-dependence of μ is actually the problem of correct definition of magnetic susceptibility.

It should be stressed that the discussions given above are meaningful only in non-chiral symmetry. The argument in the case of chiral symmetry will be given in Sect. 3.4.

3.4 Case of Chiral Symmetry: Comparison with DBF-Equations

For materials with chiral symmetry, where polar and axial vectors are indistinguishable, the conventional scheme of macroscopic M-eqs with ε and μ is not sufficient. As a symmetry argument within macroscopic regime, it was thought appropriate to add to the constitutive equations those terms which allow the electric field induced magnetization and magnetic field induced electric polarization. The generalized constitutive equations are called Drude-Born-Fedorov equations (DBF-eqs) [1]. In a homogeneous isotropic case, they are written in the form [2]

$$\boldsymbol{D} = \varepsilon(\boldsymbol{E} + \beta\nabla \times \boldsymbol{E}) , \tag{3.35}$$

$$\boldsymbol{B} = \mu(\boldsymbol{H} + \beta\nabla \times \boldsymbol{H}) . \tag{3.36}$$

The parameter β is called chiral admittance, which leads to the different phase velocities for left and right circularly polarized light in this medium, as shown below.

In Sect. 3.1.3 we have rewritten the new constitutive equation $\tilde{\boldsymbol{I}} = \chi_{\mathrm{em}}\tilde{\boldsymbol{A}}$ into $\tilde{\boldsymbol{I}} = -i\omega\tilde{\boldsymbol{P}} + i\boldsymbol{k} \times \tilde{\boldsymbol{M}}$, from which we “defined” electric polarization $\tilde{\boldsymbol{P}}$ and magnetization $\tilde{\boldsymbol{M}}$. The source fields of these induced polarizations are T electric field $(i\omega/c)\tilde{\boldsymbol{A}}$ and magnetic field $i\boldsymbol{k} \times \tilde{\boldsymbol{A}}$. In the presence of an external L electric field, we add the contributions of $\tilde{\boldsymbol{E}}^{(\mathrm{L})}$ induced terms, which leads to

$$\tilde{\boldsymbol{P}} = \chi_{\mathrm{eET}}\tilde{\boldsymbol{E}}^{(\mathrm{T})} + \chi_{\mathrm{eEL}}\tilde{\boldsymbol{E}}^{(\mathrm{L})} + \chi_{\mathrm{eB}}\tilde{\boldsymbol{B}} , \tag{3.37}$$

$$\tilde{\boldsymbol{M}} = \chi_{\mathrm{mET}}\tilde{\boldsymbol{E}}^{(\mathrm{T})} + \chi_{\mathrm{mEL}}\tilde{\boldsymbol{E}}^{(\mathrm{L})} + \chi_{\mathrm{mB}}\tilde{\boldsymbol{B}} . \tag{3.38}$$

In terms of these $\tilde{\boldsymbol{P}}$ and $\tilde{\boldsymbol{M}}$, we obtain the Ampère law in the conventional form, $\nabla \times \boldsymbol{H} = (1/c)\partial\boldsymbol{D}/\partial t$, whereby we use

$$\boldsymbol{D} = (1 + 4\pi\chi_{\mathrm{eE}})\boldsymbol{E} + 4\pi\chi_{\mathrm{eB}}\boldsymbol{B} \ \left[= (\varepsilon_0 + \chi_{\mathrm{eE}})\boldsymbol{E} + \chi_{\mathrm{eB}}\boldsymbol{B}\right]_{\mathrm{SI}} , \tag{3.39}$$

$$\boldsymbol{H} = (1 - 4\pi\chi_{\mathrm{mB}})\boldsymbol{B} - 4\pi\chi_{\mathrm{mE}}\boldsymbol{E} \left[= \left(\frac{1}{\mu_0} - \chi_{\mathrm{mB}}\right)\boldsymbol{B} - \chi_{\mathrm{mE}}\boldsymbol{E}\right]_{\mathrm{SI}} . \tag{3.40}$$

Here, we have used a short-hand notation

$$\chi_{\mathrm{eE}}\boldsymbol{E} = \chi_{\mathrm{eET}}\boldsymbol{E}^{(\mathrm{T})} + \chi_{\mathrm{eEL}}\boldsymbol{E}^{(\mathrm{L})} , \tag{3.41}$$

$$\chi_{\mathrm{mE}}\boldsymbol{E} = \chi_{\mathrm{mET}}\boldsymbol{E}^{(\mathrm{T})} + \chi_{\mathrm{mEL}}\boldsymbol{E}^{(\mathrm{L})} . \tag{3.42}$$

Equations (3.35) and (3.39) are equivalent, if we note $\tilde{\boldsymbol{B}} = -(ic/\omega)\boldsymbol{k} \times \boldsymbol{E}$. However, (3.36) and (3.40) cannot be equivalent. (If $\boldsymbol{E}$ on the r.h.s. of (3.40) were $\boldsymbol{D}$, they would be equivalent.) Therefore, the conventional DBF eqs are different from the similar expressions (3.39) and (3.40). The essential difference is that the apparently two vector equations (3.39) and (3.40) are originally a single vector equation, $\tilde{\boldsymbol{I}} = \chi_{\mathrm{em}}\tilde{\boldsymbol{A}}$, while (3.35) and (3.36) are not. This difference manifests itself in that of the dispersion equation as shown below.

The dispersion equation obtained from $\tilde{\boldsymbol{I}} = \chi_{\mathrm{em}}\tilde{\boldsymbol{A}}$ is given as (2.131), where $\chi_{\mathrm{em}}^{(\mathrm{T})}(\boldsymbol{k}, \omega)$ consists of a superposition of single poles corresponding to matter excitation energies. However, the dispersion equation derived from DBF eqs has a different behavior, as shown below. Substituting M-eqs, $\nabla \times \boldsymbol{H} = -(i\omega/c)\boldsymbol{D}$ and $\nabla \times \boldsymbol{E} = (i\omega/c)\boldsymbol{B}$ into the DBF eqs, we have

$$\varepsilon\boldsymbol{E} + \varepsilon\beta\nabla \times \boldsymbol{E} = (ic/\omega)\nabla \times \boldsymbol{H} \quad = \left[(i/\omega)\nabla \times \boldsymbol{H}\right]_{\mathrm{SI}} , \tag{3.43}$$

$$\mu\boldsymbol{H} + \mu\beta\nabla \times \boldsymbol{H} = -(ic/\omega)\nabla \times \boldsymbol{E} \quad = \left[(-i/\omega)\nabla \times \boldsymbol{E}\right]_{\mathrm{SI}} . \tag{3.44}$$

These equations can be solved for $\boldsymbol{X} = \nabla \times \boldsymbol{E}$ and $\boldsymbol{Y} = \nabla \times \boldsymbol{H}$ as

$$A_0\boldsymbol{X} = \varepsilon\mu\beta\boldsymbol{E} + i\mu(c/\omega)\boldsymbol{H} , \tag{3.45}$$

$$A_0\boldsymbol{Y} = -i\varepsilon(c/\omega)\boldsymbol{E} + \varepsilon\mu\beta\boldsymbol{H} , \tag{3.46}$$

$$= \left[\text{same expression without } c\right]_{\mathrm{SI}} ,$$

where $A_0 = (c/\omega)^2 - \varepsilon\mu\beta^2$. From (3.43), we have $\varepsilon\nabla \cdot \boldsymbol{E} = 0$, i.e., $\boldsymbol{E}$ is transverse. Taking the curl of equations (3.43) and (3.44), we have

$$\begin{aligned} \varepsilon\boldsymbol{X} + \varepsilon\beta(\nabla \times \nabla \times \boldsymbol{E}) &= i(c/\omega)\nabla \times \nabla \times \boldsymbol{H} , &(3.47)\\ &= \left[(i/\omega)\nabla \times \nabla \times \boldsymbol{H}\right]_{\mathrm{SI}} , \\ \mu\boldsymbol{Y} + \mu\beta(\nabla \times \nabla \times \boldsymbol{H}) &= -i(c/\omega)\nabla \times \nabla \times \boldsymbol{E} , &(3.48)\\ &= \left[-(i/\omega)\nabla \times \nabla \times \boldsymbol{E}\right]_{\mathrm{SI}} . \end{aligned}$$

Since both $\boldsymbol{E}$ and $\boldsymbol{H}$ are transverse, the $\nabla \times \nabla \times$ parts of these equations can be simplified as $\nabla \times \nabla \times \boldsymbol{E} = k^2\boldsymbol{E}$ and $\nabla \times \nabla \times \boldsymbol{H} = k^2\boldsymbol{H}$ (for plane waves). Substituting (3.45) and (3.46) into these equations, we get

$$\begin{aligned}\varepsilon\{\varepsilon\mu\beta\boldsymbol{E}+i\mu(c/\omega)\boldsymbol{H}\}+\varepsilon\beta k^2A_0\boldsymbol{E}&=i(c/\omega)k^2A_0\boldsymbol{H}\,, &(3.49)\\ &=\left[(i/\omega)k^2A_0\boldsymbol{H}\right]_{\mathrm{SI}}\,,\\ \mu\{\varepsilon\mu\beta\boldsymbol{H}-i\varepsilon(c/\omega)\boldsymbol{E}\}+\mu\beta k^2A_0\boldsymbol{H}&=-i(c/\omega)k^2A_0\boldsymbol{E}\,, &(3.50)\\ &=\left[(-i/\omega)k^2A_0\boldsymbol{E}\right]_{\mathrm{SI}}\,,\end{aligned}$$

which are the homogeneous linear equations for $\boldsymbol{E}$ and $\boldsymbol{H}$. The condition for the existence of nontrivial solution is the vanishing of the determinant of the coefficient 2×2 matrix, which gives

$$\left(\frac{ck}{\omega}\right)^2=\left(\frac{\sqrt{\varepsilon\mu}}{1\pm(\beta\omega/c)\sqrt{\varepsilon\mu}}\right)^2=\left[\left(\frac{c\sqrt{\varepsilon\mu}}{1\pm(\beta\omega)\sqrt{\varepsilon\mu}}\right)^2\right]_{\mathrm{SI}}\,. \tag{3.51}$$

For finite β, the refractive index ($= ck/\omega$) takes two different values, which correspond to different polarizations of the eigen modes in this medium, i.e., two different circular polarizations of this isotropic medium. In this sense, DBF eqs. give a qualitative description of an optical active medium. However, the dispersion equation is different from the first principles result. The r.h.s. of the dispersion equation given above has poles through the ω-dependence of ε and μ. From the form of the expression, all the poles of the r.h.s. are second order (or higher). This is in sharp contrast to the result (2.131) in Sect. 2.5, where the corresponding part of the equation consists of a superposition of single poles.

This result indicates that the difference between DBF eqs and the present microscopic susceptibility lies on a fundamental level, so that the former cannot be justified microscopically. In contrast, the present first-principles theory can provide a general expression of macroscopic susceptibility $\chi_{\mathrm{em}}(\boldsymbol{k},\omega)$, (2.119) in a quantum mechanical form, applicable to both chiral and non-chiral symmetry. Its $O(k^1)$ term, $k\chi_{\mathrm{em1}}$, vanishes in non-chiral symmetry, so that it plays a central role in chiral symmetry. The $O(k^0)$ and $O(k^2)$ terms are also affected by the chirality induced mixing of the eigenfunctions, but the effect is secondary. Since each element of the 2×2 matrix in the dispersion equation $\det|(c^2k^2/\omega^2)\mathbf{1}-\{\mathbf{1}+(4\pi c/\omega^2)\chi_{\mathrm{em}}^{(\mathrm{T})}(\boldsymbol{k},\omega)\}|=0$, (2.131) is at most second order in k, this dispersion equation leads to a quartic equation of k for a given ω. Further, this would become a quadratic equation of k^2 with solutions $k=\pm k_1(\omega), \pm k_2(\omega)$ in the absence of $k\chi_{\mathrm{em1}}$, i.e., in the case of non-chiral symmetry. The presence of the odd power terms of k in the quartic equation breaks this mirror symmetry (for $+k\leftrightarrow -k$). In Sect. 3.8.1, we show an example of this kind. In the neighborhood of $k=0$, dispersion curves show a k-linear behavior. The dispersion curves and the boundary conditions of EM field allow us to determine the response spectrum of the system. Thereby, the knowledge of the microscopic character of the resonance according to this scheme will be a good help for our physical interpretation of the result.

The rest of this subsection and Sect. 4.1.5 are based on the studies of chiral medium as a member of metamaterials research group, of which activity report is going to be published soon [3]. Because of its close relationship with the contents of this book, the new development is reproduced in several parts.

If one prefers to use a simpler parameterized form of constitutive equations, rather than the complicated first-principles expression, (2.125)–(2.128), we can propose an alternative of DBF eqs. Since the fundamental variables of EM field are $\boldsymbol{E}$ and $\boldsymbol{B}$, rather than $\boldsymbol{E}$ and $\boldsymbol{H}$, in the first-principles theory, we generalize the conventional constitutive eqs $\boldsymbol{D} = \varepsilon \boldsymbol{E}$ and $\boldsymbol{B} = \mu \boldsymbol{H}$ by adding $\boldsymbol{B}$-induced electric polarization and $\boldsymbol{E}$-induced magnetic polarization as

$$\boldsymbol{D} = \hat{\varepsilon}\boldsymbol{E} + i\hat{\xi}\boldsymbol{B} \,, \tag{3.52}$$

$$\boldsymbol{H} = (1/\hat{\mu})\boldsymbol{B} + i\hat{\eta}\boldsymbol{E} \,, \tag{3.53}$$

Here, ε and μ for chiral media are denoted as $\hat{\varepsilon}$ and $\hat{\mu}$, and the prefactor i for $\hat{\xi}$ and $\hat{\eta}$ is multiplied by comparing with (2.127) and (2.128).

This set of phenomenological constitutive equations are different from DBF eqs, and hereafter we call them "chiral constitutive" (ChC) eqs. The parameters of ChC eqs have correspondence to the microscopic susceptibilities χ_{eE}, χ_{eB}, χ_{mB}, χ_{mE}, (2.125)–(2.128), as

$$1 + 4\pi \chi_{\mathrm{eE}} \sim \hat{\varepsilon} \,, \qquad \chi_{\mathrm{eB}} \sim \hat{\xi} \,, \tag{3.54}$$

$$1/(1 - 4\pi \chi_{\mathrm{mB}}) \sim \hat{\mu} \,, \qquad \chi_{\mathrm{mE}} \sim \hat{\eta} \,. \tag{3.55}$$

It should be noted that the the diagonal and off-diagonal elements of χ_{eB}, χ_{mE} are interchanged from those of corresponding components of χ_{em}. This is because the tensors χ_{em} are the coefficients of $\boldsymbol{A}$ on writing induced current density, while χ_{eB} is the coefficient of $\boldsymbol{B}$ on writing the $\boldsymbol{B}$-induced electric polarization (corresponding to $-i\omega\boldsymbol{P}$ part of current density), and χ_{mE} is the coefficient of $\boldsymbol{E}$ on writing the $\boldsymbol{E}$-induced magnetization (corresponding to the $ic\boldsymbol{k} \times \boldsymbol{M}$ part of current density). Since $\boldsymbol{B} = i\boldsymbol{k} \times \boldsymbol{A}$ and $\boldsymbol{A}$ are perpendicular to each other, the diagonal (off-diagonal) elements of χ_{eB}, χ_{mE} are the off-diagonal (diagonal) elements with appropriate sign change of the corresponding tensor component of χ_{em}.

For the Cartesian framework (x, y, z) with the z-axis along $\boldsymbol{k}$, the chiral components giving rise to the different phase velocities of left- and right-circularly polarized EM waves appear on the off-diagonal, xy, yx, positions of χ_{em} in k-linear forms, while on the diagonal, xx, yy, positions of χ_{eB} and χ_{mE} in k-independent forms, for the reason mentioned above. In this way, it is possible to use scalar parameters $\hat{\xi}$, $\hat{\eta}$, as well as $\hat{\varepsilon}$, $\hat{\mu}$, for the tensor quantities shown in (3.54), (3.55) when the system is isotropic in the xy plane.

The dispersion equation arising from Maxwell eqs $\nabla \times \boldsymbol{H} = (1/c)\partial \boldsymbol{D}/\partial t$, $\nabla \times \boldsymbol{E} = (-1/c)\partial \boldsymbol{B}/\partial t$ and ChC eqs can be obtained in a similar manner as for DBF eqs. But we show here another way to the same goal. If we rewrite ChC eqs by substituting Maxwell eqs, we have a set of linear equations for $\boldsymbol{E}$, $\boldsymbol{H}$, $\nabla \times \boldsymbol{E}$, $\nabla \times \boldsymbol{H}$, which can be solved for $\nabla \times \boldsymbol{E}$, $\nabla \times \boldsymbol{H}$ as

$$\nabla \times \boldsymbol{H} = f\boldsymbol{H} + g\boldsymbol{E} \,, \tag{3.56}$$

$$\nabla \times \boldsymbol{E} = h\boldsymbol{H} + j\boldsymbol{E} \,, \tag{3.57}$$

where

$$f = (\omega/c)\hat{\xi}\hat{\mu}\ , \tag{3.58}$$

$$g = -i(\omega/c)(\hat{\varepsilon} + \hat{\mu}\hat{\xi}\eta)\ , \tag{3.59}$$

$$h = i(\omega/c)\hat{\mu}\ , \tag{3.60}$$

$$j = (\omega/c)\hat{\eta}\hat{\mu}\ . \tag{3.61}$$

From $\nabla \times \boldsymbol{H} = (1/c)\partial \boldsymbol{D}/\partial t, \boldsymbol{D}$ is transverse, i.e., $\nabla \cdot \boldsymbol{D} = 0$, which together with (3.52) and $\nabla \cdot \boldsymbol{B} = 0$, leads to $\nabla \cdot \boldsymbol{E} = 0$, and further $\nabla \cdot \boldsymbol{H} = 0$ via (3.53). Thus, for ChC eqs too, we have $\nabla \times \nabla \times \boldsymbol{E} = k^2\boldsymbol{E}$ and $\nabla \times \nabla \times \boldsymbol{H} = k^2\boldsymbol{H}$. Then, the application of $\nabla\times$ to (3.56), (3.57) gives a set of linear equations of $\boldsymbol{H}, \boldsymbol{E}$. In order for the linear equations to have non-trivial solution, we request the vanishing of the determinant of the coefficient matrix

$$\det|k^2\mathbf{1} - \mathscr{A}^2| = 0\ , \tag{3.62}$$

where $\mathscr{A}$ is a 2×2 matrix as

$$\mathscr{A} = \begin{bmatrix} f & g \\ h & j \end{bmatrix}\ . \tag{3.63}$$

We can rewrite the dispersion equation into

$$\det|k\mathbf{1} + \mathscr{A}| = 0 \quad \text{or} \quad \det|k\mathbf{1} - \mathscr{A}| = 0\ , \tag{3.64}$$

the solution of which is

$$\frac{ck}{\omega} = \pm\frac{\hat{\mu}(\hat{\xi} + \hat{\eta})}{2} \pm \frac{1}{2}\sqrt{\hat{\mu}^2(\hat{\xi} + \hat{\eta})^2 + 4\hat{\varepsilon}\hat{\mu}}\ , \tag{3.65}$$

where all the four combinations of ± are allowed. This can be rewritten as

$$\frac{ck}{\omega} = \pm\frac{2\hat{\varepsilon}\hat{\mu}}{\hat{\mu}(\hat{\xi} + \hat{\eta}) \pm \sqrt{\hat{\mu}^2(\hat{\xi} + \hat{\eta})^2 + 4\hat{\varepsilon}\hat{\mu}}}\ , \tag{3.66}$$

which facilitates the comparison with the dispersion relation for DBF eqs (3.51).

Both of the dispersion equations for DBF and ChC eqs reduce to $(ck/\omega)^2 = \varepsilon\mu (= \hat{\varepsilon}\hat{\mu})$ in the absence of chirality, as expected. However, the condition for the existence of real solution is different, i.e.,

$$\varepsilon\mu \geq 0 \tag{3.67}$$

for DBF eqs, and

$$\hat{\varepsilon}\hat{\mu} + \hat{\mu}^2\left(\frac{\hat{\xi} + \hat{\eta}}{2}\right)^2 \geq 0 \tag{3.68}$$

for ChC eqs, i.e., the condition for DBF eqs is not affected by the chirality, while the one for ChC eqs is. The latter is more relaxed in the sense that the product $\hat{\varepsilon}\hat{\mu}$ can be negative. This leads to a qualitative difference between DBF and ChC eqs, especially in the dispersion curves in the left-handed situation, as we show in Sect. 4.1.5.

In view of the fact that there exists a first-principles theory, not for DBF eqs, but for ChC eqs, ChC eqs are more reliable. However, in the field of metamaterials, DBF and ChC eqs are often regarded to be equivalent. One might think that, by adjusting the ω-dependent parameters, DBF eqs could describe the same physical situations as ChC eqs. If this were true, there must be the following relations between the parameters of DBF and ChC eqs

$$\hat{\varepsilon} = \varepsilon \ , \quad \hat{\xi} = \hat{\eta} = (\omega/c)\varepsilon\beta \ , \quad (1/\hat{\mu}) = (1/\mu) - (\omega\beta/c)^2\varepsilon \ , \tag{3.69}$$

derived from the comparison of (3.35), (3.36) and (3.52), (3.53). This relation enables us to ascribe appropriate ω-dependence to the DBF parameters, which turns out to lead to a contradiction to the assumed equivalence of the two in the behavior of the dispersion curve for a left-handed chiral medium at resonance, i.e., the linear crossing behavior of the dispersion curve of the first-principles formula is reproducible by ChC eqs, but not by DBF eqs, as discussed in Sect. 4.1.5.

From the arguments given above and in Sect. 4.1.5, there is a serious doubt about the validity of DBF eqs especially in a resonant region. They are frequently used for metamaterials research, but it seems to be little recognized that their use should be restricted to nonresonant phenomena. In fact there are examples of the use for resonant phenomena [4, 5]. The constitutive equations in these papers are $\boldsymbol{D} = \varepsilon\boldsymbol{E} - i\xi\boldsymbol{H}$, $\boldsymbol{B} = \mu\boldsymbol{H} + i\xi\boldsymbol{E}$. They are not quite DBF eqs but of the same type, in the sense that they lead to the dispersion equation $(ck/\omega) = \pm(\sqrt{\varepsilon\mu} \pm \xi)$, where the condition for real solution does not depend on chirality. Also, it gives the type of dispersion curve which is unable to reproduce the linear crossing at $k = 0$ for resonant chiral LHM. See Sect. 4.1.5 for more details.

3.5 Other Unconventional Theories

3.5.1 Single Susceptibility Theories

There are attempts by Agranovich and Ginzburg (AG) [6], and Il'inskii and Keldysh (IK) [7] to describe the macroscopic EM response of matter in terms of a single susceptibility, i.e., by using only one of the two polarization vectors $\boldsymbol{P}$ and $\boldsymbol{M}$, which had been indicated by Landau and Lifshitz (Sect. 83 of [8]). They renormalize the whole current density into displacement vector $\boldsymbol{D}$, or $\boldsymbol{P}$ via

$$\boldsymbol{P}(\boldsymbol{r}, t) = \int_{-\infty}^{t} \mathrm{d}t' \, \boldsymbol{J}(\boldsymbol{r}, t') \tag{3.70}$$

and treat it as the single dynamical variable of matter. This $\boldsymbol{P}(\boldsymbol{r}, t)$ contains both electric and magnetic polarizations, so that the generalized susceptibility defined by

$$\boldsymbol{P} = \chi^{(\mathrm{gen})} \boldsymbol{E} \tag{3.71}$$

describes all the effect of electric and magnetic polarizations. The vector field for magnetic field is $\boldsymbol{B}(= \boldsymbol{H})$.

Though we all share the motivation to give a more general formulation of macroscopic M-eqs than the conventional one, there is an essential difference between the two groups and the present author with respect to the very concept of macroscopic average. AG and IK reject LWA as a meaningful physical procedure, but the present author regards LWA, as far as its clear definition is given, as the essential step toward macroscopic description. AG and IK claim to use a statistical average in terms of Gibbs ensemble instead of LWA. There is an explicit statement about this point in Sect. 2.1 of AG to supplement their own form of macroscopic M-eqs, i.e., "The fields $\boldsymbol{E}, \boldsymbol{D}, \boldsymbol{B}$ may vary in anyway in space and time without requiring any kind of averaging (apart from quantum mechanical and statistical kinds) of the fields with respect to $\boldsymbol{r}$. Such averaging is not only unnecessary, but, generally speaking, is unfeasible in the electrodynamics of media if spatial dispersion is properly taken into account."

The present author also uses ensemble average in the case of finite temperature, but this has nothing to do with macroscopic averaging, as discussed in Sect. 1.6. In fact, an ensemble average does not erase the microscopic spatial variation of induced current density. For example, the ensemble average (for $\mathrm{T} \neq 0$) of the current density due to a discrete level of excitation will keep it discrete (apart from the increased width due to lattice vibrations), with a change only in its spectral weight. If this transition is localized, LWA is a good approximation, and it can be treated by a macroscopic theory. If, on the other hand, it has a long spatial coherence, as in the case of an exciton, the induced current density keeps its long coherence, which invalidates LWA and a macroscopic description.

If one uses only ensemble average and drop LWA for a macroscopic description, as in the comment of AG, how does one distinguish the micro- and macroscopic responses? The scheme of the present author starts from the recognition of the hierarchy of various theoretical frameworks of EM response (Sect. 1.4). The micro- and macroscopic M-eqs with the corresponding constitutive equations are those belonging to the semiclassical regime, and the approximation separating them is the LWA applied to the fundamental equations of microscopic response theory. The calculations of microscopic susceptibility in Sects. 2.2 and 5.7 is essentially same as those of AG and IK. The refusal of using LWA by AG and IK as the next step to derive macroscopic susceptibility may indicate that they do not share the understanding about the hierarchy of the micro- and macroscopic M-eqs with the present author.

As for the use of LWA, our concept is as follows. Depending on the matter system of interest and also on the quantity to be observed, LWA can be a good or bad approximation. When LWA is a good approximation, its use is the simple and logically acceptable way of macroscopic averaging, and only in this case the macroscopic description is meaningful. Whenever the quantum mechanical transitions of interests have larger coherence length than the wavelength of EM field, we cannot apply LWA. This kind of situation is frequently encountered in resonant responses, and it is the case to be handled by the microscopic nonlocal theory in Sect. 2.2. Typical cases appropriate for a macroscopic description would be the non-resonant phenomena where no particular transition make a significant contribution. The resonant phenomena for localized transitions can also be a subject for macroscopic description, if we do not want to see the precise dependence on the positions of localized states. (The resonant X-ray diffraction in Sect. 4.3 is an example of the position-sensitive case, which should be treated as a nonuniform system.)

The $\boldsymbol{k}$ dependence of susceptibility is generally called spatial dispersion effect. It will make sense to divide the $\boldsymbol{k}$ dependence into two cases, [a] the $\boldsymbol{k}$ dependence only in the numerators of susceptibility, and [b] the $\boldsymbol{k}$ dependence also in the denominators. The case [b] leads to a qualitatively new situation of multi-branch polaritons, which has been studied as ABC (additional boundary condition) problem for nearly half a century [9–11]. However, it is a problem to be treated as a microscopic response, because the $\boldsymbol{k}$ dependence of denominator arises from that of the transition energy, which means a coherently extended excited state specified by $\boldsymbol{k}$ (and material boundaries), an inappropriate situation for LWA. Thus, only the case [a] is suitable for the macroscopic description. As explicitly discussed in Sects. 2.3 and 2.4, susceptibility is expressed as a power series expansion with respect to $\boldsymbol{k}$, reflecting the Taylor expansion of current density matrix elements. Since LW means a small $|\boldsymbol{k}|$, it is quite reasonable to express LWA as a power series expansion with respect to $\boldsymbol{k}$. The merit of this expansion is that one can see the meaning of the expanded terms (moments of E1, E2, M1 transitions etc.) from the quantum mechanical expression of the matrix elements, which gives us the symmetry condition by which we keep or abandon a certain class of them.

In order to calculate a microscopic susceptibility, one uses a time dependent perturbation theory of Schrödinger or Liouville equation with appropriately defined unperturbed Hamiltonian and perturbation Hamiltonian. In this sense, all of AG, IK and the present author get similar expressions. However, IK use the "unperturbed Hamiltonian" ($\mathscr{H}_0$ in their notation) different from others. IK split vector and scalar potentials into external and induced ones, and keep the induced ones in the unperturbed Hamiltonian. It means that this unperturbed Hamiltonian does not represent a pure matter system, but a coupled one of matter and EM field. In the case of crystals, the excited states of this unperturbed Hamiltonian are polaritons, rather than excitons and/or LO phonons. This is different from the usual definition of susceptibility with respect to the total (incident plus induced) EM field, the poles of which represent, not the polariton energies, but the excitation energies of matter. Knowing this difference, IK discuss the explicit relationship between the two susceptibilities (Sect. 1.6 of [7]), and show the occurrence of polariton poles in the susceptibility defined for $\mathscr{H}_0$. This argument establishes the relationship between the two susceptibilities, so that it may seem unnecessary to worry about which susceptibility should be used.

However, there is a technical detail which becomes significant for nanostructures having strong interaction with EM field. Generally, the interaction between matter and EM field leads to the radiative correction (width and shift of excitation energies), dependent on the “size and shape” of matter and on the “state” of excitation. For appropriately designed matter systems, there can occur radiative width exceeding non-radiative one [12]. This problem can be handled straightforwardly by our formulation in Chap. 2, where the matrix elements of radiative correction appear directly in the equations to determine the selfconsistent solution. In the case of IK, this effect is formally included in the unperturbed Hamiltonian $\mathscr{H}_0$, but the recipe is missing to calculate the radiative correction for each excited level in a “size and shape” dependent way.

Though AG and IK are keen in presenting a single susceptibility scheme of macroscopic M-eqs, they do not put their result in conflict with the conventional macroscopic M-eqs. Rather than that, Agranovich et al. try to reconcile their result with the conventional one, by proposing one-to-one correspondence between the two schemes [13]. In contrast, we claim the explicit differences between our new result and the conventional M-eqs. Especially, the explicit derivation of the chirality induced components of susceptibility, χ_{eB} and χ_{mE} in Sect. 3.2, is a new result exceeding the phenomenology of DBF constitutive equations.

3.5.2 Comparison of Single Susceptibility Theories

There is a recent discussion by Chipouline et al. [14] about the possible forms of constitutive equations for macroscopic Maxwell equations in three versions, (A) Casimir form, (B) Landau-Lifshitz form, and (C) Anapole form, and their mutual relationship. This subsection deals with the comparison of these forms of M-eqs with that of this book [15].

The argument of Chipouline et al. starts by noting the arbitrariness to determine $\boldsymbol{P}$ and $\boldsymbol{I}$ from ρ by the equations

$$\rho = -\nabla \cdot \boldsymbol{P}\,, \quad \partial\rho/\partial t = -\nabla \cdot \boldsymbol{I}, \tag{3.72}$$

i.e., both $\boldsymbol{P}$ and $\boldsymbol{I}$ may contain a term of the form $\nabla \times \boldsymbol{F}$, where $\boldsymbol{F}$ is an arbitrary vector function. This allows various choices of dynamical variables to describe the change in matter state due to the external perturbations. Following this idea, they introduce the following three forms of “material equations” to describe charge and current densities in terms of the new variables:

(A) Casimir form

$$\rho = -\nabla \cdot \boldsymbol{P}_{\mathrm{C}}\,, \quad \boldsymbol{I} = -i\omega \boldsymbol{P}_{\mathrm{C}} + c\nabla \times \boldsymbol{M}_{\mathrm{C}}\,, \tag{3.73}$$

$$\left[\boldsymbol{I} = -i\omega \boldsymbol{P}_{\mathrm{C}} + \nabla \times \boldsymbol{M}_{\mathrm{C}}\right]_{\mathrm{SI}}\,, \tag{3.74}$$

(B) Landau-Lifshitz form

$$\rho = -\nabla \cdot \boldsymbol{P}_{\mathrm{LL}}\ , \quad \boldsymbol{I} = -i\omega \boldsymbol{P}_{\mathrm{LL}}\ , \tag{3.75}$$

(C) Anapole form

$$\rho = 0\ , \quad \boldsymbol{I} = \mathrm{c}\nabla \times \boldsymbol{M}_{\mathrm{A}} \quad = \left[\ \nabla \times \boldsymbol{M}_{\mathrm{A}}\ \right]_{\mathrm{SI}}\ . \tag{3.76}$$

The new variables $\boldsymbol{P}_{\mathrm{C}}$ and $\boldsymbol{M}_{\mathrm{C}}$ are the conventional electric and magnetic polarizations, respectively, while $\boldsymbol{P}_{\mathrm{LL}}$ ($\boldsymbol{M}_{\mathrm{A}}$) contains, not only electric (magnetic) polarization, but also the contribution of magnetic (electric) polarization. In cases of Landau-Lifshitz and Anapole forms, only one (vector) variable is used, while we employ two in Casimir case. The two cases (B) and (C) may be called single susceptibility theory, since there is only one susceptibility needed between the matter variable $\boldsymbol{P}_{\mathrm{LL}}$ ($\boldsymbol{M}_{\mathrm{A}}$) and source EM field.

The single susceptibility theory of this book employs, as the dynamical variable of matter, current density $\boldsymbol{I}$ itself in contrast to the cases (A)–(C), and the constitutive equation, especially the one in Sect. 3.1 rewritten in terms of $\boldsymbol{P}_{\mathrm{E,B}}$ and $\boldsymbol{M}_{\mathrm{E,B}}$ looks like the one expected from the Casimir form. However, this rewritten form is still a single susceptibility theory because the rewriting can be reversed without approximation to the original constitutive equation with a single susceptibility $\chi_{\mathrm{em}}(\boldsymbol{k}, \omega)$. In view of the fact that the conventional Casimir form is not supported by any single susceptibility scheme, we should treat these two cases as different schemes, and let us call the the one in Sect. 3.1 as (D) "natural" form.

At this stage there arises a question "how are the three single susceptibility theories (B), (C), and (D) related?" More detailed questions will be "Are there any general expressions of single susceptibility for (B) and (C) as in the case of (D)?", or "Are the dispersion equations for (B) or (C) different from that of (D)?" The effort to find the way of rewriting among (A), (B), and (C) in [14] is not quite successful. In particular, it is not possible to rewrite (C) into (A) and (B), because the variable $\boldsymbol{M}_{\mathrm{A}}$ takes care of only the transverse component of $\boldsymbol{I}$, while $\{\boldsymbol{P}_{\mathrm{C}}, \boldsymbol{M}_{\mathrm{C}}\}$ and $\{\boldsymbol{P}_{\mathrm{LL}}\}$ contain the longitudinal component of $\boldsymbol{I}$, which does not allow to rewrite the relation of the two sets of variables in a reversible way. Chipouline et al. do not intend to give the microscopic expression of the susceptibility corresponding to each form of the matter variables. Landau and Lifshitz give a symmetry argument of the response function including the contribution from the first space derivative of electric field, corresponding to the effect of E2 and M1 transitions, but the microscopic expression of linear susceptibility including all the E1, E2, and M1 transitions, like that in Sect. 3.1, is not given by them [8], neither by AG or IK [6, 7].

In order to find the way to interrelate the different forms (B), (C), and (D), we should be aware of the fact that the arbitrariness included in (3.72) is only about the T component of the vectors $\boldsymbol{P}$ and $\boldsymbol{I}$. The L components are not affected by this arbitrariness. Therefore, the choice of matter variables in each case should be

(B) $\boldsymbol{P}_{\mathrm{LL}}^{(\mathrm{T})}$ and $\boldsymbol{I}^{(\mathrm{L})}$,

(C) $\boldsymbol{M}_{\mathrm{A}}^{(\mathrm{T})}$ and $\boldsymbol{I}^{(\mathrm{L})}$,

(D) $\boldsymbol{I}^{(\mathrm{T})}$ and $\boldsymbol{I}^{(\mathrm{L})}$.

One could use $\boldsymbol{P}^{(\mathrm{L})}$ $(= -i\omega\boldsymbol{I}^{(\mathrm{L})})$ instead of $\boldsymbol{I}^{(\mathrm{L})}$ in each case, and also we may employ $\{\, \boldsymbol{P}_{\mathrm{E}}^{(\mathrm{T})}, \boldsymbol{P}_{\mathrm{B}}^{(\mathrm{T})}, \boldsymbol{M}_{\mathrm{E}}^{(\mathrm{T})}, \boldsymbol{M}_{\mathrm{B}}^{(\mathrm{T})} \,\}$, (3.12), instead of $\boldsymbol{I}^{(\mathrm{T})}$.

Noting the transversality of the variables distinguishing the three cases, we write the defining equations of the variables for (B) and (C), in terms of $\boldsymbol{I}^{(\mathrm{T})}$ for (D), as

$$\partial \boldsymbol{P}_{\mathrm{LL}}^{(\mathrm{T})}/\partial t = \boldsymbol{I}^{(\mathrm{T})} , \tag{3.77}$$

$$c\nabla \times \boldsymbol{M}_{\mathrm{A}}^{(\mathrm{T})} = \boldsymbol{I}^{(\mathrm{T})} , \tag{3.78}$$

$$\left[\, \nabla \times \boldsymbol{M}_{\mathrm{A}}^{(\mathrm{T})} = \boldsymbol{I}^{(\mathrm{T})} \,\right]_{\mathrm{SI}} . \tag{3.79}$$

These equations can be reversed, in $(\boldsymbol{k}, \omega)$ Fourier components, as

$$\boldsymbol{P}_{\mathrm{LL}}^{(\mathrm{T})} = (i/\omega)\boldsymbol{I}^{(\mathrm{T})} , \tag{3.80}$$

$$\boldsymbol{M}_{\mathrm{A}}^{(\mathrm{T})} = (i/ck^2)\boldsymbol{k} \times \boldsymbol{I}^{(\mathrm{T})} , \tag{3.81}$$

$$\left[\, \boldsymbol{M}_{\mathrm{A}}^{(\mathrm{T})} = (i/k^2)\boldsymbol{k} \times \boldsymbol{I}^{(\mathrm{T})} \,\right]_{\mathrm{SI}} . \tag{3.82}$$

This result shows how the three variables $\boldsymbol{P}_{\mathrm{LL}}^{(\mathrm{T})}, \boldsymbol{M}_{\mathrm{A}}^{(\mathrm{T})}$ and $\boldsymbol{I}^{(\mathrm{T})}$ are rewritable into each other. Thus, the knowledge of the first-principles constitutive equation in case (D), (3.23), leads to the corresponding equations for other cases, i.e.,

$$\boldsymbol{P}_{\mathrm{LL}}^{(\mathrm{T})} = (i/\omega)[\chi_{\mathrm{em}}(\boldsymbol{k}, \omega) \cdot \{\boldsymbol{A} + \frac{c}{i\omega}\boldsymbol{E}_{\mathrm{extL}}\}]^{(\mathrm{T})} , \tag{3.83}$$

$$= \left[\, \text{same expression with c replaced by 1} \,\right]_{\mathrm{SI}} , \tag{3.84}$$

$$\boldsymbol{M}_{\mathrm{A}}^{(\mathrm{T})} = (i/ck^2)\boldsymbol{k} \times [\chi_{\mathrm{em}}(\boldsymbol{k}, \omega) \cdot \{\boldsymbol{A} + \frac{c}{i\omega}\boldsymbol{E}_{\mathrm{extL}}\}]^{(\mathrm{T})} , \tag{3.85}$$

$$= \left[\, \text{same expression with c replaced by 1} \,\right]_{\mathrm{SI}} . \tag{3.86}$$

The susceptibility tensor $\chi_{\mathrm{em}}(\boldsymbol{k}, \omega)$ is derived by the LWA of its microscopic expression, and is given as (2.119) including the L component. Thus the constitutive equations given above are also the first-principles ones.

For a given medium there is a convenient orthogonal coordinate system (x, y, z) to calculate the susceptibility tensor χ_{em} according to the symmetry of the system, which is independent of the choice of the $\boldsymbol{k}$ direction. In order to solve the Maxwell and constitutive equations in $(\boldsymbol{k}, \omega)$ space, it is better to rewrite χ_{em} in another coordinate system suitable to discuss the T and L characters of the vector variables.

Let us take the Descartes coordinate system (ξ, η, ζ), where ζ axis is in the $\boldsymbol{k}$ direction, and write the components of χ_{em} as $\chi_{\xi\xi}, \chi_{\xi\eta}, \chi_{\xi\zeta}, \ldots$ etc. The rewriting from (x, y, z) to (ξ, η, ζ) systems is a simple orthogonal transformation. Remembering the T character of $\boldsymbol{A}$ and the L character of $\boldsymbol{E}_{\mathrm{ext}}$, we may write their components as

$$\boldsymbol{A} = (A_\xi, A_\eta, 0)\ , \tag{3.87}$$

$$\boldsymbol{E}_{\text{ext}} = (0, 0, E_{\text{L}})\ . \tag{3.88}$$

In terms of these components, the constitutive equation, (2.118), is written as

$$I_\xi = \chi_{\xi\xi} A_\xi + \chi_{\xi\eta} A_\eta + (c/i\omega)\chi_{\xi\zeta} E_{\text{L}}\ , \tag{3.89}$$

$$I_\eta = \chi_{\eta\xi} A_\xi + \chi_{\eta\eta} A_\eta + (c/i\omega)\chi_{\eta\zeta} E_{\text{L}}\ , \tag{3.90}$$

$$I_\zeta = \chi_{\zeta\xi} A_\xi + \chi_{\zeta\eta} A_\eta + (c/i\omega)\chi_{\zeta\zeta} E_{\text{L}}\ , \tag{3.91}$$

$$\big[\text{ same expressions with c replaced by 1 }\big]_{\text{SI}} \tag{3.92}$$

and the T components of $\boldsymbol{P}_{\text{LL}}$ and $\boldsymbol{M}_{\text{A}}$ are given as

$$P_{\text{LL}\xi} = (i/\omega) I_\xi, \quad P_{\text{LL}\eta} = (i/\omega) I_\eta, \tag{3.93}$$

$$M_{\text{A}\xi} = -(i/ck) I_\eta\ , \quad M_{\text{A}\eta} = -(i/ck) I_\xi\ , \tag{3.94}$$

$$\big[\, M_{\text{A}\xi} = -(i/k) I_\eta\ , \quad M_{\text{A}\eta} = -(i/k) I_\xi \big]_{\text{SI}}\ . \tag{3.95}$$

Dispersion relation is the condition between "ω and $\boldsymbol{k}$" of the eigenmode of coupled EM field - matter excitations. Mathematically, it is obtained by requiring the existence of the non-trivial solution of the coupled Maxwell and constitutive equations in the absence of the external EM field, i.e., for $E_{\text{L}} = 0$. In the case of (D) natural form, this process is given in Sect. 2.5. In terms of the tensor components defined above, the coupled equation after elimination of current density turns out to be

$$\begin{bmatrix} k^2 - (\omega/c)^2 & 0 \\ 0 & k^2 - (\omega/c)^2 \end{bmatrix} \begin{bmatrix} A_\xi \\ A_\eta \end{bmatrix} = \frac{4\pi}{c} \begin{bmatrix} \chi_{\xi\xi} & \chi_{\xi\eta} \\ \chi_{\eta\xi} & \chi_{\eta\eta} \end{bmatrix} \begin{bmatrix} A_\xi \\ A_\eta \end{bmatrix}\ , \tag{3.96}$$

and [the expression with the factor $4\pi/c$ replaced by μ_0] in SI units system. The condition for the existence of non-trivial solution is the vanishing of the determinant of the coefficient matrix, i.e., (2.131). The dispersion relation for the other forms (B) or (C) can be obtained from the Maxwell and constitutive equations written in terms of $\boldsymbol{P}_{\text{LL}}$ or $\boldsymbol{M}_{\text{A}}$. After eliminating these matter variables, we get the same set of equations to be satisfied by A_ξ and A_η, which leads to the same dispersion equation for (B) and (C). This is a due expectation, since the eigen modes of matter - EM field system should not depend on the matter variables to be used,as far as they are consistently defined.

It should be stressed that the tensor components $\chi_{\xi\xi}$ etc. are the functions of ω and k, as explicitly given in Sects. 2.3 and 2.4. Especially in the case of chiral symmetry, the existence of k-linear terms plays an essential role. Such terms in χ_{em} arise from the existence of the excited states with mixed E1–M1 and/or E1–E2 characters due to the absence of certain mirror symmetry in a chiral medium, and can contribute to new terms in the generalized dielectric function defined by $\boldsymbol{D} = \boldsymbol{E} + 4\pi\boldsymbol{P}_{\text{LL}} = \varepsilon_{\text{LL}}\boldsymbol{E}$. Landau and Lifshitz give some symmetry arguments about the form of this ε_{LL}. Since the new variable $\boldsymbol{P}_{\text{LL}}$ is introduced in the section about "Optical Activity" of their textbook [8], its usefulness would have been expected for the description of both single susceptibility theory and chiral response of matter. However, it is not obvious why the introduction of a new unfamiliar variable was necessary in view of the fact that the same argument can be done even more clearly in terms of the familiar variable $\boldsymbol{I}$ and the susceptibility tensor χ_{em}, as shown above.

In the argument given above, $\chi_{\xi\zeta}$ and $\chi_{\eta\zeta}$ might seem to have no contribution to the dispersion relation, in spite of the fact that they lead to a finite induced current density or charge density. Actually, the effect of induced $\boldsymbol{I}^{(\mathrm{L})}$ appears in the matter excitation energies in the poles of χ's as the additional interaction energies among the induced charge densities corresponding to $\boldsymbol{I}^{(\mathrm{L})}$. In the present formulation of theory (including the gauge invariant character in Sect. 2.2.5), the variables of EM field in the starting Hamiltonian are $\boldsymbol{A}^{(\mathrm{T})}$ and $E_{\mathrm{ext}}^{(\mathrm{L})}$, where $\boldsymbol{A}^{(\mathrm{T})}$ is the sum of incident and induced components and $E_{\mathrm{ext}}^{(\mathrm{L})}$ is the incident field itself. The induced component of $E^{(\mathrm{L})}$ is expressed as induced charge density, which can then be rewritten into $E_{\mathrm{ind}}^{(\mathrm{L})}$. In this way, the effect of induced charge (or L current) density is correctly taken into account in the dispersion equation through matter excitation energies.

In this argument, we take a standpoint that the whole effect of Coulomb interaction among internal charges is treated as a part of matter energy. However, there is another viewpoint where one considers this energy as the interaction energy between the Maxwell field (sum of incident and induced fields) and induced charge density. The definition of the source field inducing current density is then different, so that one needs a new susceptibility. These two schemes can be made consistent by treating the self interaction energy of induced current density as matter energy or matter-EM field interaction, as mentioned in the last paragraph of Sect. 2.2. For detailed argument, see [18].

3.5.3 Use of LWA on a Different Stage

LWA plays an essential role in the present derivation of macroscopic M-eqs from microscopic ground. Though it has been used also in the conventional ways of derivation, the one adopted here has a logically and mathematically clearer definition, and we believe it to be the most appropriate way of using LWA to derive macroscopic M-eqs in a general form. However, there is a proposal by Nelson [16] to use LWA in quite a different manner to derive a new scheme of macroscopic M-eqs. His intension was to build a consistent theory of dynamical response of crystalline medium from the first-principles in a unified manner, to avoid “the patchwork of phenomenologically assumed constitutive relations of so many treatments” (Preface of [16]).

As a systematic method to describe macroscopic dynamical response (including EM response) of matter, Nelson applies LWA to the Lagrangian for matter (both electronic and lattice vibrational) and EM field. In terms of the averaged Lagrangian, he discusses linear and nonlinear “optics, acoustics, and acousto-optics” of dielectric crystals. This may well be a meaningful approach to the phenomena related with these LW modes, though it has a rather unusual form.

However, if it is meant to be a general unified theory of EM response, it contains a serious drawback, i.e., it abandons the dynamical variables contributing to localized excitations of matter. By the application of LWA to Lagrangian, the dynamics of matter is described solely by the LW components of acoustic and optical phonons, excitons etc. Thus the only contribution to susceptibility is made from these LW modes of matter, i.e. the susceptibility has poles only at the frequencies of these LW modes. Since all the dynamical variables of short wavelength components are eliminated by the LWA of Lagrangian, there is no chance for localized excitations to contribute to susceptibility.

When we consider a problem, for example, of changing the refractive index of a material by adding impurities, we need to consider the macroscopic average of the contributions from localized transitions due to impurities. According to the scheme of Chap. 2 of this book, we obtain a finite contribution reflecting the density of impurities and the oscillator strength (or the magnitude of E1 transition moment) of the transition. If we use the Nelson's scheme to this problem, however, all the dynamical variables to build the localized excitations are erased out on the level of Lagrangian (and then, Hamiltonian), from which we cannot expect a finite contribution to the macroscopic susceptibility.

Optical phenomena are not always caused by the matter excitations with the wavelength comparable to the observed light. In fact, the absorption, emission and scattering of visible lights by atoms, molecules, defects and impurities are the well-known examples which built the basis of our fundamental knowledge of optical phenomena. The birth of quantum mechanics was motivated by the interpretation of the atomic spectra of hydrogen, and the earliest solid state spectroscopy was the study of color centers in alkali halides, which is a good example of macroscopic optical problems of dielectrics. These examples are all related with the interaction of localized electrons and long wavelength lights. If this group of phenomena is not covered by a theory, one would not call it a "unified" theory.

3.6 Validity Condition of LWA

The LWA in this book is a process of approximation to extract a new set of equations for the LW components of the variables of EM field ($\boldsymbol{A}$) and matter ($\boldsymbol{I}$) from the more fundamental equations containing all the wavelength components, i.e., from the microscopic M-eqs and microscopic constitutive equations. The new set of equations contains only the LW components of the dynamical variables, and hence, it is macroscopic. The description in Chap. 2 clarifies the logical and mathematical aspects of this procedure. Mathematically, we apply Taylor expansion to (the Fourier component of) the matrix element of current density for each transition (around each center coordinate), keeping a few lower order terms. These lower order terms are described by the lower order (E1, E2, M1, etc.) moments of the matrix element of current density.

As an approximation, LWA can be good or bad, depending on the case of interests. The criterion to judge it is the relative size of the wavelength (Λ) of the EM field in consideration compared with the coherence length of induced current densities. This corresponds to whether or not we can neglect the higher order terms of Taylor expansion. Since the induced current density consists of a sum of the contributions of all the excited states of matter, we cannot always assume that "all" the excitations have shorter coherence length than a given Λ. If we need the response of these modes, we should treat them, not in LWA, but microscopically. The LWA formulation in Sect. 2.3 assumes that the contribution of these modes is negligible in amplitude compared with that of remaining modes.

At this point the argument may have a subjective aspect, i.e., which physical process we want to observe or discuss. For example, an incident field may induce several different physical processes, each one of which can have different criterion for the use of LWA. An example is the inner core level excitation of a crystal, which leads to absorption and emission of light and also a resonant (X-ray) scatterings, as will be discussed in Sect. 4.3. Though the scattering intensity will be much smaller than the absorption signal, one can observe it in the specific directions of diffraction, and this requires a treatment beyond LWA. This example shows that there are cases where a subjective choice of a physical process may require a microscopic treatment of particular modes together with the macroscopic treatment of the remaining modes. The standard criterion for the use of LWA is the smallness of the signal intensity due to the LW modes, which need to be treated microscopically, in comparison with the signal due to the macroscopically averaged short wavelength modes.

The choice of Λ is connected with the physical quantity and the frequency range to be measured or discussed, and it should be noticed that this is not the wavelength in vacuum, but the one in the medium determined by the background polarization in the frequency range of interest. (If there is a resonance in this range, the contribution of this resonance should be omitted in estimating the background polarization.) Once the choice of Λ is made, one can compare it with all the candidates of excitation modes which will make the main contribution to the EM response of this system.

A reliable test of the validity of LWA for a given model would be to calculate the microscopic nonlocal response, and see whether the LW components are dominant in the response spectra. This theory gives us response spectra properly containing all the short and LW components of excitations of the matter of interest. If the amplitudes of the LW components are dominant in the response spectra, LWA is valid, and otherwise, LWA is not a good approximation. Though this kind of calculation would generally require a large scale numerical treatment, i.e., a sufficiently large size of the simultaneous linear equations of $\{F_{\mu\nu}\}$ in Sect. 2.2.3, we can obtain explicit results for simple systems (e.g., Sect. 4 of [17]), and, for larger realistic systems, we know at least the equations to solve. With this kind check of LWA, we can safely proceed to use the macroscopic M-eqs and the corresponding constitutive equation.

The derivation of a macroscopic scheme is justified when the validity condition of LWA is checked properly. Though an accurate check is generally difficult, one could develop a feeling of valid and invalid cases. In fact, the situations which allow the description in terms of the macroscopic susceptibilities are rather limited. As a macroscopic description, the ω-dependence of susceptibility may be included with poles at the energies of material excitations. As to the $\boldsymbol{k}$-dependence, on the other hand, its appearance in the excitation energies (in the pole positions of macroscopic susceptibility) is not allowed. Such a $\boldsymbol{k}$-dependence would mean that the eigenstates are coherently extended, an invalid condition for the use of LWA. Only when the band width (due to the $\boldsymbol{k}$-dispersion) is negligible in comparison with the level width (due to phonon scattering or inhomogeneity, etc.), the coherence length can be regarded as negligibly small, and the use of LWA will be allowed for the macroscopic description. In contrast, the $\boldsymbol{k}$-dependence of the numerator of susceptibility is acceptable, since the Taylor expansion, the mathematical representation of LWA, is a power series expansion of susceptibility with respect to k, as shown in Chap. 2.

When we consider the case containing non-negligible LW modes, we need an intermediate scheme between the microscopic nonlocal and fully macroscopic ones. In such a scheme, we ascribe microscopic current densities to the modes with long coherence lengths, and LWA averaged current density to the short wavelength modes. The selfconsistent motions of the long coherence modes is determined by a new scheme derived from the microscopic nonlocal one. As the examples of this case, we discuss (i) resonant X ray scattering from the inner core transitions of a crystal in Sect. 4.3, and (ii) metamaterials with long coherence modes of excitations in Sect. 4.1.3.

3.7 Boundary Conditions for EM Fields

When we determine the EM response of matter from the macroscopic M-eqs, we usually proceed as follows. First, we solve the M-eqs in- and outside the matter separately, select the incident and the response fields according to the geometry in consideration, and connect the fields across the matter boundary according to the EM boundary conditions (BC's). The physical origin of the BC's must be in the matter with a given size and shape, but the BC's are requested to the EM field. This is a peculiar aspect of the macroscopic response theory.

In the microscopic response, no BC is required to the EM field, because the microscopic nonlocal susceptibility contains all the necessary information of BC's, requested to the charged particles in matter [17]. The problem of response calculation is formulated as a scattering problem, where the response field is obtained as a convolution of incident field, the position dependent susceptibility, and the EM Green function describing the propagation of the scattered field. The information about the material boundary is included in the susceptibility in a complete form, so that the introduction of the BC for EM field is no more necessary. Based on this understanding at a fundamental level, we can connect the argument about BC's between microscopic and macroscopic response as follows.

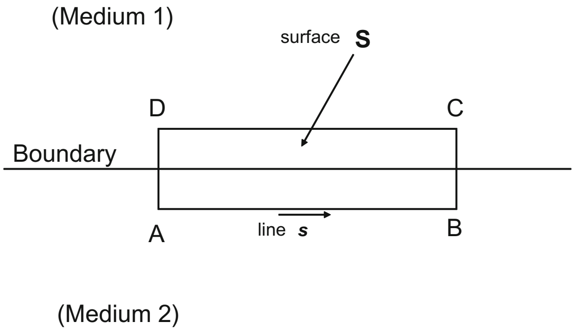


Fig. 3.1 The closed line *s* and the surface *S* enclosed by *s* for the application of Stokes theorem

The introduction of the BC's for EM field becomes necessary, when we replace the position dependent nonlocal susceptibility with the position-independent macroscopic susceptibility via the LWA of the former. Since LWA erases out the position dependence of the susceptibility, the macroscopic description mentioned above would not contain the information about the size, shape, and geometrical configuration. In order to obtain the meaningful solution for the response from such a position independent susceptibility, the BC's for EM field are introduced. This was done in a very smart way. The BC's are provided, not from an independent source, but from the macroscopic M-eqs themselves via Gauss and Stokes laws. Though the arguments are found in many textbooks, we reproduce the relevant ones here, because we use them for the new macroscopic M-eqs, i.e., the LW parts of the microscopic M-eqs.

The Faraday law is known to lead to the continuity of the tangential component of $\boldsymbol{E}$. Integrating the Faraday law (in differential form) over a closed surface S as shown in Fig. 3.1, and using the Stokes theorem to convert $\int \nabla \times \boldsymbol{E}$ into a line integral (along the line $\boldsymbol{s}$ enclosing the surface S), we have

$$\int_s \mathrm{d}\boldsymbol{s} \cdot \boldsymbol{E} = -\frac{1}{c}\frac{1}{\mathrm{d}t}\int_S \mathrm{d}\mathbf{S} \cdot \boldsymbol{B}_\mathrm{n} \quad = \Big[-\frac{1}{\mathrm{d}t}\int_S \mathrm{d}\mathbf{S} \cdot \boldsymbol{B}_\mathrm{n} \Big]_\mathrm{SI} , \tag{3.97}$$

where $\boldsymbol{B}_\mathrm{n}$ is the component of $\boldsymbol{B}$ normal to the surface S.

Let us choose S as the square ABCD across the surface of matter (at a certain point on the surface), and $\boldsymbol{s}$ as its periphery ABCD, as shown in Fig. 3.1. By taking the limit $\bar{AD}$, $\bar{BC} \to 0$, the surface integral on the r.h.s. becomes zero, because B is finite while the integration area becomes vanishing. This means that the line integral on the l.h.s. vanishes, leading to

$$\boldsymbol{E} \cdot \boldsymbol{s}_{AB} + \boldsymbol{E} \cdot \boldsymbol{s}_{CD} = 0 . \tag{3.98}$$

Since $\boldsymbol{s}_{AB} = -\boldsymbol{s}_{CD}$, this proves the continuity of the tangential component of $\boldsymbol{E}$ across the surface.

Gauss law $\nabla \cdot \boldsymbol{D} = 4\pi\rho_\mathrm{t}$ is known to lead to another type of BC. In this case, we take a volume integral of the equation for a rectangular parallelepiped in Fig. 3.2,

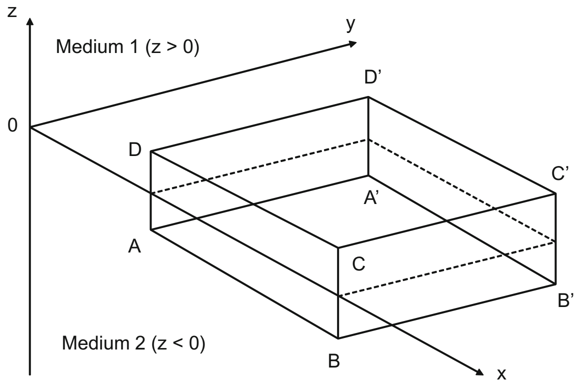


Fig. 3.2 Rectangular parallelepiped to relate the volume and surface integrals for the application of Gauss theorem

which contains the boundary surface at $z = 0$ between the two basal planes. When the hight h of this parallelepiped goes to zero, the Gauss law

$$\int_S \mathrm{d}\mathbf{S} \cdot \boldsymbol{D}_\mathrm{n} = 4\pi\rho_\mathrm{t} S h \quad = \left[\frac{\rho_\mathrm{t}}{\varepsilon_0} S h\right]_\mathrm{SI} \tag{3.99}$$

leads to

$$\boldsymbol{D}_z^{(1)} - \boldsymbol{D}_z^{(2)} = 4\pi\rho_\mathrm{t} h\big|_{h\to 0} = 4\pi\rho_\mathrm{s} \quad = \left[\frac{\rho_\mathrm{t}}{\varepsilon_0}\right]_\mathrm{SI}, \tag{3.100}$$

because the contribution of the side surface of the parallelepiped becomes zero. The superfix 1, 2 denote the two media divided by the surface at $z = 0$. The quantity $\rho_\mathrm{t} h \ \ (h \to 0)$ represent the surface charge density ρ_s in the macroscopic sense. If $\rho_\mathrm{s} = 0$, the r.h.s. becomes zero, which means the continuity of the normal component of $\boldsymbol{D}$ across the surface. If, on the other hand, there is a finite surface charge density ρ_s, the normal components of $\boldsymbol{D}$ have a finite difference across the boundary by the amount $4\pi\rho_\mathrm{s}$.

In Sect. 3.2 we showed that the new macroscopic constitutive equation (for $\boldsymbol{k}, \omega$ Fourier component) $\boldsymbol{I} = \chi_\mathrm{em}\boldsymbol{A} + (c/i\omega)\chi_\mathrm{em}\boldsymbol{E}_\mathrm{extL}$ can be rewritten as $\boldsymbol{I} = -i\omega(\boldsymbol{P}_\mathrm{ET} + \boldsymbol{P}_\mathrm{EL} + \boldsymbol{P}_\mathrm{B}) + i\boldsymbol{k} \times (\boldsymbol{M}_\mathrm{B} + \boldsymbol{M}_\mathrm{ET} + \boldsymbol{M}_\mathrm{EL})$, which allows us to rewrite the microscopic Ampère law into the well-known conventional form $i\boldsymbol{k} \times \boldsymbol{H} = (4\pi/c)\boldsymbol{I}^{(\mathrm{T})} - i(\omega/c)\boldsymbol{D}$ via the definition $\boldsymbol{H} = \boldsymbol{B} - 4\pi(\boldsymbol{M}_\mathrm{B} + \boldsymbol{M}_\mathrm{ET} + \boldsymbol{M}_\mathrm{EL})$ and $\boldsymbol{D} = \boldsymbol{E} + 4\pi(\boldsymbol{P}_\mathrm{ET} + \boldsymbol{P}_\mathrm{EL} + \boldsymbol{P}_\mathrm{B})$. Since this manipulation gives us a formally same set of macroscopic M-eqs as the conventional one, we may also expect the same set of BC's in terms of the newly defined $\{\boldsymbol{E}, \boldsymbol{D}, \boldsymbol{B}, \boldsymbol{H}\}$. However, in view of the fact that the direct form of the selfconsistent response is obtained in terms of $\boldsymbol{A}$ (and $\boldsymbol{I}$), we provide the BC's in a form easily rewritable into those for $\boldsymbol{A}$ and $\boldsymbol{E}_\mathrm{extL}$. For this purpose, it is useful to write the macroscopic M-eqs for T and L components separately in the following form.

The L components arise from the Gauss's law of electric charges, and also from the Ampère law as

$$\nabla \cdot \boldsymbol{E}^{(\mathrm{L})} = 4\pi\rho \quad = \left[\frac{\rho_{\mathrm{t}}}{\varepsilon_0}\right]_{\mathrm{SI}} , \tag{3.101}$$

$$\frac{4\pi}{c}\boldsymbol{I}^{(\mathrm{L})} + \frac{1}{c}\frac{\partial \boldsymbol{E}^{(\mathrm{L})}}{\partial t} = 0 \quad = \left[\boldsymbol{I}^{(\mathrm{L})} + \varepsilon_0 \frac{\partial \boldsymbol{E}^{(\mathrm{L})}}{\partial t}\right]_{\mathrm{SI}} . \tag{3.102}$$

As easily seen by taking the divergence of the second equation, this relation holds identically in the presence of continuity equation and the first equation. Therefore, we need to consider only the first one as a macroscopic equation of L component. The charge density ρ is not the one in the microscopic M-eqs, which determines the quantum mechanical details of the matter eigen states. Rather, it is the charge density to be calculated from the macroscopically averaged current density $\tilde{\boldsymbol{I}}$ via the continuity equation

$$\rho_{\mathrm{m}}(\boldsymbol{r}, \omega) = -\frac{i}{\omega}\nabla \cdot \tilde{\boldsymbol{I}}_{\mathrm{m}}(\boldsymbol{r}, \omega) . \tag{3.103}$$

(For the static case, i.e., $\omega = 0$, ρ_{m} is calculated from the induced electric polarization $\boldsymbol{P}^{(\mathrm{L})}$, explicitly given in Sect. 5.7.2.)

The T components arise from the Ampère law and Faraday law as

$$\nabla \times \boldsymbol{H} = \frac{4\pi}{c}\boldsymbol{I}^{(\mathrm{T})} + \frac{1}{c}\frac{\partial \boldsymbol{D}^{(\mathrm{T})}}{\partial t} \quad = \left[\boldsymbol{I}^{(\mathrm{T})} + \frac{\partial \boldsymbol{D}^{(\mathrm{T})}}{\partial t}\right]_{\mathrm{SI}} ,$$

$$\nabla \times \boldsymbol{E}^{(\mathrm{T})} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t} \quad = \left[-\frac{\partial \boldsymbol{B}}{\partial t}\right]_{\mathrm{SI}} . \tag{3.104}$$

Applying the Gauss theorem to (3.101), we obtain

$$\boldsymbol{n} \cdot (\boldsymbol{E}^{(\mathrm{L1})} - \boldsymbol{E}^{(\mathrm{L2})}) = 4\pi\rho_{\mathrm{s}} \quad = \left[\frac{\rho_{\mathrm{s}}}{\varepsilon_0}\right]_{\mathrm{SI}} , \tag{3.105}$$

where $\boldsymbol{n}$ is a surface normal unit vector at the point to consider BC, and ρ_{s} is the surface density of the total charge, defined in a similar way as in (3.100). Thus the BC for the L-field is given as the difference of the surface normal components by the surface charge density (times 4π). Using the same manipulation as in (5.164) of Sect. 5.7.2, we have

$$\tilde{\boldsymbol{E}}^{(\mathrm{L})} = -\frac{4\pi}{i\omega}\tilde{\boldsymbol{I}}^{(\mathrm{L})} \quad = \left[-\frac{1}{i\omega\varepsilon_0}\tilde{\boldsymbol{I}}^{(\mathrm{L})}\right]_{\mathrm{SI}} . \tag{3.106}$$

Since $\tilde{\boldsymbol{I}}^{(\mathrm{L})}$ is given as

$$\begin{aligned}\tilde{\boldsymbol{I}}^{(\mathrm{L})} &= \sum_{\tau} \chi_{\mathrm{em}}^{(\zeta\tau)} \tilde{\boldsymbol{A}}_{\tau} + \frac{c}{i\omega} \chi_{\mathrm{em}}^{(\zeta\zeta)} \tilde{\boldsymbol{E}}_{\mathrm{extL}} \;, \\ &= \left[\text{same expression without } c\right]_{\mathrm{SI}} \;,\end{aligned} \tag{3.107}$$

the boundary condition (3.105) can be written in terms of $\tilde{\boldsymbol{A}}$ and $\tilde{\boldsymbol{E}}_{\mathrm{extL}}$.

The application of the Stokes theorem to (3.104) leads to the boundary conditions for the surface tangential components as

$$\boldsymbol{n} \times (\boldsymbol{H}^{(1)} - \boldsymbol{H}^{(2)}) = \frac{4\pi}{c} \boldsymbol{I}_{\mathrm{s}}^{(\mathrm{T})} \;\; \left[= \boldsymbol{I}_{\mathrm{s}}^{(\mathrm{T})} \right]_{\mathrm{SI}} \;, \tag{3.108}$$

$$\boldsymbol{n} \times (\boldsymbol{E}^{(1)} - \boldsymbol{E}^{(2)}) = 0 \;, \tag{3.109}$$

where the surface current density (of T character) $\boldsymbol{I}_{\mathrm{s}}^{(\mathrm{T})}$ is defined as

$$\boldsymbol{I}_{\mathrm{s}}^{(\mathrm{T})} = \boldsymbol{I}_{\mathrm{m}}^{(\mathrm{T})} h \;, \quad (h \to 0) \;. \tag{3.110}$$

To rewrite these BC's in terms of $\tilde{\boldsymbol{A}}$ and $\tilde{\boldsymbol{E}}_{\mathrm{extL}}$, we make use of $\boldsymbol{E}^{(\mathrm{T})} = (i\omega/c)\boldsymbol{A}$ and

$$\boldsymbol{H} = \boldsymbol{B} - 4\pi \boldsymbol{M}_{\mathrm{B}} - 4\pi \boldsymbol{M}_{\mathrm{ET}} - 4\pi \boldsymbol{M}_{\mathrm{EL}} \;, \tag{3.111}$$

$$= (1 - 4\pi\,\chi_{\mathrm{mB}}) \nabla \times \boldsymbol{A} - 4\pi\,\chi_{\mathrm{mET}} \frac{i\omega}{c} \boldsymbol{A} - 4\pi\,\chi_{\mathrm{mEL}} \boldsymbol{E}_{\mathrm{extL}} \tag{3.112}$$

$$= \left[\left(\frac{1}{\mu_0} - \chi_{\mathrm{mB}} \right) \nabla \times \boldsymbol{A} - \chi_{\mathrm{mET}} \frac{i\omega}{c} \boldsymbol{A} - \chi_{\mathrm{mEL}} \boldsymbol{E}_{\mathrm{extL}} \right]_{\mathrm{SI}} \;.$$

In this way we can write all the BC's in terms of $\tilde{\boldsymbol{A}}$ and $\tilde{\boldsymbol{E}}_{\mathrm{extL}}$. The general case described by these BC's become simplified when the symmetry of the system does not mix T and L modes ($\chi_{\mathrm{em}}^{(\zeta\tau)} = 0,\; \chi_{\mathrm{mEL}} = 0;\; \tau = \xi, \eta$). Further, if the system is non-chiral, $\chi_{\mathrm{mET}} = 0$, so that the conventional relation $\boldsymbol{H} = (1 - 4\pi\,\chi_{\mathrm{mB}})\boldsymbol{B}$ (in terms of the magnetic susceptibility defined with respect to $\boldsymbol{B}$) is recovered.

3.8 Some Examples of Application

In this section, we show how to use the new macroscopic M-eqs, taking simple examples. The general procedure to calculate the response of a given macroscopic medium is quite similar to the conventional case. The whole space is occupied by materials with different susceptibility tensors and/or vacuum. First we solve the dispersion equation in each region, which generally give plural number of solutions. In order to obtain the response of the system for a given incident EM field, we need to make linear combinations of these solutions in all the regions plus the incident and response field, which satisfy the boundary condition at each interface of the different regions. From the solution of these simultaneous equations, we can express the response field as a function of the incident field.

The difference from the conventional procedure is that the susceptibility of matter and the dispersion equation are different from the usual one, and that the field variable to be used are $\boldsymbol{J}$, $\boldsymbol{A}$ and $\boldsymbol{E}_{\text{extL}}$.

3.8.1 Dispersion Relation in Chiral and Non-chiral Cases

The dispersion relation of the EM waves in the macroscopic medium averaged via LWA is determined by

$$\det\left|\frac{c^2k^2}{\omega^2}\mathbf{1}-\left\{\mathbf{1}+\frac{4\pi c}{\omega^2}\,\chi_{\text{em}}^{(\text{T})}(\boldsymbol{k},\omega)\right\}\right|=0\,, \tag{3.113}$$

$$\left[\text{same expression with } 4\pi c \text{ replaced by } 1/\varepsilon_0\right]_{\text{SI}}$$

of Sect. 2.5 giving the condition for the existence of finite amplitude solution of $\boldsymbol{A}$ and induced current density $\boldsymbol{I}^{(\text{T})}$ in the absence of incident field $\boldsymbol{A}_0$, i.e., the eigen modes of coupled EM wave and T current density. The susceptibility $\chi_{\text{em}}^{(\text{T})}$ is given as a power series expansion with respect to k, i.e., the sum of $O(k^0)$, $O(k^1)$ and $O(k^2)$ … terms, which consist of (E1, E1), {(E1, M1 + E2), (M1 + E2, E1)}, (M1 + E2, M1 + E2), … transitions, respectively, as discussed in Sect. 2.4.

In view of the fact that the new susceptibility is obtained by LWA, we may generally expect that the $O(k^0)$, $O(k^1)$, and $O(k^2)$ terms have decreasing magnitudes in this order. This will be true except for the resonant region of χ_{em1} and χ_{em2}, where a particular term of them can become resonantly large. Unless we concentrate on the resonances of the weaker components, we may generally expect that the principal contribution is made by χ_{em0}.

Let us use the Cartesian coordinate system (ξ,η,ζ), where ζ axis is parallel to $\boldsymbol{k}$. Then, $\chi_{\text{em}}^{(\text{T})}$ is the 2×2 matrix in the (ξ,η) space. If we keep only the leading order term $O(k^0)$, $\chi_{\text{em}}^{(\text{T})}$ is a k-independent 2×2 matrix. Choosing a new coordinate system (ξ',η') in the (ξ,η) space, which diagonalize the 2×2 matrix $\chi_{\text{em}}^{(\text{T})}$, we can decompose the dispersion equation into two components

$$\left(\frac{ck}{\omega}\right)^2=1+\frac{4\pi c}{\omega^2}\,\chi_{\text{em}}^{(\xi',\xi')}(\omega)\,, \tag{3.114}$$

$$\left[\text{same expression with } 4\pi c \text{ replaced by } 1/\varepsilon_0\right]_{\text{SI}}$$

$$\left(\frac{ck}{\omega}\right)^2=1+\frac{4\pi c}{\omega^2}\,\chi_{\text{em}}^{(\eta',\eta')}(\omega)\,. \tag{3.115}$$

$$\left[\text{same expression with } 4\pi c \text{ replaced by } 1/\varepsilon_0\right]_{\text{SI}}$$

The (ξ', η') axes constitute an oblique coordinate system in general. Since the r.h.s. of these equations does not contain k, it is easy to solve them in the form $k = \pm k_\xi(\omega)$ and $k = \pm k_\eta(\omega)$.

In using these dispersion equations, where we neglect magnetization, we should also neglect the magnetization induced current density ($c\nabla \times \boldsymbol{M}$) in considering the boundary conditions, i.e., the boundary condition $\boldsymbol{n} \times (\tilde{\boldsymbol{H}}^{(1)} - \tilde{\boldsymbol{H}}^{(2)}) = (4\pi/c)\ \tilde{\boldsymbol{I}}_{\rm s}$ is simplified as $\boldsymbol{n} \times (\tilde{\boldsymbol{B}}^{(1)} - \tilde{\boldsymbol{B}}^{(2)}) = (4\pi/c)\ \tilde{\boldsymbol{I}}_{\rm s}$. The l.h.s. of this equation can be expressed in terms of $\boldsymbol{E}$'s by using the Faraday law $\boldsymbol{B} = (c/\omega)\boldsymbol{k} \times \boldsymbol{E}$. This allows us to write all the BC's in the form of simultaneous linear equations of $\boldsymbol{E}$'s or $\boldsymbol{A}$'s, which can easily be solved. These processes are applicable to both resonant and non-resonant case of $\chi_{\rm em0}(\omega)$.

As discussed in Sects. 2.4 and 2.5, $O(k^1)$ terms are non-zero in the case of chiral symmetry. Let us consider the case of $\rm T_d$ symmetry for $\boldsymbol{k}$ in one of the cubic axis (z-axis), i.e., $\xi = x, \eta = y, \zeta = z$. Typical transitions contributing to E1 transition are those between an s-like state and (p_x, p_y, p_z)-like states (in the usual notation for a hydrogen-like atom). In the $\rm T_d$ symmetry, an s-like state has a mixed component of xyz-like state, and p_x-like state has a mixing with yz-like state. This means that the transition between the s- and p_x-like states has non-zero matrix element, not only for the operator $\hat{p}_x$, but also for $\hat{z}\hat{p}_y$. Namely, this transition is active both as E1 and (M1 + E2) transitions. Therefore, the current density produced by an y-polarized light propagating along z-axis (for which $\boldsymbol{k} \cdot \boldsymbol{rp} \cdot \boldsymbol{A}$ is $k\hat{z}\hat{p}_y A_y$) can have an x-component as

$$\chi_{\rm em} \cdot \boldsymbol{A} \sim \langle s + xyz|\hat{p}_x|x + yz\rangle\langle x + yz|k\hat{z}\hat{p}_y A_y|s + xyz\rangle \tag{3.116}$$

This corresponds to the element $\chi_{\rm em}^{(xy)}(k_z, \omega)$, i.e., the $O(k^1)$ term, of $\chi_{\rm em}$. Thus, if we consider up to $O(k^1)$ term, the components of $\chi_{\rm em}^{\rm (T)}$ are given as

$$(\chi_{\rm em}^{(xx)}, \chi_{\rm em}^{(xy)}) = (a, ibk)\ , \tag{3.117}$$

$$(\chi_{\rm em}^{(yx)}, \chi_{\rm em}^{(yy)}) = (-ibk, a)\ , \tag{3.118}$$

where a, b are the ω-dependent factors representing the diagonal and non-diagonal components of the susceptibility $\chi_{\rm em}^{\rm (T)}$ with the contributions from the (E1, E1) and (E1, M1 + E2) transitions, respectively.

The dispersion equation reduces to

$$\left(\frac{ck}{\omega}\right)^2 = 1 + \frac{4\pi c}{\omega^2}(a \pm |b|k)\ , \tag{3.119}$$

$$\left[\text{same expression with } 4\pi c \text{ replaced by } 1/\varepsilon_0\right]_{\rm SI}\ ,$$

which can be solved in the form $k = \pm k_+,\ \pm k_-$ where

$$\frac{ck_\pm}{\omega} = \frac{1}{2}\{\pm\tilde{\beta} + (\tilde{\beta}^2 + 4\tilde{\varepsilon})^{1/2}\}\ , \tag{3.120}$$

given in terms of simplified notations $\tilde{\beta} = 4\pi|b|/\omega$ $[= |b|/c\omega\varepsilon_0]_{\mathrm{SI}}$ and $\tilde{\varepsilon} = 1 + (4\pi ca/\omega^2)$ $[= 1 + (a/\omega^2\varepsilon_0)]_{\mathrm{SI}}$. From the form of the matrix χ_{eB}, (3.14), the eigen vectors of these solutions are $\sim A_x \pm iA_y$, i.e., right and left circularly polarized waves. The difference in k_+ and k_- leads to the optical activity of the medium, i.e., the phase velocity is different for the two circularly polarized lights, which is a well-known properties of chiral medium (of cubic symmetry). This effect appears already in non-resonant spectral region, and in a resonant region it will be enhanced through the resonant behavior of $\tilde{\beta}$. As mentioned already in Sect. 3.3, the present result and that from the DBF-eqs show a qualitative difference in the dispersion curve in a resonant region (because of the different order of pole).

3.8.2 Transmission Window in Left-Handed Materials: A Test of New and Conventional Schemes

Let us consider a simple case of non-chiral symmetry. This corresponds to the susceptibility without $O(k^1)$ term, i.e.,

$$\chi_{\mathrm{em}}(\boldsymbol{k}, \omega) = \chi_{\mathrm{em0}}(\omega) + k^2 \chi_{\mathrm{em2}}(\hat{\boldsymbol{k}}, \omega) \ . \tag{3.121}$$

Let us also assume that χ_{em} is a diagonal tensor giving two orthogonal directions of polarization. The dispersion equation split into two independent components for two polarizations, each of which has the form

$$\begin{aligned}\left(\frac{ck}{\omega}\right)^2 &= 1 + \frac{4\pi c}{\omega^2}[\chi_{\mathrm{em0}}(\omega) + k^2 \chi_{\mathrm{em2}}(\omega)] \ , \\ &= \big[\text{same expression with } 4\pi c \text{ replaced by } 1/\varepsilon_0\big]_{\mathrm{SI}}\end{aligned} \tag{3.122}$$

with $\{\chi_{\mathrm{em0}}(\omega), \chi_{\mathrm{em2}}(\omega)\}$ dependent on each polarization. As discussed in Sect. 2.4, χ_{em0} and χ_{em2} represent the E1 and {E2, M1} transitions, respectively, so that $1 + (4\pi c/\omega^2)\chi_{\mathrm{em0}}$ is essentially ε in the conventional M-eqs. If a M1 type resonance of χ_{em2} occurs in the frequency range where $1 + (4\pi c/\omega^2)\chi_{\mathrm{em0}} < 0$, a LHM feature is expected to emerge. In the following, we neglect the E2 component for simplicity.

The dispersion equation can be rearranged in the form

$$\begin{aligned}\left(\frac{ck}{\omega}\right)^2 &= \frac{1 + (4\pi c/\omega^2)\chi_{\mathrm{em0}}(\omega)}{1 - (4\pi/c)\chi_{\mathrm{em2}}(\omega)} \ . \\ &= \big[\text{same expression with } 4\pi c \text{ replaced by } 1/\varepsilon_0\big]_{\mathrm{SI}}\end{aligned} \tag{3.123}$$

In order for the real k solution to exist, the r.h.s. must be positive. For the resonance of χ_{em2} expressed as

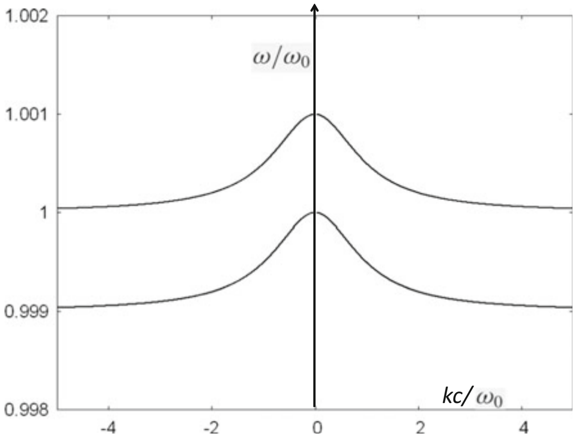


Fig. 3.3 Dispersion curves of non-chiral LHM for (A) the conventional χ_{m} and (B) the new χ_{em}. The frequency and wave number are normalized by ω_0 and ω_0/c, respectively. The parameter values are: $\varepsilon = -1.0$, $\delta = 4\pi\beta/\omega_0 = 4\pi\beta'/\omega_0 = 0.001$

$$\chi_{\mathrm{em2}}(\omega) = \frac{c\beta}{\omega_0 - \omega - i0^+}, \quad (\beta > 0) \tag{3.124}$$

in the frequency range where $1 + (4\pi c/\omega^2)\chi_{\mathrm{em0}} < 0$, the real k solution appears for ω satisfying $\omega_0 - (4\pi\beta) < \omega < \omega_0$. This is in contrast with the situation in the conventional scheme based on $(ck/\omega)^2 = \varepsilon\mu$, which gives the real k solution in the frequency range higher than ω_0. Assuming the magnetic susceptibility χ_{m} in the form $\chi_{\mathrm{m}}(\omega) = \beta'/(\omega_0 - \omega - i0^+)$, one can rewrite the condition $\mu = 1 + 4\pi\chi_{\mathrm{m}} < 0$ for the appearance of LHM behavior as $\omega_0 < \omega < \omega_0 + 4\pi\beta'$. Figure 3.3 shows the two dispersion curves mentioned above. Though their forms are very similar, their positions with respect to the resonance at ω_0 are just opposite. This is a very fundamental problem, which requires experimental tests or some theoretical explanation.

As a simple experimental test, it would be appropriate to measure the spectrum of transmission window due to the propagating mode of Fig. 3.3 for normal incidence of light on a slab. Measuring the resonant frequency (ω_0) of the magnetic susceptibility of the same sample independently, we can compare the relative positions of the transmission window and the pole ω_0. This will be a simple, but definitive check from the experimental side.

The spectrum of the transmission window according to the new scheme can be calculated as follows. The slab occupies the region $0 \le z \le d$ in vacuum, and the incident field is polarized along x-axis. For χ_{em} with a diagonal form with respect to (x, y) axes, all the $\boldsymbol{E}$ fields are x-polarized. The field amplitudes of incident (E_{i}), reflected (E_{r}), transmitted (E_{t}), and the two waves in the medium (E_1, E_2) are defined as in Fig. 3.4. The reference point (z-coordinate) of each field is marked by a solid dot in the figure. The arrows for E_1 and E_2 indicate the direction of the group velocity (or that of the decay of their amplitudes).

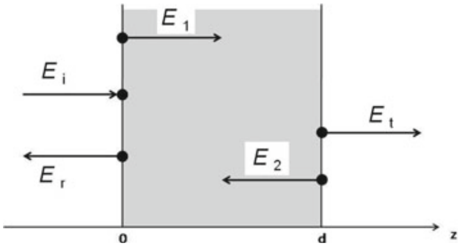


Fig. 3.4 Configuration of relevant wave components. The reference point (z-coordinate) of each wave is marked by a solid dot in the figure

The solution of the dispersion equation is given as

$$k_\ell = \pm\frac{\omega}{c}\sqrt{\frac{1+(4\pi c/\omega^2)\chi_{\rm em0}}{1-(4\pi/c)\chi_{\rm em2}}}\ , \tag{3.125}$$

$$= \big[\text{same expression with } 4\pi \text{ replaced by } 1/c\varepsilon_0\big]_{\rm SI}$$

where $\ell = 1$ and $\ell = 2$ corresponds to the roots with the positive and negative imaginary parts, respectively. As discussed in Sect. 4.1.1, the solution with positive (negative) real part has negative (positive) group velocity, and negative (positive) imaginary part, which is the peculiar point of LHM.

As the boundary conditions, we require the continuity of $\boldsymbol{E}$ and $\boldsymbol{H}$ across $z = 0$ and $z = d$, as discussed in Sect. 3.7. The continuity of $\boldsymbol{H}(= \boldsymbol{B} - 4\pi\boldsymbol{M})$ can be rewritten as that of $[1 - (4\pi/c)\chi_{\rm em2}]B_y$, which is further rewritten as $(ck/\omega)[1 - (4\pi/c)\chi_{\rm em2}]E_x$ by using Faraday law $\boldsymbol{k} \times \boldsymbol{E} = (\omega/c)\boldsymbol{B}$. The boundary conditions at $z = 0$ are

$$E_{\rm i} + E_{\rm r} = E_1 + f_2 E_2 \tag{3.126}$$

$$E_{\rm i} - E_{\rm r} = n_1 E_1 + n_2 f_2 E_2 \tag{3.127}$$

where

$$f_2 = \exp(-ik_2 d)\ , \quad n_\ell = \frac{ck_\ell}{\omega}\left(1 - \frac{4\pi}{c}\chi_{\rm em2}\right)\ , \quad (\ell = 1, 2) \tag{3.128}$$

Similarly the boundary conditions at $z = d$ are

$$E_{\rm t} = f_1 E_1 + E_2 \tag{3.129}$$

$$E_{\rm t} = n_1 f_1 E_1 + n_2 E_2 \tag{3.130}$$

where $f_1 = \exp(ik_1 d)$. The factors f_1 and f_2 are defined in such a way that they go to zero for $d \to \infty$.

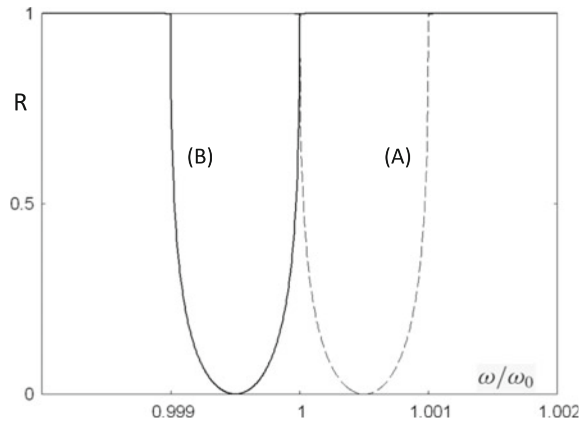


Fig. 3.5 Reflectivity spectrum with a transmission window due to the left-handed mode of Fig. 3.3 calculated by (A) the conventional χ_{m}, and (B) the new χ_{em}

From the four equations of boundary conditions, we obtain the reflection amplitude of the form $E_{\mathrm{r}}/E_{\mathrm{i}} = a(-1 + f_1 f_2)/(b + cf_1 f_2)$, where

$$a = (n_1 - 1)(n_2 - 1), \quad b = (n_1 + 1)(n_2 - 1), \quad c = (n_1 - 1)(n_2 + 1) \,. \tag{3.131}$$

In the limit of $d \to \infty$, $E_{\mathrm{r}}/E_{\mathrm{i}} = -a/b = -(n_1 - 1)/(n_1 + 1)$, which is the reflection amplitude for a semi-infinite medium. It should be noted that, though $\mathrm{Re}[k_1]$ is negative, $\mathrm{Re}[n_1]$ is positive in the frequency region of the dispersion branch, so that $|E_{\mathrm{r}}/E_{\mathrm{i}}|^2 = |(n_1 - 1)/(n_1 + 1)|^2 \leq 1$. Namely, it is guaranteed that the reflectivity never exceeds unity. It is also worth noting that, for $d = \infty$, the incident wave is connected with, not the k_2, but the k_1 branch which has negative real part and positive imaginary part, i.e., positive group velocity, corroborating the LHM nature of this system. Figure 3.5 shows the transmission window due to this propagating mode. For comparison, the result of the conventional method is also given. Corresponding to the curves in Fig. 3.3, the transmission window opens in the lower (higher) frequency region of ω_0 by the new (conventional) method. The calculation by the conventional method is very similar to the one given above, except for the replacement of n_1 and n_2 with

$$n_\ell \to n'_\ell = \frac{1}{1 + 4\pi\chi_{\mathrm{m}}} \frac{ck_\ell}{\omega} \,, \quad (\ell = 1, 2) \tag{3.132}$$

The experiment proposed above would be a crucial test of the two definitions, $\boldsymbol{M} = \chi_{\mathrm{m}}\boldsymbol{H}$ or $\boldsymbol{M} = \chi_{\mathrm{B}}\boldsymbol{B}$ with the interpretation of the poles of the susceptibility as magnetic excitation energies. The arguments about the definition of matter Hamiltonian (Sect. 2.2) and the rewriting of χ_{cd} (Sects. 2.4, 3.1) obviously prefer the latter definition. Since, however, the use of the former definition is still the main trend today, and since a correct theory should have an experimental support, it is desirable for the proposed experiment to be performed.

References

1. Drude, P.: Lehrbuch der Optik, Leipzig, S. Hirzel (1912); Born, M.: Optik, J. Springer, Heidelberg (1933); Fedorov, F.I.: Opt. Spectrosc. **6** 49 (1959); ibid. **6** 237 (1959)
2. Band, Y.B.: Light and Matter, p. 142. Wiley, New Jersey (2006)
3. Cho, K.: Electromagnetic Metamaterials: A Modern Problem of Macroscopic Electromagnetic Fiels. Sakoda, K. (ed.), Chap. 12. Springer, Berlin
4. Luan, P.-G., Wang, Y.-T., Zhang, S., Zhang, X.: Opt. Lett. **36**, 675 (2011)
5. Tomita, S., Sawada, K., Porokhnyuk, A., Ueda, T.: Phys. Rev. Lett. **113**, 235501 (2014)
6. Agranovich, V.M., Ginzburg, V.L.: Crystal Optics with Spatial Dispersion, and Excitons, Sect. 6. Springer, Berlin (1984)
7. Il'inskii, Yu.A., Keldysh, L.V.: Electromagnetic Response of Material Media. Plenum Press, New York (1994)
8. Landau, L.D., Lifshitz, E.M.: Electromagnetics of Continuous Media. Pergamon Press, Oxford (1960)
9. Pekar, S.I.: Zh. Eksp. Teor. Fiz. **33**, 1022 (1957) [Sov. Phys. JETP **6**, 785 (1957)]
10. Hopfield, J.J., Thomas, D.G.: Phys. Rev. **132**, 563 (1963)
11. Birman, J.L.: Excitons. In: Rashba, E.I., Sturge, M.D. (eds.) p. 72, North Holland (1982); Halevi, P.: Spatial Dispersion in Solids and Plasmas, p. 339, In Halevi, P. Elsevier, Amsterdam (1992)
12. Ikawa, T., Cho, K.: Phys. Rev. B**66**, 085338 (2002); Hübner, M., Prineas, J.P., Ell, C., Brick, P., Lee, E.S., Khitrova, G., Gibbs, H.M., Koch, S.W.: Phys. Rev. Lett. **83** 2841 (1999)
13. Agranovich, V.M., Shen, Y.R., Baughman, R.H., Zakhidov, A.A.: Phys. Rev. B **69**, 165112 (2004)
14. Chipouline, A., Simovski, K., Tretyakov, S.: Metamaterials **6**, 77–120 (2012)
15. Cho, K.: arXiv:1303.0355; Proceedings of the International Congress on Advanced Electromagnetic Materials in Microwave and Optics, https://doi.org/10.1109/Metamaterials.2013.6809094, Bordeaux (2013)
16. Nelson, D.F.: Electric, Optic, and Acoustic Interactions in Dielectrics. Wiley, New York (1979)
17. Cho, K.: Optical Response of Nanostructures: Microscopic Nonlocal Theory. Springer, Heidelberg (2003)
18. Cho, k.: J. Phys. Soc. Jpn. **68**, 683 (1999)

Chapter 4
Further Considerations

Abstract More detailed discussions are given about the problems of metamaterials from both formal and specific points of view. The latter is the dispersion curves in left-handed chiral medium in resonant region, which critically distinguishes DBF equations from the present (ChC) equations. As a peculiar example of problems with invalid LWA, resonant Bragg scattering is discussed, which enables us to distinguish left- from right-handed quartz crystals. How to treat the interaction energy of induced charge densities is shown to lead to different standpoints on choosing pure external field or full Maxwell field as the source of induced current density. This may cause a problem in homogenizing metamaterials as to the validity of LWA. A brief discussion is given about the extension to nonlinear processes.

4.1 Consequences to the Metamaterials Studies

4.1.1 Definition of Left-Handed Materials (LHM)

For the conventional definition of LHM, "$\varepsilon < 0,\ \mu < 0$", one needs two independent susceptibilities. If we describe the same physical situation in terms of a single susceptibility, we obviously need a different definition. The common language for this purpose is, not the susceptibility, but dispersion curve, as explained below. We give a conventional description of LHM in the first half of this subsection, and in the latter half, we rephrase the same (but inequivalent) physics by the new single susceptibility scheme.

The first proposal of LHM by Veselago was made as a medium with $\varepsilon < 0,\ \mu < 0$ [1]. The dispersion equation $(ck/\omega)^2 = \varepsilon\mu$ for a plane wave $\sim \exp(i\boldsymbol{k} \cdot \boldsymbol{r} - i\omega t)$ in the conventional macroscopic M-eqs has real solutions $k = \pm(\omega/c)\sqrt{\varepsilon\mu}$ in this case. Both ε and μ are functions of ω, and can take positive and negative values. If $\varepsilon\mu < 0$ in a frequency region, the medium is totally reflecting, because the dispersion equation allows only evanescent waves. In a frequency region where $\varepsilon < 0,\ \mu < 0$, the medium becomes transmissive due to the existence of propagating modes with real wave number.

K. Cho, *Reconstruction of Macroscopic Maxwell Equations*,
Springer Tracts in Modern Physics 237,
https://doi.org/10.1007/978-3-662-58424-8_4

The ω dependence of ε and μ is generally written as sums of single poles according to the lowest order time-dependent perturbation calculation of quantum mechanics. Except for the very neighborhood of the poles (corresponding to the excitation energies of matter), both $\varepsilon(\omega)$ and $\mu(\omega)$ are increasing functions of ω between neighboring poles. (This is due to the positiveness of the residue of each pole, which is generally the case for materials in equilibrium.)

If we increase ω starting from a certain frequency where $\varepsilon < 0, \ \mu < 0$, both ε and μ increase toward zero. This means that the product $\varepsilon\mu$ is a positive, decreasing function of ω, eventually crosses zero and becomes negative. The propagating modes are allowed only while the product is positive. If we combine this fact with the dispersion relation $k = \pm(\omega/c)\sqrt{\varepsilon\mu}$, the frequency ω_1 for which $\varepsilon\mu = 0$ corresponds to $k = 0$ is a local maximum of the dispersion curve, i.e., as ω decreases from ω_1, the corresponding $|k|$ increases. Altogether, the dispersion curve is convex toward higher ω, and the lower bound of this branch corresponds to the closest resonance frequency of ε or μ on the lower ω side (see an example in Fig. 3.3a, which shows a LHM branch for a resonance of μ in the broad range of negative ε). For the positive k side, the group velocity $v_g = \mathrm{d}\omega/\mathrm{d}k$ is negative. This is a typical example of the dispersion curves representing the LHM character.

If we send an incident light in the frequency range of this branch, we can excite this mode. To determine the amplitude of this mode, we need to apply boundary conditions to the relevant waves in- and outside the boundary. For normal incidence of light on a semi-infinite slab, we have a plane wave with one of the wave vectors $k = \pm(\omega/c)\sqrt{\varepsilon\mu}$. For normal (right handed) system, we know that the choice of positive sign leads to correct answer. What is the underlying reason for it and what is the correct choice in the case of LHM?

The right answer is obtained from the consideration of the spatial and temporal decay of this wave. The (non-radiative) decay occurs through the excitation of phonons and other electronic transitions. Since the heat bath system consists of infinitely many degrees of freedom, the direction of energy flow must be from the EM field to the medium (heat bath). Therefore, the amplitude of the induced (matter - EM field coupled) mode should be decreasing from the incident surface to the interior. The change in the phase (and amplitude) of the wave after a distance d is $\exp(ikd)$, so that we need $\mathrm{Im}[k] > 0$ in order for this change to be a spatial decay, i.e., a decreasing function of d. Therefore the correct choice is the k with positive imaginary part when we allow damping effect.

The damping effect in the time region is expressed by considering a positive imaginary part to ω. This is understood by a simple example of damped oscillator to calculate polarizability (Lorentz oscillator model). Suppose we have an electric oscillators with mass m_0, charge q_0, resonant frequency ω_0 exposed in an electric field $E_0(t)$. The Newton equation of motion of this oscillator is

$$m_0 \frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = q_0 E_0(t) - Kx - m_0 \gamma \frac{\mathrm{d}x}{\mathrm{d}t} \tag{4.1}$$

where $-Kx$ is the restoring force (Hook's law, $K = m_0\omega_0^2$), and the last term on the r.h.s. is the damping force proportional to velocity ($\gamma > 0$). The solution of this equation is obtained by Fourier expansion, which leads, for frequency ω, to the induced polarization as

$$\tilde{P}(\omega) = \frac{N_0 q_0^2/m_0}{\omega_0^2 - \omega^2 - i\gamma\omega}\,\tilde{E}_0(\omega) \tag{4.2}$$

where N_0 is the number density of the oscillators. If $E_0(t)$ is a delta function at $t = 0$, i.e., $E_0(t) = E_0 \int \mathrm{d}\omega \exp(-i\omega t)$, the induced polarization as a function of t

$$P(t) = \frac{1}{2\pi}\int \mathrm{d}\omega \tilde{P}(\omega)\exp(-i\omega t) \tag{4.3}$$

is evaluated by the residue at the pole of $\tilde{P}(\omega)$, which leads to the time dependence $\sim\exp(-i\omega_0 - \gamma t/2)$. This shows that $\gamma > 0$ leads to the damping of $P(t)$ in the positive t direction, which of course leads to the same damping behavior of the induced field by this $P(t)$.

The general expressions of induced current density, (2.38) and (2.107), have also the similar pole structure with the imaginary part 0^+, so that its temporal response to $\delta(t)$ like incident field is the (very slow) time decay with $0^+/2$. Thus, the analytic continuation of real (k, ω) dispersion to the complex ω with positive imaginary part gives a correct behavior of temporal damping in general.

For the calculation of the spectral response (for real ω), we need to choose the appropriate wave number(s) satisfying the dispersion equation, and to set up the boundary conditions on each relevant surface/interface. If we consider a semi-infinite slab and a normally incident light propagating in the positive z direction, for simplicity, we need to consider which of the two solutions $k = \pm k(\omega)$ should be chosen as the wave inside the slab induced by the incident light. From the argument given above, we should choose the branch with positive Im$[k]$. This corresponds to the branch with positive group velocity v_g [2]. The reason is as follows. From the relation $\mathrm{d}k/\mathrm{d}\omega = 1/v_g$, or $\Delta\omega = v_g\Delta k$, where $\Delta\omega$ and Δk are the small increments from the real (ω, k) solution, the relative sign of $\Delta\omega$ and Δk is the same as the sign of v_g. For the correct temporal decay, it is required that $\Delta\omega$ represents a positive imaginary part. Then, $v_g\Delta k$ must also give a positive imaginary part. In order for both $\Delta\omega$ and Δk to give positive imaginary part, $v_g = \Delta\omega/\Delta k$ must be positive.

The dispersion curves $k = \pm(\omega/c)\sqrt{\varepsilon\mu}$ consist of positive and negative k branches. As seen from Fig. 3.3, the positive v_g occurs on the negative k branch. If the medium of this LHM behavior occupies the semi-infinite space $z = 0 \sim \infty$, the right mode to be connected to the incident field in the positive direction ($\exp[ikz]$) is this mode on the negative k branch with positive v_g. This choice gives us the occurrence of a transmission window in the total reflection range, which is a general feature of LHM. Since the convex dispersion curve toward higher ω and the occurrence of transmission window arise also in the new macroscopic scheme without using ε and μ, we can use this feature as a new definition of LHM.

In the non-chiral case, the maximum of the dispersion curve occurs at $k = 0$, and the dispersion curve is symmetric for the ± directions of k. When such a dispersion curve is degenerate for two polarizations, the introduction of chiral symmetry leads to the k-linear splitting. This gives rise to the lifting of the degeneracy, and in the neighborhood of $k = 0$ two branches cross linearly with positive and negative group velocities. Nevertheless, the convex character toward higher ω is kept for these dispersion curves. An example of this kind will be discussed in Sect. 4.1.5, by modifying the model of Fig. 3.3.

Thus the alternative definition of LHM without depending on the use of ε and μ, or χ_e and χ_m, would be "a medium with dispersion curves of convex form toward higher ω". Though this feature is common to both conventional and the new macroscopic schemes of M-eqs, the relative position of the resonance frequency and the dispersion curve is different in these two schemes, as described in Sect. 3.8.2, which can be a simple test to decide the consistency of the schemes.

4.1.2 Use of (ε, μ) and Homogenization

Today's popularity of metamaterials study seems to be driven by the idea of free designing of ε and μ beyond the hitherto accepted range of these parameter values. Typical examples are the case of LHM [3], where one needs an exotic situation "$\varepsilon < 0,\ \mu < 0$", and the case of cloaking [4], where the spatially varying values of ε according to the form of a body makes the body invisible.

Since these are all man-made substances consisting of an array of the unit structures, each one of which can be made smaller than the wavelength of EM field. In order to make theoretical analysis simpler, the response of such a system is replaced by a uniformly homogenized material obtained from the original one. Usually, the homogenization (or LWA) is justified by claiming the smallness of the unit structure in comparison with the wavelength of EM field. As explained in Sect. 3.6, this justification is not always correct, since the interaction among the induced charge densities on the unit structures may produce excitations with long spatial coherence. If we are interested in the resonant behavior of such an artificial structure, we need to take account of this possibility, because it may well invalidate the homogenization, unless the non-radiative scattering mechanism is strong enough.

In the narrow definition of metamaterials, it is said or is taken for granted that they are uniform materials obtained by homogenization. If one takes this definition, a rather large group of material systems will be omitted from "metamaterials, because, among the possible man-made substances, the condition for the homogenization will not be generally satisfied. There is, on the other hand, a broader definition of metamaterials. In fact, the metamaterials made of circuit elements (L, C, R) are interested in their dispersion behavior [5], i.e., the eigen frequencies of an extended circuit array depending on the phase difference between neighboring circuit elements (which is equivalent to the wave number). For this group of researchers, the existence of large spatial coherence in metamaterials is an important subject for metamaterials, which for example may make an antenna emitting microswaves in a wide angle by changing frequency [6]. In view of this type of activity, and also of the fact that “the metamaterials with non-homogenized components” is also theoretically tractable, as will be mentioned in Sects. 4.1.4 and 4.3, it is not necessary to include homogenization as the necessary condition for metamaterials.

4.1.3 “Microscopic”, “Semi-macroscopic” and “Electric Circuit” Approaches

Within the semiclassical framework of EM response theory, macroscopic M-eqs are derived from the microscopic M-eqs by assuming the validity of LWA. This requires a comparison of the coherent extension of induced current densities and a relevant wavelength of EM field. Since the former depends on each quantum transition of matter, it is not rare that the condition for LWA is not satisfied. If such a transition is off-resonant with the incident frequency, one may rather safely neglect its microscopic contribution. If, on the contrary, it is resonant, LWA is certainly a bad approximation to handle the contribution of such a transition. There will be rather many cases of this kind in both natural and artificial materials. From the viewpoint of microscopic nonlocal theory, it is usual to treat some group of resonant transitions microscopically and the rest as a background medium with uniform dielectric constant [7]. This is a mixed use of micro- and macroscopic responses. Both of these examples show the existence of the matter systems to be theoretically treated by a mixed use of micro- and macroscopic responses.

Since the mixed use is rather undeveloped from the side of macroscopic response, we describe two examples in this book. One is the nonlocal response of metamaterials in the next subsection, which may be a new concept in a system consisting of semi-macroscopic unit structures such as SRR. The other is the resonant Bragg scattering due to inner core excitations in Sect. 4.3, which turns out to lead to a general expression of refraction including chiral systems. It will theoretically be more reasonable to leave a room for the definition of metamaterials so as to allow a partially microscopic character.

The macroscopic response is described in terms of susceptibility χ_{em} in this book. In the low frequency regime of EM response, there is a well established way of describing the response in terms of electric circuit elements, such as capacity C, inductance L, and resistivity R. In principle, these constants of circuit elements can be calculated from the knowledge of the material constants of a bulk matter. The resistance is a bulk material constant (reciprocal of conductivity) times the length divided by the cross section, of a wire. The capacity C is the coefficients of the electric potential as functions of accumulated charge on a sample, and the inductance L the coefficient of magnetic potential as a flowing current through the element. For a simple geometry, this type of calculation will not be too difficult. For a realistic structure, however, this would require a large scale numerical calculations.

4.1.4 Nonlocal Response of Metamaterials

Taking a broader definition of metamaterials, we consider arrays of unit structures, which are electronically separated but are interacting with one another through the induced charge densities. This interaction may cause a long coherence length in the excitations of matter, which does not justify the use of LWA or homogenization procedure.

The unit structure can be anything, a quantum dot, a fine particle, or a man-made piece of matter with particular shape and size. Each one of them will have its own excited levels contributing to some resonant response to EM field. If the interaction among unit structures is not important, one could treat the response of the whole system in terms of homogenized macroscopic susceptibilities. It is essentially the response of a single unit structure multiplied by the number density. If the interaction is strong, however, some of the excited states may have long spatial coherence, which should be treated microscopically.

Because of the large microscopic degrees of freedom of quantum dots, fine particles, and SRR's, a complete description of the quantum mechanical motions of matter will be quite difficult. But a fair description will be possible by concentrating on special modes of excitation, which have strong interaction with EM field. In the case of a quantum dot, confined excitons are such modes. In the case of a SRR specified by a conductivity, shape and size, numerical simulations (such as FDTD or else) give us resonant frequencies and the corresponding current densities [8], from which we make use of the knowledge of the eigen energy of excited states (with damping), $E_n\ (-i\Gamma_n)$, and the corresponding current densities, $\boldsymbol{J}_n(\boldsymbol{r})$, where $n = 1, 2, 3, \ldots$ represents the mode number of the excited levels. Putting the unit structures of this kind in a regular lattice, we look for the EM response of this system. The induced current density of this array is written as a linear combination of $\boldsymbol{J}_n(\boldsymbol{r})$, and the selfconsistent equations for the expansion coefficients $\{F_{n0}\}$ can be built in terms of the eigen energies $E_n - i\Gamma_n$ and the matrix elements of Coulomb interaction and radiative correction ((2.86) and (2.87) of [7]).

The interaction between the induced current densities on a different unit cells takes place via the T and L components of EM field. An induced current density, generally consisting of T and L vector fields, produces T and L EM fields around it according to the microscopic M-eqs, and they interact with the T and L components, respectively, of the current density on the other cell. This interaction can be calculated in the following way.

The interaction between two current densities $\boldsymbol{I}_1(\boldsymbol{r}, \omega)$ and $\boldsymbol{I}_2(\boldsymbol{r}, \omega)$ is mediated by both T and L components of EM field. Using the EM Green function described in detail in Sect. 5.7.1, we can write the electric field produced by $\boldsymbol{I}_1$ as

$$\boldsymbol{E}(\boldsymbol{r}, \omega) = \frac{i\omega}{c^2} \int \mathrm{d}\boldsymbol{r}'\, \tilde{\boldsymbol{G}}_q(\boldsymbol{r}, \boldsymbol{r}', \omega) \cdot \boldsymbol{I}_1(\boldsymbol{r}', \omega)\ , \tag{4.4}$$

$$= \left[\text{same expression with } 1/c^2 \text{ replaced by } \mu_0/4\pi\right]_{\mathrm{SI}}\ .$$

This Green function is the sum of T and L components, $\tilde{\boldsymbol{G}}_q^{(\mathrm{T})}$ and $\tilde{\boldsymbol{G}}_q^{(\mathrm{L})}$ defined in Sect. 5.7.1, which produce T and L fields, respectively, by taking a convolution with $\boldsymbol{I}_1(\boldsymbol{r}, \omega)$.

Let us divide the induced electric field and current density into T and L components as

$$\boldsymbol{E}(\boldsymbol{r}, \omega) = \boldsymbol{E}^{(\mathrm{T})}(\boldsymbol{r}, \omega) + \boldsymbol{E}^{(\mathrm{L})}(\boldsymbol{r}, \omega)\ , \quad \boldsymbol{I}(\boldsymbol{r}, \omega) = \boldsymbol{I}^{(\mathrm{T})}(\boldsymbol{r}, \omega) + \boldsymbol{I}^{(\mathrm{L})}(\boldsymbol{r}, \omega)\ . \tag{4.5}$$

The T component of $\boldsymbol{E}$ and the L component of $\boldsymbol{I}$ can be rewritten as

$$\boldsymbol{E}^{(\mathrm{T})}(\boldsymbol{r}, \omega) = iq\boldsymbol{A}(\boldsymbol{r}, \omega)\big[= i\omega\boldsymbol{A}\big]_{\mathrm{SI}} \quad \boldsymbol{I}^{(\mathrm{L})}(\boldsymbol{r}, \omega) = -i\omega\boldsymbol{P}^{(\mathrm{L})}(\boldsymbol{r}, \omega)\ , \tag{4.6}$$

because of the Coulomb gauge and the T character of magnetization induced current density. The (time averaged) interaction energy between the EM field ($\boldsymbol{E}^{(\mathrm{T})} + \boldsymbol{E}^{(\mathrm{L})}$) due to $\boldsymbol{I}_1(\boldsymbol{r}, t)$ and a current density $\boldsymbol{I}_2(\boldsymbol{r}, t)$ is written as

$$\begin{aligned} E_{\mathrm{int}} &= -\frac{2\pi}{c}\int_{-\infty}^{\infty} \mathrm{d}\omega \int \mathrm{d}\boldsymbol{r} \\ &\qquad \big[\boldsymbol{A}(\boldsymbol{r}, -\omega)\cdot\boldsymbol{I}_2^{(\mathrm{T})}(\boldsymbol{r}, \omega) + c\boldsymbol{E}^{(\mathrm{L})}(\boldsymbol{r}, -\omega)\cdot\boldsymbol{P}_2^{(\mathrm{L})}(\boldsymbol{r}, \omega)\big]\ , \\ &= \big[\text{same expression without } 1/c\big]_{\mathrm{SI}}\ , \end{aligned} \tag{4.7}$$

which can be rewritten in terms of $\boldsymbol{E}$ and $\boldsymbol{I}_2$ as

$$E_{\mathrm{int}} = \frac{-2\pi i}{\omega}\int_{-\infty}^{\infty} \mathrm{d}\omega \int \mathrm{d}\boldsymbol{r}\,\big[\boldsymbol{E}^{(\mathrm{T})}(\boldsymbol{r}, -\omega)\cdot\boldsymbol{I}_2^{(\mathrm{T})}(\boldsymbol{r}, \omega) + \boldsymbol{E}^{(\mathrm{L})}(\boldsymbol{r}, -\omega)\cdot\boldsymbol{I}_2^{(\mathrm{L})}(\boldsymbol{r}, \omega)\big]. \tag{4.8}$$

Thus, by substituting the second term of (4.4), the interaction energies mediated by T and L fields are expressed as

$$E_{\mathrm{int}}^{(\mathrm{Y})} = -\frac{2\pi q^2}{\omega^2}\int_{-\infty}^{\infty} \mathrm{d}\omega \int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r}'\ \boldsymbol{I}_2^{(\mathrm{Y})}(\boldsymbol{r}, \omega)\cdot\tilde{\boldsymbol{G}}_q^{(\mathrm{Y})}(\boldsymbol{r} - \boldsymbol{r}', \omega)\cdot\boldsymbol{I}_1^{(\mathrm{Y})}(\boldsymbol{r}', -\omega)\ , \tag{4.9}$$

where Y = T or L. This expression is valid for arbitrary $\boldsymbol{I}_1$ and $\boldsymbol{I}_2$ including $\boldsymbol{I}_1 = \boldsymbol{I}_2$. The term $E_{\mathrm{int}}^{(\mathrm{L})}$ is the Coulomb interaction energy between the induced charge densities accompanying $\boldsymbol{I}_1$ and $\boldsymbol{I}_2$ as

$$E_{\mathrm{int}}^{(\mathrm{L})} = 2\pi\int_{-\infty}^{\infty} \mathrm{d}\omega \int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r}'\ \frac{\rho_2(\boldsymbol{r}, \omega)\ \rho_1(\boldsymbol{r}, -\omega)}{|\boldsymbol{r} - \boldsymbol{r}'|}\ , \tag{4.10}$$

which can be easily seen by rewriting the current density via continuity equation $\nabla\cdot\boldsymbol{I} + \partial\rho/\partial t = 0$ and the explicit form of $\tilde{\boldsymbol{G}}_q^{(\mathrm{L})}$ in Sect. 5.7.1. On the other hand, $E_{\mathrm{int}}^{(\mathrm{T})}$ is the radiative correction, and plays an important role in the equations to determine the expansion coefficients $F_{\mu\nu}(\omega)$, (2.48), of current density. This interaction energy is complex even for the diagonal element, i.e., for $\boldsymbol{I}_1 = \boldsymbol{I}_2$, giving the shift and width to the resonant energies of matter [7].

The simultaneous linear equations of $\{F_{\mu 0}, F_{0\mu}\}$ mentioned at the end of Sect. 2.2.3 has a clear physical meaning. If we rewrite the set of equations in terms of a new set of variables $\{X_{\mu 0} = g_\mu(\omega) F_{\mu 0}\}$ and $\{X_{0\mu} = h_\mu(\omega) F_{0\mu}\}$, the coefficient matrix is the sum of material excitation energies (plus or minus ω) and the radiative correction. Since the basis of the matrix is chosen as the eigen functions of matter excitation, the material excitation energies are diagonalized, while the radiative correction contains both diagonal and off-diagonal elements. The essential point is that the solution of the coupled linear equations of $\{X_{\mu 0}\}$ and $\{X_{0\mu}\}$ have resonances at the matter excitation energies with radiative shifts and widths. If we keep only the resonant part for simplicity, the set of the linear equations to determine $\{X_{\mu 0}\}$ is

$$F^{(0)}_{\mu 0} = \sum_\nu \{(E_{\nu 0} - \hbar\omega)\delta_{\mu\nu} + \mathscr{A}_{0\mu,\nu 0}\}\ X_{\nu 0} \tag{4.11}$$

where $\mathscr{A}$ is the matrix element of the radiative correction, and the current density (resonant part) is

$$\boldsymbol{I}(\boldsymbol{r}, \omega) = \frac{1}{c} \sum_\nu X_{\nu 0}\, \boldsymbol{I}_{0\nu}(\boldsymbol{r}) \tag{4.12}$$

This scheme can be combined with the problem of regular arrays of unit structures (metamaterials), in the following way. On each unit cell, we have a set of local current densities. Their eigen frequencies (with damping) and the corresponding spatial structures can be prepared by a numerical calculation for a single unit structure, i.e., a single SRR for example. The effect of $E^{(\mathrm{T})}_{\mathrm{int}}$ and $E^{(\mathrm{L})}_{\mathrm{int}}$ for a single cell will be contained, but the inter-cell components are not included in a single cell calculation. The linear equations to determine $\{X_{\mu 0}\}$ is given in a matrix form as

$$F^{(0)}_{\mu 0} = \sum_\nu [\ E^{(\mathrm{T})}_{\mathrm{int}} +\ E^{(\mathrm{L})}_{\mathrm{int}} + (E_n - i\Gamma_n)\mathbf{1} - \hbar\omega\mathbf{1}]_{\mu\nu} X_{\nu 0}\ . \tag{4.13}$$

The suffices μ, ν of the matrix elements contain both lattice site index and the sub level index (n) of a unit structure. The matrix elements of $E^{(\mathrm{T})}_{\mathrm{int}}$ and $E^{(\mathrm{L})}_{\mathrm{int}}$ are calculated from (4.9), where $\boldsymbol{I}_1, \boldsymbol{I}_2$ represent the current densities in each unit structure, distinguished by the cell number ℓ and the internal quantum number n of each unit structure. Therefore, the input information is $\{E_n,\ \Gamma_n,\ \boldsymbol{I}_{\ell,n}(\boldsymbol{r})\}$. (Since we assume a same unit structure in each cell, $\{E_n,\ \Gamma_n\}$ do not depend on ℓ.) If the eigen energy is calculated with the effects of the interaction energies $E^{(\mathrm{T})}_{\mathrm{int}}$ and $E^{(\mathrm{L})}_{\mathrm{int}}$, then we should omit the corresponding contributions in the matrix elements of (4.13). Once we have prepared the coefficient matrix and initial condition $\mathbf{F}^{(0)}$, we just invert the matrix to obtain

$$X_{\nu 0} = \sum_\mu [\ E^{(\mathrm{T})}_{\mathrm{int}} +\ E^{(\mathrm{L})}_{\mathrm{int}} + (E_n - i\Gamma_n)\mathbf{1} - \hbar\omega\mathbf{1}]^{-1}_{\nu\mu}\ F^{(0)}_{\mu 0}\ , \tag{4.14}$$

which gives the induced current density via (4.12). From this result, we can further calculate the induced EM field via the M-eqs.

When the interaction $E_{\rm int}^{\rm (T)}$ and $E_{\rm int}^{\rm (L)}$ are large for the inter-cell components, it will lead to the spatial dispersion effect in the resonance energy. Then, the response will be delicately dependent on the geometry to calculate the response spectra, which is an aspect missing in homogenized metamaterials systems. Since the above-mentioned scheme does not have an essential difficulty to prevent the procedure, we may claim that the nonlocal response of metamaterials can also be treated in this fashion.

Of course, there are some additional aspects to be discussed about how one takes the effects of non-resonant components into account. A standard way to treat the effect is to ascribe a background dielectric constant to the resonant part of susceptibility. If we further assume that this background dielectric is extended to the infinity, we could renormalize the effect into the EM Green function rather easily. This will change the estimate of $E_{\rm int}^{\rm (T)}$ and $E_{\rm int}^{\rm (L)}$. If, however, we want to treat this background dielectric as a finite confined object, which may cause a cavity effect, we need to prepare a more complex renormalized EM Green function [9]. The preparation of this renormalized EM Green function is feasible for simple geometries, such as a multilayer slab or a multi-layer sphere [10]. For such a case, the procedure mentioned above can be carried out just by replacing $\mathscr{G}$ with the renormalized EM Green function.

4.1.5 Dispersion Curves in Chiral LHM: Difference Between DBF and ChC eqs

In Sect. 3.4 we have compared the dispersion equations corresponding to the first-principles constitutive eq. (2.118) and phenomenological DBF eqs. (3.35), (3.36) in the case of chiral symmetry. Because of the low symmetry, $\boldsymbol{P}$ ($\boldsymbol{M}$) is induced by, not only electric (magnetic) field, but also magnetic (electric) field. We have also considered another parameterized (phenomenological) chiral constitutive (ChC) equations (3.52)–(3.53) corresponding to the first-principles ones. The dispersion eqs to be compared are (2.131), (3.52) and (3.65).

In this subsection, we discuss how DBF and ChC eqs are different in reproducing the dispersion curves of the first-principles result. For this demonstration, we choose a special situation, where a resonance with mixed characters of E1 and M1 transitions lies in the frequency region of negative background dielectric function ε. If the resonance is purely M1 character, this is a typical example of left-handed materials (LHM), where a transmission window opens up in the frequency range of total reflection, as discussed in Sect. 3.8.2 with respect to Figs. 3.3, and 3.5.

For this model of chiral LHM, we choose the transverse part of the first-principles macroscopic susceptibility $\chi_{\rm em}^{\rm (T)}$ appearing in the dispersion equation (2.131) in the following form

$$\mathbf{1} + \frac{4\pi c}{\omega^2}\chi_{\rm em}^{\rm (T)} = (\varepsilon_{\rm b} + a' + c'k^2)\mathbf{1} + \begin{bmatrix} 0 & ib'k \\ -ib'k & 0 \end{bmatrix}, \tag{4.15}$$

where ε_b, assumed to be negative, is the background dielectric constant due to all the other resonances, and a', b', c' represent the contributions of a resonant level of mixed (E1, M1) character at frequency ω_0. Namely,

$$(a',\ b',\ c') = \left(\frac{a''}{\omega_0 - \omega},\ \frac{b''}{\omega_0 - \omega},\ \frac{c''}{\omega_0 - \omega} \right) \tag{4.16}$$

where a'', b'' and c'' are the strengths of E1, mixed E1–M1, and M1 components, respectively.

In terms of

$$\bar{\beta} = \omega b'/c\ , \tag{4.17}$$

$$\bar{\varepsilon} = \varepsilon_b + a'\ , \tag{4.18}$$

$$\bar{\mu} = 1/[1 - (\omega/c)^2 c']\ , \tag{4.19}$$

the dispersion equation (2.131) can be put in the form

$$\left(\frac{ck}{\omega} \right)^2 = \bar{\varepsilon}\bar{\mu} \pm \bar{\beta}\bar{\mu}\, \frac{ck}{\omega}\ , \tag{4.20}$$

and its solution is

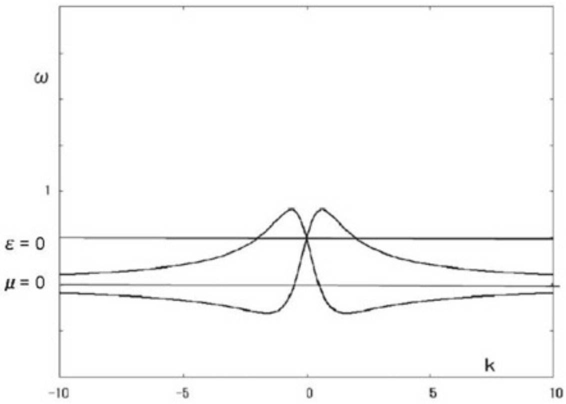
$$\frac{ck}{\omega} = \pm \frac{1}{2} \left[\pm \bar{\beta}\bar{\mu} + \sqrt{\bar{\beta}^2 \bar{\mu}^2 + 4\bar{\varepsilon}\bar{\mu}} \right], \tag{4.21}$$

This is the same type of dispersion relation as (3.65) for ChC eqs, as seen from the appearance of chiral parameters $\bar{\beta}$ and $\hat{\bar{\xi}} + \hat{\eta}$ in the same position of the equations. In contrast, the dispersion relation (3.51) of DBF eqs is a qualitatively different type, since the chiral parameter β does not appear in the square root. This means that β does not affect the condition for the existence of real dispersion curves, while $\bar{\beta}$ of ChC eqs does. The consequence of this difference manifests itself in the presence or absence of linear crossing in the dispersion curves of DBF and ChC eqs, as we discuss below.

Figure 4.1 shows an example of dispersion curve based on this model of ChC eqs. The four combinations of ± give two branches of dispersion curves with linear crossing at $k = 0$ for $\varepsilon_b < 0$. The frequency of the linear crossing is the one satisfying $\bar{\varepsilon} = 0$, which is understood in the following manner. The r.h.s. of (4.21) is a function of ω alone, so that the linear crossing at $k = 0$ means that the r.h.s. is a linear function of ω in the neighborhood of the crossing. To see this, we need to examine the ω-dependence of the two factors $\bar{\beta}\bar{\mu}$ and $\bar{\varepsilon}\bar{\mu}$ when the r.h.s. is approaching zero. The only condition for the r.h.s. to become zero is $\bar{\varepsilon}\bar{\mu} = 0$. Since $\bar{\mu} = (\omega_0 - \omega)/[\omega_0 - \omega - (\omega^2/c^2)c'']$, its zero point at $\omega = \omega_0$ is canceled in the product $\bar{\varepsilon}\bar{\mu}$. Thus, the only condition for realizing $\bar{\varepsilon}\bar{\mu} = 0$ is

$$\bar{\varepsilon} = \varepsilon_b + a' = 0\ , \tag{4.22}$$

Fig. 4.1 Dispersion curves of a chiral left-handed medium for the model in the text. Both ordinate and abscissa are normalized by the frequency of the pole as ω/ω_0 and ck/ω_0. Two horizontal lines show the frequncies of $\varepsilon = 0$ and $\mu = 0$



i.e., $\omega = \omega_0 + a''/\varepsilon_b = \omega_1$, which is less than ω_0 for $\varepsilon_b < 0$. At this frequency the factor $\bar{\beta}\bar{\mu}$ is a finite analytic function, so that $\bar{\varepsilon}\bar{\mu}$ is proportional to $\omega - \omega_1$ in this neighborhood. This leads to the linear crossing of dispersion curves at ($k = 0, \omega = \omega_1$).

The frequency for diverging k ($\to \pm\infty$) arises from $\bar{\mu} = \infty$, i.e. at $\omega = \omega_2$ satisfying $\omega_0 - \omega - (\omega^2/c^2)c'' = 0$. (The root in the close neighborhood of the resonance is $\omega_2 \sim \omega_0 - (\omega_0^2/c^2)c''$.) The r.h.s. of (4.21) has two sets of values ($-\infty, \Delta k$) and ($+\infty, -\Delta k$) for $\bar{\mu} = \infty$, where Δk is a positive finite number. The finite values are understood by rewriting (4.21) as

$$\frac{ck}{\omega} = \pm \frac{2\bar{\varepsilon}\bar{\mu}}{\mp\bar{\beta}\bar{\mu} + \sqrt{\bar{\beta}^2\bar{\mu}^2 + 4\bar{\varepsilon}\bar{\mu}}} \,. \tag{4.23}$$

For $\bar{\mu} \to \infty$ the main factor in the square root becomes $\bar{\beta}^2\bar{\mu}^2$. Then the factors $\bar{\mu}$ on the denominator and numerator are canceled out, and for one of the ± combination the quotient becomes a finite value.

As for the reproducibility of the linear crossing by the dispersion curves based on ChC eqs, nearly the same argument holds by noting the correspondence among the parameters in dispersion equations (3.65) and (4.21)

$$\bar{\varepsilon} \leftrightarrow \hat{\varepsilon} \,, \quad \bar{\mu} \leftrightarrow \hat{\mu} \,, \quad \bar{\beta} \leftrightarrow (\hat{\xi} + \hat{\eta}) \,. \tag{4.24}$$

Ascribing the appropriate ω-dependence to the ChC parameters based on the single pole nature of all the components of linear susceptibility, we can expect the same behavior of the dispersion equation (3.65) as the microscopic model.

In order to carry out a similar argument about DBF eqs, we need to consider how to ascribe the ω-dependence to the DBF parameters, which does not exist in the literature to the author's knowledge. Apart from the details, a single pole character should be ascribed to each ε, μ, $\beta\varepsilon$, $\beta\mu$, since they are linear susceptibilities. For the chiral components, the single pole character should be assigned, not to β and

ε, μ separately, but to each of the products $\beta\varepsilon$, $\beta\mu$. In this sense, we write the DBF dispersion equation as

$$\frac{ck}{\omega} = \pm\frac{\sqrt{\varepsilon\mu}}{1 \pm (\omega/c)\sqrt{(\beta\varepsilon)(\beta\mu)}} \tag{4.25}$$

Generally, the r.h.s. vanishes at the zero of $\varepsilon\mu$ or at the poles of $\beta\varepsilon$ and $\beta\mu$. Since the zeros of the former do not generally coincide with the zeros and poles of the latter, we may discuss the two cases independently. Since ε and μ do not simultaneously become zero in general, and since their zero points are not singular point, they are linear functions of ω near the zero points ω_z, which leads to $k \sim \sqrt{|\omega - \omega_z|}$. Therefore, no linear crossing occurs.

The pole of $\beta\varepsilon$ and $\beta\mu$ arises from the chiral nature of the resonance, which has the mixed character of E1 and M1 transitions in the case of the microscopic model considered above, This means that the same resonance contributes to ε and μ, too. If all of the four parameters have the same pole, its effect is canceled in the quotient, leading to the finite value of the r.h.s. of (4.25).

In (3.55) we showed the relation $\hat{\mu} \sim 1/(1 - 4\pi\chi_{\mathrm{mB}})$, which means that the pole of $\hat{\mu}$ is shifted from that of χ_{mB}. If one extend this argument to the relation between μ and χ_{mB}, the canceling of the chiral pole position is expected between ε and $\beta\varepsilon\beta\mu$. Then, there remains the square root of the inverse of the pole, $\sqrt{\omega_0 - \omega}$, on the r.h.s. of (4.25). This again fails to show the linear crossing.

We also showed the relations (3.69) between the parameters of DBF and ChC eqs based on the assumption of the equivalence of DBF and ChC eqs. We also want to check the consequence of this assumption, whether or not this relation applied to the DBF dispersion relation can reproduce the linear crossing. From these relations, we can derive

$$\frac{1}{\mu} = \frac{1}{\hat{\mu}} + \frac{\hat{\xi}^2}{\hat{\varepsilon}} \tag{4.26}$$

and

$$\frac{\omega^2}{c^2}\beta^2\varepsilon\mu = 1 - \frac{\mu}{\hat{\mu}}\ . \tag{4.27}$$

For the pole structure of the ChC parameters $\hat{\varepsilon}$, $\hat{\mu}$, $\hat{\xi}$, $\hat{\eta}$ corresponding to the microscopic model discussed above, the first equation shows that the pole of $1/\mu$ occurs at the same position as that of $1/\hat{\mu}$, i.e., at $\omega = \omega_0$. Then, the r.h.s. of the second equation loses the same pole because of the cancellation. Since the l.h.s. of the second equation appears in the denominator of the DBF dispersion equation (4.25), we can rule out the possibility of the divergence of the denominator at $\omega = \omega_0$. Then, the only possibility for the r.h.s. of this dispersion equation to approach zero is the vanishing of its numerator $\sqrt{\varepsilon\mu}$. In this case, the behavior of the dispersion curves near $k = 0$ is $k \sim \sqrt{|\omega - \omega_z|}$ as discussed above, which is different from linear crossing. Thus, the equivalence of DBF and ChC eqs is an invalid assumption. From all the arguments given above, we can conclude that DBF dispersion curves do not reproduce the linear crossing for the left-handed chiral resonance.

For all the reasons mentioned above, including the different conditions for real solution of the dispersion equation (3.67) versus (3.68), it is strongly recommended to use ChC eqs rather than DBF eqs for the discussion of resonant chiral LHM, if one prefers to stay in phenomenology.

4.2 Spatial Dispersion in Macro- Versus Microscopic Schemes

The wave vector ($\boldsymbol{k}$) dependence of ε and μ in the conventional scheme, or χ_{em} in the present one, is generally called spatial dispersion effect. The $\boldsymbol{k}$-dependence may occur both in the denominators and in the numerators, but, from the viewpoint of the physics involved, we should distinguish the $\boldsymbol{k}$-dependence (a) in the denominators (and numerators) and (b) only in the numerators. The underlying physics is as follows. If the $\boldsymbol{k}$-dependence appears in the microscopic susceptibility, it reflects the translational symmetry of the microscopic system in consideration. Unless the k-dependence, especially of the denominator, is negligible, LWA is not a good approximation, so that we need to stay in the regime of microscopic (nonlocal) response as described in Sect. 2.2. On the other hand, if the microscopic system has no translational symmetry and if LWA is a good approximation, the macroscopic average of this susceptibility can be expressed as a macroscopic susceptibility χ_{em} with a $\boldsymbol{k}$-dependence only in the numerator. Therefore, this is the only $\boldsymbol{k}$-dependence allowed in the macroscopic description.

If we consider the Taylor expansion of each component of the microscopic susceptibility up to the $O(k^2)$ terms, as we explicitly show for $\chi_{\mathrm{em}}(\boldsymbol{k}, \omega)$ in Sect. 2.3, the dispersion equation of the coupled waves of matter and EM field, (2.131), is the quartic equation of k for a given frequency ω. Since the four waves correspond to the forward and backward propagating waves for two polarizations, there arises no problem of "additional waves" as in the next case of resonant spatial dispersion described below. In this case, the standard treatment of macroscopic boundary conditions, given in Sect. 3.7, is enough to determine the response uniquely. The k-linear term in the dispersion equation may lead to a complex situation involving the mixing of polarizations, but the number of the boundary conditions does not increase in comparison with the conventional case of non-spatially dispersive media.

An essentially new situation arises, when the $\boldsymbol{k}$-dependence appears in the denominator of microscopic susceptibility. Though this is the case outside the macroscopic response, we give an outline of the physics involved in this situation. The essential point here is that the microscopic eigenstates of the medium are the coherent waves specified by $\boldsymbol{k}$, which does not allow the use of LWA. The coherence effect appears not only in the denominator of susceptibility via excitation energies, but also in the numerators through the corresponding eigenfunctions. Because of the $\boldsymbol{k}$-dependence in the denominator, the dispersion equation becomes a polynomial equation higher than the quartic equation of $\boldsymbol{k}$. In the first example of this category discussed by Hopfield in early days [11], the $\boldsymbol{k}$-dependence in the denominator was considered as the $O(k^2)$ dependence of exciton energy, which leads to the quadratic equation of k^2 as the dispersion equation for a given polarization. This equation gives four solution for k (for a given polarization), i.e., two waves in a given direction (forward or backward). Therefore, there is an additional wave in each direction of propagation in this medium, which gives rise to a famous problem of additional boundary condition, ABC problem, to determine the relative amplitude of the waves and, then, the response of the matter uniquely.

How to determine the form ABC for a given medium with such a spatial dispersion effect has been a long debated problem in the physics of excitons [12]. There have been both phenomenological and first-principles approaches to this problem. An essential progress has been made by the latter through considering the susceptibility of the medium as that in the presence of surface, which breaks the translational symmetry of the medium. The solution of M-eqs in terms of such a susceptibility can determine the form of ABC, which (in principle) reflects the details of the surface contribution to the susceptibility. Also, it was noticed that the same M-eqs can be solved without referring to ABC [13], which was an essential seed of the microscopic nonlocal response theory given in Sect. 2.2. The details of this development is described in Sect. 3.8 of [7].

Since the spatial dispersion effect in the denominator of susceptibility has a much more profound meaning than that in the numerator, we should specify which case is meant on mentioning spatial dispersion effect. To summarize this section, it should be noted that the only spatial dispersion effect compatible with LWA is the $\boldsymbol{k}$-dependence in the numerator.

4.3 Resonant Bragg Scattering from Inner-Core Excitations

The arguments in the main formulation in Sect. 2.3 are all based on the assumption that all the excited states of matter can be treated in LWA. As discussed in Sect. 3.6, there are various cases where this assumption is not valid, which, however, does not mean our incapability of handling such cases. In this section, we show an example of this kind, for which we can present a useful framework to analyze some relevant experimental results.

If we irradiate a crystal with an X ray which can excite the inner shell of its constituent atoms, we can expect a resonant diffraction of X ray, which is mediated by the inner shell excitations. The scattering process reflects how the resonant atoms are arranged in the crystal lattice. The clearest signal of X ray scattering is that the change in the wave vector $\boldsymbol{k}$ is equal to one of the reciprocal lattice vectors $\{\boldsymbol{G}\}$. It is a linear process in the sense that the signal amplitude is linear in the incident field amplitude. Since an inner shell excitation is localized on each atom, which has much smaller spatial extension than the X ray, it seems to be all right to apply LWA to the microscopic susceptibility of this process. However, the LWA averaged susceptibility $\chi_{\mathrm{em}}(\boldsymbol{k}, \omega)$ obviously does not describe the diffraction process, because $\chi_{\mathrm{em}}(\boldsymbol{k}, \omega)$ is the susceptibility for a given wave vector without any change before and after the interaction with matter.

The key to solve this discomfort is (2.109) of Sect. 2.3, which shows that the LWA averaged induced current density can have a different wave vector $\boldsymbol{k}$ from that of the incident EM field $\boldsymbol{k}'$. It should be reminded that we picked up only the scattered fields with $\boldsymbol{k} = \boldsymbol{k}'$ by assuming the spatial uniformity of the LWA averaged macroscopic medium. At this point, we should realize the possibility that the LWA average of a localized inner shell excitation does not necessarily mean the smeared out distribution of the similar excitations. In other words, we keep the meaning of (2.109) as it is, and note that the summation index ν contains the positions of inner shell atomic excitation in a regular lattice, which leads us to the selection rule $\boldsymbol{k} - \boldsymbol{k}' = \boldsymbol{G}$. The explicit formulation goes as follows.

The linear susceptibility describing the diffraction process can be obtained in the following manner. Since we are interested in the resonant excitation of inner shell transition, we keep only the resonant terms in the microscopic susceptibility. Dividing the summation index ν into the atomic position $\boldsymbol{R}$ and the quantum number $\bar{\nu}$ for the resonant transition of a particular species of atoms in the crystal, we obtain the resonant terms

$$\begin{aligned}\chi_{\mathrm{cd}}(\boldsymbol{r}, \boldsymbol{r}', \omega) &= \frac{1}{c} \sum_{\bar{\nu}} \sum_{\boldsymbol{R}} g_{\bar{\nu}}(\omega) \boldsymbol{I}_{0\bar{\nu}}(\boldsymbol{r} - \boldsymbol{R}) \boldsymbol{I}_{\bar{\nu}0}(\boldsymbol{r}' - \boldsymbol{R}) \,. \qquad (4.28)\\ &= \left[\text{same expression without } 1/c\right]_{\mathrm{SI}}\end{aligned}$$

Though the final state of the transition is affected by the surrounding atoms or the band structure in the corresponding energy range, the induced current density of the transition is well-localized because of the strong localization of the core state wave function. When the crystal consists of sublattices, we may write $\boldsymbol{R} = \bar{\boldsymbol{r}} + \boldsymbol{\tau}$, where $\boldsymbol{\tau}$ is the vector defining the position of sublattices in a unit cell, and $\bar{\boldsymbol{r}}$ is the vector of Bravais lattice, for which reciprocal lattice vectors $\{\boldsymbol{G}\}$ are defined as $\bar{\boldsymbol{r}} \cdot \boldsymbol{G} = 2\pi \times$ integer.

The $(\boldsymbol{k}, \omega)$ Fourier component of the current density induced by this χ_{cd} is

$$\begin{aligned}\tilde{\boldsymbol{I}}(\boldsymbol{k}, \omega) &= \frac{V^2}{8\pi^3 c} \sum_{\bar{\nu}} \sum_{\boldsymbol{R}} \mathrm{e}^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot \boldsymbol{R}}\, g_{\bar{\nu}}(\omega)\, \bar{\boldsymbol{I}}_{0\bar{\nu},\boldsymbol{\tau}}(\boldsymbol{k}) \sum_{\boldsymbol{k}'} \bar{\boldsymbol{I}}_{\bar{\nu}0,\boldsymbol{\tau}}(-\boldsymbol{k}') \cdot \boldsymbol{A}(\boldsymbol{k}', \omega),\\ &= [\text{same expression without } 1/c]_{\mathrm{SI}} \qquad (4.29)\end{aligned}$$

where the $\boldsymbol{k}$ Fourier component of the matrix element at the site $\boldsymbol{R} = \bar{\boldsymbol{r}} + \boldsymbol{\tau}$ is defined as

$$\bar{\boldsymbol{I}}_{0\bar{\nu},\boldsymbol{\tau}}(\boldsymbol{k}) = \frac{1}{V} \int \mathrm{d}\boldsymbol{r} \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \boldsymbol{I}_{0\bar{\nu},\boldsymbol{\tau}}(\boldsymbol{r} - \boldsymbol{R}) = \frac{1}{V} \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{R}} \int \mathrm{d}\boldsymbol{r}'\, \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{r}'} \boldsymbol{I}_{0\bar{\nu},\boldsymbol{\tau}}(\boldsymbol{r}') \quad (4.30)$$

to extract the position dependent phase factor from $\tilde{\boldsymbol{I}}_{0\bar{\nu},\boldsymbol{\tau}}(\boldsymbol{k})$. We attach the $\boldsymbol{\tau}$-dependence explicitly to the matrix element of current density, since a same atomic transition $0 \rightarrow \bar{\nu}$ can give different results for different sublattices because of the difference in the surroundings.

The microscopic current density given above could be used as the source term of the M-eqs for vector potential. However, in view of the short localization length of the induced current density at each site, we can apply LWA to the microscopic current density. Using the result (2.116), we have

$$\begin{aligned} \tilde{\boldsymbol{I}}_{0\bar{\nu},\boldsymbol{\tau}}(\boldsymbol{k}) &= \frac{1}{V} \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{R}} \{\bar{\boldsymbol{J}}_{0\bar{\nu},\boldsymbol{\tau}} - i\boldsymbol{k} \cdot \bar{\boldsymbol{Q}}^{(\mathrm{e}2)}_{0\bar{\nu},\boldsymbol{\tau}} + ic\boldsymbol{k} \times \bar{\boldsymbol{M}}_{0\bar{\nu},\boldsymbol{\tau}}\} , \\ &= [\text{same expression without } c]_{\mathrm{SI}} . \end{aligned} \tag{4.31}$$

In terms of this LWA expression of $\tilde{\boldsymbol{I}}_{0\bar{\nu},\boldsymbol{\tau}}(\boldsymbol{k})$, we obtain the current density as a function of $\boldsymbol{A}(\boldsymbol{k}', \omega)$ as

$$\begin{aligned} \tilde{\boldsymbol{I}}(\boldsymbol{k}, \omega) &= \frac{1}{c} \sum_{\bar{\nu}} \sum_{\boldsymbol{R}} \mathrm{e}^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}}\, g_{\bar{\nu}}(\omega) \\ &\quad \times \{\bar{\boldsymbol{J}}_{0\bar{\nu},\boldsymbol{\tau}} - i\boldsymbol{k} \cdot \bar{\boldsymbol{Q}}^{(\mathrm{e}2)}_{0\bar{\nu},\boldsymbol{\tau}} + ic\boldsymbol{k} \times \bar{\boldsymbol{M}}_{0\bar{\nu},\boldsymbol{\tau}}\} \\ &\quad \times \sum_{\boldsymbol{k}} \{\bar{\boldsymbol{J}}_{\bar{\nu}0,\boldsymbol{\tau}} + i\boldsymbol{k}' \cdot \bar{\boldsymbol{Q}}^{(\mathrm{e}2)}_{\bar{\nu}0,\boldsymbol{\tau}} - ic\boldsymbol{k}' \times \bar{\boldsymbol{M}}_{\bar{\nu}0,\boldsymbol{\tau}}\} \cdot \boldsymbol{A}(\boldsymbol{k}', \omega) , \\ &= \left[\text{same expression without } 1/c\right]_{\mathrm{SI}} \end{aligned} \tag{4.32}$$

The Bravais lattice part of phase factor $\exp[i(\boldsymbol{k} - \boldsymbol{k}') \cdot \bar{\boldsymbol{r}}]$ becomes unity for the wave vector transfer by a reciprocal lattice vector $\boldsymbol{k} - \boldsymbol{k}' = \boldsymbol{G}$. Thus, (4.29) gives the induced current density satisfying the Bragg condition for an arbitrary incident X ray with wave vector $\boldsymbol{k}'$. The amplitude of scattered X ray is calculated from the Maxwell equation with this current density as a source term, which is rewritten as a set of linear equations for the variables $\tilde{\boldsymbol{I}}(\boldsymbol{k}' + \boldsymbol{G}, \omega)$ and $\boldsymbol{A}(\boldsymbol{k}' + \boldsymbol{G}, \omega)$, containing the incident field $\boldsymbol{A}_0(\boldsymbol{k}', \omega)$ as a parameter. The number of $\boldsymbol{G}$'s to be considered depends on the strength of interaction. Since X-ray scattering is usually a weak process, even at a resonance, it will be a reasonable approximation to treat it kinematically, i.e., to consider single scattering processes alone. This approximation corresponds to the use of (4.32) with the $\boldsymbol{A}(\boldsymbol{k}', \omega)$ replaced with the incident wave $\boldsymbol{A}_0(\boldsymbol{k}', \omega)$. The scattered wave is polarized perpendicular to $\boldsymbol{k}$, and the amplitudes of each polarized component is determined by the projection of $\tilde{\boldsymbol{I}}(\boldsymbol{k}, \omega)$ on the (unit) polarization vector $\hat{\boldsymbol{e}}(\boldsymbol{k})$ ($\perp \boldsymbol{k}$).

The scattering amplitude for a given incident field $\boldsymbol{A}_0(\boldsymbol{k}')$ is

$$\begin{aligned} \boldsymbol{A}(\boldsymbol{k}, \omega) &= \frac{4\pi}{\omega^2 - c^2k^2} \sum_{\bar{\nu}} \sum_{\boldsymbol{R}} \exp\{-i(\boldsymbol{k} - \boldsymbol{k}') \cdot \boldsymbol{R}\}\, g_{\bar{\nu}}(\omega) \\ &\quad \times \{\bar{\boldsymbol{J}}_{0\bar{\nu},\boldsymbol{\tau}} - i\boldsymbol{k} \cdot \bar{\boldsymbol{Q}}^{(\mathrm{e}2)}_{0\bar{\nu},\boldsymbol{\tau}} + ic\boldsymbol{k} \times \bar{\boldsymbol{M}}_{0\bar{\nu},\boldsymbol{\tau}}\} \end{aligned}$$

$$\times \sum_{\boldsymbol{k}'} \{\bar{\boldsymbol{J}}_{\bar{\nu}0,\boldsymbol{\tau}} + i\boldsymbol{k}' \cdot \overline{\boldsymbol{Q}}^{(\mathrm{e}2)}_{\bar{\nu}0,\boldsymbol{\tau}} - ic\boldsymbol{k}' \times \bar{\boldsymbol{M}}_{\bar{\nu}0,\boldsymbol{\tau}}\} \cdot \boldsymbol{A}_0(\boldsymbol{k}', \omega) \,, \tag{4.33}$$

$$= \left[\text{same expression with } 4\pi \text{ replaced by } c^2\mu_0\right]_{\mathrm{SI}}$$

where $\boldsymbol{k} = \boldsymbol{k}' + \boldsymbol{G}$.

The uniqueness of this result is that it contains the case of chirality-induced Bragg scattering in a general form applicable to any inner shell transition and any symmetry of crystal. Usually X ray scattering is said to be unable to distinguish left (L-) and right (R-) handed chirality, which is a conjecture derived from the intensity of the allowed beams. Recently, there has appeared a paper reporting the successful distinction of L- and R-handed quartzes by means of the forbidden beams enhanced by the resonance with inner shell transitions [14]. The expression obtained above is suitable to such a description, as shown below.

At low temperatures, quartz crystals show L- and R-handed distortion around its trigonal (c-) axis. Both of them consist of triangular sublattices stacked along the trigonal axis with three layers as a unit. The lattice points of the three layers are arranged in a three fold rotation symmetry with a non-primitive translation by $c/3$, so that the lattice vector characterizing the three sublattices are $\boldsymbol{\tau}_1 = (a_1, b_1, 0)$, $\boldsymbol{\tau}_2 = (a_2, b_2, c/3)$, $\boldsymbol{\tau}_3 = (a_3, b_3, 2c/3)$, where the 2 dimensional vectors (a_1, b_1), (a_2, b_2), (a_3, b_3) are related with one another by the three fold rotation around the c-axis. The reciprocal lattice vector in the c-direction is written as $\boldsymbol{G}_c = (2\pi/c)(0, 0, \ell)$ for an arbitrary integer ℓ. For any lattice point $\boldsymbol{R} = \bar{\boldsymbol{r}} + \boldsymbol{\tau}_j$, we have

$$\boldsymbol{R} \cdot \boldsymbol{G}_c = 2\pi \times \text{integer} + \frac{2\pi\ell}{3}(j-1) \,, \tag{4.34}$$

which will be used below.

In a simple theory of X-ray diffraction, we assume a regular array of spherical scatterers. To calculate the amplitude of scattered waves, we sum up the contributions from all the scatterers, which is the product of "atomic scattering factor" times the sum of phase difference of all the scatterers. The latter is

$$\sum_{\boldsymbol{R}} \exp(i\boldsymbol{R} \cdot \boldsymbol{G}) \,, \tag{4.35}$$

which in the case of (4.34) is zero except for $\ell =$ multiples of 3. This means that the diffraction with $\boldsymbol{G}_c = (2\pi/c)(0, 0, \ell)$ is forbidden except for $\ell =$ multiples of 3.

The result given just above is due to the $\boldsymbol{\tau}$ independence of the "atomic scattering factor". The forbidden character of the scattering is generally relaxed, when we use a resonance condition, since it picks up a detailed electronic structure of the resonant state, which can be dependent on the sublattices. More specifically, this relaxation occurs through the lowering of the symmetry of the induced current densities contributing the resonant scattering. In the case of an isolated atom, there always exist a set of degenerate excited states belonging to a irreducible representation of electric dipole (E1) character. Because of the degeneracy, we can choose any Cartesian framework to express the dipole moments, so that this set of states act to the EM field as a spherical scatterer. This argument can be checked by (4.33). If we can assume spherical symmetry at each site $\boldsymbol{R}$, we can choose the basis for { $|\bar{\nu}\rangle$'s } independent of $\boldsymbol{\tau}$. Then, the summation over $\boldsymbol{\tau}$ acts only on the phase factor $\exp(i\boldsymbol{G}\cdot\boldsymbol{\tau})$, leading to zero. Thus, non-zero scattering amplitude is due to the deviation of the site symmetry from a spherical one. This deviation is obvious, since lattice structures have always lower symmetry than spherical. Even in cubic symmetry, $\ell = 2$ (or higher) angular momentum states are no more completely degenerate, so that the contribution from these states will give non-spherical effect. If the symmetry is lower, even E1 transitions ($\ell = 1$ angular momentum states) will split into several levels in a different way for each sublattice $\boldsymbol{\tau}$. The eigenfunctions are also affected by this splitting, giving a $\boldsymbol{\tau}$-dependence to the matrix elements. All these effects will preserve the $\boldsymbol{\tau}$-dependence of the matrix elements $\bar{\boldsymbol{J}}_{0\bar{\nu}}$, $\bar{\boldsymbol{Q}}^{(\mathrm{e2})}_{0\bar{\nu}}$, $\bar{\boldsymbol{M}}_{0\bar{\nu}}$ in (4.33), leading to the relaxation of the forbidden character of the scattered beams with $\boldsymbol{G}_c$.

4.4 Renormalization of L Current Density into $\boldsymbol{E}_{\mathrm{L}}$

4.4.1 Use of $\boldsymbol{E}_{\mathrm{L}}$ as External Field

As mentioned in Sect. 2.2, the main part of this new formulation is made according to the scheme where matter Hamiltonian contains the complete Coulomb interaction. This means that the interaction between the induced L electric field and matter polarization, which can be rewritten as the Coulomb interaction energy among the induced charge densities (see below), is, not a part of interaction Hamiltonian, but a part of matter Hamiltonian. This energy appears as a part of matter excitation energies defining the poles of susceptibility, and has been called LT splitting energy, electron-hole exchange interaction, or depolarization energy. Thus the EM field inducing matter polarization is $\boldsymbol{E}_{\mathrm{T}}$ alone, while $\boldsymbol{E}_{\mathrm{L}}$ is an internal quantity. For a T-field incidence, the physical variables to be determined selfconsistently are $\boldsymbol{A}$ and $\boldsymbol{I}_{\mathrm{T}}$, i.e., the T-components of EM field and induced current density. In this process there is no need of considering $\boldsymbol{E}_{\mathrm{L}}$ and $\boldsymbol{P}_{\mathrm{L}}$, as long as the Coulomb potential is properly handled in the quantum mechanical calculation. If one dare to know $\boldsymbol{E}_{\mathrm{L}}$ and $\boldsymbol{P}_{\mathrm{L}}$, they can be calculated by using the selfconsistently determined values of $\boldsymbol{A}$ and the T-L mixing components of χ_{em}. The case of L-field incidence caused by an external charge density is described in Sect. 5.7.2.

There is an alternative scheme to treat the induced $\boldsymbol{E}_{\mathrm{L}}$ as, not an internal, but an "external" field even in the absence of external charge density. In this case, $-\int \boldsymbol{P}\cdot\boldsymbol{E}_{\mathrm{L}}\mathrm{d}\boldsymbol{r}$ is the interaction between "external" field $\boldsymbol{E}_{\mathrm{L}}$ and matter. This point of view requires a change in the definition of matter Hamiltonian and matter-EM field interaction, as discussed below. The interaction energy $-\int \boldsymbol{P}\cdot\boldsymbol{E}_{\mathrm{L}}\mathrm{d}\boldsymbol{r}$ can be rewritten, in terms of the induced charge density ρ, as

$$-\int \boldsymbol{P}\cdot\boldsymbol{E}_\mathrm{L}\mathrm{d}\boldsymbol{r} = \int \mathrm{d}\boldsymbol{r}\int \mathrm{d}\boldsymbol{r}'\ \boldsymbol{P}\cdot\nabla\frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} \tag{4.36}$$

$$= -\int \mathrm{d}\boldsymbol{r}\int \mathrm{d}\boldsymbol{r}'\ \nabla\cdot\boldsymbol{P}(\boldsymbol{r})\frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} \tag{4.37}$$

$$= \int \mathrm{d}\boldsymbol{r}\int \mathrm{d}\boldsymbol{r}'\ \frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} \tag{4.38}$$

$$\equiv H_\mathrm{cc} \tag{4.39}$$

which is the Coulomb (self-) interaction energy of induced charge density $\mathrm{H_{cc}}$. Since this energy appears in the presence of induced polarization or induced charge density, i.e., in the excited states of matter, it is a part of the Coulomb interaction among the charged particles of matter. If we treat this energy as the interaction between matter and EM field, we have to subtract this part of Coulomb interaction from the (original) matter Hamiltonian in order to keep the consistency within the total Hamiltonian.

Since this affects the eigenvalues and eigen functions of matter, this new choice requires a certain modification of EM response theory, which we describe in this subsection with a stress on the difference compared with the scheme used in the description of Chaps. 2 and 3. If we use them properly, the two schemes should produce the same response. However, there can arise a difference in judging the validity of LWA to derive macroscopic M-eqs, which is discussed in the next subsection.

In the scheme of Chap. 2, the matter Hamiltonian $\mathrm{H^{(0)}}$, (2.17), is the sum of kinetic energy and full Coulomb potential (and relativistic corrections), and the interaction H_int, (2.24), contains only T-field $\boldsymbol{A}$. In the new scheme of EM response, the matter Hamiltonian is

$$\tilde{H}^{(0)} = H^{(0)} - H_\mathrm{cc} \tag{4.40}$$

and the matter - EM field interaction is

$$\tilde{H}_\mathrm{int} = H_\mathrm{int} - \int \mathrm{d}\boldsymbol{r}\ \boldsymbol{P}_\mathrm{L}\cdot\boldsymbol{E}_\mathrm{L} \tag{4.41}$$

where $\boldsymbol{E}_\mathrm{L}$ is the sum of incident and induced L-fields as

$$\boldsymbol{E}_\mathrm{L}(\boldsymbol{r}) = -\nabla\int \mathrm{d}\boldsymbol{r}'\ \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} + \boldsymbol{E}_\mathrm{extL}\ . \tag{4.42}$$

In the interaction Hamiltonian, the first part $\mathrm{H_{int}}$ takes care of the response of the T-components, and the second part $-\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{P}_\mathrm{L}\cdot\boldsymbol{E}_\mathrm{L}$ that of the L-components. In chiral symmetry, $\boldsymbol{A}$ can induce $\boldsymbol{P}_\mathrm{L}$ (or $\boldsymbol{J}_\mathrm{L}$), as well as $\boldsymbol{J}_\mathrm{T}$, and $\boldsymbol{E}_\mathrm{extL}$ can induce $\boldsymbol{J}_\mathrm{T}$, as well as $\boldsymbol{P}_\mathrm{L}$.

The calculation of microscopic susceptibilities and the application of LWA go in a very similar way as in Chap. 2 (for T-response) and Sect. 5.7.2 (for L-response). What we need anew is to use the eigenvalues and eigenfunctions of the new matter

Hamiltonian $\tilde{H}^{(0)}$, which changes the positions of poles and their intensities. The constitutive equations in this case are

$$\boldsymbol{I}(\boldsymbol{k},\omega) = \tilde{\chi}_{\mathrm{em}}(\boldsymbol{k},\omega)\ \boldsymbol{A}(\boldsymbol{k},\omega) + \tilde{\chi}_{\mathrm{JEL}}(\boldsymbol{k},\omega)\ \boldsymbol{E}_{\mathrm{L}}(\boldsymbol{k},\omega)\ , \tag{4.43}$$

where $\tilde{\chi}_{\mathrm{em}}$ and $\tilde{\chi}_{\mathrm{JEL}}$ are defined in the same way as χ_{em} and χ_{JEL} in terms of the energy eigen values and eigen functions of $\tilde{H}^{(0)}$.

In terms of the new susceptibilities and the source fields $\{\tilde{\chi}_{\mathrm{em}}, \tilde{\chi}_{\mathrm{JEL}}, \boldsymbol{A}, \boldsymbol{E}_{\mathrm{L}}\}$, we can make a similar selfconsistent scheme as that in terms of $\{\chi_{\mathrm{em}}, \chi_{\mathrm{JEL}}, \boldsymbol{A}, \boldsymbol{E}_{\mathrm{extL}}\}$. For high symmetry cases where L and T modes do not mix, the first and second terms on the r.h.s. describes the T and L response, respectively. When LT mixing occurs, we pick up the T and L components of the response as follows. The T-component of $\boldsymbol{I}(\boldsymbol{k},\omega)$, needed for the M-eq of $\boldsymbol{A}$, is obtained by applying the projection operator $(\mathbf{1} - \hat{\boldsymbol{k}}\hat{\boldsymbol{k}}) \cdot \boldsymbol{I}(\boldsymbol{k},\omega)$, and the L-component of $\boldsymbol{P}$, to be used in (4.42) via $\nabla \cdot \boldsymbol{P} = -\rho$, is $\boldsymbol{P}_{\mathrm{L}}(\boldsymbol{k},\omega) = (i/\omega)\hat{\boldsymbol{k}} \cdot \boldsymbol{I}(\boldsymbol{k},\omega)$.

The change in the pole positions due to the change in the matter Hamiltonian is reflected in the conditions for the eigen modes of L character. In the case of χ_{em} and χ_{JEL}, the poles represent the transition energies of $\mathrm{H}^{(0)}$ containing H_{cc}, so that the T and L modes energies are directly included in the pole position. In fact, $\boldsymbol{I}^{(\mathrm{L})} = \chi_{\mathrm{JEL}} \boldsymbol{E}_{\mathrm{extL}}$ indicates the presence of finite amplitude L-mode current density $\boldsymbol{I}^{(\mathrm{L})}$ for vanishing $\boldsymbol{E}_{\mathrm{extL}}$ when χ_{IEL} goes to infinity, i.e. at the excitation frequency of a L-mode. On the other hand, the poles of $\tilde{\chi}_{\mathrm{JEL}}$ do not have the contribution of H_{cc}, but the equation

$$\boldsymbol{E}_{\mathrm{L}} = -4\pi \boldsymbol{P}_{\mathrm{L}} + \boldsymbol{E}_{\mathrm{extL}}\ , \tag{4.44}$$

which is an extension of (5.164) by including the incident L field, indicates that the condition for the existence of finite amplitude solution of $\boldsymbol{E}_{\mathrm{L}}$ in the absence of $\boldsymbol{E}_{\mathrm{extL}}$ is

$$1 + 4\pi \tilde{\chi}_{\mathrm{JEL}} = 0\ . \tag{4.45}$$

This means that, though the susceptibility $\tilde{\chi}_{\mathrm{JEL}}$ does not have the poles at L mode excitations, it provides the eigen mode condition for them. It should be noted that this is the same condition as $\varepsilon_{\mathrm{L}} = 0$ in the conventional macroscopic M-eqs, which is the direct consequence of the conventional definition of χ_{e} as $\boldsymbol{P} = \chi_{\mathrm{e}}\boldsymbol{E}$ and $\varepsilon = 1 + 4\pi\chi_{\mathrm{e}}$, where $\boldsymbol{E}$ contains $\boldsymbol{E}_{\mathrm{L}}$. In this sense, the conventionally defined χ_{e} should be calculated from the matter Hamiltonian $\tilde{H}^{(0)}$.

4.4.2 Difference in the Criterion for LWA

The validity condition of LWA is that, among various quantum mechanical excitations of a matter system, we can neglect the contribution of those with long range coherence in the spectral range of interest. A simple example allowing the use of LWA is an assembly of impurities (of a single species, for simplicity). The supporting argument is that the spatial extension of the wave functions of the transition is well localized in comparison with the wavelength of the EM field corresponding to the transition energy. In this argument, we usually neglect the dipole-dipole interaction between different impurities, which is the leading term of H_{cc}. If we consider this interaction among the impurities, it will cause reorganized energy levels with a certain distribution, or, if the impurities are in a regular lattice, a band structure. In any case, it will lead to a band of new energy eigen values. If this band width is larger than the (non-radiative) width of the impurity levels, we cannot neglect the LW coherence of matter excitations, so that LWA is not a good approximation. The validity condition of LWA is therefore the larger (non-radiative) width of the impurity levels than the band width due to dipole-dipole interaction.

However, if we replace such impurities with split-ring resonators (SRR's) to make metamaterials of desired resonant frequency, we need to reconsider the validity of LWA in homogenizing the contributions of SRR's to obtain an effective macroscopic susceptibility. Especially, if we want to get a high resonant frequency, one uses small structures containing coherent motions of electrons, or plasmons, which produce large amplitude charge densities on each SRR.

The various resonances of a SRR are accompanied by different modes of such charge densities. A high resonant frequency is caused by a strong restoring force, which means that the mode is accompanied by a large amplitude of charge density. This contributes, not only to the high resonant frequency, but also to the large inter-SRR interaction through the long range nature of the Coulomb interaction between charge densities. (The overlap of wave functions is not needed for this interaction.) This inter-SRR interaction leads to the formation of the coherent excited states among the SRR's. Then, each of the coherent state has different eigen frequency with a fully extended wave function over the positions of all the SRR's. These new eigenstates have a band of eigen frequencies. (If SRR's are arranged in a periodic lattice, there arises a energy band structure for the excited states.) If the band width is larger than the width of each level, one cannot neglect the coherence. In other words, LWA or homogenization is not applicable to this system.

Now, if we take the scheme with "$\tilde{H}^{(0)}$, $\tilde{H}_{\rm int}$", the eigenstates of matter are constructed without H_{cc}, so that the coherence of induced polarization due to the interaction of charge densities is not brought in the eigenstates. Therefore, if we judge the validity of LWA for these eigenstates, an important factor will be missing. Thus, there is a possibility in this scheme to make a mistake in judging the validity of LWA.

In contrast, the scheme with "$H^{(0)}$, $H_{\rm int}$" contains H_{cc} in $H^{(0)}$, so that the long range coherence of induced polarization is determined solely by $H^{(0)}$. Since all the elements contributing to the coherence of eigenstates are considered in $H^{(0)}$, we can make a correct decision about the validity of LWA, in contrast to the scheme with "$\tilde{H}^{(0)}$, $\tilde{H}_{\rm int}$".

4.5 Extension to Nonlinear Response

The higher rank theory which we use for the derivation of macroscopic M-eqs is the microscopic nonlocal response theory consisting of the microscopic M-eqs and the microscopic constitutive equation. The latter is given as a power series expansion with respect to $\boldsymbol{A}(\boldsymbol{r}, \omega)$ in the form of integrals containing various susceptibilities as integral kernels. Since these kernels are separable, the Nth order nonlinear induced current density is an Nth order polynomial of the factor

$$F_{\mu\nu}(\omega) = \int \mathrm{d}\boldsymbol{r} \ \langle \mu | \boldsymbol{I}(\boldsymbol{r}) | \nu \rangle \cdot \boldsymbol{A}(\boldsymbol{r}, \omega) \tag{4.46}$$

for various combinations of μ, ν and ω's, including the linear case $N = 1$. (see Sect. 2.6 of [7].)

For example, one of the eight components of the third order nonlinear current density is given, for the field components with frequencies $\omega_1, \omega_2, \omega_3$, as ((2.119) of [7])

$$\boldsymbol{J}^{(3)}(\omega_1 + \omega_2 + \omega_3) = \frac{-1}{(\hbar c)^3} \sum_{\mu} \sum_{\nu} \sum_{\sigma} \frac{F_{0\mu}(\omega_3) F_{\mu\nu}(\omega_2) F_{\nu\sigma}(\omega_1) \langle \sigma | \boldsymbol{I}(\boldsymbol{r}) | 0 \rangle}{(\omega_{0\mu} - \Omega_3)(\omega_{0\nu} - \Omega_2)(\omega_{0\sigma} - \Omega_1)} \tag{4.47}$$

where $\Omega_3 = \omega_3 + i0^+, \Omega_2 = \omega_3 + \omega_2 + i0^+, \Omega_1 = \omega_3 + \omega_2 + \omega_1 + i0^+$, and $\omega_{\mu\nu} = (E_\mu - E_\nu)/\hbar$. The EM field components included in the $F_{\mu\nu}(\omega)$'s have various frequencies and polarizations, and the solution of such (microscopic, nonlinear) constitutive equations and M-eqs turn out to be the solution of simultaneous cubic equations of $F_{\mu\nu}(\omega)$'s. In general, it is possible to rearrange the integral equations for the Nth order nonlinear problem into a set of Nth order polynomial equations of $F_{\mu\nu}(\omega)$'s [7]. By including enough number of the transitions in the calculation, the microscopic spatial structure of the EM response is reproduced by such a calculation.

For the macroscopic description, we need LWA averaged microscopic constitutive equation. Since the coordinates included in various factors $F_{\mu\nu}(\omega)$'s are independent, LWA can be done for each $F_{\mu\nu}(\omega)$'s separately. Thus, the process of LWA is equivalent to the Taylor expansion of each factor $F_{\mu\nu}(\omega)$ where we retain only a few leading terms representing the moments of $\boldsymbol{I}_{\mu\nu}(\boldsymbol{r})$ with E1, M1, E2, ... characters.

In the case of linear response, we have the factors $F_{\mu\nu}(\omega)$ appear as a product of the form $\boldsymbol{I}_{0\nu}(\omega) F_{\nu 0}(\omega)$ or $\boldsymbol{I}_{\nu 0}(\omega) F_{0\nu}(\omega)$ in the susceptibility, which contains a single intermediate step $|\nu\rangle$ between the initial and final states. On the other hand, the Nth order nonlinear susceptibility contains N different intermediate steps. For $N = 3$, we have $|\mu\rangle, |\nu\rangle, |\sigma\rangle$ as shown in the example given above. Via Taylor expansion, each factor $F_{\mu\nu}(\omega)$ is expressed as a linear combination of the moments of $\boldsymbol{I}_{\mu\nu}(\boldsymbol{r})$. Generally, the lowest moment (E1 transition) is the main contribution, so that a large contribution will occur for the process connecting the initial and final states ($|0\rangle$) via the E1 transitions alone. If the frequency of the EM field inducing each E1 transition is resonant to the corresponding transition, the intensity of the whole nonlinear process will become large. If some of the transitions are E1 forbidden, then M1 and/or E2 components will contribute to the finite amplitude of the whole process. This mixture of E1 and (M1, E2) characters can occur for any system including non-chiral case. This feature is different from the linear response, where the mixture of E1 and (M1, E2) transitions is expected only in chiral symmetry.

The merit of the present approach to the macroscopic description of nonlinear response may not be so obvious because of the complexity of the macroscopic nonlinear susceptibilities in terms of the quantum mechanical eigenvalues and eigen functions. For non-resonant processes such a representation does not have much meaning. If one focuses on a particular resonant nonlinear process containing E1 forbidden transitions, there will be a chance for this kind of scheme to show its merit, because it allows the precise description of the resonant nonlinear process with explicit evaluation of E1, E2, and M1 components.

References

1. Veselago, V.G.: Sov. Phys. Uspekh **10**, 509 (1968)
2. Suga, S., Cho, K., Niji, Y., Merle, J.C., Sauder, T.: Phys. Rev. B **22**, 4931 (1980)
3. Ramakrishna, S.A.: Rep. Prog. Phys. **68**, 449 (2005)
4. Schurig, D., Mock, J.J., Justice, B.J., Cummer, S.A., Pendry, J.B., Starr, A.F., Smith, D.R.: Science **314**, 977–980 (2006)
5. Caloz, C., Itoh, T.: Electromagnetic Metamaterials: Transmission Line Theory and Microwave Applications. Wiley, Hoboken (2006)
6. Matsuzawa, S., Sato, K., Inoue, Y., Nomura, T.: IEICE Trans. Electron. **E89-C**, 1337–1344 (2006)
7. Cho, K.: Optical Response of Nanostructures: Microscopic Nonlocal Theory. Springer, Heidelberg (2003)
8. Rockstuhl, C., Zentgraf, T., Guo, H., Liu, N., Etrich, C., Loa, I., Syassen, K., Kuhl, J., Lederer, F., Giessen, H.: Appl. Phys. B **84**, 219–227 (2006)
9. Cho, K.: J. Phys. Condens. Matter **16**, S3695–S3702 (2004)
10. Chew, W.C.: Waves and Fields in Inhomogeneous Media. Van Nostrand Reinhold, New York (1990)
11. Hopfield, J.J., Thomas, D.G.: Phys. Rev. **132**, 563 (1963)
12. Birman, J.L.: Excitons. In: Rashba, E.I., Sturge, M.D. (eds.), p. 72. North Holland, Amsterdam (1982); Halevi, P.: Spatial Dispersion in Solids and Plasmas, p. 339. Elsevier, Amsterdam (1992)
13. Cho, K.: J. Phys. Soc. Jpn. **55**, 4113 (1986)
14. Tanaka, Y., Takeuchi, T., Lovesey, S.W., Knight, K.S., Chainani, A., Takata, Y., Oura, O., Senba, Y., Ohashi, H., Shin, S.: Phys. Rev. Lett. **100**, 145502 (2008)

Chapter 5
Mathematical Details and Additional Physics

Abstract In the previous sections, some mathematical details and additional physics are omitted for the purpose of showing the central line of description straightforwardly. In this chapter, the omitted subjects are given in detail. Each section is independent, and the related subjects are given in subsections. Some of the problems can be found in other books or papers, but they are reproduced here for the sake of self-containing description.

5.1 Continuity Equation and Operator Forms of *P* and *M* in Particle Picture

The continuity equation (1.5) or (1.16) represents the charge conservation during the motion of charges. Therefore, unless we consider those phenomena such as electron-positron pair production by photon in the relativistic regime, it is expected to be valid for usual EM phenomena in non-relativistic regime, both in the continuum and the particle picture of matter. But it will be of some interests to see its validity in particle picture by explicit mathematics. This is shown by the Fourier representation of the properly defined forms of $\rho(\boldsymbol{r})$ and $\boldsymbol{J}(\boldsymbol{r})$ [1]. In the same manner, the operator forms of $\boldsymbol{P}$ and $\boldsymbol{M}$ satisfying the relations expected in macroscopic M-eqs are given.

The expressions of $\rho(\boldsymbol{r})$ and $\boldsymbol{J}(\boldsymbol{r})$ in particle picture

$$\rho(\boldsymbol{r}) = \sum_{\ell} \mathrm{e}_{\ell}\, \delta(\boldsymbol{r} - \boldsymbol{r}_{\ell}) \tag{5.1}$$

$$\boldsymbol{J}(\boldsymbol{r}) = \sum_{\ell} \mathrm{e}_{\ell} \boldsymbol{v}_{\ell}\, \delta(\boldsymbol{r} - \boldsymbol{r}_{\ell}), \tag{5.2}$$

are Fourier decomposed as

$$\rho(\boldsymbol{r}) = \frac{V}{8\pi^3} \int \mathrm{d}\boldsymbol{k}\ \rho_{\boldsymbol{k}} \exp(i\boldsymbol{k} \cdot \boldsymbol{r}) \ , \quad \boldsymbol{J}(\boldsymbol{r}) = \frac{V}{8\pi^3} \int \mathrm{d}\boldsymbol{k}\ \boldsymbol{J}_{\boldsymbol{k}} \exp(i\boldsymbol{k} \cdot \boldsymbol{r}) \ , \tag{5.3}$$

K. Cho, *Reconstruction of Macroscopic Maxwell Equations*,
Springer Tracts in Modern Physics 237,
https://doi.org/10.1007/978-3-662-58424-8_5

where V is the volume for periodic boundary condition to define discrete $\boldsymbol{k}$'s and is supposed to take the limiting value ∞, i.e., $\Sigma_{\boldsymbol{k}} \to (V/8\pi^3)\int \mathrm{d}\boldsymbol{k}$ in the continuum limit. Their Fourier components are

$$\rho_{\boldsymbol{k}} = \sum_\ell \mathrm{e}_\ell \exp(-i\boldsymbol{k}\cdot\boldsymbol{r}_\ell)\ , \quad \boldsymbol{J}_{\boldsymbol{k}} = \sum_\ell \mathrm{e}_\ell \boldsymbol{v}_\ell \exp(-i\boldsymbol{k}\cdot\boldsymbol{r}_\ell)\ . \tag{5.4}$$

The time evolution of ρ occurs through that of each particle, so that we have

$$\frac{\partial\rho}{\partial t} = \sum_\ell \boldsymbol{v}_\ell \cdot \frac{\partial\rho}{\partial\boldsymbol{r}_\ell} = -\frac{iV}{8\pi^3}\sum_\ell \int \mathrm{d}\boldsymbol{k}\ e_\ell(\boldsymbol{v}_\ell\cdot\boldsymbol{k})\ \exp[i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}_\ell)]\ , \tag{5.5}$$

and

$$\nabla\cdot\boldsymbol{J} = \frac{iV}{8\pi^3}\int \mathrm{d}\boldsymbol{k}\ (\boldsymbol{J}_{\boldsymbol{k}}\cdot\boldsymbol{k})\ \exp(i\boldsymbol{k}\cdot\boldsymbol{r}) = -\frac{\partial\rho}{\partial t}\ . \tag{5.6}$$

In this way, the continuity equation is explicitly shown to be valid in particle picture.

The microscopic definition of electric and (orbital) magnetic polarizations, $\boldsymbol{P}$ and $\boldsymbol{M}$, respectively, is given as (Sect. IV.C of [1])

$$\boldsymbol{P}(\boldsymbol{r}) = \int_0^1 \mathrm{d}u \sum_\ell e_\ell \boldsymbol{r}_\ell\ \delta(\boldsymbol{r}-u\boldsymbol{r}_\ell)\ , \tag{5.7}$$

$$\boldsymbol{M}(\boldsymbol{r}) = \frac{1}{c}\int_0^1 u\ \mathrm{d}u \sum_\ell e_\ell \boldsymbol{r}_\ell \times \boldsymbol{v}_\ell\ \delta(\boldsymbol{r}-u\boldsymbol{r}_\ell)$$

$$\left[= \int_0^1 u\ \mathrm{d}u \sum_\ell e_\ell \boldsymbol{r}_\ell \times \boldsymbol{v}_\ell\ \delta(\boldsymbol{r}-u\boldsymbol{r}_\ell)\right]_{\mathrm{SI}}\ . \tag{5.8}$$

This definition satisfies the expected relations

$$\nabla\cdot\boldsymbol{P} = -\rho\ , \tag{5.9}$$

$$\boldsymbol{J} = \frac{\partial\boldsymbol{P}}{\partial t} + c\nabla\times\boldsymbol{M}\ \left[= \frac{\partial\boldsymbol{P}}{\partial t} + \nabla\times\boldsymbol{M}\right]_{\mathrm{SI}} \tag{5.10}$$

for charge neutral systems.

The Fourier component of $\nabla\cdot\boldsymbol{P}(\boldsymbol{r})$ is

$$(\nabla\cdot\boldsymbol{P})_{\boldsymbol{k}} = \int_0^1 \mathrm{d}u \sum_\ell e_\ell(i\boldsymbol{k}\cdot\boldsymbol{r}_\ell)\exp(-iu\boldsymbol{k}\cdot\boldsymbol{r}_\ell)\ , \tag{5.11}$$

$$= -\sum_\ell e_\ell \exp(-i\boldsymbol{k}\cdot\boldsymbol{r}_\ell) + \sum e_\ell\ , \tag{5.12}$$

$$= -(\rho)_{\boldsymbol{k}} + \sum e_\ell\ , \tag{5.13}$$

which shows the validity of $\nabla \cdot \boldsymbol{P} = -\rho$ for charge neutral systems ($\sum e_\ell = 0$).

The Fourier component of $\partial \boldsymbol{P}_{\boldsymbol{k}}/\partial t$ is

$$\begin{aligned}\frac{\partial \boldsymbol{P}_{\boldsymbol{k}}}{\partial t} &= \frac{\partial}{\partial t}\int_0^1 \mathrm{d}u \sum_\ell e_\ell \boldsymbol{r}_\ell \exp(-iu\boldsymbol{k}\cdot\boldsymbol{r}_\ell) \\ &= \int_0^1 \mathrm{d}u \sum_\ell e_\ell \left\{\boldsymbol{v}_\ell - iu\boldsymbol{r}_\ell(\boldsymbol{k}\cdot\boldsymbol{v}_\ell)\right\} \exp(-iu\boldsymbol{k}\cdot\boldsymbol{r}_\ell)\end{aligned} \tag{5.14}$$

The Fourier component of $c\nabla \times \boldsymbol{M}$ is

$$\begin{aligned}(c\nabla \times \boldsymbol{M})_{\boldsymbol{k}} &= i\int_0^1 u\mathrm{d}u \sum_\ell e_\ell \boldsymbol{k} \times (\boldsymbol{r}_\ell \times \boldsymbol{v}_\ell) \exp(-iu\boldsymbol{k}\cdot\boldsymbol{r}_\ell)\ , \\ &\Big[= (\nabla \times \boldsymbol{M})_{\boldsymbol{k}}\Big]_{\mathrm{SI}}\end{aligned} \tag{5.15}$$

The vector triple product is rewritten as

$$\boldsymbol{k} \times (\boldsymbol{r}_\ell \times \boldsymbol{v}_\ell) = (\boldsymbol{k}\cdot\boldsymbol{v}_\ell)\,\boldsymbol{r}_\ell - (\boldsymbol{k}\cdot\boldsymbol{r}_\ell)\,\boldsymbol{v}_\ell \tag{5.16}$$

The contribution of the first term on the r.h.s. cancels the second term in the braces of $\partial \boldsymbol{P}_{\boldsymbol{k}}/\partial t$, and the remaining contribution is rewritten as

$$-\,i\int_0^1 u\mathrm{d}u \sum_\ell e_\ell(\boldsymbol{k}\cdot\boldsymbol{r}_\ell)\,\boldsymbol{v}_\ell \exp(-iu\boldsymbol{k}\cdot\boldsymbol{r}_\ell) = \sum_\ell e_\ell \boldsymbol{v}_\ell \int_0^1 u\mathrm{d}u\,\frac{\mathrm{d}}{\mathrm{d}u}\exp(-iu\boldsymbol{k}\cdot\boldsymbol{r}_\ell) \tag{5.17}$$

which, via partial integration, leads to

$$\sum_\ell e_\ell \boldsymbol{v}_\ell \left\{\exp(-i\boldsymbol{k}\cdot\boldsymbol{r}_\ell) - \int_0^1 \mathrm{d}u\ \exp(-iu\boldsymbol{k}\cdot\boldsymbol{r}_\ell)\right\}\ . \tag{5.18}$$

The second term on the r.h.s. cancels the remaining first term on the r.h.s. of (5.14), and the final result is

$$\begin{aligned}\frac{\partial \boldsymbol{P}_{\boldsymbol{k}}}{\partial t} + ic\boldsymbol{k} \times \boldsymbol{M}_{\boldsymbol{k}} &= \sum_\ell e_\ell \boldsymbol{v}_\ell\ \exp(-i\boldsymbol{k}\cdot\boldsymbol{r}_\ell) = \boldsymbol{J}_{\boldsymbol{k}}\ , \\ &\left[= \frac{\partial \boldsymbol{P}_{\boldsymbol{k}}}{\partial t} + i\boldsymbol{k} \times \boldsymbol{M}_{\boldsymbol{k}} = \boldsymbol{J}_{\boldsymbol{k}}\right]_{\mathrm{SI}}\end{aligned} \tag{5.19}$$

which is the Fourier representation of (5.10).

5.2 Equations of Motion Obtained from Lagrangian *L*

The Lagrangian for interacting EM field and charged particles in general is

$$L = \sum_{\ell} \left\{ \frac{1}{2} m_\ell v_\ell^2 - e_\ell \phi(\boldsymbol{r}_\ell) + \frac{e_\ell}{c} \boldsymbol{v}_\ell \cdot \boldsymbol{A}(\boldsymbol{r}_\ell) \right\} + L_{\text{EM}}$$

$$\left[= \sum_{\ell} \left\{ \frac{1}{2} m_\ell v_\ell^2 - e_\ell \phi(\boldsymbol{r}_\ell) + e_\ell \boldsymbol{v}_\ell \cdot \boldsymbol{A}(\boldsymbol{r}_\ell) \right\} + L_{\text{EM}} \right]_{\text{SI}} \tag{5.20}$$

where

$$L_{\text{EM}} = \int \mathrm{d}\boldsymbol{r}\, \frac{1}{8\pi} \left\{ \left(\frac{1}{c} \frac{\partial \boldsymbol{A}}{\partial t} + \nabla \phi \right)^2 - (\nabla \times \boldsymbol{A})^2 \right\},$$

$$\left[= \frac{1}{2} \int \mathrm{d}\boldsymbol{r} \left\{ \varepsilon_0 \left(\frac{\partial \boldsymbol{A}}{\partial t} + \nabla \phi \right)^2 - \frac{1}{\mu_0} (\nabla \times \boldsymbol{A})^2 \right\} \right]_{\text{SI}} \tag{5.21}$$

is the Lagrangian for vacuum EM field. In such a system, each charged particle feels the Lorentz force acting at its position, and the EM field should be determined by the charge and current densities of matter. The explicit forms of the equations to describe such situations are derived from the least action principle for the Lagrangian, or the Lagrange equations. The generalized coordinates for this derivation are the coordinates of the particles $\{\boldsymbol{r}_\ell\}$, vector potential $\boldsymbol{A}(\boldsymbol{r})$ and scalar potential $\phi(\boldsymbol{r})$.

The action for a Lagrangian is defined as

$$S = \int \mathrm{d}t\, L \tag{5.22}$$

which is a functional of the generalized coordinates. To consider a change in the action for a generalized coordinate $q(t)$, we allow a small variation of $q(t)$ between a certain time interval, but fix the values of $q(t)$ at the both ends of the interval. Denoting the physically allowed path as $\bar{q}(t)$ and the deviation from it as $\delta q(t)$, we request that the difference $S[\bar{q} + \delta q] - S[\bar{q}]$ should vanish in the first order of δq. This requirement leads to an equation fulfilled by $\bar{q}$, which is the Lagrange equation for the generalized coordinate q.

5.2.1 Newton Equation for a Charged Particle Under Lorentz Force

First of all, let us take x_i as $\bar{q}$, i.e., the x coordinate of the ith particle. Then, the difference $S[\bar{q} + \delta q] - S[\bar{q}]$ consists of three terms, those due to the kinetic energy term δS_{kin}, scalar potential δS_{sp}, and vector potential δS_{vp}. The first two are easily calculated, to the first order in δx_i, as

$$\begin{aligned}\delta S_{\text{kin}} &= \int \mathrm{d}t \frac{m_i}{2} \left[\left\{ \frac{d}{dt}(x_i + \delta x_i) \right\}^2 - \left(\frac{dx_i}{dt} \right)^2 \right] = \int \mathrm{d}t \left[m_i \frac{dx_i}{dt} \frac{d\delta x_i}{dt} \right] , \\ &= -m_i \int \mathrm{d}t \, \frac{d^2 x_i}{dt^2} \, \delta x_i \ , \end{aligned} \tag{5.23}$$

$$\delta S_{\text{sp}} = -e_i \int \mathrm{d}t \left[\phi(x_i + \delta x_i) - \phi(x_i) \right] = -e_i \int \mathrm{d}t \, \frac{\partial \phi}{\partial x_i} \, \delta x_i \ . \tag{5.24}$$

The last equation for δS_{kin} is obtained via partial integration. (Since $\delta x_i(t)$ is zero for the both ends of integration, no term appears from the boundaries of integral.) The third one δS_{vp} is a little complicated. The increment of $\sum_\ell (e_\ell/c)\boldsymbol{v}_\ell \cdot \boldsymbol{A}$, when only one coordinate component is changed from x_i to $x_i + \delta x_i$, is

$$\begin{aligned} &\frac{e_i}{c} \left[\left\{ \frac{d}{dt}(x_i + \delta x_i) \right\} A_x(x_i + \delta x_i) + \frac{dy_i}{dt} A_y(x_i + \delta x_i) + \frac{dz_i}{dt} A_z(x_i + \delta x_i) \right. \\ &\qquad \left. - \left\{ \frac{dx_i}{dt} A_x(x_i) + \frac{dy_i}{dt} A_y(x_i) + \frac{dz_i}{dt} A_z(x_i) \right\} \right] \\ &= \frac{e_i}{c} \frac{d\delta x_i}{dt} A_x(x_i) + \frac{e_i}{c} \left\{ \frac{dx_i}{dt} \frac{\partial A_x}{\partial x_i} + \frac{dy_i}{dt} \frac{\partial A_y}{\partial x_i} + \frac{dz_i}{dt} \frac{\partial A_z}{\partial x_i} \right\} \delta x_i \ , \\ &\left[= e_i \frac{d\delta x_i}{dt} A_x(x_i) + e_i \left\{ \frac{dx_i}{dt} \frac{\partial A_x}{\partial x_i} + \frac{dy_i}{dt} \frac{\partial A_y}{\partial x_i} + \frac{dz_i}{dt} \frac{\partial A_z}{\partial x_i} \right\} \delta x_i \right]_{\text{SI}} \end{aligned} \tag{5.25}$$

where, the arguments of $\boldsymbol{A}$ without variation are not explicitly written, i.e., $A_x(x_i + \delta x_i) = A_x(x_i + \delta x_i, y_i, z_i, t)$, $A_y(x_i + \delta x_i) = A_x(x_i + \delta x_i, y_i, z_i, t)$, etc., and the derivatives are evaluated at (x_i, y_i, z_i, t). The time integral of the first term on the r.h.s. in calculating δS_{vp} gives

$$\int \mathrm{d}t \frac{d\delta x_i}{dt} A_x = -\int \mathrm{d}t \, \frac{dA_x}{dt} \delta x_i \tag{5.26}$$

where we made partial integration with the fixed values of $x_i(t)$ at the lower and upper ends of the integration. Note that the time evolution of $A_x(x_i, y_i, z_i, t)$ occurs through the explicit t-dependence of A_x and that of x_i, y_i, z_i, i.e.,

$$\frac{dA_x}{dt} = \frac{\partial A_x}{\partial t} + \left(\frac{dx_i}{dt} \frac{\partial A_x}{\partial x_i} + \frac{dy_i}{dt} \frac{\partial A_x}{\partial y_i} + \frac{dz_i}{dt} \frac{\partial A_x}{\partial z_i} \right) . \tag{5.27}$$

Using these preliminary results, we can calculate

$$\begin{aligned}\delta S_{\mathrm{vp}} = & -\int \mathrm{d}t \frac{e_i}{c} \frac{\partial A_x}{\partial t} \delta x_i \\ & + \int \mathrm{d}t \frac{e_i}{c} \delta x_i \left[\left\{ \frac{dx_i}{dt}\frac{\partial A_x}{\partial x_i} + \frac{dy_i}{dt}\frac{\partial A_y}{\partial x_i} + \frac{dz_i}{dt}\frac{\partial A_z}{\partial x_i} \right\} \right. \\ & \left. - \left(\frac{dx_i}{dt}\frac{\partial A_x}{\partial x_i} + \frac{dy_i}{dt}\frac{\partial A_x}{\partial y_i} + \frac{dz_i}{dt}\frac{\partial A_x}{\partial z_i} \right) \right] \qquad (5.28)\end{aligned}$$

$$\begin{aligned} = & -\int \mathrm{d}t \frac{e_i}{c} \frac{\partial A_x}{\partial t} \delta x_i \\ & + \int \mathrm{d}t \frac{e_i}{c} \delta x_i \left[\frac{dy_i}{dt} \left(\frac{\partial A_y}{\partial x_i} - \frac{\partial A_x}{\partial y_i} \right) - \frac{dz_i}{dt} \left(\frac{\partial A_x}{\partial z_i} - \frac{\partial A_z}{\partial x_i} \right) \right] \qquad (5.29)\end{aligned}$$

$$\begin{aligned} = & -\int \mathrm{d}t \frac{e_i}{c} \frac{\partial A_x}{\partial t} \delta x_i \\ & + \int dt \frac{e_i}{c} \delta x_i \left[\frac{dy_i}{dt} (\nabla \times \boldsymbol{A})_z - \frac{dz_i}{dt} (\nabla \times \boldsymbol{A})_y \right] \qquad (5.30)\end{aligned}$$

$$= -\int dt \frac{e_i}{c} \frac{\partial A_x}{\partial t} \delta x_i + \int dt \frac{e_i}{c} (\boldsymbol{v}_i \times \boldsymbol{B})_x \, \delta x_i \qquad (5.31)$$

$$\left[\delta S_{\mathrm{vp}} = e_i \int \mathrm{d}t \, \frac{\partial A_x}{\partial t} \delta x_i + e_i \int \mathrm{d}t \, (\boldsymbol{v}_i \times \boldsymbol{B})_x \, \delta x_i \right]_{\mathrm{SI}} \qquad (5.32)$$

Denoting the sum $\delta S_{\mathrm{kin}} + \delta S_{\mathrm{sp}} + \delta S_{\mathrm{vp}}$ as δS_{all}, we have

$$\begin{aligned}\delta S_{\mathrm{all}} &= \int \mathrm{d}t \, \delta x_i \left[-m_i \frac{d^2 x_i}{dt^2} + e_i \left(-\frac{\partial \phi}{\partial x_i} - \frac{1}{c}\frac{\partial A_x}{\partial t} \right) + \frac{e_i}{c} (\boldsymbol{v}_i \times \boldsymbol{B})_x \right], \\ &= \int \mathrm{d}t \, \delta x_i \left[-m_i \frac{d^2 x_i}{dt^2} + e_i \left(\boldsymbol{E} + \frac{\boldsymbol{v}_i}{c} \times \boldsymbol{B} \right)_x \right], \\ & \left[\delta S_{\mathrm{all}} = \int \mathrm{d}t \, \delta x_i \left\{ -m_i \frac{d^2 x_i}{dt^2} + e_i (\boldsymbol{E} + \boldsymbol{v}_i \times \boldsymbol{B})_x \right\} \right]_{\mathrm{SI}} \qquad (5.33)\end{aligned}$$

In order for this increment to vanish for arbitrary δx_i, the $[\cdots]$ part of the integrand should be zero, i.e.,

$$\begin{aligned} m_i \frac{d^2 x_i}{dt^2} &= e_i \left(\boldsymbol{E} + \frac{\boldsymbol{v}_i}{c} \times \boldsymbol{B} \right)_x \\ & \left[= e_i (\boldsymbol{E} + \boldsymbol{v}_i \times \boldsymbol{B})_x \right]_{\mathrm{SI}} \qquad (5.34)\end{aligned}$$

which is the Newton equation of motion for a charged particle under Lorentz force.

5.2.2 Equations of Motion for ϕ and A

To calculate the variations for field variables ϕ and $\boldsymbol{A}$, we rewrite the interaction terms as

$$L_{\text{int}} = \sum_{\ell} \left\{ -e_\ell \phi(\boldsymbol{r}_\ell) + \frac{e_\ell}{c} \boldsymbol{v}_\ell \cdot \boldsymbol{A}(\boldsymbol{r}_\ell) \right\} = \int \mathrm{d}\boldsymbol{r} \, \left\{ -\rho(\boldsymbol{r})\phi(\boldsymbol{r}) + \frac{1}{c} \boldsymbol{J}(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r}) \right\}$$
$$\left[= \sum_{\ell} \{ -e_\ell \phi(\boldsymbol{r}_\ell) + e_\ell \boldsymbol{v}_\ell \cdot \boldsymbol{A}(\boldsymbol{r}_\ell) \} = \int \mathrm{d}\boldsymbol{r} \, \left\{ -\rho(\boldsymbol{r})\phi(\boldsymbol{r}) + \boldsymbol{J}(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r}) \right\} \right]_{\text{SI}} \tag{5.35}$$

where the charge and current densities, ρ and $\boldsymbol{J}$ are defined as (1.14) and (1.15). The action for the "ϕ, $\boldsymbol{A}$" related part of the Lagrangian is a double integral over t and $\boldsymbol{r}$. The small variation δq is an arbitrary function of t and $\boldsymbol{r}$ except that they are fixed to zero at the upper and lower limits of the integrations.

Let us first consider the variation of scalar potential from ϕ to $\phi + \delta\phi$. The increment of the action δS_ϕ due to this variation is given from the action for $L_{\text{int}} + L_{\text{EM}}$ as

$$\delta S_\phi = \int \mathrm{d}t \int \mathrm{d}\boldsymbol{r} \, \left[-\rho \, \delta\phi + \frac{1}{4\pi c} \, \nabla\delta\phi \cdot \frac{\partial \boldsymbol{A}}{\partial t} + \frac{1}{4\pi} \, \nabla\phi \cdot \nabla\delta\phi \right] \tag{5.36}$$
$$= - \int \mathrm{d}t \int \mathrm{d}\boldsymbol{r} \, \left[\rho + \frac{1}{4\pi c} \nabla \cdot \frac{\partial \boldsymbol{A}}{\partial t} + \frac{1}{4\pi} \, \nabla^2 \phi \right] \, \delta\phi$$
$$\left[\delta S_\phi = - \int \mathrm{d}t \int \mathrm{d}\boldsymbol{r} \, \left[\rho + \varepsilon_0 \nabla \cdot \frac{\partial \boldsymbol{A}}{\partial t} + \varepsilon_0 \, \nabla^2 \phi \right] \, \delta\phi \right]_{\text{SI}} \tag{5.37}$$

The condition for this increment to vanish for arbitrary $\delta\phi$ is the vanishing of $[\cdots]$ in the integrand

$$\rho + \frac{1}{4\pi} \, \nabla^2 \phi + \frac{1}{4\pi c} \nabla \cdot \frac{\partial \boldsymbol{A}}{\partial t} = 0 \, ,$$
$$\left[\rho + \varepsilon_0 \left(\nabla^2 \phi + \nabla \cdot \frac{\partial \boldsymbol{A}}{\partial t} \right) = 0 \right]_{\text{SI}} \tag{5.38}$$

which is $\nabla \cdot \boldsymbol{E} = 4\pi\rho$, (1.3). For the Coulomb gauge, $\nabla \cdot \boldsymbol{A} = 0$, it is the Poisson equation

$$\nabla^2 \phi = -4\pi\rho \quad \left[= -\frac{1}{\varepsilon_0} \rho \right]_{\text{SI}} \tag{5.39}$$

and for the Lorentz gauge, $\nabla \cdot \boldsymbol{A} + (1/c)\partial\phi/\partial t = 0$, it is the wave equation for ϕ

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi = 4\pi\rho \quad \left[= \frac{1}{\varepsilon_0} \rho \right]_{\text{SI}} \, . \tag{5.40}$$

The variation of vector potential leads to the change in action

$$\begin{aligned}
\delta S_{\mathrm{A}} = \int \mathrm{d}t \int \mathrm{d}\boldsymbol{r} \Bigg[& \frac{1}{c}\boldsymbol{J}\cdot\delta\boldsymbol{A} + \frac{1}{4\pi c}\nabla\phi\cdot\frac{\partial\delta\boldsymbol{A}}{\partial t} \\
& + \frac{1}{8\pi c^2}\left\{\left(\frac{\partial(\boldsymbol{A}+\delta\boldsymbol{A})}{\partial t}\right)^2 - \left(\frac{\partial\boldsymbol{A}}{\partial t}\right)^2\right\} \\
& - \frac{1}{8\pi}\{(\nabla\times\boldsymbol{A}+\nabla\times\delta\boldsymbol{A})^2 - (\nabla\times\boldsymbol{A})^2\}\Bigg], \qquad (5.41) \\
= \int \mathrm{d}t \int \mathrm{d}\boldsymbol{r} \Bigg[& \frac{1}{c}\boldsymbol{J}\cdot\delta\boldsymbol{A} - \frac{1}{4\pi c}\frac{\partial\nabla\phi}{\partial t}\delta\boldsymbol{A} \\
& + \frac{1}{4\pi c^2}\frac{\partial\boldsymbol{A}}{\partial t}\cdot\frac{\partial\delta\boldsymbol{A}}{\partial t} - \frac{1}{4\pi}\nabla\times\boldsymbol{A}\cdot(\nabla\times\delta\boldsymbol{A})\Bigg] \\
= \int \mathrm{d}t \int \mathrm{d}\boldsymbol{r} \Bigg[& \frac{1}{c}\boldsymbol{J} - \frac{1}{4\pi c}\frac{\partial\nabla\phi}{\partial t} - \frac{1}{4\pi c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2} - \frac{1}{4\pi}\nabla\times\nabla\times\boldsymbol{A}\Bigg]\cdot\delta\boldsymbol{A}
\end{aligned}$$

$$\left[\delta S_{\mathrm{A}} = \int \mathrm{d}t \int \mathrm{d}\boldsymbol{r} \left[\boldsymbol{J} - \varepsilon_0\frac{\partial\nabla\phi}{\partial t} - \varepsilon_0\frac{\partial^2\boldsymbol{A}}{\partial t^2} - \frac{1}{\mu_0}\nabla\times\nabla\times\boldsymbol{A}\right]\cdot\delta\boldsymbol{A}\right]_{\mathrm{SI}} \qquad (5.42)$$

where we have used partial integration for t and $\boldsymbol{r}$ variables to rewrite the terms containing $\partial\nabla\phi/\partial t$, $\partial\delta\boldsymbol{A}/\partial t$, and $\nabla\times\delta\boldsymbol{A}$. The condition for δS_{A} to be zero for arbitrary $\delta\boldsymbol{A}$ is the vanishing of the $[\cdots]$ part of the integrand

$$\frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2} - \nabla^2\boldsymbol{A} + \nabla\left(\nabla\cdot\boldsymbol{A} + \frac{1}{c}\frac{\partial\phi}{\partial t}\right) = \frac{4\pi}{c}\boldsymbol{J}, \qquad (5.43)$$

$$\left[\frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2} - \nabla^2\boldsymbol{A} + \nabla\left(\nabla\cdot\boldsymbol{A} + \frac{1}{c^2}\frac{\partial\phi}{\partial t}\right) = \mu_0\,\boldsymbol{J}\right]_{\mathrm{SI}} \qquad (5.44)$$

where we have used $\nabla\times\nabla\times\boldsymbol{A} = \nabla\nabla\cdot\boldsymbol{A} - \nabla^2\boldsymbol{A}$. For the Coulomb gauge, it is

$$\frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2} - \nabla^2\boldsymbol{A} + \nabla\frac{1}{c}\frac{\partial\phi}{\partial t} = \frac{4\pi}{c}\boldsymbol{J} \quad \left[\, = \mu_0\,\boldsymbol{J}\right]_{\mathrm{SI}} \qquad (5.45)$$

and for the Lorentz gauge it is a simple wave equation for $\boldsymbol{A}$

$$\frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2} - \nabla^2\boldsymbol{A} = \frac{4\pi}{c}\boldsymbol{J} \quad \left[\, = \mu_0\,\boldsymbol{J}\right]_{\mathrm{SI}} \qquad (5.46)$$

5.2.3 Generalized Momenta and Hamiltonian

The generalized momentum p conjugate to the generalized coordinate q is defined by $p = \partial L/\partial \dot{q}$, where $\dot{q} = \mathrm{d}q/\mathrm{d}t$, and the Hamiltonian is defined as

$$H = \sum p \frac{\mathrm{d}q}{\mathrm{d}t} - L \,, \tag{5.47}$$

where we omitted the suffices of the generalized coordinates and momenta to distinguish the particle number and Cartesian components, and the summation is meant for these suffices. According to this rule of analytic mechanics, the momenta $\boldsymbol{p}_\ell$, P_ϕ, P_A conjugate to the variables (generalized coordinates) $\boldsymbol{r}_\ell$, ϕ, and $\boldsymbol{A}$, respectively, are

$$\begin{aligned} &\boldsymbol{p}_\ell = m_\ell \boldsymbol{v}_\ell + \frac{e_\ell}{c}\boldsymbol{A}(\boldsymbol{r}_\ell), \quad P_\phi = 0, \quad \boldsymbol{\Pi} = \frac{1}{4\pi c}\left(\frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} + \nabla\phi\right) \,, \\ &\left[\boldsymbol{p}_\ell = m_\ell \boldsymbol{v}_\ell + e_\ell \boldsymbol{A}(\boldsymbol{r}_\ell), \quad P_\phi = 0, \quad \boldsymbol{\Pi} = \varepsilon_0 \left(\frac{\partial \boldsymbol{A}}{\partial t} + \nabla\phi\right)\right]_{\mathrm{SI}} \end{aligned} \tag{5.48}$$

and the Hamiltonian is

$$\begin{aligned} H &= \sum_\ell \boldsymbol{p}_\ell \cdot \boldsymbol{v}_\ell + \frac{1}{4\pi c}\int \mathrm{d}\boldsymbol{r} \, \left(\frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} + \nabla\phi\right) \frac{\partial \boldsymbol{A}}{\partial t} - L \,, \\ &= \sum_\ell \frac{m_\ell v_\ell^2}{2} + \frac{1}{8\pi}\int \mathrm{d}\boldsymbol{r} \, \left\{\left(\frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} + \nabla\phi\right)^2 + (\nabla \times \boldsymbol{A})^2\right\} \\ &\left[= \sum_\ell \frac{m_\ell v_\ell^2}{2} + \frac{1}{2}\int \mathrm{d}\boldsymbol{r} \left\{\varepsilon_0 \left(\frac{\partial \boldsymbol{A}}{\partial t} + \nabla\phi\right)^2 + \frac{1}{\mu_0}(\nabla \times \boldsymbol{A})^2\right\}\right]_{\mathrm{SI}} \end{aligned} \tag{5.49}$$

The manipulation from the first to the second line is made by rewriting the second term of the first line as

$$\frac{1}{4\pi}\int \mathrm{d}\boldsymbol{r} \left(\frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} + \nabla\phi\right)^2 - \frac{1}{4\pi}\int \mathrm{d}\boldsymbol{r} \left(\frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} + \nabla\phi\right) \cdot \nabla\phi \tag{5.50}$$

$$= \frac{1}{4\pi}\int \mathrm{d}\boldsymbol{r}[\boldsymbol{E}^2 + \boldsymbol{E} \cdot \nabla\phi] \tag{5.51}$$

$$= \frac{1}{4\pi}\int \mathrm{d}\boldsymbol{r}[\boldsymbol{E}^2 - \nabla \cdot \boldsymbol{E} \, \phi] \tag{5.52}$$

$$= \frac{1}{4\pi}\int \mathrm{d}\boldsymbol{r}\boldsymbol{E}^2 - \int \mathrm{d}\boldsymbol{r}\rho\phi \tag{5.53}$$

The second term on the r.h.s. cancels the corresponding term in L, and a half of the first term cancels the vacuum E field energy. The first term on the r.h.s. of (5.49) is

$$\sum_\ell \boldsymbol{p}_\ell \cdot \boldsymbol{v}_\ell = \sum_\ell \left\{ m_\ell \boldsymbol{v}_\ell^2 + \frac{e_\ell}{c} \boldsymbol{A}(\boldsymbol{r}_\ell) \cdot \boldsymbol{v}_\ell \right\} . \tag{5.54}$$

The second term and a half of the first term cancel the corresponding terms of L, and the remaining terms gives the second line of (5.49).

The last integral of (5.49) is the energy of EM field. Its T and L components are

$$\begin{aligned} H_{\mathrm{EM}}^{(\mathrm{T})} &= \frac{1}{8\pi} \int \mathrm{d}\boldsymbol{r}\ \{(4\pi c \boldsymbol{\Pi}^{(\mathrm{T})})^2 + (\nabla \times \boldsymbol{A}^{(\mathrm{T})})^2\} \\ &\left[= \frac{1}{2} \int \mathrm{d}\boldsymbol{r}\ \left\{ \frac{1}{\varepsilon_0} [\boldsymbol{\Pi}^{(\mathrm{T})}]^2 + \frac{1}{\mu_0} (\nabla \times \boldsymbol{A}^{(\mathrm{T})})^2 \right\} \right]_{\mathrm{SI}} \end{aligned} \tag{5.55}$$

$$\begin{aligned} H_{\mathrm{EM}}^{(\mathrm{L})} &= \frac{1}{8\pi} \int \mathrm{d}\boldsymbol{r}\ \left(\frac{1}{c} \frac{\partial \boldsymbol{A}^{(\mathrm{L})}}{\partial t} + \nabla\phi \right)^2 = \frac{1}{8\pi} \int \mathrm{d}\boldsymbol{r}\ \{\boldsymbol{E}^{(\mathrm{L})}\}^2 \\ &\left[= \frac{\varepsilon_0}{2} \int \mathrm{d}\boldsymbol{r}\ \left(\frac{\partial \boldsymbol{A}^{(\mathrm{L})}}{\partial t} + \nabla\phi \right)^2 = \frac{\varepsilon_0}{2} \int \mathrm{d}\boldsymbol{r}\ \{\boldsymbol{E}^{(\mathrm{L})}\}^2 \right]_{\mathrm{SI}} \end{aligned} \tag{5.56}$$

Using the Gauss law $\nabla \cdot \boldsymbol{E}^{(\mathrm{L})} = 4\pi\rho$, we can show that the L component $H_{\mathrm{EM}}^{(\mathrm{L})}$ is the Coulomb potential among the particles (Sect. 1.2), i.e.,

$$H_{\mathrm{EM}}^{(\mathrm{L})} = \frac{1}{2} \sum_\ell \sum_{\ell'} \frac{e_\ell e_{\ell'}}{|\boldsymbol{r}_\ell - \boldsymbol{r}_{\ell'}|} \quad \left[= \frac{1}{8\pi\varepsilon_0} \sum_\ell \sum_{\ell'} \frac{e_\ell e_{\ell'}}{|\boldsymbol{r}_\ell - \boldsymbol{r}_{\ell'}|} \right]_{\mathrm{SI}} \tag{5.57}$$

Then, we have the total Hamiltonian in the following form

$$\begin{aligned} H &= \sum_\ell \frac{1}{2m_\ell} \left\{ \boldsymbol{p}_\ell - \frac{e_\ell}{c} \boldsymbol{A}(\boldsymbol{r}_\ell) \right\}^2 + \frac{1}{2} \sum_\ell \sum_{\ell'} \frac{e_\ell e_{\ell'}}{|\boldsymbol{r}_\ell - \boldsymbol{r}_{\ell'}|} \\ &\quad + \frac{1}{8\pi} \int \mathrm{d}\boldsymbol{r}\ [\{4\pi c\ \boldsymbol{\Pi}^{(\mathrm{T})}(\boldsymbol{r})\}^2 + \{\nabla \times \boldsymbol{A}^{(\mathrm{T})}(\boldsymbol{r})\}^2] , \\ &\left[= \sum_\ell \frac{1}{2m_\ell} \{\boldsymbol{p}_\ell - e_\ell \boldsymbol{A}(\boldsymbol{r}_\ell)\}^2 + \frac{1}{8\pi\varepsilon_0} \sum_\ell \sum_{\ell'} \frac{e_\ell e_{\ell'}}{|\boldsymbol{r}_\ell - \boldsymbol{r}_{\ell'}|} \right. \\ &\quad \left. + \frac{1}{2} \int \mathrm{d}\boldsymbol{r}\ \left\{ \frac{1}{\varepsilon_0} [\boldsymbol{\Pi}^{(\mathrm{T})}]^2 + \frac{1}{\mu_0} (\nabla \times \boldsymbol{A}^{(\mathrm{T})})^2 \right\} \right]_{\mathrm{SI}} \end{aligned} \tag{5.58}$$

which is valid for any gauge. The L field is contained in both $\boldsymbol{A}$ of the first term and the Coulomb potential, and the remaining part of EM field is written by the conjugate variables of the T components, $\boldsymbol{A}^{(\mathrm{T})}$ and $\boldsymbol{\Pi}^{(\mathrm{T})}$.

It should be noted that a gauge transformation determines the way to divide the L field into the contributions of $\boldsymbol{A}$ and ϕ without affecting the T field. This allows us to make a gauge independent definition of the Hamiltonians of matter and (T) EM field in a usual way, i.e., the sum of kinetic energy and Coulomb potential for matter, and $H_{\mathrm{EM}}^{(\mathrm{T})}$ for the (T) EM field. This definition is very common to most studies in non-relativistic regime. The choice of gauge is made to facilitate the treatment of the interaction between matter and L field. The Coulomb gauge is simple in the sense that T and L field is cleanly separated as $\boldsymbol{A}$ and ϕ, respectively, and, if an incident L field does not exist, the L field is considered automatically by the proper treatment of Coulomb potential. In this sense, we adopted the Coulomb gauge in most part of this book. The case of L field incidence is treated in Sect. 5.7.

The Coulomb potential in (1.33) contains the summation over $\ell = \ell'$, too. It is the self-interaction energy of each charged particle. For a point charge, it is an infinitely large quantity, and it is finite if a particle size is finite. In the non-relativistic treatment, we just neglect these terms, since it is a (large) quantity attached to each particle separately, independent of the inter-particle behavior. In this way, we arrive at the usual form of Coulomb potential term for a charged particle system

$$U_{\mathrm{C}} = \frac{1}{2}\sum_{\ell}\sum_{\ell'\neq\ell} \frac{e_\ell e_{\ell'}}{|\boldsymbol{r}_\ell - \boldsymbol{r}_{\ell'}|} \quad \left[= \frac{1}{8\pi\varepsilon_0}\sum_{\ell}\sum_{\ell'\neq\ell} \frac{e_\ell e_{\ell'}}{|\boldsymbol{r}_\ell - \boldsymbol{r}_{\ell'}|} \right]_{\mathrm{SI}} . \tag{5.59}$$

5.3 Form of Interaction Term

5.3.1 Another Set of Lagrangian and Hamiltonian

In the main text, we used the Hamiltonians for matter, radiation, and their mutual interaction in the Coulomb gauge, as given in Sect. 5.2. It is assumed that there is no external charge density, so that the treatment applies only the external excitation by T field. (The case of the external excitation by L field is given in Sect. 5.7.) The interaction is described by the current density and vector potential, so that any linear response due to this interaction gives an induced change linear in $\boldsymbol{A}$. In order to calculate the conventional electric and magnetic susceptibilities, $\chi_{\mathrm{e}},\ \chi_{\mathrm{m}}$, via $\boldsymbol{P} = \chi_{\mathrm{e}}\boldsymbol{E},\ \boldsymbol{M} = \chi_{\mathrm{e}}\boldsymbol{H}$, we obviously need a new set of Lagrangian and Hamiltonian, where the interaction term is linear in $\boldsymbol{E}$ and $\boldsymbol{H}$ explicitly. (In the case of L field excitation, the interaction between the internal charge density and $\boldsymbol{E}^{(\mathrm{L})}$ is already written as $-\int \boldsymbol{E}^{(\mathrm{L})} \cdot \boldsymbol{P}^{(\mathrm{L})} \mathrm{d}\boldsymbol{r}$, as shown in Sect. 5.7, so that the consideration of this case is omitted from the argument given below.)

Knowing that the Lagrangian in Sect. 2.2 is the sound basis for general systems of interacting matter-EM field, we would need a unitary transformation which rewrites the interaction term $\boldsymbol{J} \cdot \boldsymbol{A}$ into the types like $\boldsymbol{P} \cdot \boldsymbol{E}$ and/or $\boldsymbol{M} \cdot \boldsymbol{H}$. However, no such a transformation is known as a rigorous theory. It is known that the interaction Hamiltonian $-\boldsymbol{E} \cdot \int \mathrm{d}\boldsymbol{r}\ \boldsymbol{P}$ is obtained via a unitary transformation based on the electric dipole approximation, or LWA, (see e.g., p. 304 of [1]). Because of the LWA assuming the uniformity of $\boldsymbol{A}$ (i.e., $\nabla \times \boldsymbol{A} = 0$) one cannot extend this argument to determine the magnetic counter part $-\boldsymbol{H} \cdot \int \mathrm{d}\boldsymbol{r}\ \boldsymbol{M}$.

A hint to proceed is obtained by the following argument. If we use the operator identity $\boldsymbol{J} = \partial \boldsymbol{P}/\partial t + c\nabla \times \boldsymbol{M}$, the interaction term $\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{J} \cdot \boldsymbol{A}$ is rewritten as

$$\frac{1}{c}\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{J} \cdot \boldsymbol{A} = \frac{1}{c}\int \mathrm{d}\boldsymbol{r}\ \frac{\partial \boldsymbol{P}}{\partial t} \cdot \boldsymbol{A} + \int \mathrm{d}\boldsymbol{r}\ (\nabla \times \boldsymbol{M}) \cdot \boldsymbol{A}$$

$$\left[\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{J} \cdot \boldsymbol{A} = \int \mathrm{d}\boldsymbol{r}\ \frac{\partial \boldsymbol{P}}{\partial t} \cdot \boldsymbol{A} + \int \mathrm{d}\boldsymbol{r}\ (\nabla \times \boldsymbol{M}) \cdot \boldsymbol{A}\right]_{\mathrm{SI}} \tag{5.60}$$

The first and second terms on the r.h.s. are written as the invariants from the inner products of polar and axial vectors, respectively, which are distinguishable for systems with inversion symmetry. The second term on the r.h.s. is rewritten, by partial integration, into $\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{M} \cdot \boldsymbol{B}$. Therefore, the magnetic interaction term is $\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{M} \cdot \boldsymbol{B}$, rather than $\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{M} \cdot \boldsymbol{H}$. Though the first term is not $\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{P} \cdot \boldsymbol{E}$, it is the same type of invariant made from polar vectors. In this sense, the interaction $(1/c)\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{J} \cdot \boldsymbol{A}$ is divided into two independent terms in inversion symmetric systems, and one of them is the magnetic interaction linear in $\boldsymbol{B}$.

In order to study this point more in detail, we consider the Power-Zienau-Woolley (PZW) transformation [1]. In analytical mechanics, it is well known that the addition of a total time derivative of an arbitrary function F (of generalized coordinates and time) to a Lagrangian L does not change the condition for least action. In PZW transformation we use

$$F = -\frac{1}{c}\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{P}(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r}, t)\ \ \left[= -\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{P}(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r}, t)\right]_{\mathrm{SI}} \tag{5.61}$$

where the operator form of $\boldsymbol{P}(\boldsymbol{r})$ is explicitly given in Sect. 5.1. Then,

$$\frac{\mathrm{d}F}{\mathrm{d}t} = -\frac{1}{c}\int \mathrm{d}\boldsymbol{r}\ \left\{\frac{\partial \boldsymbol{P}(\boldsymbol{r})}{\partial t} \cdot \boldsymbol{A}(\boldsymbol{r}, t) + \boldsymbol{P}(\boldsymbol{r}) \cdot \frac{\partial \boldsymbol{A}(\boldsymbol{r}, t)}{\partial t}\right\},$$

$$\left[= -\int \mathrm{d}\boldsymbol{r}\ \left\{\frac{\partial \boldsymbol{P}(\boldsymbol{r})}{\partial t} \cdot \boldsymbol{A}(\boldsymbol{r}, t) + \boldsymbol{P}(\boldsymbol{r}) \cdot \frac{\partial \boldsymbol{A}(\boldsymbol{r}, t)}{\partial t}\right\}\right]_{\mathrm{SI}} \tag{5.62}$$

where $\partial \boldsymbol{P}/\partial t = \sum_\ell \boldsymbol{v}_\ell(\partial \boldsymbol{P}/\partial \boldsymbol{r}_\ell)$ corresponds to the current density due to $\boldsymbol{P}$ in (5.10). The second term in the integral is $\boldsymbol{P} \cdot \boldsymbol{E}^{(\mathrm{T})}$.

The old Lagrangian can be written as

$$L = \sum_\ell \frac{m_\ell \boldsymbol{v}_\ell^2}{2} - U_C + \frac{1}{c}\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{J}(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r})$$
$$+\frac{1}{8\pi}\int \mathrm{d}\boldsymbol{r}\ \left\{\left(\frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}\right)^2 - (\nabla \times \boldsymbol{A})^2\right\}$$

$$\left[= \sum_\ell \frac{m_\ell \boldsymbol{v}_\ell^2}{2} - U_C + \int \mathrm{d}\boldsymbol{r}\ \boldsymbol{J}(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r})\right.$$

$$+\frac{1}{2}\int \mathrm{d}\boldsymbol{r}\ \left\{\varepsilon_0\left(\frac{\partial \boldsymbol{A}}{\partial t}\right)^2-\frac{1}{\mu_0}\left(\nabla\times\boldsymbol{A}\right)^2\right\}\Bigg]_{\mathrm{SI}} \tag{5.63}$$

where the terms related with the L field (or scalar potential) are written in terms of the inter-particle Coulomb potential U_C, (1.33). The EM field described by $\boldsymbol{A}$ is purely T field. The conjugate momenta for $\boldsymbol{r}_\ell$ and $\boldsymbol{A}(\boldsymbol{r})$ are $m_\ell \boldsymbol{v}_\ell + (e_\ell/c)\boldsymbol{A}(\boldsymbol{r}_\ell)$ and $(1/4\pi c^2)\dot{\boldsymbol{A}}(= -\boldsymbol{E}^{(\mathrm{T})}/4\pi c)$, respectively, which should be compared with those for the new Lagrangian, (5.68) and (5.69).

We can rewrite the sum of $\mathrm{d}F/\mathrm{d}t$ and the interaction as

$$\frac{1}{c}\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{J}(\boldsymbol{r})\cdot\boldsymbol{A}(\boldsymbol{r})+\frac{\mathrm{d}F}{\mathrm{d}t}=\int \mathrm{d}\boldsymbol{r}\ \left\{\boldsymbol{M}\cdot\boldsymbol{B}+\boldsymbol{P}\cdot\boldsymbol{E}^{(\mathrm{T})}\right\},$$

$$\left[\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{J}(\boldsymbol{r})\cdot\boldsymbol{A}(\boldsymbol{r})+\frac{\mathrm{d}F}{\mathrm{d}t}=\int \mathrm{d}\boldsymbol{r}\ \left\{\boldsymbol{M}\cdot\boldsymbol{B}+\boldsymbol{P}\cdot\boldsymbol{E}^{(\mathrm{T})}\right\}\right]_{\mathrm{SI}} \tag{5.64}$$

by the use of (5.10), partial integration, and

$$\boldsymbol{B}=\nabla\times\boldsymbol{A},\quad \boldsymbol{E}^{(\mathrm{T})}=-\frac{1}{c}\frac{\partial\boldsymbol{A}}{\partial t},\quad \left[\boldsymbol{B}=\nabla\times\boldsymbol{A},\quad \boldsymbol{E}^{(\mathrm{T})}=-\frac{\partial\boldsymbol{A}}{\partial t}\right]_{\mathrm{SI}}. \tag{5.65}$$

The (orbital) magnetization $\boldsymbol{M}(\boldsymbol{r})$ is defined by (5.8), so that its product with $\boldsymbol{B}$ can be written as

$$\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{M}\cdot\boldsymbol{B}=\frac{1}{c}\int_0^1 u\ \mathrm{d}u\sum_\ell\ e_\ell\left\{\boldsymbol{B}(u\boldsymbol{r}_\ell)\times\boldsymbol{r}_\ell\right\}\cdot\boldsymbol{v}_\ell$$

$$\left[=\int_0^1 u\ \mathrm{d}u\sum_\ell\ e_\ell\left\{\boldsymbol{B}(u\boldsymbol{r}_\ell)\times\boldsymbol{r}_\ell\right\}\cdot\boldsymbol{v}_\ell\right]_{\mathrm{SI}} \tag{5.66}$$

The new Lagrangian is

$$L'=\sum_\ell\frac{m_\ell\boldsymbol{v}_\ell^2}{2}-U_\mathrm{C}+\int \mathrm{d}\boldsymbol{r}\ \{\boldsymbol{M}\cdot\boldsymbol{B}+\boldsymbol{P}\cdot\boldsymbol{E}^{(\mathrm{T})}\}$$

$$+\frac{1}{8\pi}\int \mathrm{d}\boldsymbol{r}\ \{E^{(\mathrm{T})2}-B^2\}$$

$$\Bigg[=\sum_\ell\frac{m_\ell\boldsymbol{v}_\ell^2}{2}-U_\mathrm{C}+\int \mathrm{d}\boldsymbol{r}\ \{\boldsymbol{M}\cdot\boldsymbol{B}+\boldsymbol{P}\cdot\boldsymbol{E}^{(\mathrm{T})}\}$$

$$+\frac{1}{2}\int \mathrm{d}\boldsymbol{r}\ \left\{\varepsilon_0 E^{(\mathrm{T})2}-\frac{1}{\mu_0}B^2\right\}\Bigg]_{\mathrm{SI}}. \tag{5.67}$$

Since $\boldsymbol{P}$ appears as an inner product with $\boldsymbol{E}^{(\mathrm{T})}$, only its T component $\boldsymbol{P}^{(\mathrm{T})}$ contributes to the integral.

In order to derive the corresponding Hamiltonian, we calculate the conjugate momenta $\partial L'/\partial \boldsymbol{v}_\ell$ for $\boldsymbol{r}_\ell$ and $\partial L'/\partial \dot{\boldsymbol{A}}$ for $\boldsymbol{A}$, where $\dot{\boldsymbol{A}} = \partial \boldsymbol{A}/\partial t$. These generalized momenta $\bar{\boldsymbol{p}}_\ell$ and $\bar{\Pi}$, respectively, are given as

$$\bar{\boldsymbol{p}}_\ell = m_\ell \boldsymbol{v}_\ell + \frac{1}{c}\int_0^1 u\mathrm{d}u\ e_\ell\ \boldsymbol{B}(u\boldsymbol{r}_\ell) \times \boldsymbol{r}_\ell$$

$$\left[= m_\ell \boldsymbol{v}_\ell + \int_0^1 u\mathrm{d}u\ e_\ell\ \boldsymbol{B}(u\boldsymbol{r}_\ell) \times \boldsymbol{r}_\ell \right]_{\mathrm{SI}} \tag{5.68}$$

$$\bar{\Pi} = -\frac{1}{4\pi c}(\boldsymbol{E}^{(\mathrm{T})} + 4\pi \boldsymbol{P}^{(\mathrm{T})}) = -\frac{1}{4\pi c}\boldsymbol{D}^{(\mathrm{T})}$$

$$\left[= -(\varepsilon_0 \boldsymbol{E}^{(\mathrm{T})} + \boldsymbol{P}^{(\mathrm{T})}) = -\boldsymbol{D}^{(\mathrm{T})} \right]_{\mathrm{SI}} \tag{5.69}$$

The new Hamiltonian $H_{L'}$ is obtained according to the standard rule as

$$H_{L'} = \sum_\ell \bar{\boldsymbol{p}}_\ell \cdot \boldsymbol{v}_\ell + \int \mathrm{d}\boldsymbol{r}\ \bar{\Pi}(\boldsymbol{r}) \cdot \dot{\boldsymbol{A}}(\boldsymbol{r}) - L' , \tag{5.70}$$

$$= H_{0(L')} + H_{R(L')} + H_{\mathrm{int}(L')}$$

$$H_{0(L')} = \sum_\ell \frac{\bar{\boldsymbol{p}}_\ell^2}{2m_\ell} + U_{\mathrm{C}} + 2\pi \int \mathrm{d}\boldsymbol{r} \boldsymbol{P}^{(\mathrm{T})}(\boldsymbol{r})^2 , \tag{5.71}$$

$$\left[= \sum_\ell \frac{\bar{\boldsymbol{p}}_\ell^2}{2m_\ell} + U_{\mathrm{C}} + \frac{1}{2\varepsilon_0} \int \mathrm{d}\boldsymbol{r} \boldsymbol{P}^{(\mathrm{T})}(\boldsymbol{r})^2 \right]_{\mathrm{SI}} , \tag{5.72}$$

$$H_{R(L')} = \frac{1}{8\pi} \int \mathrm{d}\boldsymbol{r}([\boldsymbol{D}^{(\mathrm{T})}]^2 + \boldsymbol{B}^2) ,$$

$$\left[= \frac{1}{2} \int \mathrm{d}\boldsymbol{r} \left(\frac{1}{\varepsilon_0}[\boldsymbol{D}^{(\mathrm{T})}]^2 + \frac{1}{\mu_0}\boldsymbol{B}^2 \right) \right]_{\mathrm{SI}} , \tag{5.73}$$

$$H_{\mathrm{int}(L')} = -\int \mathrm{d}\boldsymbol{r} \left\{ \boldsymbol{M}' \cdot \boldsymbol{B} + \boldsymbol{P}^{(\mathrm{T})} \cdot \boldsymbol{D}^{(\mathrm{T})} \right\} + \sum_\ell \frac{e_\ell^2}{2m_\ell c^2} \tilde{\boldsymbol{A}}_\ell^2$$

$$\left[= -\int \mathrm{d}\boldsymbol{r} \left\{ \boldsymbol{M}' \cdot \boldsymbol{B} + \frac{1}{\varepsilon_0} \boldsymbol{P}^{(\mathrm{T})} \cdot \boldsymbol{D}^{(\mathrm{T})} \right\} + \sum_\ell \frac{e_\ell^2}{2m_\ell} \tilde{\boldsymbol{A}}_\ell^2 \right]_{\mathrm{SI}} \tag{5.74}$$

where $\boldsymbol{M}'$ in $H_{\mathrm{int}(L')}$ is the $\boldsymbol{B}$-independent part of the orbital magnetization operator (5.8), i.e., the one with $\boldsymbol{v}_\ell$ replaced by $\bar{\boldsymbol{p}}_\ell/m_\ell$, and

$$\tilde{\boldsymbol{A}}_\ell = \int_0^1 u\mathrm{d}u\ \boldsymbol{B}(u\boldsymbol{r}_\ell) \times \boldsymbol{r}_\ell . \tag{5.75}$$

In this form, the matter Hamiltonian is $H_{0(L')}$, the vacuum EM field Hamiltonian is $H_{R(L')}$, and the linear and quadratic interaction terms $H_{\mathrm{int}(L')}$. In particular, we should note that the linear interaction term is $-\int \mathrm{d}\boldsymbol{r}[\boldsymbol{M}\cdot\boldsymbol{B}+\boldsymbol{P}^{(\mathrm{T})}\cdot\boldsymbol{D}^{(\mathrm{T})}]$, and that the matter Hamiltonian $H_{0(L')}$ contains an additional term $\propto \int \mathrm{d}\boldsymbol{r}\, \boldsymbol{P}^{(\mathrm{T})}(\boldsymbol{r})^2$ in comparison with $H_{0(L)}$.

If we dare to write the interaction terms as $-\int \mathrm{d}\boldsymbol{r}[\boldsymbol{M}\cdot\boldsymbol{H}+\boldsymbol{P}^{(\mathrm{T})}\cdot\boldsymbol{D}^{(\mathrm{T})}]$, we have to add the difference $-4\pi\int \mathrm{d}\boldsymbol{r}\boldsymbol{M}(\boldsymbol{r})^2$ to the matter Hamiltonian $H_{0(L')}$. Similarly, if we dare to write the interaction terms as $-\int \mathrm{d}\boldsymbol{r}[\boldsymbol{M}\cdot\boldsymbol{B}+\boldsymbol{P}^{(\mathrm{T})}\cdot\boldsymbol{E}^{(\mathrm{T})}]$, we have to add the difference $-4\pi\int \mathrm{d}\boldsymbol{r}\boldsymbol{P}^{(\mathrm{T})}(\boldsymbol{r})^2$ to the matter Hamiltonian $H_{0(L')}$. Thus, the attempts to rewrite the interaction Hamiltonian as $-\boldsymbol{M}\cdot\boldsymbol{H}$ or $-\boldsymbol{P}^{(\mathrm{T})}\cdot\boldsymbol{E}^{(\mathrm{T})}$ must always face to the corresponding change in the matter Hamiltonian. This means that the poles of the susceptibilities calculated by such a matter Hamiltonian are different from those of χ_{cd}, (2.39) because of the difference in the matter Hamiltonians. In the conventional definition of χ_{e} and χ_{m}, such a change in matter Hamiltonian is not included. Moreover, the rearranged interaction is no more written in terms of the conjugate variables $\{\boldsymbol{A}, \bar{\boldsymbol{\Pi}}\}$, i.e., it is no more a linear combination of photon creation and annihilation operators (when quantized). This will bring about a new complex situation.

The argument in this section shows the difficulty to derive a linear interaction term proportional to (the T components of) $\boldsymbol{E}$ and $\boldsymbol{H}$ as an exact theory. The principle of analytical mechanics allows us to use different sets of dynamical variables to describe a given system, leading to another sets of "matter, interaction, and EM-field Hamiltonians". On the other hand, there is a natural choice of matter Hamiltonian, i.e., the sum of the kinetic energy and Coulomb potential of the charged particles, written in terms of the masses and charges of the particles in a very simple way. The argument of PZW transformation shows an example of a different set of "matter, interaction, and EM-field Hamiltonians", which however does not look a very useful tool. From this consideration, the only reasonable choice as an exact treatment is the scheme based on the Lagrangian L, with the matter Hamiltonian $H_{0(L)}$ and the linear interaction $H_{\mathrm{int}(L)}$. The use of the susceptibilities χ_{e} and χ_{m}, rather than χ_{em} of Sect. 2.4, in the conventional macroscopic M-eqs does not have a sound foundation in the sense mentioned above. (Note, however, that the interaction with L-field is well described by χ_{e}. See Sect. 5.7.)

Even if we admit the use of interaction Hamiltonian proportional to electric and magnetic fields, their mutual interference in the case of chiral symmetry does not allow the simple use of χ_{e} and χ_{m}. In this case, there has been a phenomenological treatment called Drude–Born–Fedorov constitutive equations [2], which requires additional "chiral admittances". However, as the discussion in Sect. 3.4 shows, this kind of phenomenology cannot be supported from the first-principles theory.

5.3.2 *Velocity Gauge Versus Length Gauge*

In the long history of quantum mechanical description of matter-EM field interaction, there has been a lot of arguments since early days [3] as to the form of interaction between charged particle and EM field, either $-e\boldsymbol{E}\cdot\boldsymbol{r}$ or $(-e/mc)\boldsymbol{p}\cdot\boldsymbol{A}$, which are usually called length gauge and velocity gauge, respectively. As this naming indicates, the main point of the arguments is why these two forms, which are (supposed to be) related via gauge transformation, do not lead to the same observable result in various systems from a hydrogen atom to crystals. In this subsection, a comment is given to this problem from a different point of view.

As we have discussed in the previous subsection, these two forms of interaction are related by Power-Zienau-Woolley (PZW) transformation in the Lagrangian [1], which is specified by the addition, to the Lagrangian in the Coulomb gauge L_{C}, of the total time derivative of a function of general coordinates

$$F = -\frac{1}{c}\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{P}(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r}) \ = \left[-\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{P}(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r}) \right]_{\mathrm{SI}} . \tag{5.76}$$

Because of the independence of minimum action principle on the total time derivative term in Lagrangian, the new Lagrangian $L' = L_{\mathrm{C}} + \mathrm{d}F/\mathrm{d}t$ equally serves to the quantum mechanical description of the system. Adding $\mathrm{d}F/\mathrm{d}t$ to the interaction term of L_{C}

$$L_{\mathrm{int}} = \int \mathrm{d}\boldsymbol{r}\ \boldsymbol{J}(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r}) \ , \tag{5.77}$$

we obtain the new interaction term

$$L'_{\mathrm{int}} = \int \mathrm{d}\boldsymbol{r}\{\boldsymbol{P}^{(\mathrm{T})}(\boldsymbol{r}) \cdot \boldsymbol{E}^{(\mathrm{T})}(\boldsymbol{r}) + \boldsymbol{M}(\boldsymbol{r}) \cdot \boldsymbol{B}(\boldsymbol{r})\} \ , \tag{5.78}$$

where we have used the identity $\boldsymbol{J} = (\partial \boldsymbol{P}/\partial t) + \nabla \times \boldsymbol{M}$. From this new Lagrangian, we can derive the Hamiltonian whose (linear) interaction term is $\boldsymbol{M} \cdot \boldsymbol{B} + \boldsymbol{P} \cdot \boldsymbol{D}$ type as shown in (5.74). This can further be rewritten into $\boldsymbol{M} \cdot \boldsymbol{B} + \boldsymbol{P} \cdot \boldsymbol{E}$ type or $\boldsymbol{M} \cdot \boldsymbol{H} + \boldsymbol{P} \cdot \boldsymbol{E}$ type. The interaction term $\boldsymbol{J} \cdot \boldsymbol{A}$ of L leads to the Hamiltonian written in "velocity gauge", while L'_{int} allows the Hamiltonian in "length gauge". Thus, the general principle of analytical mechanics allows the both forms for the description of the interacting charged particles and EM field. However, it should be noted that the "matter" Hamiltonian is also changed by the PZW transformation, as mentioned in the paragraphs below (5.75). Therefore, for the consistent description, we should consider, not only the form of interaction term, but also that of matter Hamiltonian. Proper consideration of this change in the matter Hamiltonian will lead to the result independent of the "length gauge" and "velocity gauge".

There is another point to be mentioned about the naming of this subject. For this purpose, it is instructive to compare the total time derivative term $\mathrm{d}F/\mathrm{d}t$ with the change in the Lagrangian δL caused by the gauge transformation defined by (1.20), (1.21). The change caused in the minimal coupling Lagrangian is

$$\delta L = \frac{\mathrm{d}}{\mathrm{d}t}\int \mathrm{d}\boldsymbol{r}\ \rho\bar{\chi} = \frac{\mathrm{d}}{\mathrm{d}t}\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{P}^{(\mathrm{L})} \cdot \nabla\bar{\chi} \ . \tag{5.79}$$

This looks very similar to $\mathrm{d}F/\mathrm{d}t$. However, they cannot be same for any choice of $\bar{\chi}$, because the T component $\boldsymbol{P}^{(\mathrm{T})}$ exists in F but not in the latter. This arises from the fact that the non-uniqueness of a Lagrangian with respect to the addition of an arbitrary total time derivative term is a broader concept than the one due to gauge transformation. In other words, the general framework of analytical mechanics should work also in systems which have nothing to do with electromagnetism. For this reason, "$-\boldsymbol{E} \cdot \boldsymbol{r}$ or $(-e/m)\boldsymbol{p} \cdot \boldsymbol{A}$" is not a problem of gauge, and the naming of velocity gauge and length gauge is misleading.

5.4 Derivation of Constitutive Equation from Density Matrix

In Chap. 2, we have calculated the induced current density from the matter Hamiltonian $H^{(0)}$, (2.17), and the matter-EM field interaction H_{int}, (2.24). Thereby, it is necessary to fix the initial condition of matter state, and we assumed that the matter state was in its ground state in the remote past ($t \to \infty$). The result is therefore dependent on the initial condition of matter. Though the one we used in Chap. 2 is a standard one, one could use different conditions, too. A typical one is the use of ensemble for the description of matter states, where the matter states are quantum statistical ensemble. The time evolution of such an ensemble is described by density matrix $\hat{\rho}^{(m)}$, which obeys the equation of motion (Liouville equation)

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}^{(m)} = [H^{(0)} + H_{\mathrm{int}}, \hat{\rho}^{(m)}] \tag{5.80}$$

where $[a, b] = ab - ba$ represents a commutator of two operators. Once we know the solution of this equation $\hat{\rho}^{(m)}(t)$, we can calculate the statistical average of arbitrary physical quantity $\hat{b}$ at time t as a diagonal sum (Trace) of the following form

$$\langle b \rangle_{(t)} = \mathrm{Tr}\{\hat{\rho}^{(m)}(t)\, \hat{b}\} = \sum_{\nu} \langle \nu | \hat{\rho}^{(m)}(t)\, \hat{b} | \nu \rangle\ . \tag{5.81}$$

In the case of our interest, $\hat{b}$ is the current density $\boldsymbol{I}(\boldsymbol{r})$, (2.27), or (1.15).

The solution of (5.80) contains a density matrix corresponding to the initial condition of the matter state. A typical case of such an initial condition is the canonical ensemble at temperature T, which assumes the initial state of matter as a superposition of various (ground and excited) states with the weight of the Boltzmann factor $\exp(-E_{\nu}/k_{\mathrm{B}}T)$. This initial ensemble allows the EM excitations among the excited levels, as well as between the ground and excited levels, with the probability of the Boltzmann factor for the initial quantum level, which is not included in the calculation of Chap. 2. In this section, we show how this element is incorporated in the final result.

To solve the equation for the density matrix (5.80) in the lowest order of $H_{\rm int}$, we first rewrite it in the interaction representation

$$\hat{\rho}^{(m)}(t) = \exp(-iH^{(0)}t/\hbar)\ \hat{\rho}^{(\rm int)}(t)\ \exp(iH^{(0)}t/\hbar)\ . \tag{5.82}$$

Substituting this form into (5.80), we obtain the equation for $\hat{\rho}^{(\rm int)}$ as

$$i\hbar\,\frac{\rm d}{{\rm d}t}\hat{\rho}^{(\rm int)} = [H'_{\rm int}(t), \hat{\rho}^{(\rm int)}]\ , \tag{5.83}$$

where

$$H'_{\rm int}(t) = \exp(iH^{(0)}t/\hbar)\ H_{\rm int}\ \exp(-iH^{(0)}t/\hbar)\ . \tag{5.84}$$

One can solve (5.83) by iteration, assuming an initial condition

$$\hat{\rho}^{(\rm int)}(t \to \infty) = \hat{\rho}_0\ . \tag{5.85}$$

The first order solution satisfies

$$i\hbar\,\frac{\rm d}{{\rm d}t}\hat{\rho}^{(\rm int)} = [H'_{\rm int}(t), \hat{\rho}_0]\ , \tag{5.86}$$

where the $\hat{\rho}^{(\rm int)}$ on the r.h.s. is replaced by the initial condition $\hat{\rho}_0$. Since $\hat{\rho}_0$ is a known quantity, we immediately have the solution

$$\hat{\rho}^{(\rm int)}(t) = \hat{\rho}_0 - \frac{i}{\hbar}\int_{-\infty}^{t} {\rm d}t_1\ [H'_{\rm int}(t_1), \hat{\rho}_0]\exp(\gamma t_1) \tag{5.87}$$

satisfying the initial condition. Here also we assume the adiabatic switching of matter-EM field interaction as in (2.31), described by the infinitesimal positive constant $\gamma (= 0^+)$.

The initial density matrix $\hat{\rho}_0$ represents the matter state (as a statistical ensemble) in the absence of $H_{\rm int}$, it should be a stationary solution of (5.80) for $H_{\rm int} = 0$. Thus, $\hat{\rho}_0$ should satisfy the condition of stationarity

$$[H^{(0)},\ \hat{\rho}_0] = H^{(0)}\ \hat{\rho}_0 - \hat{\rho}_0\ H^{(0)} = 0\ . \tag{5.88}$$

This commutability of $H^{(0)}$ and $\hat{\rho}_0$ will be used to write the statistical average in a compact form later. As a typical model of $\hat{\rho}_0$, we take the canonical ensemble

$$\hat{\rho}_0 = \frac{1}{Z_0}\exp\left(-\frac{H^{(0)}}{k_{\rm B}T}\right) = \sum_{\nu} |\nu\rangle\ W_{\nu}\ \langle\nu| \tag{5.89}$$

where

$$W_\nu = \frac{1}{Z_0}\ \exp\left(-\frac{E_\nu}{k_\mathrm{B}T}\right)\ , \tag{5.90}$$

and Z_0 is the normalization factor (partition function)

$$Z_0 = \sum_\nu \exp\left(-\frac{E_\nu}{k_\mathrm{B}T}\right). \tag{5.91}$$

Thus, the matrix element of $\hat{\rho}_0$ is generally

$$\langle\mu|\hat{\rho}_0|\mu'\rangle = W_\mu\ \delta_{\mu,\mu'} \tag{5.92}$$

where $\delta_{\mu,\mu'}$ ($= 1$ for $\mu = \mu'$, and $= 0$ for $\mu \neq \mu'$) is the Kronecker's delta,

The $\boldsymbol{A}$-linear terms of the statistical average Trace $\{\boldsymbol{I}(\boldsymbol{r})\ \hat{\rho}^{(m)}(t)\}$ arise from the two sources, as already discussed in relation with (2.38). One ($\boldsymbol{I}_1$) is the contribution of $\hat{\rho}_0$ in (5.87) combined with the $\boldsymbol{A}$-linear term, $(-1/c)\hat{N}(\boldsymbol{r})\boldsymbol{A}(\boldsymbol{r})$, of $\boldsymbol{I}(\boldsymbol{r})$, (2.14), and the other ($\boldsymbol{I}_2$) is the contribution of the H'_int induced term of (5.87). In this term, the linear $\boldsymbol{A}$ dependence is already included in H'_int, so that we use the $\boldsymbol{A}$ independent part of the operator $\boldsymbol{I}(\boldsymbol{r})$. Their explicit forms are

$$\begin{aligned}\boldsymbol{I}_1 &= -\frac{1}{c}\ \mathrm{Tr}\{\hat{N}(\boldsymbol{r})\ \exp(-iH^{(0)}t/\hbar)\ \hat{\rho}_0\ \exp(iH^{(0)}t/\hbar)\}\boldsymbol{A}(\boldsymbol{r})\\ &= -\frac{1}{c}\sum_\mu\ W_\mu\ \langle\mu|\hat{N}(\boldsymbol{r})|\mu\rangle\ \boldsymbol{A}(\boldsymbol{r})\end{aligned} \tag{5.93}$$

$$\begin{aligned}\boldsymbol{I}_2 &= -\frac{i}{\hbar}\int_{-\infty}^{t} \mathrm{d}t_1\ \exp(\gamma t_1)\\ &\quad \sum_\nu \langle\nu|\exp(-iH^{(0)}t/\hbar)\ [H'_\mathrm{int}(t_1),\hat{\rho}_0]\exp(iH^{(0)}t/\hbar)\ \boldsymbol{I}(\boldsymbol{r})|\ \nu\rangle\\ &= -\frac{i}{\hbar}\int_{-\infty}^{t} \mathrm{d}t_1\ \exp(\gamma t_1)\ \sum_\nu \langle\nu|\exp[-iH^{(0)}(t-t_1)/\hbar]\\ &\qquad \{H_\mathrm{int}\hat{\rho}_0 - \hat{\rho}_0 H_\mathrm{int}\}\exp[iH^{(0)}(t-t_1)/\hbar]\ \boldsymbol{I}(\boldsymbol{r})\ |\nu\rangle\ ,\\ &= -\frac{i}{\hbar}\int_{-\infty}^{t} \mathrm{d}t_1\ \exp(\gamma t_1)\ \sum_\nu\sum_\mu \exp[-i(E_\nu - E_\mu)(t-t_1)/\hbar]\\ &\qquad \langle\nu|\{H_\mathrm{int}\hat{\rho}_0 - \hat{\rho}_0 H_\mathrm{int}\}|\mu\rangle\langle\mu|\boldsymbol{I}(\boldsymbol{r})\ |\nu\rangle\ ,\end{aligned} \tag{5.94}$$

For the last transformation, we have used the commutability of $H^{(0)}$ and $\hat{\rho}_0$. Assuming that the vector potential inducing current density has frequency ω, i.e.,

$$H_\mathrm{int} = -\frac{1}{c}\int \mathrm{d}\boldsymbol{r}'\boldsymbol{I}(\boldsymbol{r}')\cdot\boldsymbol{A}(\boldsymbol{r}',\omega)\exp(-i\omega t_1) \tag{5.95}$$

we further rewrite $\boldsymbol{I}_2$ as

$$\boldsymbol{I}_2 = \frac{i}{c\hbar}\sum_\nu\sum_\mu \int_{-\infty}^{t} \mathrm{d}t_1 \; \exp[-i(E_\nu - E_\mu)(t-t_1)/\hbar] \int \mathrm{d}\boldsymbol{r}' \exp(\gamma t_1)$$
$$(W_\mu - W_\nu)\langle\mu|\boldsymbol{I}(\boldsymbol{r})\,|\nu\rangle\,\langle\nu|\boldsymbol{I}(\boldsymbol{r}')|\mu\rangle \cdot \boldsymbol{A}(\boldsymbol{r}',\omega)\exp(-i\omega t). \quad (5.96)$$

Carrying out the time integration over t_1 and changing the summation induces μ, ν in one of the summands with the factor W_μ or W_ν, we finally obtain

$$\boldsymbol{I}_2 = \frac{1}{c\hbar}\exp(-i\omega t + \gamma t)\,\sum_\nu\sum_\mu W_\mu \int \mathrm{d}\boldsymbol{r}'\big[g_{\nu\mu}(\omega)\langle\mu|\boldsymbol{I}(\boldsymbol{r})|\nu\rangle\,\langle\nu|\boldsymbol{I}(\boldsymbol{r}')|\mu\rangle$$
$$+h_{\nu\mu}(\omega)\langle\nu|\boldsymbol{I}(\boldsymbol{r})|\mu\rangle\,\langle\mu|\boldsymbol{I}(\boldsymbol{r}')|\nu\rangle\big]\cdot\boldsymbol{A}(\boldsymbol{r}',\omega)\,, \quad (5.97)$$

where

$$g_{\nu\mu}(\omega) = \frac{1}{\omega_{\nu\mu} - \omega - i\gamma}\,, \quad h_{\nu\mu}(\omega) = \frac{1}{\omega_{\nu\mu} + \omega + i\gamma}\,, \quad (5.98)$$

and $\omega_{\nu\mu} = (E_\nu - E_\mu)/\hbar$.

The sum of $\boldsymbol{I}_1$ and $\boldsymbol{I}_2$ gives the total induced current density. Writing the sum in the form of $\boldsymbol{I}(\boldsymbol{r},\omega)$, we have

$$\boldsymbol{I}(\boldsymbol{r},\omega) = \int \mathrm{d}\boldsymbol{r}'\;\chi_{\mathrm{can}}(\boldsymbol{r},\boldsymbol{r}',\omega)\cdot\boldsymbol{A}(\boldsymbol{r}',\omega) \quad (5.99)$$

where

$$\chi_{\mathrm{can}}(\boldsymbol{r},\boldsymbol{r}',\omega) = -\frac{1}{c}\sum_\mu W_\mu\langle\mu|\hat{N}(\boldsymbol{r})|\mu\rangle\,\delta(\boldsymbol{r}-\boldsymbol{r}')$$
$$+\frac{1}{c}\sum_\mu W_\mu\sum_\nu\big[g_{\nu\mu}(\omega)\langle\mu|\boldsymbol{I}(\boldsymbol{r})|\nu\rangle\,\langle\nu|\boldsymbol{I}(\boldsymbol{r}')|\mu\rangle$$
$$+h_{\nu\mu}(\omega)\langle\nu|\boldsymbol{I}(\boldsymbol{r})|\mu\rangle\,\langle\mu|\boldsymbol{I}(\boldsymbol{r}')|\nu\rangle\big]\,. \quad (5.100)$$

If we take the limit of T = 0° K, i.e., $W_\mu = \delta_{\mu,0}$, $\chi_{\mathrm{can}}(\boldsymbol{r},\boldsymbol{r}',\omega)$ reduces to $\chi_{\mathrm{cd}}(\boldsymbol{r},\boldsymbol{r}',\omega)$, (2.39) of Chap. 2. The result obtained in this section is a simple extension of χ_{cd} by allowing the initial states of excitation at all the excited, as well as the ground, states of $H^{(0)}$, with the probability W_μ (Boltzmann factor).

5.5 Rewriting the $\langle 0|\hat{N}(\boldsymbol{r})|0\rangle$ Term

In the induced microscopic current density, (2.38), the term proportional to $\langle 0|\hat{N}(\boldsymbol{r})|0\rangle$ has a peculiar form, i.e., it represents a local response in contrast to the remaining terms. However, there is a useful way of rewriting this term in the following manner, which facilitates the formulation of microscopic nonlocal response theory. We will show that the following relation

$$\langle 0|\hat{N}(\boldsymbol{r})|0\rangle\ \boldsymbol{A}(\boldsymbol{r}) = \sum_{\nu} \frac{1}{E_{\nu 0}} \big[\boldsymbol{I}_{0\nu}(\boldsymbol{r}) F_{\nu 0}(\omega) + \boldsymbol{I}_{\nu 0}(\boldsymbol{r}) F_{0\nu}(\omega)\big] \tag{5.101}$$

holds as a good approximation, when [a] the relativistic correction in $H^{(0)}$ is negligible in comparison with the main term, and [b] LWA is valid. This expression allows us to rewrite the microscopic susceptibility χ_{cd} into a compact form (2.44). Though an essentially same argument is given in [4], we reproduce it here with some more details.

The relevant term appears as a part of induced current density arising from the $\boldsymbol{A}$ dependent term of the current density operator

$$\frac{1}{c}\ \hat{N}(\boldsymbol{r})\ \boldsymbol{A}(\boldsymbol{r})\ , \tag{5.102}$$

where

$$\hat{N}(\boldsymbol{r}) = \sum_{\ell} \frac{e_\ell^2}{m_\ell}\ \delta(\boldsymbol{r} - \boldsymbol{r}_\ell)\ . \tag{5.103}$$

The operator $\boldsymbol{I}(\boldsymbol{r})$ is the $\boldsymbol{A}$-independent part of the current density operator,

$$\boldsymbol{I}(\boldsymbol{r}) = \sum_{\ell} \frac{e_\ell}{2m_\ell} \big[\boldsymbol{p}_\ell \delta(\boldsymbol{r} - \boldsymbol{r}_\ell) + \delta(\boldsymbol{r} - \boldsymbol{r}_\ell)\boldsymbol{p}_\ell\big]\ . \tag{5.104}$$

The spin dependent terms are neglected, since the relativistic correction is assumed to be small.

We introduce one more operator

$$\hat{\boldsymbol{R}}(\boldsymbol{r}) = \sum_{\ell} e_\ell \boldsymbol{r}_\ell \delta(\boldsymbol{r} - \boldsymbol{r}_\ell)\ , \tag{5.105}$$

$$= \boldsymbol{r} \sum_{\ell}\ e_\ell\ \delta(\boldsymbol{r} - \boldsymbol{r}_\ell)\ . \tag{5.106}$$

Now we evaluate the commutators $[\hat{R}_\xi, H^{(0)}]$ and $[\hat{R}_\xi(\boldsymbol{r}), \hat{I}_\eta(\boldsymbol{r}')]$, where ξ, η are Cartesian coordinates. We begin with

$$[\hat{R}_\xi(\boldsymbol{r}), H^{(0)}] = r_\xi \sum_{\ell} \frac{e_\ell}{2m_\ell} [\delta(\boldsymbol{r} - \boldsymbol{r}_\ell), {\boldsymbol{p}_\ell}^2] \tag{5.107}$$

where the relativistic correction terms are neglected in $H^{(0)}$. For the evaluation of the commutators we use the relation

$$\boldsymbol{p}_\ell\ \delta(\boldsymbol{r} - \boldsymbol{r}_\ell) = -\boldsymbol{p}\ \delta(\boldsymbol{r} - \boldsymbol{r}_\ell)\ , \tag{5.108}$$

which allows us to move $\boldsymbol{p}$ to the outside of the summation over ℓ.

The commutator in (5.107) is expanded as

$$\begin{aligned}
&[\delta(\boldsymbol{r}-\boldsymbol{r}_\ell),\, {\boldsymbol{p}_\ell}^2] = \delta(\boldsymbol{r}-\boldsymbol{r}_\ell){\boldsymbol{p}_\ell}^2 - {\boldsymbol{p}_\ell}^2\delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\\
&= \delta(\boldsymbol{r}-\boldsymbol{r}_\ell){\boldsymbol{p}_\ell}^2 - \boldsymbol{p}_\ell\cdot\{\boldsymbol{p}_\ell\delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\} - \boldsymbol{p}_\ell\,\delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\cdot\boldsymbol{p}_\ell\\
&= \delta(\boldsymbol{r}-\boldsymbol{r}_\ell){\boldsymbol{p}_\ell}^2 + \boldsymbol{p}_\ell\cdot\{\boldsymbol{p}\,\delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\} - \{\boldsymbol{p}_\ell\,\delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\}\cdot\boldsymbol{p}_\ell - \delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\,{\boldsymbol{p}_\ell}^2\\
&= \boldsymbol{p}\cdot\{\boldsymbol{p}_\ell\delta(\boldsymbol{r}-\boldsymbol{r}_\ell) + \delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\,\boldsymbol{p}_\ell\} \qquad (5.109)
\end{aligned}$$

where (5.108) is used twice. Substituting this result into (5.107), we obtain

$$[\hat{R}_\xi(\boldsymbol{r}),\, H^{(0)}] = r_\xi\,\boldsymbol{p}\cdot\boldsymbol{I}(\boldsymbol{r})\;. \qquad (5.110)$$

Another commutator $[\hat{R}_\xi(\boldsymbol{r}),\, \hat{I}_\eta(\boldsymbol{r}')]$ is evaluated as

$$\begin{aligned}
&[\hat{R}_\xi(\boldsymbol{r}),\, \hat{I}_\eta(\boldsymbol{r}')] = r_\xi\sum_\ell\frac{e_\ell^2}{2m_\ell}[\delta(\boldsymbol{r}-\boldsymbol{r}_\ell),\, p_{\ell\eta}\,\delta(\boldsymbol{r}'-\boldsymbol{r}_\ell) + \delta(\boldsymbol{r}'-\boldsymbol{r}_\ell)\,p_{\ell\eta}]\\
&= r_\xi\sum_\ell\frac{e_\ell^2}{2m_\ell}\{\delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\,p_{\ell\eta}\,\delta(\boldsymbol{r}'-\boldsymbol{r}_\ell) - p_{\ell\eta}\,\delta(\boldsymbol{r}'-\boldsymbol{r}_\ell)\,\delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\\
&\qquad +\delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\,\delta(\boldsymbol{r}'-\boldsymbol{r}_\ell)\,p_{\ell\eta} - \delta(\boldsymbol{r}'-\boldsymbol{r}_\ell)\,p_{\ell\eta}\,\delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\}\\
&= r_\xi\sum_\ell\frac{e_\ell^2}{2m_\ell}\big[\delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\,\{p_{\ell\eta}\,\delta(\boldsymbol{r}'-\boldsymbol{r}_\ell)\} + \delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\,\delta(\boldsymbol{r}'-\boldsymbol{r}_\ell)\,p_{\ell\eta}\\
&\qquad -\{p_{\ell\eta}\,\delta(\boldsymbol{r}'-\boldsymbol{r}_\ell)\}\,\delta(\boldsymbol{r}-\boldsymbol{r}_\ell) - \delta(\boldsymbol{r}'-\boldsymbol{r}_\ell)\{p_{\ell\eta}\,\delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\}\\
&\qquad -\delta(\boldsymbol{r}'-\boldsymbol{r}_\ell)\,\delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\,p_{\ell\eta} + \delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\,\delta(\boldsymbol{r}'-\boldsymbol{r}_\ell)\,\boldsymbol{p}_{\ell,\eta}\\
&\qquad -\delta(\boldsymbol{r}'-\boldsymbol{r}_\ell)\{p_{\ell\eta}\,\delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\} - \delta(\boldsymbol{r}'-\boldsymbol{r}_\ell)\,\delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\,p_{\ell\eta}\big]\\
&= r_\xi\sum_\ell\frac{e_\ell^2}{2m_\ell}\big[-p'_\eta\delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\,\delta(\boldsymbol{r}'-\boldsymbol{r}_\ell) + p'_\eta\delta(\boldsymbol{r}'-\boldsymbol{r}_\ell)\,\delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\\
&\qquad +p_\eta\delta(\boldsymbol{r}'-\boldsymbol{r}_\ell)\,\delta(\boldsymbol{r}-\boldsymbol{r}_\ell) + p_\eta\delta(\boldsymbol{r}'-\boldsymbol{r}_\ell)\,\delta(\boldsymbol{r}-\boldsymbol{r}_\ell)\big]\\
&= r_\xi\;p_\eta\sum_\ell\frac{e_\ell^2}{m_\ell}\delta(\boldsymbol{r}'-\boldsymbol{r}_\ell)\,\delta(\boldsymbol{r}-\boldsymbol{r}_\ell) \qquad (5.111)\\
&= r_\xi\{p_\eta\delta(\boldsymbol{r}-\boldsymbol{r}')\}\hat{N}(\boldsymbol{r}') \qquad (5.112)
\end{aligned}$$

Let us define two operators

$$\hat{Q}(\omega) = \int \mathrm{d}\boldsymbol{r}\;\hat{\boldsymbol{R}}(\boldsymbol{r})\cdot\boldsymbol{A}(\boldsymbol{r},\omega) \qquad (5.113)$$

$$\hat{F}(\omega) = \int \mathrm{d}\boldsymbol{r}\;\hat{\boldsymbol{I}}(\boldsymbol{r})\cdot\boldsymbol{A}(\boldsymbol{r},\omega)\;, \qquad (5.114)$$

in terms of which (5.110) and (5.112) are rewritten as

$$\left[\hat{Q}(\omega), H^{(0)}\right] = -i\hbar \int \mathrm{d}\boldsymbol{r}\ \boldsymbol{r}\cdot\boldsymbol{A}(\boldsymbol{r},\omega)\nabla\cdot\hat{\boldsymbol{I}}(\boldsymbol{r}) \tag{5.115}$$

$$\left[\hat{Q}(\omega), \hat{I}_\eta(\boldsymbol{r}')\right] = -i\hbar \int \mathrm{d}\boldsymbol{r}\ \boldsymbol{r}\cdot\boldsymbol{A}(\boldsymbol{r},\omega)\left[\frac{\partial}{\partial r_\eta}\delta(\boldsymbol{r}-\boldsymbol{r}')\right]\hat{N}(\boldsymbol{r}')\ . \tag{5.116}$$

These two integrals can be rewritten via partial integration into

$$\left[\hat{Q}(\omega), H^{(0)}\right] = i\hbar \int \mathrm{d}\boldsymbol{r}\ \nabla\{\boldsymbol{r}\cdot\boldsymbol{A}(\boldsymbol{r},\omega)\}\ \cdot\hat{\boldsymbol{I}}(\boldsymbol{r}) \tag{5.117}$$

$$\left[\hat{Q}(\omega), \hat{I}_\eta(\boldsymbol{r}')\right] = i\hbar \int \mathrm{d}\boldsymbol{r}\ \frac{\partial}{\partial r_\eta}\{\boldsymbol{r}\cdot\boldsymbol{A}(\boldsymbol{r},\omega)\}\delta(\boldsymbol{r}-\boldsymbol{r}')\hat{N}(\boldsymbol{r}')\ . \tag{5.118}$$

Both of them contain the following factor in the integrand

$$\frac{\partial}{\partial r_\eta}\{\boldsymbol{r}\cdot\boldsymbol{A}(\boldsymbol{r},\omega)\} = A_\eta + \sum_\xi r_\xi \frac{\partial A_\xi}{\partial r_\eta}\ , \tag{5.119}$$

which can be approximated as $A_\eta(\boldsymbol{r},\omega)$ when LWA is a good approximation. In this case, these two commutators can be written as

$$\left[\hat{Q}(\omega), H^{(0)}\right] = i\hbar \int \mathrm{d}\boldsymbol{r}\ \boldsymbol{A}(\boldsymbol{r},\omega)\cdot\hat{\boldsymbol{I}}(\boldsymbol{r}) \tag{5.120}$$

$$\left[\hat{Q}(\omega), \hat{\boldsymbol{I}}(\boldsymbol{r}')\right] = i\hbar\ \boldsymbol{A}(\boldsymbol{r}',\omega)\ \hat{N}(\boldsymbol{r}')\ . \tag{5.121}$$

Equation (5.101) is the $\langle 0|\cdots|0\rangle$ matrix element of (5.121), i.e.,

$$\langle 0|\hat{N}(\boldsymbol{r})|0\rangle\ \boldsymbol{A}(\boldsymbol{r},\omega) = \frac{-i}{\hbar}\sum_\nu \left[\langle 0|[\hat{Q}(\omega)|\nu\rangle\langle\nu|\hat{\boldsymbol{I}}(\boldsymbol{r})|0\rangle - \langle 0|[\hat{\boldsymbol{I}}(\boldsymbol{r})|\nu\rangle\langle\nu|\hat{Q}(\omega)|0\rangle\right]. \tag{5.122}$$

To evaluate $\langle\nu|\hat{Q}(\omega)|\mu\rangle$, we take the $\langle\nu|\cdots|\mu\rangle$ matrix element of (5.120) as

$$(E_\mu - E_\nu)\ \langle\nu|\hat{Q}|\mu\rangle = i\hbar F_{\nu\mu} \tag{5.123}$$

Thus, we obtain the desired result

$$\langle 0|\hat{N}(\boldsymbol{r})|0\rangle\ \boldsymbol{A}(\boldsymbol{r},\omega) = \sum_\nu \frac{1}{E_{\nu 0}}\left[F_{0\nu}(\omega)\boldsymbol{I}_{\nu 0}(\boldsymbol{r}) + F_{\nu 0}(\omega)\boldsymbol{I}_{0\nu}(\boldsymbol{r})\right], \tag{5.124}$$

with $E_{\nu 0} = E_\nu - E_0$.

The corresponding expression for the case of canonical ensemble is obtained as

$$\sum_\mu W_\mu\ \langle\mu|\hat{N}(\boldsymbol{r})|\mu\rangle\ \boldsymbol{A}(\boldsymbol{r},\omega) = \sum_\mu\sum_\nu \frac{W_\mu}{E_{\nu\mu}}\left[F_{\mu\nu}(\omega)\boldsymbol{I}_{\nu\mu}(\boldsymbol{r}) + F_{\nu\mu}(\omega)\boldsymbol{I}_{\mu\nu}(\boldsymbol{r})\right]. \tag{5.125}$$

where $E_{\nu\mu} = E_\nu - E_\mu$, and this allows us to rewrite the susceptibility (5.100) into

$$\chi_{\rm can}(\boldsymbol{r}, \boldsymbol{r}', \omega) = \frac{1}{c}\sum_\mu W_\mu \sum_\nu \big[\bar{g}_{\nu\mu}(\omega)\langle\mu|\boldsymbol{I}(\boldsymbol{r})|\nu\rangle\ \langle\nu|\boldsymbol{I}(\boldsymbol{r}')|\mu\rangle \\ +\bar{h}_{\nu\mu}(\omega)\langle\nu|\boldsymbol{I}(\boldsymbol{r})|\mu\rangle\ \langle\mu|\boldsymbol{I}(\boldsymbol{r}')|\nu\rangle\big]\,, \tag{5.126}$$

where

$$\bar{g}_{\nu\mu}(\omega) = g_{\nu\mu}(\omega) - \frac{1}{E_{\nu\mu}}\,, \tag{5.127}$$

$$\bar{h}_{\nu\mu}(\omega) = h_{\nu\mu}(\omega) - \frac{1}{E_{\nu\mu}}\,. \tag{5.128}$$

5.6 Division of $\bar{\boldsymbol{Q}}_{\mu\nu}$ into E2 and M1 Components

The Taylor expansion of the current density matrix element $\tilde{\boldsymbol{I}}_{\mu\nu}$ leads to the sum of various moments, as in (2.112). The second term $\bar{\boldsymbol{Q}}_{\mu\nu}$ is the first order moment of the orbital current density. As discussed in Sect. 5.1, the orbital current density operator is the sum of the contributions of electric polarization and orbital magnetization, which induce E2 and M1 transitions, respectively. From this viewpoint, it is interesting to divide the matrix element $\bar{\boldsymbol{Q}}_{\mu\nu}$ into E2 and M1 components.

For this purpose, we write $\hat{\boldsymbol{k}} \cdot \bar{\boldsymbol{Q}}_{\mu\nu}$ as

$$\hat{\boldsymbol{k}} \cdot \bar{\boldsymbol{Q}}_{\mu\nu} = \sum_\ell \frac{e_\ell}{2m_\ell}\int \mathrm{d}\boldsymbol{r} \\ \hat{\boldsymbol{k}} \cdot \{\langle\mu|(\boldsymbol{r}_\ell - \bar{\boldsymbol{r}})\boldsymbol{p}_\ell\, \delta(\boldsymbol{r}_\ell - \boldsymbol{r}) + \delta(\boldsymbol{r}_\ell - \boldsymbol{r})\ (\boldsymbol{r}_\ell - \bar{\boldsymbol{r}})\boldsymbol{p}_\ell|\nu\rangle\}\,. \tag{5.129}$$

Since $\bar{\boldsymbol{r}}$ plays no important role in this discussion, we put $\bar{\boldsymbol{r}} = 0$ for the moment. We consider a particular ℓ and omit ℓ from $\boldsymbol{r}_\ell$ and $\boldsymbol{p}_\ell$. Rewriting $\hat{\boldsymbol{k}} \cdot \boldsymbol{r}\boldsymbol{p}$ as

$$(\hat{\boldsymbol{k}} \cdot \boldsymbol{r}\ \boldsymbol{p})_x = \hat{k}_x x p_x + \hat{k}_y y p_x + \hat{k}_z z p_x \\ = x(\hat{k}_x p_x + \hat{k}_y p_y + \hat{k}_z p_z) + \hat{k}_y(y p_x - x p_y) + \hat{k}_z(z p_x - x p_z)\,, \tag{5.130}$$

we find

$$\hat{\boldsymbol{k}} \cdot \boldsymbol{r}\ \boldsymbol{p} = \boldsymbol{r}(\hat{\boldsymbol{k}} \cdot \boldsymbol{p}) - \hat{\boldsymbol{k}} \times (\boldsymbol{r} \times \boldsymbol{p})\,. \tag{5.131}$$

Since $\boldsymbol{r} \times \boldsymbol{p}$ is the orbital angular momentum $\boldsymbol{L}$ (of each particle), this one-particle operator induces M1 transition, while the remaining term $\boldsymbol{r}(\hat{\boldsymbol{k}} \cdot \boldsymbol{p})$ has the electric quadrupole character contributing to E2 transition. The factor $\boldsymbol{k} \cdot \boldsymbol{r}\ \ \boldsymbol{p} \cdot \boldsymbol{A}$, which appears in the variable $F_{\mu\nu}(\omega)$, can be rewritten as

$$\boldsymbol{k}\cdot\boldsymbol{r}\ \boldsymbol{p}\cdot\boldsymbol{A}=(\boldsymbol{k}\cdot\boldsymbol{p})(\boldsymbol{r}\cdot\boldsymbol{A})+\boldsymbol{L}\cdot(\boldsymbol{k}\times\boldsymbol{A})\ . \tag{5.132}$$

Since the factor $\boldsymbol{k}\times\boldsymbol{A}$ is the $\boldsymbol{k}$ Fourier component of $-i\nabla\times\boldsymbol{A}$ $(=-i\boldsymbol{B})$, this term is proportional to the orbital Zeeman energy.

Thus, we have the desired division of $\boldsymbol{k}\cdot\bar{\boldsymbol{Q}}_{\mu\nu}$ into $\boldsymbol{k}\cdot\bar{\boldsymbol{Q}}^{(\mathrm{e2})}_{\mu\nu}-c\hat{\boldsymbol{k}}\times\bar{\boldsymbol{M}}^{(\mathrm{orb})}_{\mu\nu}$, where

$$\begin{aligned}\bar{\boldsymbol{Q}}^{(\mathrm{e2})}_{\mu\nu}=\sum_{\ell}\frac{e_\ell}{2m_\ell}\int \mathrm{d}\boldsymbol{r}\\ \{\langle\mu|(\boldsymbol{r}_\ell-\bar{\boldsymbol{r}})\hat{\boldsymbol{k}}\cdot\boldsymbol{p}_\ell\,\delta(\boldsymbol{r}_\ell-\boldsymbol{r})+\delta(\boldsymbol{r}_\ell-\boldsymbol{r})\,(\boldsymbol{r}_\ell-\bar{\boldsymbol{r}})\hat{\boldsymbol{k}}\cdot\boldsymbol{p}_\ell|\nu\rangle\}\ ,\end{aligned} \tag{5.133}$$

and

$$\bar{\boldsymbol{M}}^{(\mathrm{orb})}_{\mu\nu}=\sum_{\ell}\frac{e_\ell}{2m_\ell c}\int \mathrm{d}\boldsymbol{r}\ \{\langle\mu|\boldsymbol{L}_\ell(\bar{\boldsymbol{r}})\,\delta(\boldsymbol{r}_\ell-\boldsymbol{r})+\delta(\boldsymbol{r}_\ell-\boldsymbol{r})\,\boldsymbol{L}_\ell(\bar{\boldsymbol{r}})|\nu\rangle\}\ . \tag{5.134}$$

The angular momentum of the ℓth particle is defined as $\boldsymbol{L}_\ell(\bar{\boldsymbol{r}})=(\boldsymbol{r}_\ell-\bar{\boldsymbol{r}})\times\boldsymbol{p}_\ell$, i.e., around the center position $\bar{\boldsymbol{r}}$.

5.7 Problems of Longitudinal (L) Field

5.7.1 T and L Character of Induced Field

In terms of vector and scalar potentials, the microscopic M-eqs are

$$\nabla^2\phi=-4\pi\rho\ , \tag{5.135}$$

$$\frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2}-\nabla^2\boldsymbol{A}+\nabla\frac{1}{c}\frac{\partial\phi}{\partial t}=\frac{4\pi}{c}\boldsymbol{J}\ , \tag{5.136}$$

in the Coulomb gauge, and

$$\frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2}-\nabla^2\phi=4\pi\rho\ , \tag{5.137}$$

$$\frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2}-\nabla^2\boldsymbol{A}=\frac{4\pi}{c}\boldsymbol{J}\ , \tag{5.138}$$

in the Lorentz gauge. This $\boldsymbol{J}$ represents the orbital contribution, $\boldsymbol{J}_{\mathrm{orb}}$, and, in both cases, we could add the relativistic correction term (spin induced current density) $\boldsymbol{J}_{\mathrm{s}}$ to the r.h.s., as discussed in deriving (2.27), and this does not change the following arguments.

The equation for $\boldsymbol{A}$ in the Coulomb gauge can be rewritten as

$$\frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} - \nabla^2 \boldsymbol{A} = \frac{4\pi}{c}\boldsymbol{J}^{(\mathrm{T})} , \tag{5.139}$$

indicating that this is the equation only for T components. This rewriting is done by substituting the solution of the Poisson equation

$$\phi(\boldsymbol{r}) = \int \mathrm{d}\boldsymbol{r}' \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \tag{5.140}$$

into (5.136), and replacing the $\partial\rho/\partial t$ with $-\nabla \cdot \boldsymbol{J}$ (continuity equation). This gives

$$\frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} - \nabla^2 \boldsymbol{A} = \frac{4\pi}{c}\boldsymbol{J}(\boldsymbol{r}) + \frac{1}{c}\nabla \int \mathrm{d}\boldsymbol{r}' \frac{\nabla' \cdot \boldsymbol{J}(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} . \tag{5.141}$$

The r.h.s. of this equation is $(4\pi/c)\boldsymbol{J}^{(\mathrm{T})}$, because, if we apply divergence from the left, it becomes zero by using $\nabla^2(1/|\boldsymbol{r} - \boldsymbol{r}'|) = -4\pi\delta(\boldsymbol{r} - \boldsymbol{r}')$. Namely, the quantity

$$-\frac{1}{4\pi}\nabla \int \mathrm{d}\boldsymbol{r}' \frac{\nabla' \cdot \boldsymbol{C}(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \tag{5.142}$$

for a general vector field $\boldsymbol{C}$ is its L component.

The solutions of the M-eqs for $\boldsymbol{A}$ in the Coulomb and Lorentz gauges are given in terms of the scalar EM Green function defined by, for $q = \omega/c$,

$$(-\nabla^2 - q^2)\, G_q(\boldsymbol{r} - \boldsymbol{r}') = 4\pi\ \delta(\boldsymbol{r} - \boldsymbol{r}') , \tag{5.143}$$

where its special solution is $G_q(\boldsymbol{r}) = \exp(iq|\boldsymbol{r}|)/|\boldsymbol{r}|$. The solution of the M-eqs for $\boldsymbol{A}$ in the Coulomb gauge is the T field as

$$\boldsymbol{A}(\boldsymbol{r}, \omega) = \frac{1}{c}\int \mathrm{d}\boldsymbol{r}'\ G_q(\boldsymbol{r} - \boldsymbol{r}')\ \boldsymbol{J}^{(\mathrm{T})}(\boldsymbol{r}', \omega) , \tag{5.144}$$

and the solution in the Lorentz gauge is

$$\boldsymbol{A}(\boldsymbol{r}, \omega) = \frac{1}{c}\int \mathrm{d}\boldsymbol{r}'\ G_q(\boldsymbol{r} - \boldsymbol{r}')\ \boldsymbol{J}(\boldsymbol{r}', \omega) , \tag{5.145}$$

containing both T and L components. Obviously, the T component of the latter agrees with the solution in the Coulomb gauge. The solution for ϕ in the Lorentz gauge has a similar form

$$\phi(\boldsymbol{r}, \omega) = \int \mathrm{d}\boldsymbol{r}'\ G_q(\boldsymbol{r} - \boldsymbol{r}')\rho(\boldsymbol{r}', \omega) . \tag{5.146}$$

If we rewrite the Lorentz condition in terms of these solutions as

$$\nabla \cdot \boldsymbol{A} + \frac{1}{c}\frac{\partial \phi}{\partial t} = \frac{1}{c}\int \mathrm{d}\boldsymbol{r}'\, G_q(\boldsymbol{r}-\boldsymbol{r}')\{\nabla' \cdot \boldsymbol{J}(\boldsymbol{r}') - i\omega\rho\} = 0\,, \tag{5.147}$$

its validity is guaranteed by the continuity equation.

The form of the induced L-field is the one due to $\boldsymbol{J}^{(\mathrm{L})}(\boldsymbol{r})$ propagated via the scalar Green function G_q, but there is an alternative way of description, i.e., the one due to the "whole" current density $\boldsymbol{J}(\boldsymbol{r})$ propagated via the L component of the tensor Green function. For this purpose, we rewrite the induced L-field as

$$-\frac{1}{c}\int \mathrm{d}\boldsymbol{r}'\, G_q(\boldsymbol{r}-\boldsymbol{r}')\boldsymbol{J}^{(\mathrm{L})}(\boldsymbol{r}') = -\frac{1}{4\pi c}\int \mathrm{d}\boldsymbol{r}' \int \mathrm{d}\boldsymbol{r}''\, G_q(\boldsymbol{r}-\boldsymbol{r}')\frac{\nabla''\nabla'' \cdot \boldsymbol{J}(\boldsymbol{r}'')}{|\boldsymbol{r}'-\boldsymbol{r}''|} \tag{5.148}$$

where partial integration is made to convert ∇' to ∇''. From the equation $(-\nabla^2 - q^2)G_q = 4\pi\delta(\boldsymbol{r}-\boldsymbol{r}')$ and that for G_0, we get $G_q = (-1/q^2)\nabla^2[G_q - G_0]$. Substituting this expression into (5.148), and performing the partial integration about ∇^2, we can rewrite the r.h.s. of (5.148) as

$$\frac{1}{4\pi q^2 c}\int \mathrm{d}\boldsymbol{r}' \int \mathrm{d}\boldsymbol{r}''\, [G_q(\boldsymbol{r}-\boldsymbol{r}') - G_0(\boldsymbol{r}-\boldsymbol{r}')]\, \nabla'^2\, \frac{1}{|\boldsymbol{r}'-\boldsymbol{r}''|}\, \nabla''\nabla'' \cdot \boldsymbol{J}(\boldsymbol{r}'')\,. \tag{5.149}$$

which reduces, via $\nabla'^2(1/|\boldsymbol{r}'-\boldsymbol{r}''|) = -4\pi\delta(\boldsymbol{r}'-\boldsymbol{r}'')$, to

$$-\frac{1}{c}\int \mathrm{d}\boldsymbol{r}'\, G_q(\boldsymbol{r}-\boldsymbol{r}')\boldsymbol{J}^{(\mathrm{L})}(\boldsymbol{r}') = -\frac{1}{c}\int \mathrm{d}\boldsymbol{r}'\, \tilde{\mathbf{G}}_q^{(\mathrm{L})}(\boldsymbol{r}-\boldsymbol{r}') \cdot \boldsymbol{J}(\boldsymbol{r}')\,. \tag{5.150}$$

The tensor Green function describing the L part of the induced field is

$$\tilde{\mathbf{G}}_q^{(\mathrm{L})}(\boldsymbol{r}-\boldsymbol{r}') = \frac{1}{q^2}\, G_0(\boldsymbol{r}-\boldsymbol{r}')\, \nabla'\nabla' \tag{5.151}$$

which produces a L field by operating on a full current density (with T and L components). The counterpart, i.e., the tensor Green function describing the T part of the induced field is obtained by subtracting this L part from the total one $G_q\mathbf{1}$ as

$$\tilde{\mathbf{G}}_q^{(\mathrm{T})}(\boldsymbol{r}-\boldsymbol{r}') = G_q(\boldsymbol{r}-\boldsymbol{r}')\mathbf{1} + \frac{1}{q^2}\, [G_q(\boldsymbol{r}-\boldsymbol{r}') - G_0(\boldsymbol{r}-\boldsymbol{r}')]\, \nabla'\nabla' \tag{5.152}$$

which produces a T field by operating on a full current density. Thus the T field induced by $\boldsymbol{J}$ can be expressed in the following two ways, i.e.,

$$\boldsymbol{A}_{\mathrm{T}}(\boldsymbol{r},\omega) = \int \mathrm{d}\boldsymbol{r}'\, G_q(\boldsymbol{r}-\boldsymbol{r}')\, \boldsymbol{J}^{(\mathrm{T})}(\boldsymbol{r}',\omega) = \int \mathrm{d}\boldsymbol{r}'\, \tilde{\mathbf{G}}_q^{(\mathrm{T})}(\boldsymbol{r}-\boldsymbol{r}')\, \cdot \boldsymbol{J}(\boldsymbol{r}',\omega)\,. \tag{5.153}$$

5.7.2 Excitation by an External L Field

When matter is excited by an external EM field, there arises an induced current density which may also be described as charge density, electric polarization, or magnetization. The eigenmodes of these induced polarizations correspond to the quantum mechanical excited states of the matter, and can be classified into the L, T, and LT-mixed modes according to their symmetry properties. The external EM field inducing matter excitations as an incident field has also T and L characters. Typically, light is a T field, and the field due to external charged particles is regarded as a L field. (However, a moving charge produces T field as well as L field, which is known as Cerenkov radiation [5] and Smith–Purcell radiation (SPR) [6]. Cerenkov radiation is the propagating T waves produced by a moving charge when the particle velocity exceed the light velocity in a dielectric medium. Below the critical velocity, there arise evanescent waves of T character, which, together with the (evanescent) L components, interact with the periodic crystals, producing scattered (propagating) light modes of T character i.e., SPR.)

The incident T field can excite “T and LT mixed” modes of matter, and the incident L field with “L and LT mixed” modes of matter excitations. If the matter excitations are purely T and purely L modes in a given geometry, they can be detected by the spectroscopy using incident field of T- (light) and L- (charged particles) characters, respectively. As a propagating wave, T mode is polarized perpendicular to the wave vector of the mode, so that there are two independent directions of polarization, while L mode, polarized along the wave vector, has only one direction of the polarization. Therefore, we need two different polarizations to detect both types of the T modes.

When the symmetry of matter is low, there arises a mixing between the T and L modes. These LT mixed modes can generally be detected by either L or T incident field. When this mixing occurs, there is no purely L modes from symmetry ground, while there can still be another, purely T modes, which do not mix with L modes. Depending on the symmetry, we can classify all the modes into (a) LT-mixed modes alone, (b) LT-mixed modes and purely T-modes, (c) L-modes and two types of T modes. The treatment in the main text restricting the incident EM field to the T character can cover the most cases except for the matter excitations of purely L-character in case (c).

The interaction Hamiltonians for the T and L modes derived from the standard Hamiltonian of the coupled matter-EM field system in the Coulomb gauge are different, i.e.,

$$H_{\mathrm{intL}} = -\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{P} \cdot \boldsymbol{E}\ , \tag{5.154}$$

$$H_{\mathrm{intT}} = -\frac{1}{c}\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{J}_0 \cdot \boldsymbol{A}\ , \tag{5.155}$$

where J_0 is defined in (2.13) and $O(A^2)$ term is omitted in H_{intT}. The second term H_{intT} is H_{int} defined in (2.24). The difference is due to their different sources, i.e.,

H_{intL} is derived from the Coulomb potential U_{C} (as shown below) and H_{intT} from the kinetic energy term $\sum(1/2m)\{\boldsymbol{p}-(e/c)\boldsymbol{A}\}^2$. As discussed in detail in Sect. 5.3, there is no exact way to rewrite H_{intT} in terms of $\boldsymbol{E}$ and $\boldsymbol{P}$ without changing the matter Hamiltonian consisting of the sum of the kinetic energy and Coulomb potential (plus relativistic correction). In the conventional theory of macroscopic M-eqs, this distinction is not severely recognized, and very often the form $-\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{P}\cdot\boldsymbol{E}$ is used as the interaction Hamiltonian for both T and L modes. However, as the careful consideration in this book shows, we should distinguish the form of interaction Hamiltonian for T and L modes.

As a missing part of the main text, we give a description here about the matter excitation by an EM field of L character. This contains the cases of electron energy loss spectroscopy and the application of static electric field. Another example would be the use of a "probe" to measure the response of a "sample", as in the case of scanning near-field optical microscopy (SNOM), where both probe and sample consist of charged particles interacting via the EM field of L, as well as T, character. More generally, if one separates matter into two parts, i.e., sample part and the rest, these two parts can generally interact via the Coulomb interaction, even if they are electronically separated. In these cases, the interaction between $\boldsymbol{E}_{\mathrm{L}}$ and $\boldsymbol{P}$ serves, on the one hand, to detect the L response of matter (or sample), and contributes, on the other, to the resonance energy of the response spectrum.

In the presence of the external potential $\phi_{\mathrm{ext}}(\boldsymbol{r},t)$ due to an external charge density $\rho_{\mathrm{ext}}(\boldsymbol{r},t)$, i.e.,

$$\phi_{\mathrm{ext}}(\boldsymbol{r},t)=\int \mathrm{d}\boldsymbol{r}'\ \frac{\rho_{\mathrm{ext}}(\boldsymbol{r}',t)}{|\boldsymbol{r}-\boldsymbol{r}'|}\ , \tag{5.156}$$

we need to consider the interaction between $\phi_{\mathrm{ext}}(\boldsymbol{r},t)$ and the internal charge density, i.e., "matter" charge density $\rho(\boldsymbol{r})$

$$H_{\mathrm{intL}}=\int \mathrm{d}\boldsymbol{r}\ \rho(\boldsymbol{r})\ \phi_{\mathrm{ext}}(\boldsymbol{r},t)\ , \tag{5.157}$$

$$=\int \mathrm{d}\boldsymbol{r}\int \mathrm{d}\boldsymbol{r}'\ \frac{\rho(\boldsymbol{r},t)\ \rho_{\mathrm{ext}}(\boldsymbol{r}',t)}{|\boldsymbol{r}-\boldsymbol{r}'|}\ , \tag{5.158}$$

$$=-\int \mathrm{d}\boldsymbol{r}\ \boldsymbol{P}(\boldsymbol{r})\cdot\boldsymbol{E}_{\mathrm{extL}}(\boldsymbol{r},t)\ , \tag{5.159}$$

where we have used $\boldsymbol{E}_{\mathrm{extL}}(\boldsymbol{r},t)=-\nabla\phi_{\mathrm{ext}}(\boldsymbol{r},t)$, $\nabla\cdot\boldsymbol{P}(\boldsymbol{r})=-\rho(\boldsymbol{r})$, partial integration, and the assumption that the matter system is charge neutral, i.e., $\sum_{\ell} e_{\ell}=0$.

Generally speaking, an external field may contain both L and T components. In that case, we need to add the interaction Hamiltonian H_{intT} also to H_{intL}. This will lead to the complete expression of linear response of a given matter system. However, we just give only the contribution of H_{intL} below, since the consequence of H_{intT} is discussed in detail in Chap. 2.

The current density induced by $\boldsymbol{E}_{\text{extL}}$ is calculated in a similar manner as in Sect. 2.2 by the time dependent perturbation theory for the matter Hamiltonian $H^{(0)}$ and the matter-EM field interaction H_{intL}. The expectation value of current density is

$$\boldsymbol{I}(\boldsymbol{r},\omega) = \sum_{\nu}\int \mathrm{d}\boldsymbol{r}'\big[g_{\nu}(\omega)\boldsymbol{I}_{0\nu}(\boldsymbol{r})\boldsymbol{P}_{\nu 0}(\boldsymbol{r}') + h_{\nu}(\omega)\boldsymbol{I}_{\nu 0}(\boldsymbol{r})\boldsymbol{P}_{0\nu}(\boldsymbol{r}')\big]\cdot\boldsymbol{E}_{\text{extL}}(\boldsymbol{r}',\omega)\,, \tag{5.160}$$

$$= \int \mathrm{d}\boldsymbol{r}'\chi_{\text{JEL}}(\boldsymbol{r},\boldsymbol{r}',\omega)\cdot\boldsymbol{E}_{\text{extL}}(\boldsymbol{r}',\omega)\,, \tag{5.161}$$

where we introduce the susceptibility for the induced current density due to the external L field. Since we do not consider the presence of vector potential $\boldsymbol{A}$ in this calculation, the term due to the $\boldsymbol{A}$-dependent term in particle velocity

$$-\frac{1}{c}\langle 0|\hat{N}(\boldsymbol{r})|0\rangle\ \boldsymbol{A}(\boldsymbol{r},\omega) \tag{5.162}$$

does not exist in the expectation value. The products of the matrix elements of $\boldsymbol{I}$ and $\boldsymbol{P}$ can be rewritten by those of two $\boldsymbol{I}$'s, as shown in Sect. 3.2, which allows us to unify the expressions of induced current densities by T and L fields.

The induced current density contains the components of both electric polarization $-i\omega\langle\boldsymbol{P}\rangle$ and magnetization $c\nabla\times\langle\boldsymbol{M}\rangle$. Since the former is zero for $\omega=0$, one may prefer $\langle\boldsymbol{P}\rangle$ to $\langle\boldsymbol{I}\rangle$ as an induced change of matter which is non-zero for $\omega=0$. This is calculated in the same manner as

$$\boldsymbol{P}(\boldsymbol{r},\omega) = \sum_{\nu}\int \mathrm{d}\boldsymbol{r}'\big[g_{\nu}(\omega)\boldsymbol{P}_{0\nu}(\boldsymbol{r})\boldsymbol{P}_{\nu 0}(\boldsymbol{r}') + h_{\nu}(\omega)\boldsymbol{P}_{\nu 0}(\boldsymbol{r})\boldsymbol{P}_{0\nu}(\boldsymbol{r}')\big]\cdot\boldsymbol{E}_{\text{extL}}(\boldsymbol{r}',\omega)\,. \tag{5.163}$$

The L electric field produced by this polarization is

$$\begin{aligned}\boldsymbol{E}^{(\text{L})}(\boldsymbol{r},\omega) &= \nabla\int\frac{\nabla'\cdot\boldsymbol{P}^{(\text{L})}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\mathrm{d}\boldsymbol{r}' = \int\frac{\nabla'\nabla'\cdot\boldsymbol{P}^{(\text{L})}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\mathrm{d}\boldsymbol{r}'\\ &= -4\pi\boldsymbol{P}^{(\text{L})}(\boldsymbol{r},\omega)\,,\end{aligned} \tag{5.164}$$

where we used partial integration, $\nabla\nabla\cdot = \nabla^2 + \nabla\times\nabla\times$, $\nabla\times\boldsymbol{P}^{(\text{L})}=0$, and $\nabla^2(1/|\boldsymbol{r}-\boldsymbol{r}'|) = -4\pi\delta[\boldsymbol{r}-\boldsymbol{r}']$.

Though the argument given above may seem to be necessary to calculate induced polarization, especially for $\omega=0$, within the framework of the macroscopic M-eqs, we can show that the expression of the induced current density contains the static component of $\boldsymbol{P}$ correctly in the limit of $\omega\to 0$.

Now we consider the induced current density as a response to external electric field as a function of ω, and examine whether its limiting value for $\omega\to 0$ agrees with the one calculated in static regime. The current density induced by the external electric field $\boldsymbol{E}_{\text{extL}}(\boldsymbol{r}, \omega)$ is given in (5.160). If we rewrite the matrix element of $\boldsymbol{P}^{(\text{L})}_{\mu\nu}(\boldsymbol{r})$ by the help of $-i\hbar\langle\mu|\boldsymbol{J}^{(\text{L})}(\boldsymbol{r})|\nu\rangle = (E_\mu - E_\nu)\ \langle\mu|\boldsymbol{P}^{(\text{L})}(\boldsymbol{r})|\nu\rangle$, (3.25), the susceptibility χ_{JEL} is given as

$$\chi_{\text{JEL}}(\boldsymbol{r}, \boldsymbol{r}'.\omega) = -i\hbar \sum_\nu \frac{1}{E_{\nu 0}}[g_\nu(\omega)\boldsymbol{J}_{0\nu}(\boldsymbol{r})\boldsymbol{J}^{(\text{L})}_{\nu 0}(\boldsymbol{r}') - h_\nu(\omega)\boldsymbol{J}_{\nu 0}(\boldsymbol{r})\boldsymbol{J}^{(\text{L})}_{0\nu}(\boldsymbol{r}')] \tag{5.165}$$

where $E_{\nu 0} = E_\nu - E_0$. This is the result for microscopic response. For macroscopic use, we apply LWA. In the lowest order of LWA, we only need to replace the current density matrix element with the first order moment, i.e., electric dipole moment defined by (2.113). This leads us to the L component of the susceptibility as

$$\chi^{\zeta\zeta}_{\text{JEL}}(\omega) = -i\hbar \sum_\nu \frac{1}{E_{\nu 0}}[g_\nu(\omega)\bar{J}^\zeta_{0\nu}\bar{J}^\zeta_{\nu 0} - h_\nu(\omega)\bar{J}^\zeta_{\nu 0}\bar{J}^\zeta_{0\nu}] \tag{5.166}$$

where ζ-axis is parallel to $\boldsymbol{E}_{\text{extL}}$.

The matrix element $\bar{\boldsymbol{J}}_{\nu 0}$ is a vector representing the E1 transition dipole moment. Denoting its magnitude as $\bar{I}^{(0)}_{\nu 0}$ and the angle between its direction and the ζ-axis as θ_ν, we can write

$$\bar{J}^\zeta_{\nu 0} = \bar{I}^{(0)}_{\nu 0} \cos\theta_\nu\ . \tag{5.167}$$

Using this result, we have

$$\chi^{\zeta\zeta}_{\text{JEL}}(\omega) = i\hbar \sum_\nu \frac{|\bar{I}^{(0)}_{\nu 0}|^2 \cos^2\theta_\nu}{E_{\nu 0}}\{g_\nu(\omega) - h_\nu(\omega)\}\ , \tag{5.168}$$

$$= -2i\hbar^2\omega \sum_\nu \frac{|\bar{I}^{(0)}_{\nu 0}|^2 \cos^2\theta_\nu}{E_{\nu 0}[E^2_{\nu 0} - (\hbar\omega + i0^+)^2]}\ . \tag{5.169}$$

Multiplying $\boldsymbol{E}^{(\text{L})}(\omega)$ to this susceptibility, we have the ζ component of induced current density, which is also written as $-i\omega P$. Thus, the L component of the induced static electric polarization is, in the limit of $\omega \to 0$, given as

$$\bar{P}^\zeta = 2\hbar^2 \sum_\nu \frac{|\bar{I}^{(0)}_{\nu 0}|^2 \cos^2\theta_\nu}{E^3_{\nu 0}} E_{\text{extL}}(\omega = 0)\ , \tag{5.170}$$

$$= 2 \sum_\nu \frac{|\bar{P}^{(0)}_{\nu 0}|^2 \cos^2\theta_\nu}{E_{\nu 0}} E_{\text{extL}}(\omega = 0)\ , \tag{5.171}$$

which is a finite value. The last expression is in good coincidence with the static limit of (5.163).

If we write the T components of the vector $\bar{\boldsymbol{J}}_{0\nu}$ in terms of its polar and azimuthal angles (θ_ν, ϕ_ν), as

$$\bar{J}^{\xi}_{0\nu} = \bar{I}^{(0)}_{\nu 0} \sin\theta_\nu \cos\phi_\nu \ , \quad \bar{J}^{\eta}_{0\nu} = \bar{I}^{(0)}_{\nu 0} \sin\theta_\nu \sin\phi_\nu \ , \tag{5.172}$$

we have the T components of the static induced polarizations as

$$\bar{P}^{\xi} = 2\sum_\nu \frac{|\bar{P}^{(0)}_{\nu 0}|^2 \cos\theta_\nu \sin\theta_\nu \cos\phi_\nu}{E_{\nu 0}} E_{\mathrm{extL}}(\omega = 0) \ , \tag{5.173}$$

$$\bar{P}^{\eta} = 2\sum_\nu \frac{|\bar{P}^{(0)}_{\nu 0}|^2 \cos\theta_\nu \sin\theta_\nu \sin\phi_\nu}{E_{\nu 0}} E_{\mathrm{extL}}(\omega = 0) \ . \tag{5.174}$$

The angles (θ_ν, ϕ_ν) are determined by the symmetry and quantum transition $(\nu \leftrightarrow 0)$. In a high symmetry situation, where all the transitions are classified into T and L modes with respect to the external electric field, these T components of $\bar{\boldsymbol{P}}$ vanish.

As a conclusion, the current density in terms of χ_{em} describes the static limit of induced electric polarization correctly, so that it is not necessary to derive an extra expression for the induced electric polarization.

5.7.3 L and T Field Produced by a Moving Charge

An external charge density has been treated as a source of L electric field in the previous subsections. When the charge density is moving, however, it can induce T-, as well as L-field. Such a component is related with Cerenkov radiation and Smith–Purcell effect, as mentioned in Sect. 5.7.2. In this subsection, we calculate the T- and L-field produced by a charged particle moving with a constant velocity.

Let us consider a particle with a charge Q moving in the x-direction with a velocity v. Following Yamaguti et al. [6], we write the associated charge density as

$$\rho(\boldsymbol{r}, t) = Q\ \delta(x - vt)\ \delta(y)\ \delta(z) \ , \tag{5.175}$$

and the current density due to this moving particle as

$$\boldsymbol{J}(\boldsymbol{r}, t) = Qv\ \hat{x}\ \delta(x - vt)\ \delta(y)\ \delta(z) \ . \tag{5.176}$$

Here we use the definition of Fourier and its reverse transforms as

$$f(t) = \frac{1}{2\pi}\int \mathrm{d}\omega\ \mathrm{e}^{-i\omega t}\ \bar{f}(\omega) \ , \quad \bar{f}(\omega) = \int \mathrm{d}t\ \mathrm{e}^{i\omega t}\ f(t) \tag{5.177}$$

$$g(x) = \frac{1}{2\pi}\int \mathrm{d}k\ \mathrm{e}^{ikx}\ \bar{g}(k) \ , \quad \bar{g}(k) = \int \mathrm{d}x\ \mathrm{e}^{-ikx}\ g(x) \ . \tag{5.178}$$

Then, the ω Fourier components of ρ and $\boldsymbol{J}$ are

$$\rho(\boldsymbol{r}, \omega) = \frac{Q}{v}\, \mathrm{e}^{ik_x x}\ \delta(y)\,\delta(z)\,, \tag{5.179}$$

$$\boldsymbol{J}(\boldsymbol{r}, \omega) = Q\hat{x}\, \mathrm{e}^{ik_x x}\ \delta(y)\,\delta(z)\,, \tag{5.180}$$

where the wave number in the x-direction is defined as $k_x = \omega/v$. Obviously these definitions of ρ and $\boldsymbol{J}$ satisfy the continuity equation as seen from

$$\frac{\partial \rho}{\partial t} = -i\omega\rho \tag{5.181}$$

$$\nabla \cdot \boldsymbol{J} = i\frac{Q\omega}{v}\, \mathrm{e}^{ik_x x}\ \delta(y)\,\delta(z)\,. \tag{5.182}$$

The L field $\boldsymbol{E}_{\mathrm{L}}$ is the solution of $\nabla \cdot \boldsymbol{E}_{\mathrm{L}} = 4\pi\rho$. In the form of potential defined by $\boldsymbol{E}_{\mathrm{L}} = -\nabla\phi$, the solution is

$$\phi(\boldsymbol{r}, \omega) = \int \mathrm{d}\boldsymbol{r}'\, \frac{\rho(\boldsymbol{r}', \omega)}{|\boldsymbol{r} - \boldsymbol{r}'|}\,. \tag{5.183}$$

By the Fourier expansion of $1/|\boldsymbol{r} - \boldsymbol{r}'|$, we obtain

$$\begin{aligned}
\phi(\boldsymbol{r}, \omega) &= \frac{1}{8\pi^3}\int \mathrm{d}\boldsymbol{q} \int \mathrm{d}\boldsymbol{r}'\, \frac{4\pi}{q^2}\, \rho(\boldsymbol{r}', \omega) \exp[i\boldsymbol{q}\cdot(\boldsymbol{r} - \boldsymbol{r}')] && (5.184)\\
&= \frac{Q}{\pi v}\int \mathrm{d}\boldsymbol{q}\, \frac{\exp[i\boldsymbol{q}\cdot\boldsymbol{r}]}{|\boldsymbol{q}|^2}\delta(q_x - k_x) && (5.185)\\
&= \frac{Q}{\pi v}\int \mathrm{d}q_y \int \mathrm{d}q_z \frac{\exp(ik_x x + iq_y y + iq_z z)}{k_x^2 + q_y^2 + q_z^2} && (5.186)\\
&= \frac{Q}{v}\int \mathrm{d}q_y \frac{1}{\gamma_0}\exp(ik_x x + iq_y y - \gamma_0|z|)\,, && (5.187)
\end{aligned}$$

where $\gamma_0 = \sqrt{k_x^2 + q_y^2}$. In evaluating the third equation, we used Cauchy theorem. Thus we obtain

$$\boldsymbol{E}_{\mathrm{L}}(\boldsymbol{r}, \omega) = i\frac{Q\omega}{v^2}\int \mathrm{d}q_y \frac{1}{\gamma_0}\left(1, \frac{q_y}{k_x}, \pm i\frac{\gamma_0}{k_x}\right)\exp[ik_x x + iq_y y - \gamma_0|z|]\,. \tag{5.188}$$

The ± signs for the z-component mean that (−) sign for $z > 0$, and (+) for $z < 0$. This is the plane wave expansion in the (x, y) plane, which leads to the evanescent L-field in the $|z|$-direction with the decay constant dependent on (k_x, q_y).

The T field $\boldsymbol{E}_{\mathrm{T}}$ can be calculated as $\boldsymbol{E} - \boldsymbol{E}_{\mathrm{L}}$. The total electric field $\boldsymbol{E}$ is easily obtained from the equation for the vector potential $\boldsymbol{A}$ in Lorentz gauge

$$\left(\nabla^2 + \frac{\omega^2}{c^2}\right)\boldsymbol{A}(\boldsymbol{r}, \omega) = -\frac{4\pi}{c}\boldsymbol{J}(\boldsymbol{r}, \omega) \tag{5.189}$$

together with the relation between $\boldsymbol{A}$ and $\boldsymbol{E}$

$$\boldsymbol{E}(\boldsymbol{r}, \omega) = \frac{ic}{\omega}\left[\frac{\omega^2}{c^2} + \nabla\nabla\cdot\right]\boldsymbol{A}(\boldsymbol{r}, \omega) \,. \tag{5.190}$$

The solution of (5.189) is obtained via Fourier expansion as

$$\boldsymbol{A}(\boldsymbol{r}, \omega) = -\frac{Q}{c\pi}\hat{x}\int \mathrm{d}q_y \mathrm{d}q_z \frac{\exp(ik_x x + iq_y y + iq_z z)}{(\omega/c)^2 - k_x^2 - q_y^2 - q_z^2} \,. \tag{5.191}$$

Performing the q_z-integration via Cauchy theorem, we obtain

$$\boldsymbol{E}(\boldsymbol{r}, \omega) = i\frac{\omega Q}{v^2}\int \mathrm{d}q_y \left(1 - \beta^2, \frac{q_y}{k_x}, \pm i\frac{\Gamma_0}{k_x}\right)\frac{\exp(ik_x x + iq_y y - \Gamma_0|z|)}{\Gamma_0} \,, \tag{5.192}$$

where

$$\beta = v/c \,. \tag{5.193}$$

The decay constant of the total $\boldsymbol{E}$ in the z-direction is

$$\Gamma_0 = \sqrt{(\omega/v)^2 - (\omega/c)^2 + q_y^2} \,, \tag{5.194}$$

which is smaller than that of $\boldsymbol{E}_{\mathrm{T}}$

$$\gamma_0 = \sqrt{(\omega/v)^2 + q_y^2} \tag{5.195}$$

This leads to the expression of $\boldsymbol{E}_{\mathrm{T}}$ as

$$\begin{aligned}\boldsymbol{E}_{\mathrm{T}}(\mathbf{r}, \omega) = i\frac{\omega Q}{v^2}\mathrm{e}^{ik_x x}\int \mathrm{d}q_y \Bigg[&\left(-\beta^2 + 1, \frac{q_y}{k_x}, \pm i\frac{\Gamma_0}{k_x}\right)\frac{\exp(iq_y y - \Gamma_0|z|)}{\Gamma_0} \\ &- \left(1, \frac{q_y}{k_x}, \pm i\frac{\gamma_0}{k_x}\right)\frac{\exp(iq_y y - \gamma_0|z|)}{\gamma_0}\Bigg] \,.\end{aligned} \tag{5.196}$$

For $v \to 0$, E_{T} is smaller than E_{L} by the factor β^2.

5.8 Dimension of the Susceptibilities in SI and cgs Gauss Units

One of the tedious aspects of SI units system is the different dimensions of $\boldsymbol{E}$, $\boldsymbol{B}$, $\boldsymbol{D}$, $\boldsymbol{H}$, and hence, various susceptibilities. In writing the SI expressions of the formulas, especially in Sect. 3.1, we need to pay particular attention to this point.

In this subsection, we present some consideration on this problem. We denote the dimension of a physical quantity U as $[U]$, and those of length, time, electric charge, and energy as L, T, $[e]$, and $\mathscr{E}$, respectively.

From the Faraday law in SI units, we have $L^{-1}[\boldsymbol{E}] = T^{-1}[\boldsymbol{B}]$, so that

$$[\boldsymbol{E}] = LT^{-1}[\boldsymbol{B}] \,. \tag{5.197}$$

Similarly, from the decomposition of current density as $\boldsymbol{I} = \partial \boldsymbol{P}/\partial t + \nabla \times \boldsymbol{M}$, we have

$$[\boldsymbol{M}] = LT^{-1}[\boldsymbol{P}] \,. \tag{5.198}$$

The dimension of the matrix elements in the expressions of $\chi_{\rm eE}$, $\chi_{\rm eB}$, $\chi_{\rm mB}$, $\chi_{\rm mE}$ in Sect. 3.1 are

$$[\bar{J}_{\nu 0}] = [e]LT^{-1}, \quad \bar{Q}^{(\rm e2)}_{\nu 0} = [e]L^2T^{-1}, \quad [\bar{M}_{\nu 0}] = [e]L^2T^{-1} \,. \tag{5.199}$$

From the form of Coulomb potential, the square of electric charge has the dimension $[e^2] = \mathscr{E}L[\varepsilon_0]$. Using these results, we can evaluate the dimension of the susceptibilities. For example,

$$\begin{aligned}[\chi_{\rm eE}] &= [1/\omega^2 V][\bar{g}_\nu][\bar{J}^2_{\nu 0}] = T^2L^{-3}\mathscr{E}^{-1}[e^2]L^2T^{-2} \\ &= TL^{-3}\mathscr{E}^{-1}\mathscr{E}L[\varepsilon_0]L^2T^{-2} = [\varepsilon_0] \,.\end{aligned} \tag{5.200}$$

In the similar way, we obtain

$$[\chi_{\rm eB}] = [1/\omega V]\mathscr{E}^{-1}\mathscr{E}L[\varepsilon_0]L^3T^{-2} = [\varepsilon_0]LT^{-1} \,, \tag{5.201}$$

$$[\chi_{\rm mE}] = [1/\omega V]\mathscr{E}^{-1}\mathscr{E}L[\varepsilon_0]L^3T^{-2} = [\varepsilon_0]LT^{-1} \,, \tag{5.202}$$

$$[\chi_{\rm mB}] = [1/V]\mathscr{E}^{-1}\mathscr{E}L[\varepsilon_0]L^4T^{-2} = [\varepsilon_0]L^2T^{-2} \,. \tag{5.203}$$

This leads to

$$[\chi_{\rm eE}\boldsymbol{E}] = [\chi_{\rm eB}\boldsymbol{B}] = [\varepsilon_0][\boldsymbol{E}] \,, \tag{5.204}$$

$$[\mu_0\chi_{\rm mE}] = [\mu_0\varepsilon_0]LT^{-1} = L^{-1}T \,, \tag{5.205}$$

$$[\mu_0\chi_{\rm mB}] = [\mu_0\varepsilon_0]L^2T^{-2} = 1 \,, \tag{5.206}$$

which can be used to judge the correct combinations of different quantities from the dimensional viewpoint. For example, let us consider the case of rewriting the microscopic Ampère law into macroscopic form in SI units. Substituting $\boldsymbol{I} = \partial \boldsymbol{P}/\partial t + \nabla \times \boldsymbol{M}$ ($\boldsymbol{P} = \chi_{\rm eE}\boldsymbol{E} + \chi_{\rm eB}\boldsymbol{B}$, $\boldsymbol{M} = \chi_{\rm mE}\boldsymbol{E} + \chi_{\rm mB}\boldsymbol{B}$) into the microscopic Ampère law, we have

$$\frac{1}{\mu_0}\nabla \times (\boldsymbol{B} - \mu_0 \boldsymbol{M}) = \frac{\partial}{\partial t}(\varepsilon_0 \boldsymbol{E} + \boldsymbol{P}) \, , \tag{5.207}$$

where the dimension of

$$\mu_0 \boldsymbol{M} = \mu_0 \chi_{\mathrm{mE}} \boldsymbol{E} + \mu_0 \chi_{\mathrm{mB}} \boldsymbol{B} \tag{5.208}$$

is same as that of $\boldsymbol{B}$, and the dimension of

$$\boldsymbol{P} = \chi_{\mathrm{eE}} \boldsymbol{E} + \chi_{\mathrm{eB}} \boldsymbol{B} \tag{5.209}$$

is $[\varepsilon_0][\boldsymbol{E}]$, so that the combinations $\boldsymbol{B} - \mu_0 \boldsymbol{M}$ and $\varepsilon_0 \boldsymbol{E} + \boldsymbol{P}$ are seen to be dimensionally correct.

Contrary to the SI units system, we have much simpler relationship among the fields $\boldsymbol{E}$, $\boldsymbol{B}$, $\boldsymbol{D}$, $\boldsymbol{H}$ in the cgs Gauss units system. From the Faraday law, we have $[\boldsymbol{E}] = [\boldsymbol{B}]$ and from Ampère law $[\boldsymbol{P}] = [\boldsymbol{M}]$, so that all the fields have the same dimension, i.e., $[\boldsymbol{E}] = [\boldsymbol{B}] = [\boldsymbol{D}] = [\boldsymbol{H}] = [\boldsymbol{P}] = [\boldsymbol{M}]$, and all the linear susceptibilities $\chi_{\mathrm{eE}}, \chi_{\mathrm{eB}}, \chi_{\mathrm{mB}}, \chi_{\mathrm{mE}}$ and (ε, μ) are non-dimensional.

References

1. Cohen-Tannoudji, C., Dupont-Roc, J., Grynberg, G.: Photons and Atoms, Sect. IV.C. Wiley Interscience, New York (1989)
2. Drude, P.: Lehrbuch der Optik. S. Hirzel, Leipzig (1912); Born, M.: Optik J. Springer, Heidelberg (1933); Fedorov, F.I.: Opt. Spectrosc. **6**, 49 (1959); ibid. **6**, 237 (1959)
3. Lamb Jr., W.E.: Phys. Rev. **85**, 259 (1952); Yang, K.-H.: Ann. Phys. (N.Y.) **101**, 62 (1976); Lamb Jr., W.E., Schlichter, R.R., Scully, M.O.: Phys. Rev. A **36**, 2763 (1987); Del Sole, R., Girlanda, R.: Phys. Rev. B **48**, 11789 (1993); Scully, M.O., Zubairy, M.S.: Quantum Optics, Appendix 5A. Cambridge University Press, Cambridge (1997)
4. Cho, K.: Optical Response of Nanostructures: Microscopic Nonlocal Theory. Springer, Heidelberg (2003)
5. Landau, L.D., Lifshitz, E.M.: Electromagnetics of Continuous Media. Pergamon Press, Oxford (1960)
6. Smith, S.J., Purcell, E.M.: Phys. Rev. **92**, 1069 (1953); Yamaguti, S., Inoue, J., Haeberle, O., Ohtaka, K.: Phys. Rev. B **66**, 195202 (2002)

Index

K. Cho, *Reconstruction of Macroscopic Maxwell Equations*,
Springer Tracts in Modern Physics 237,
https://doi.org/10.1007/978-3-662-58424-8